The Meritocratic Consensus and Stratification in Higher Education

Zachary Bleemer and Jesse Rothstein *

August 2026

Abstract

American colleges and universities are highly stratified by pre-college academic achievement, family background, and institutional resources. We study the meritocratic consensus in American higher education: colleges that high-testing students (who are generally also from high-income families) attend spend dramatically more on instruction than do those that enroll lower-testing students. Stratification by test scores has been largely stable since the 1960s, but the stratification of instructional resources has risen sharply since 1970 at both private and public institutions. Non-academic admissions criteria like athletics, legacy, and affirmative action are second-order in determining the allocation of students to universities. Potential economic justifications for the positive association of instructional expenditures with student prior achievement – q-complementarity between achievement and resources, convex social returns to high human capital, and incentives to invest in learning prior to college – have little empirical support. Resource stratification across universities has not increased in the past decade, largely due to increased public funding of universities that enroll lower-testing students through financial aid programs like California's CalGrant, but stratification within institutions is now rising swiftly.
JEL Codes: I23, I24, N32, Z13

*Zachary Bleemer is Assistant Professor of Economics, Princeton University, Princeton, New Jersey. Jesse Rothstein is Professor of Public Policy and Economics and David Gardner Chair of Higher Education, University of California at Berkeley, Berkeley, California. Thanks to Sarah Turner for helpful comments and Dylan Black, Ivy Chen, and Shinrea Su for excellent research assistance.

1 Introduction

The modern American higher education ecosystem came into being in the period following World War II. Increased federal scientific research funding made some institutions much more research-intensive, with dramatic increases in revenues and scientific activity. Simultaneously, the college-going rate rose substantially, and states expanded and created a large number of teaching-focused public colleges and universities. An implicit question arose: Which students would attend which institutions?

Most states arrived quickly at the same answer: students with greater pre-college academic achievement (using traditional metrics like test scores and grade point averages) would be grouped together at better-funded and higher-quality colleges and universities. Students with weaker pre-college achievement would also be grouped, but at lower-funded teaching-focused institutions. We characterize this idea, taken for granted as the organizing principle of U.S. higher education, as the "meritocratic consensus." In this essay, we review the meritocratic consensus and consider potential justifications for it.

First, we demonstrate that higher education is highly stratified. There are substantial disparities across institutions in expenditures per student, and high-testing students – who tend to be from families with high socioeconomic status – are concentrated at the higher-spending institutions.

Second, we provide a short history of how this has evolved over the last 125 years. As late as 1960, most US higher education institutions remained effectively open-access and variation in per-student expenditures was relatively low. Since 1960, stratification has steadily intensified. In 1969, the top quarter of institutions by selectivity (enrolling the highest-testing students) had per-student instructional expenditures about double those at the institutions in the bottom quarter of test scores. By 1993, this had risen to about 2.5 times, and by 2019 to three times.

Third, we discuss some exceptions to the "meritocratic" aspect of the consensus. While high-resource institutions conduct selective admissions that select heavily on academic achievement, they also consider non-academic criteria like athletic talent, potential donations, and various dimensions of student diversity. These objectives are often described as dimensions of 'merit,' but they tend to cut against the admission of the most academically

accomplished students. We argue that relatively few students are shifted toward higher-expenditure institutions by policies like athletic, donor, and legacy preferences. While race-based affirmative action was once a larger factor, it never accounted for more than a few percent of university admissions, and was banned outright by the 2023 US Supreme Court decision in *Students for Fair Admissions (SFFA) v. Harvard* (600 US 181 [2023]).

With this background in place, we discuss three potential economic justifications for the meritocratic consensus. There are clear reasons why one might want to group students together by prior achievement, so we focus on the strong correlation between student achievement and institutional resources. (As we make clear, this is not simply a market outcome, and it appears in the large public sector as well as the smaller private sector.) We consider several reasons that it could be socially desirable to spend more money on the instruction of high-achieving students than on that of students with worse high school achievement. One is that pre-college academic preparation and college resources may be q-complements in educational production: that is, higher-achieving students may benefit relatively more from smaller classes and from access to cutting-edge research. A second is that there may be large social value in providing the highest-quality training to students who are most likely to become scientists, successful entrepreneurs, creative geniuses, and others who can expand the boundaries of human knowledge and creativity. Third, meritocratic selection may create useful incentives for high school students to invest in skill development. While all three models are plausible, we find that there is little or no empirical evidence in support of any of them, and some evidence pointing against them.

In the conclusion, we discuss recent developments and the future of the meritocratic consensus. After several decades in which the association between selectivity and per-student university expenditures steadily strengthened, stratification across universities has leveled out since 2010. During this same period, however, there has been a rise in stratification *inside* colleges and universities, particularly at public institutions. These institutions have increasingly created "honors colleges" and other programs designed to separate and provide additional resources to their highest-achieving students. In addition, more than 40 percent of lucrative majors at public research universities – like engineering and business – have adopted supplemental achievement-based admissions policies over the past 50 years. Because these majors tend to enroll the highest-achieving students from the

Table 1: Characteristics of Four-Year Colleges by Tier

Tier	Share of Enr. (%)	Admit Rate (%)	Avg. SAT	Mean Parental Inc. Rank	Expenditures per Student ($) Inst.	Res.	Admin.	Students per Faculty	Grad. Rate (%)
Ivy League	0.9	7.4	1502	80.5	68,800	38,300	81,700	10.0	96.1
Other Top-50 Priv.	3.3	15.2	1466	80.0	61,200	32,300	41,800	10.4	92.0
State Flagship	15.2	64.9	1275	71.7	22,000	15,000	14,800	15.9	78.7
Other Public R1	15.2	63.9	1232	68.4	19,800	11,700	14,400	19.5	72.5
Less-Sel. Private	27.4	67.1	1168	68.5	16,400	1,400	19,800	22.9	64.6
Less-Sel. Public	36.9	73.1	1105	62.3	12,000	1,700	9,700	23.6	54.6
For-Profit	1.2	70.6	-	44.3	7,800	0	15,900	141.5	40.3

Note: See the text for college tier definitions, which partition four-year institutions. For each tier, the table reports the share of total full-time first-year enrollment in four-year institutions and several enrollment-weighted means: the undergraduate admissions rate; SAT score (except for for-profits, which generally do not report average scores); parental income rank; per-student instructional, research, and other administrative expenditures; student-to-faculty ratio; and six-year graduation rate. See Appendix A for variable definitions. Six-year graduation rate is measured for the 2017 first-year cohort; parental income rank is measured for the 1989-1991 birth cohort by students' modal enrollment institution; and all other statistics are measured in 2019.
Source: IPEDS and Chetty et al. (2020).

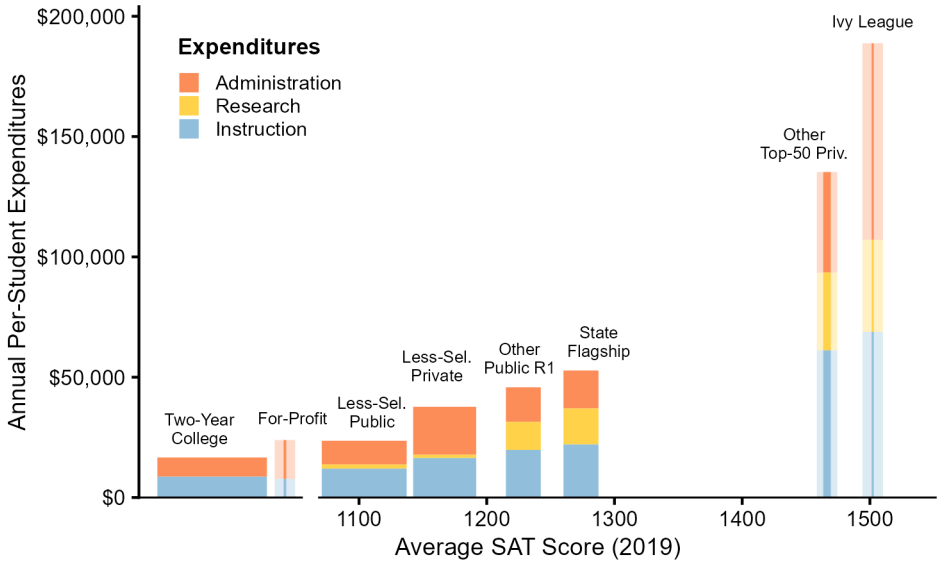
most advantaged backgrounds, increasing stratification between college majors parallels the across-institution stratification that we document.

2 The Heterogeneity of Higher Education

About three-fifths of students who graduate high school enroll directly in college the following fall. While not all of those students ultimately graduate, most get some kind of credential: More than half of Americans aged 25-29 have at least a two-year Associate's degree, while 40 percent have a four-year Bachelor's degree or more (NCES, 2021).

Students receive these credentials from institutions that differ widely. Table 1 shows summary statistics for seven tiers of four-year colleges and universities: the eight Ivy League institutions, the other 42 Top-50 private institutions (with a 25th percentile SAT score in 2019 of 1340 or higher), the 50 state flagship public universities, the 56 non-flagship "R1" public universities classified as having very high research activity (The Carnegie Foundation for the Advancement of Teaching, 2000), other four-year private

Figure 1: Universities' 2019 Expenditures by Tier



Note: This figure shows average 2019 expenditures per full-time equivalent student (counting each part-time student as one-third of an FTE) by tier of institution and type of expenditure. See Appendix A for variable definitions. Width of solid bars are proportional to FTE enrollment in the tier; semi-transparent bars are widened relative to that for visibility. Average SAT scores are not available for two-year and for-profit colleges, where many students do not take admissions tests; these tiers are shown here separately for comparison. All dollar amounts are CPI-adjusted to 2025. Source: IPEDS.

and public universities, and for-profit institutions.[1] The top-50 private universities with the highest average SAT scores together enroll only about 4 percent of four-year college students. In contrast, public R1 universities, including "flagships" like the University of California, Berkeley, the University of Wisconsin, and the University of Virginia, as well as others like Utah State University and the University of Pittsburgh, enroll about 30 percent of students. About 64 percent of students attend less-selective public or private institutions. The for-profit share of all first-time first-year college enrollment (including two-year schools not shown in Table 1) was 2.6 percent in 2019, though it was much higher, nearly 10 percent, in the late 2000s.

These tiers of higher education institutions also differ in the educational experiences

[1] See Appendix A for definitions and details on data construction.

they offer. The more selective institutions have much higher expenditures, more students per faculty member, and higher graduation rates.

Figure 1 shows average instructional, research, and administrative expenditures per pupil at institutions in each tier. These estimates follow university reporting guidelines and are admittedly imperfect. For example, faculty salaries are counted as part of "instruction," but of course some of that amount is payment for doing research.[2] Similarly, libraries and student services may contribute to instruction but are included here as administrative expenditures. But such issues will not overshadow the enormous differences shown here: the most selective schools spend many times as much as less-selective and two-year colleges, both overall and in dedicated instructional expenditures.

This is the "meritocratic consensus": selective schools have much higher spending and revenues per student. In 2019, universities' per-student instructional expenditures were 29 percent higher per 100 average SAT points of their enrolled students. The pattern is similar among public and private universities; at public universities, for example, per-student instructional expenditures were 24 percent higher for each 100 average SAT points. Across states, the gradient of per-student instructional expenditures with respect to SAT scores ranges from 10 percent in Illinois and Mississippi to 50 percent in West Virginia and Washington.[3]

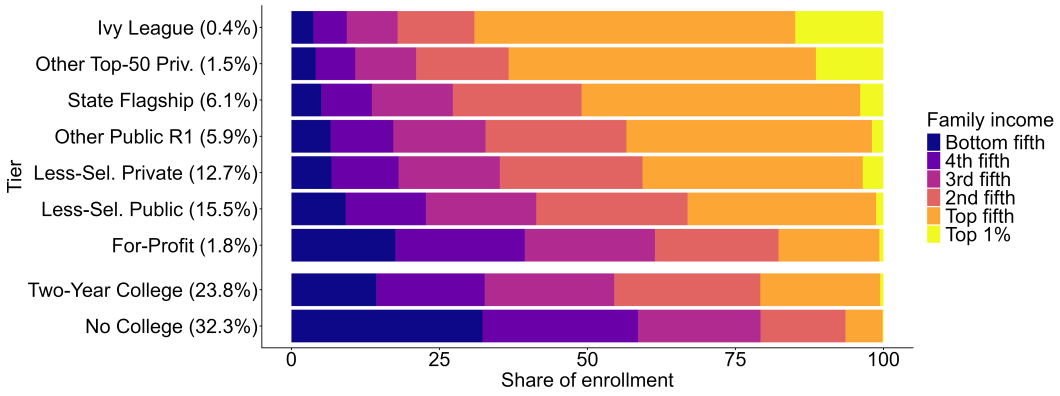
Government educational appropriations – including both direct appropriations, which are largely for research purposes, and financial aid programs (like Pell and state grant programs) – are roughly evenly distributed across more- and less-selective institutions. The larger expenditures at selective institutions are financed by substantially higher net tuition funding (gross tuition payments minus gross financial aid), private funding (including alumni donations, affiliated groups like booster clubs, and endowment returns), and government research contracts (see Figure A-1).

The institutions that enroll higher-testing students also tend to enroll students from higher socioeconomic backgrounds, as shown in Table 1, which shows average parental income ranks for students in each tier. Figure 2 provides more detail, dividing students into

[2]Research time may contribute to instruction if it is a compensating amenity for teachers or if it is complementary with teaching.

[3]See Figures A-3, A-4, and A-5.

Figure 2: College Enrollment by Tier and Family Income



Note: This figure shows the parental income distribution of students enrolled at colleges in each of the tiers used in Table 1. Tiers are augmented by additional categories representing two-year colleges (including public, private, and for-profit institutions) and students who do not go to college after high school (including those who do not graduate high school). Tier labels include (in parentheses) the ratio of total US-resident first-time first-year enrollment in 2000 to the number of high school graduates in 2000 (from NCES, 2021); the 'no college' share is the residual. Parental income is measured among the 1980-1982 birth cohorts, assigning children to their modal institution and parents to quintiles (splitting out the top 1% from the top quintile). See Appendix A for variable definitions. Source: IPEDS and Chetty et al. (2020).

five quintiles by family income and separating the top 1% from the rest of the top quintile. About two-thirds of students at Ivy League and other top-50 private universities – and more than half of students at state flagships – come from families in the top income quintile. The income distributions of students at less-selective public and two-year colleges are closer to those of the country as a whole. More than half of students who do not go to college come from families in the bottom 40 percent of the income distribution.[4]

Many researchers have investigated the relationship between university selectivity and value-added (Dale and Krueger, 2002; Hoekstra, 2009; Cohodes and Goodman, 2014; Bleemer, 2022). Results have been mixed, without clear evidence of a strong relationship. A smaller literature examines the relationship between university expenditures and value-added to students. This is clearer, showing a consistent positive association (Chetty, Deming and Friedman, 2026; Bleemer, 2021; Mountjoy and Hickman, 2021).[5] Deming

[4]This figure and much of our discussion below exclude non-resident international students at US universities, who accounted for about 3 percent of first-time full-time undergraduates in 2019.

[5]Of particular note is Figure 5 in Mountjoy and Hickman (2021), which reproduces the Dale and Krueger

and Walters (2014) provide suggestive evidence that increasing university expenditures improves students' educational outcomes, suggesting that the correlation is causal and paralleling evidence that increases in K-12 education spending raise student test scores and improve longer-run outcomes (Jackson and Mackevicius, 2024; Lafortune, Rothstein and Schanzenbach, 2018; Rothstein and Schanzenbach, 2022). If this result is correct, the patterns seen here suggest a pattern of self-reinforcing economic inequality across generations: high-income families provide greater support for their children to achieve greater traditional pre-college academic accomplishment ('merit'), leading those children to enroll at the highest-resource institutions, and thereby conveying an important relative advantage in lifetime income.

3 The Rise of the Meritocratic Consensus

Where did the meritocratic consensus come from? This section traces its slow evolution in American higher education over the past 125 years.[6]

3.1 The Early Landscape: 1900-1945

At the beginning of the twentieth century, between 5 and 10 percent of young people went to college (NCES, 2021). Three-quarters of college students were at private institutions, with 10 percent attending the colleges that would soon become known as the 'Ivy League' (Goldin and Katz, 1999; Bleemer and Quincy, 2025a). College admission – even at the Ivy League – was generally available to any student who had completed the requisite number of subject-specific high school units. The influential 1892 Committee of Ten – which included Harvard University president Charles Eliot and US Commissioner of Education

(2002) finding of no relationship between value-added and selectivity but finds a positive relationship with institutional resources. Several studies show that decreased public university revenues lead to worse quality and lower graduation rates (Kane, Orszag and Apostolov, 2005; Bound and Turner, 2007; Bound, Lovenheim and Turner, 2010).

[6]We generally rely on enrolled students' average SAT scores (or ACT scores converted to the SAT scale) as our measure of institutional selectivity due to limited availability of high school grade point average data. Figure A-7 shows that the SAT rankings of institutions have been largely stable since at least the mid-1980s, and we sometimes use the post-1980 SAT as a proxy for hard-to-measure earlier ordinal selectivity.

William Harris among others – recommended that "any successful graduate of a good secondary school should be free to present himself at the gates of the college or scientific school of his choice" (Committee of Ten, 1892). Public universities in many states directly accredited their states' high school curricula as a way of enforcing minimum pre-college academic preparation.

Some universities additionally "administer[ed] admission examinations to weed out the clearly incompetent," but this was used only to set a floor, not to conduct competitive admissions. At Harvard, for example, "the size of the Freshman Class was determined ... by the number of qualified applicants" (Bender, 1952). Neither high school grades nor standardized tests were regularly considered in public or private university admissions (Hinckley, 1941), and public universities were, in the 1907 words of University of Illinois president Andrew Draper, oriented toward "put[ting] away all exclusiveness and dedicat[ing] themselves to universal public service" (Veysey, 1965). This openness was also manifested in prices: average net tuition and fees at public and private universities in 1923 were $1,390 and $3,150, respectively, in 2025 dollars (Bleemer and Quincy, 2025a).

Universities maintained academic standards by evaluating classroom performance rather than standardized test scores; many low-performing students were dismissed. This often resulted in graduation rates below 50 percent at even the most prestigious institutions like Harvard University and Brown University (Thelin, 2010).

There was already substantial between-university variation in annual per-student instructional expenditures in the early 20th century, but far less than in subsequent years. Table 2 shows how cross-university variation in per-student institutional expenditures has risen over the past 100 years.[7] Panel A shows average instructional expenditures per student (converted to 2025 dollars), while panel B shows the coefficient of variation (that is, the standard deviation divided by the mean). At the 111 public institutions that awarded Bachelor's degrees in 1929, average per-student expenditures were about $4,400 and the standard deviation of per-student instructional expenditures was about 38 percent of that, or $1,700 (in 2025 dollars). Flagship public universities spent about 21 percent more per student than non-flagships in the same state, while the soon-to-be-named "Ivy League"

[7]We use a smaller number of categories than in earlier tabulations to permit comparability over time.

Table 2: Distribution of Instructional Expenditures Per Student Over Time

	1929	1969	1984	1996	2008	2019
Panel A: Average Instructional Expenditures per Student ($2025)						
All Four-Year Institutions	4,100	8,500	10,700	13,400	14,700	16,800
All Public Four-Year Institutions	4,400	8,100	10,800	12,300	14,100	15,800
Ivy League	8,300	23,100	29,500	44,200	77,200	75,000
State Flagship	5,000	10,800	12,900	15,800	20,200	21,900
Public Non-Flagship	3,800	7,700	10,300	11,500	12,600	14,200
Ivy vs. Flagship Gap (%)	68	115	128	181	281	242
Flagship vs. Non-Flagship Gap (%)	21	35	35	43	54	50
Panel B: Coefficient of Variation of Per-Student Instructional Expenditures (%)						
Public Institutions	38	30	49	55	50	49
All Four-Year Institutions	53	42	50	59	60	60

Note: Means and standard deviations are weighted by enrollment. All series are winsorized at 5 percent to limit the contribution of any single university. Enrollment is measured as total enrollment in 1929 and FTE enrollment in other years (counting part-time enrollees as 1/3). The flagship/non-flagship gap is measured by year-specific OLS regressions of log expenditures on a flagship indicator, including state fixed effects to isolate within-state comparisons. See Appendix A for variable definitions.
Source: Biennial Survey of Education, 1928-1930 (Cooper, 1931), HEGIS, and IPEDS.

spent 68 more than the flagship publics.[8]

3.2 Growing Enrollment and Selectivity: 1945-1990

The share of young men who earned college degrees approximately doubled (from 5 to 10 percent) between 1905 and 1940; it then doubled again in the next decade (Goldin, Katz and Kuziemko, 2006; Stanley, 2003). This shock in student demand forced supply-constrained universities to implement novel rationing strategies, including new admissions reviews that often relied on standardized test scores (Karabel, 2005; Lemann, 2000). States also opened or expanded many public colleges – or new branches of existing colleges – to relieve the pressure on the flagship institutions.

[8]Table A-1 shows that gaps in total per-student revenues were similar but slightly smaller in magnitude.

Rutgers, The State University of New Jersey, can serve as a paradigmatic example. Rutgers was an early adopter of standardized testing. In 1939, any student who satisfactorily completed the curriculum of any high school in New Jersey (or many schools in neighboring states) *or* achieved a minimum standardized test score was admitted to Rutgers' New Brunswick campus (Meder, 1939). In 1946, Rutgers implemented a new admissions policy following "unprecedented demand for admission": "other things being equal, preference is given to men having the best high school and aptitude test records" (Clothier, 1946). Qualified but lower-performing applicants were offered admission to evening courses at the university, or eventually to Rutgers's new campuses at Newark and Camden (created in 1946 and 1950, respectively).

While these new 'meritocratic' admissions policies were often intended as temporary responses to a short-run demand shock, the continued growth in college demand motivated their permanent codification. One popular blueprint was the 1960 California Master Plan for Higher Education. The Master Plan laid out a three-tiered higher education system – the research universities of the University of California; the teaching-oriented, bachelor's-only California State University campuses; and two-year community colleges – and suggested using applicants' pre-college academic achievement to track students between the segments. The admissions criteria were valued for their predictive "validity" in identifying students who were likely to successfully complete their studies (Spindt et al., 1961).[9] Each tier massively expanded mid-century; the University of California campuses at Davis, Irvine, Riverside, San Diego, San Francisco, Santa Barbara, and Santa Cruz were all established between 1954 and 1965, in some cases by converting existing institutions, and eight new California State University campuses were opened.The authors of the Master Plan nevertheless recognized that rising demand among qualified students for certain institutions – especially UC Berkeley and UCLA – would soon outstrip capacity, and suggested two remedies: (1) campus-level selective admissions policies designed such that only "the best students should be granted their first choice," which were implemented for the first time in California public universities in 1974 (Karabel, 1989), and (2) divert-

[9]Three studies on the predictive validity of College Board standardized tests were published in 1960. By 1970, over 30 scholarly publications per year were evaluating the SAT's 'predictive validity' (as indexed by Google Scholar); by 2010, this number had risen above 150 (for example, Westrick et al., 2019).

ing demand to other campuses, via "increased prestige of the junior colleges" and "more personal instruction, a richer student life, and superior housing and parking facilities" at the less-oversubscribed University of California and California State University campuses (Coons et al., 1960). The latter has proven difficult to achieve.

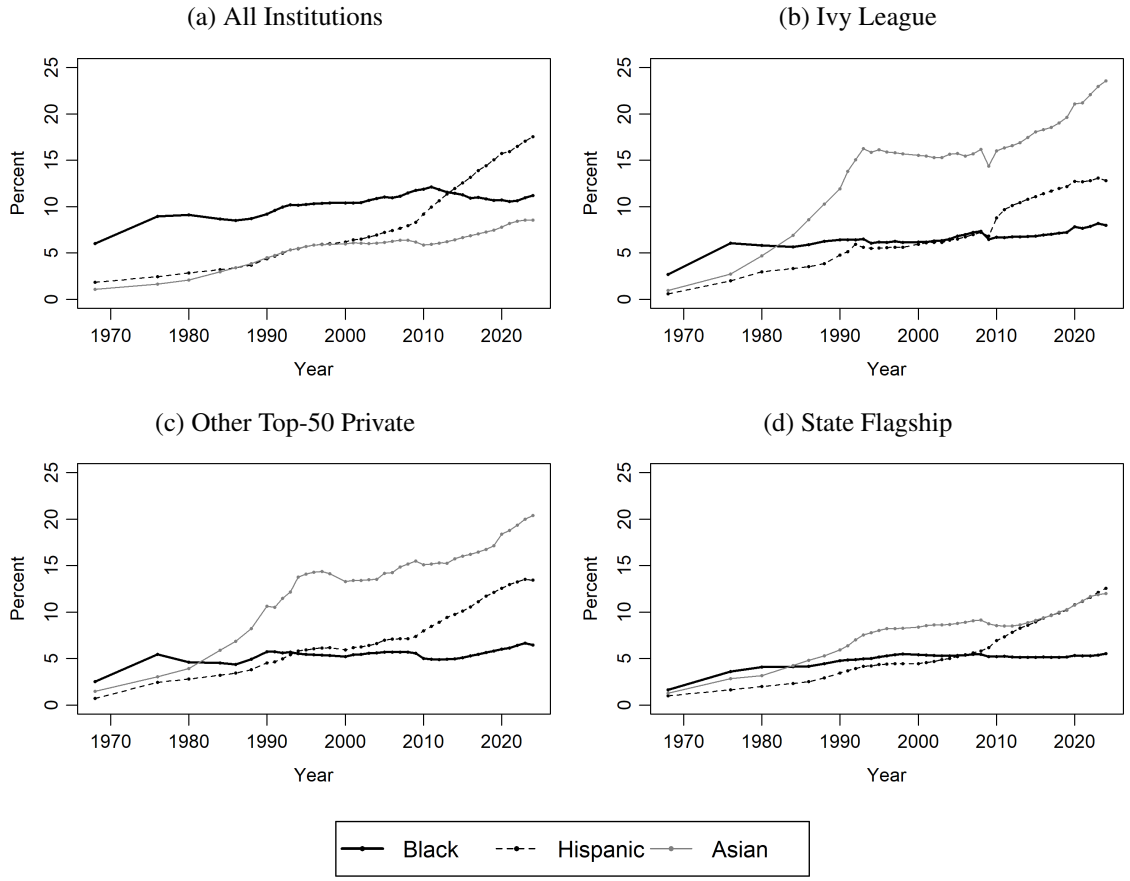
Circa 1940, fewer than 5 percent of higher education institutions relied upon tests for admission, but this rose steadily over the 50 years following 1940. Private institutions led the way, in part because their greater geographic market integration made standardized measures of pre-college academic achievement more valuable to evaluate students from distant and unfamiliar high schools (in this journal Hoxby, 2009). By 1962, over 50 percent of private universities, but only 28 percent of public institutions required tests in admissions. By the early 1980s, almost all private colleges and universities required test scores, and public institutions had caught up by the end of that decade (Figure A-6).[10]

The use of standardized tests in admissions, coupled with strong student preferences that are correlated with university resources, results in strong assortative matching between high-resourced universities and high-scoring students. Students' preferences could arise either because students have direct preferences for high instructional spending or because they prefer other characteristics that are correlated with this, perhaps simply academically-strong (or high-family-income) peers (Hendricks, Herrington and Schoellman, 2021; Rothstein, 2006). Since 1970, deviations from that assortative matching "overwhelmingly result from choices made by students and their families, not choices made by college admissions offices" (Dillon and Smith, 2017).

One factor that partially offset the stratifying effect of test-based admissions was growing concern about diversity. Both public and private universities broadened their scopes in the 1960s and 1970s, with 67 institutions admitting their first female students between 1960 and 1975 (Truffa and Wong, 2025) and many institutions admitting non-negligible numbers of Black undergraduates for the first time. However, schools that wanted to admit more Black students but that relied on standardized tests for admissions faced a problem: Due to the legacy of racial discrimination and segregation, very few Black students met

[10]While several studies have emphasized universities' race to enroll high-testing students as being partly motivated by the skewed incentives of public reputational rankings (Espeland and Sauder, 2007; Bleemer et al., 2023), this transition notably predates such rankings: US News and World report released its first reputational ranking of universities in 1983.

Figure 3: Undergraduate Racial Composition by Tier Over Time



Note: Share of full-time undergraduate students who are Black, Hispanic, or Asian at all four-year universities and by university tier. Tables A-3 and A-4 identify the top-50 private and state flagship universities. 'All four-year universities' includes some two-year institutions in 1968 (namely, those that do not have 'junior' or 'community' in their name) due to data limitations. Source: McVay (1969), HEGIS, and IPEDS.

selective colleges' usual admissions bar. Colleges began to implement affirmative action programs, whereby Black students were admitted despite not having the pre-college academic achievement that would have been required of a white applicant. Figure 3 shows that the share of Black students at top-50 private universities more than doubled between 1968 and 1976, with a similar proportional gain (from a lower base) at flagship public uni-

versities.[11] Affirmative action also provides somewhat smaller admissions preferences to Hispanic students (Bleemer, 2022), and their share at selective colleges also grew rapidly between 1968 and 1976 (though this is somewhat confounded by their growing population share). Figure 3 shows that Asian students were already over-represented at selective universities in 1968 (when the Asian population share was 0.8 percent) and are substantially overrepresented at those schools today relative to the US population. Affirmative action programs never grew proportionally larger than in this 1970s period – indeed, the Black enrollment share at highly-selective private schools hardly changed between 1976 and 2019 – but were highly controversial. We discuss such programs further below.

3.3 Growing Resource Gaps: Since 1970

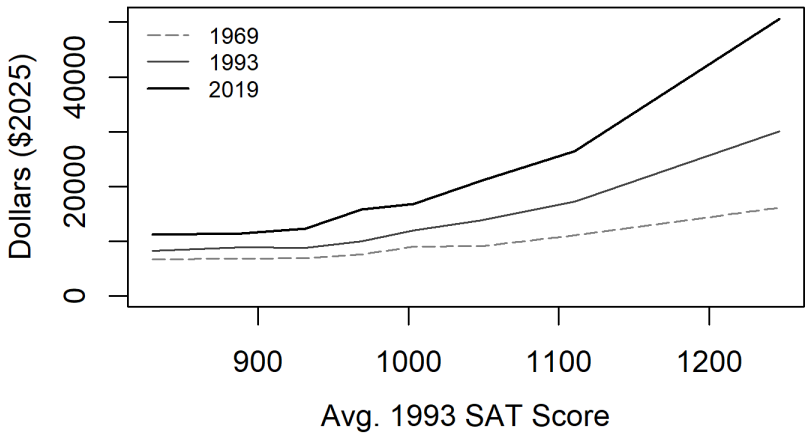
The cross-university test score distribution had largely stabilized by 1970: for example, the mean test score gap between the bottom and top deciles across universities has remained between 2 and 2.5 national standard deviations (Figure A-7). The share of elite college students from high-income families declined between 1940 and 1980, but this had reverted by 2000 (Abramitzky et al., 2024). See also Bleemer and Quincy (2025a).

But long-standing funding gaps between universities grew sharply in the last half-century. The coefficient of variation (again, the standard deviation divided by the mean) of public universities' per-student instructional expenditures rose from 0.30 in 1969 to over 0.5 in the mid-1990s (as shown earlier in Table 2). Flagship public universities were spending about 35 percent more on instruction than non-flagships in 1969; by 2008 the gap had grown to 54 percent. As one illustration of what this meant, consider salaries at California's public universities. In 1960, the year of the California Master Plan, the gap in mean wages between full professors in the higher-testing University of California system and the lower-testing California State University system was 23 percent (CCHE, 1966). By 2023, the gap had grown to 124 percent, with CSU wages closely tracking inflation and UC wages growing at a much faster pace (Freedman, 2023; University of California, 2023).

Figure 4 shows average instructional expenditures per student by institutional average

[11] See Table A-2 for numerical values.

Figure 4: Instructional Expenditures Per Student by Institutions' Average 1993 SAT Score



Note: Average freshman-enrollment-weighted instructional expenditures per FTE student by eight quantile bins of institutions' 1993 average SAT score, among the 559 four-year institutions with reported 1993 test scores and 1969, 1993, and 2019 enrollments and instructional expenditures. See Appendix A for variable definitions. Source: HEGIS, IPEDS, and the College Board Annual Survey of Colleges.

SAT score in 1969, 1993, and 2019. To focus on changes in resources at the institution level, institutions' SAT scores are held fixed here at their 1993 level.[12] In 1969, per-student instructional expenditures at the four-year institutions in the top testing quartile were already just over double those of institutions in the bottom quartile. This national gap widened over time, to almost 150 percent higher in 1993 and 200 percent higher in 2019. Stratification in instructional expenditures has risen more in the private university sector (resulting in a 310 percent expenditure gap between the top and bottom testing terciles in 2019) than the public sector (where the gap is 130 percent), and has been mirrored by rising stratification in both research and administration expenditures. These gaps grew simultaneously with substantial growth in the demand for high-testing institutions (Bound, Hershbein and Long, 2009) and in those institutions' observational wage value-added (Bleemer and Quincy, 2025a).

An examination of the revenue changes that financed the growth in the expenditure

[12]The Appendix includes a number of figures (A-7 to A-11) that offer robustness checks and additional perspectives on the issues discussed here.

gap between top- and bottom-test-quartile institutions from 1993 to 2019 reveals evidence of the causes of that growth. The per-student instructional expenditure gap grew by approximately $10,000 over this period. The high-testing institutions saw relative increases in net tuition revenues (by $7,900 in the top-quartile relative to bottom-quartile schools), while relative per-student government appropriations actually declined (by $2,100). This was a period of substantial state disinvestment from public higher education in terms of states' block funding allocations (Ehrenberg, 2012), and higher-testing universities were relatively more successful at privatizing and raising greater replacement revenue from their students. The remainder of the gap is largely explained by the portions of increased revenue from private gifts and grants (at least $1,200) and government research contracts ($7,800) that subsidize the universities' instructional activities, suggesting that rising federal scientific research investment and philanthropy at high-testing institutions has meaningfully spilled over into those schools' instructional expenditures.

The stratification of instructional resources by standardized test score would look even more extreme if two-year colleges were included in these trends. Community colleges have always had lower average instructional expenditures than their states' non-flagship public universities, by about 40 percent since at least the 1970s, and community colleges' wage value-added is considerably lower (Mountjoy, 2022). Two-year college enrollment grew from less than 10 percent of four-year enrollment in 1930 – at junior colleges in over 400 cities (Bleemer and Quincy, 2025b) – to about half of four-year enrollment in the 2010s. A small share of high-demand community college programs implement 'meritocratic' admissions policies (Grosz, 2021), but their negligible tuition costs and open-access admissions policies generally make them broadly available, which has resulted in increasingly disproportionate enrollment among lower-income and lower-testing students (Kane and Rouse, 1999; Bleemer and Quincy, 2025a).[13]

In sum, these trends reveal that over the past century, American higher education has become dramatically more stratified by student achievement, with increasing relative expenditures at the institutions that high-achieving students attend. This dual trend is what

[13] Of course, our noting that community colleges have lower value-added than four-year colleges should not be taken to suggest that they are not worthwhile for students unable or unprepared to attend a four-year college (Kane and Rouse, 1999).

we call the "meritocratic consensus."

4 Deviations from "Meritocratic" Admissions

While we've told a story of rising stratification of students by academic achievement and growing funding disparities across colleges and universities, there are some exceptions to these trends. Some highly visible admissions priorities – for diversity (as discussed above), athletes, or children of alumni and donors – cut against the increasing selectivity on student academic backgrounds. Universities often finesse these admission choices by redefining "merit" to include them, but for our purposes they are importantly distinct. However, as we argue below, none of these deviations from academic understandings of merit have been quantitatively large enough to significantly alter our understanding of admissions goals, at least until recently.

4.1 Athletics and Other Special Skills

Many selective universities set aside admissions slots for students with special non-academic skills that are valued for their contributions to campus life. The most prominent such example is collegiate athletics. Athletes typically have poorer pre-college academic achievement than their non-athlete peers at selective universities. For example, the median SAT score of 2014 freshman athletes at UC Berkeley was at the 10th percentile of the non-athlete SAT distribution (as calculated by the authors, see Bleemer, 2018).

At a few small colleges, athletes are a substantial share of the student body – 36 percent at Bowdoin and 47 percent at Claremont McKenna in 2023, for example (U.S. Department of Education, 2025). However, the athlete share of selective institutional enrollment as a whole is much smaller. Only about 7 percent of college students were intercollegiate athletes in 2023, and only 2 percent were NCAA Division I athletes eligible for preferential tuition pricing and other benefits. Fewer than 10 percent of students at research-intensive R1 public universities or at top-50 private universities are athletes. As a result, athletics contributes very little to the overall allocation of students to US universities.

Some universities also set aside slots for other special talents, such as music. There

are many fewer of these slots, and academic standards are not lowered as much (if at all) for them. For example, about 20 freshmen and sophomores (0.3 percent) are members of UC Berkeley's University Symphony Orchestra each year, but their test scores were at the 74th percentile of the non-musician SAT distribution.

4.2 University Community and Finances

Over 80 percent of the top-50 private institutions provide admissions advantages to the relatives of alumni, donors, and faculty (Hurwitz, 2011), though these shares have declined in recent years. The justifications offered for these policies are as benefits for institutional stakeholders and incentives for financial support for the institution (Castilla and Poskanzer, 2022).

At Harvard in 2014, children of alumni, donors, and faculty were 5.2, 6.9, and 4.1 times more likely, respectively, to be admitted than other academically-similar applicants (Arcidiacono, Kinsler and Ransom, 2022). Fully 14 percent of admitted students were the children of alumni (Arcidiacono, 2018). This vastly overstates the size of the alumni preference, however, as "legacy" applicants also have relatively strong academic credentials. Chetty, Deming and Friedman (2026) report that, due to preferences for the children of alumni, an unnamed Ivy League institution enrolls about 3 percentage points more students from the top 1 percent of the parental income distribution than it would have if admissions were conducted strictly using standardized test scores.

Fewer than 10 percent of enrollment-weighted research-intensive R1 public universities employ such preferences (Hurwitz, 2011). As a result, legacies are relatively limited in their overall impact except at a small number of mostly-private institutions.

4.3 Affirmative Action

Since the 1960s, many selective universities have provided admissions advantages to students with characteristics like race or class that would otherwise be underrepresented relative to the local or national population. Such policies are most commonly justified by the claim that traditional metrics understate these students' "merit" as they do not take into

account the students' poorer pre-college academic opportunity and circumstances (for example, Walton, Spencer and Erman, 2013).

The most prominent example of diversity-based admissions preferences is race-based affirmative action. Standard measures of pre-college academic achievement differ sharply by race (Grodsky, Warren and Felts, 2008). If every Black high school graduate with at least a 1400 SAT score were to attend one of the 50 American colleges and universities where at least 75 percent of enrollees had a 1400 SAT in 2023, those colleges would be just 3 percent Black, less than one-fifth of the 16 percent Black share of high school graduates (Lemann, 2024). In fact, 8 percent of those first-year students at those colleges were Black. Among universities that rejected at least 85 (50) percent of their applicants in 2023, 76 (47) percent provided race-based preferences in 2023 (Reber, Nagashima and Goodman, 2023). Preferences had substantial effects on Black and Hispanic students' admissions chances: One study found that in the 1990s, applicants from underrepresented minorities to UC Berkeley and UCLA were twice as likely to be admitted as other academically-comparable applicants (Bleemer, 2023).

Ten states (including California) banned race-based preferences between 1997 and 2020, and the Supreme Court banned them nationwide in 2023 (*Students for Fair Admissions v. Harvard*, 600 US 181 [2023]). As a result, colleges no longer consider race explicitly in admissions (though they may consider other factors that correlate with it).

Some universities have also used affirmative action to manage their gender distributions. Female academic performance generally outpaces that of males in high school (Reeves, 2022). As a result, gender-blind admissions would typically yield a class that is much more than 50 percent female at many highly-selective colleges. Baum and Goodstein (2005) find evidence of male admissions preferences in the late 1990s at small liberal arts colleges (though see also Bielby et al. (2014)). The impact of male affirmative action is quantitatively too small to play a significant role in the aggregate allocation of students to universities (see Appendix B).

4.4 Holistic Review and Alternative Metrics

Some less-selective colleges use formulas or cutoffs for admissions (Zimmerman, 2014; Mountjoy, 2024; Miller, 2025), but nearly all highly-selective public and private universities conduct admissions by "holistic review" (Bastedo et al., 2018): that is, rather than admitting students using a fixed-weight ranking of students by academic criteria (such as a minimum SAT threshold or SAT-GPA combination), admissions decisions are made by a "trained evaluator or set of evaluators [who] craft a single score for the applicant based upon a combination of [academic and non-academic] criteria" (BOARS, 2012).

Holistic review reallocates admissions weight from test scores to essays and other components. The latent weights that admissions readers place on the different application components in holistic review likely vary across institutions, but in most cases it appears that scoring remains based primarily on an assessment of academic achievement (Mare, 2014; Arcidiacono, 2018; Bleemer and Friedman, 2026). Zitzewitz (2025) shows that switching to holistic review can affect the admissions decisions of 10-15 percent of admitted students and result in a decline in average test scores. However, this probably overstates the effect on selection by academic achievement, as in many cases holistic review merely replaces students with high standardized test scores with others whose applications show other achievement indicators (like strong high school grades).

Five states' public university systems combine test-based or holistic review with "top percent" policies, which provide an admissions advantage to the highest-ranking (by high school grade point average) graduates from every eligible high school, regardless of their test scores: Texas, California, Florida, Georgia, and New York.[14] The net effect of these policies is to provide admissions advantages to high-achieving students from low-performing high schools, because top students at high-performing schools generally have high test scores and would be admitted under traditional rules (Black, Denning and Rothstein, 2023). They have been found to slightly increase enrollment of underrepresented minorities (by 1-2 percentage points; Kapor, 2024; Bleemer, 2021) and enrollment from families with below-median income at participating selective universities (as much as 1

[14]These policies have a precedent in the "highest seventh plan" implemented by Harvard and other private Massachusetts colleges in the 1920s (Hinckley, 1941).

percentage point; Bleemer, 2023).

Clearly, admissions to selective US colleges involves more than standardized test scores. Some students are admitted for their athletic ability or personal connections, while others get boosts for increasing diversity. But most college and university admissions decisions, while not formulaic, remain based on demonstrated academic achievement.

5 Is Merit-Based Stratification Efficient?

The American higher education system provides substantially greater educational resources to students with relatively stronger pre-college academic achievement. Is this allocation efficient, in the sense of allocating students and resources to schools in a way that maximizes some understanding of social welfare? This cannot be assumed; the higher education sector is dominated by imperfect information, reputation effects, non-price rationing, and non-profit and state actors, and no welfare theorem ensures that the market outcome is efficient. There might well be other allocations of students to schools, or of resources to students, that would achieve higher social welfare, and perhaps even that could be sustained as equilibria given the behavior of students and institutions.

Higher education can be seen as a production process where institution-provided inputs like classrooms, dormitories, and instructors are combined with students to produce learning output. We set aside the question of the relevant dimensions of output – is the goal to maximize earnings, or are there other forms of learning that should be pursued? – and assume an agreed-upon, unidimensional output. Each university produces this output by combining institutional inputs with students, who we assume can also be summarized by a single index, their average pre-college academic achievement (setting aside questions considered above about how to measure this). The production function can be quite general, allowing for various interactions. In particular, the dependence of output on student characteristics allows for a range of types of peer effects, where the outcomes of one student depend on the characteristics of classmates.

One initial question is whether it is efficient for students to be so stratified at all by pre-college academic achievement. In many countries (like Canada and Germany) there is much less stratification, with students generally attending their local college. There is little

evidence on the efficiency of 'tracking' in higher education; the large literature on K-12 tracking suggests that increasing the variance of student achievement increases inequality but has little effect on average outcomes (Betts, 2011; Terrin and Triventi, 2023).

We focus on a second question: Given substantial stratification, it is efficient for resources to be allocated preferentially to the highest-achieving students? We consider three theories about why it might be.

5.1 q-Complementarity

Two inputs in production are q-complements, where q refers to 'quantity,' if increasing the amount of one raises the marginal product of the other. The most straightforward rationalization of the meritocratic consensus would be if pre-college academic achievement and per-student instructional spending are q-complements in the production of human capital. This model has some plausibility: perhaps higher-achieving students benefit more from small classes (which facilitate individualized exploration), professors at the research frontier who can teach more difficult course material, and more up-to-date lab facilities. If so, social gains can arise from segregating those high-achieving students from others and directing substantial resources in their direction (Sallee, Resch and Courant, 2008). This creates an equity-efficiency tradeoff: it is efficient to direct more resources to the highest-achieving students, but because high-achievement students tend to come from advantaged families, such an allocation is regressive.

Benabou (2000) notes that in the presence of q-complementarity, the meritocratic consensus may be in effect by *anti*-meritocratic if meritocracy is understood as maximizing the share of income variation arising from variation in innate ability (as opposed to family background). A traditionally-meritocratic allocation would direct higher instructional spending to the schools that high-ability, disadvantaged students attend, which might be schools with low average pre-college academic achievement if that is sufficiently correlated with family background. See related discussion in Roemer (1998).

An alternative view of the higher education production process, in which pre-college educational attainment and per-student spending are *q-substitutes* rather than q-complements, is also plausible. Perhaps weaker students benefit more from individualized attention, out-

of-classroom advising, tutoring, and student support services, while stronger students can learn nearly as effectively with only a lecture and a textbook. Elementary and secondary education funding policies are often based on an assumption like this about the production process; for example, "weighted student funding" formulas direct more resources to schools serving disadvantaged students (for example, Chingos and Blagg, 2017).

If student aptitude or preparedness is a q-substitute for school resources, it might still make sense to segregate stronger from weaker students (depending on the nature of peer effects), but the efficient allocation would direct *more* resources to the schools where weaker students enroll. In public policy terms, such a policy might involve more generous funding for community colleges and open access institutions, and lower funding – at least for instruction – at state flagships. However, there is little sign that policymakers prefer such funding arrangements.

5.2 Convex Social Welfare

A second model in which the observed allocations might be efficient is one in which the social welfare function puts very high weight on the human capital of the highest achievers – in technical terms, if the social welfare function is convex in students' post-college human capital. A possible justification for such a model would be if these top achievers are responsible for a disproportionate share of scientific breakthroughs, entrepreneurial successes, and artistic masterpieces. In this case, it could then be optimal to concentrate the best students with the highest educational resources, not because they receive the highest human capital benefit from these resources but because any human capital increment that they receive is disproportionately valuable to society.

5.3 High School Incentivization

A third model emphasizes educational production in high school rather than in college. Students decide how much effort to exert on their high school studies, and policies that increase the reward for high achievement can increase student effort (Coate and Loury, 1993; Fryer, Loury and Yuret, 2008; Akhtari, Bau and Laliberte, 2024). Collegiate resources can be such a reward, even if they are substitutes for student preparedness in the

college production process. This follows the same logic as meritocracy in competitive labor markets: when firms hire and promote the most productive workers, the resulting wage dispersion encourages effort and allocative efficiency (e.g. Holmström, 1979). The incentive effects of large rewards for high school achievement could theoretically induce sufficient effort in high school to justify even inefficient allocations in college.

5.4 Evidence on the Theories

What does the empirical evidence tell us about these theories? There is a fair amount of available evidence regarding the q-complementarity theory. Many scholars have studied whether the wage return to *more-selective* university enrollment is higher for higher-testing students, focusing in particular on whether some lower-testing students would be "mismatched" with the demanding curriculum of more-selective universities and thus better off enrolling elsewhere. For example, an early series of observational and structural studies of race-based affirmative action – which admits lower-testing underrepresented minority students to more-selective universities – provided mixed evidence that the policy 'mismatched' students (for example, Loury and Garman, 1995; Arcidiacono and Lovenheim, 2016; Rothstein and Yoon, 2008). A more recent quasi-experimental literature has found that, if anything, the benefits of university selectivity *decrease* in pre-college academic achievement for reasonably-prepared students (Bleemer, 2022; Barahona, Dobbin and Otero, 2023). This finding is mirrored in broader studies of the return to university selectivity: lower-testing students generally benefit as much as (Mountjoy and Hickman, 2021; Chetty, Deming and Friedman, 2026) or more than (Dale and Krueger, 2014; Bleemer, 2021) higher-testing students from attending more selective colleges (though see Dillon and Smith (2020) for a contrary conclusion). However, these studies, whatever their findings, generally do not distinguish between peer groups and resources as determinants of college value-added.[15]

[15]Models of dynamic complementarity ("skill begets skill") in K-12 educational production (Cunha and Heckman, 2007; Johnson and Jackson, 2019) support the view that resources are complementary with prior achievement. However, several studies have found that small classes are more beneficial for lower-achieving students (Krueger, 1999; Summers and Wolfe, 1977), where dynamic complementarity would imply the opposite. In any case, the postsecondary production function may differ from that of K-12.

Other evidence comes from specific program evaluations. Several resource-intensive programs have yielded large benefits for college students with low prior achievement. The ASAP program provides financial, academic, and personal support to students at the City University of New York (CUNY) and significantly raises student success (Scrivener et al., 2015). Similar evidence comes from the Opening Doors demonstration project, aimed at improving higher education outcomes for community college students (Scrivener et al., 2008; Scrivener and Coghlan, 2011). However, it is difficult to compare the benefits of these programs to the potential benefits of devoting similar resources to high-prior-achievement students. Similarly, evidence from "massively open online courses" (MOOCs), which have attracted attention as a potentially much less resource intensive way to deliver postsecondary instruction, has found that this model was less effective for students with low prior achievement – which again suggests that resources and preparedness may be q-substitutes rather than complements (Hansen and Reich, 2015; Kofoed et al., 2024).

Turning to the other two models, there is very little evidence on the convexity of the relationship between human capital and social value. There is a positive convex relationship between childhood test scores and innovation, as measured by patenting (Bell et al., 2019), but this could be a direct effect of individual characteristics or family background rather than a consequence of school-constructed human capital. The relationship between entrepreneurial success and SAT scores, on the other hand, is positive but slightly *concave* (Chetty et al., 2025), even in an environment of highly stratified college resources. This suggests that high-testing students are unlikely to become disproportionately more likely to found successful firms as a result of enrolling at high-resource institutions. Finally, there is little available evidence on the potential magnitude of incentive effects of college quality on high school effort or achievement.

Considering these three possible justifications for the meritocratic consensus as a whole, there is no compelling evidence that it is efficient to direct the most resources to the most highly prepared students. What evidence there is offers weak support for the view that resources can substitute for preparedness and suggests that, at current margins, overall learning would be increased by shifting resources to schools that enroll students with lower prior achievement.

6 The Future of the Meritocratic Consensus

Several recent developments seem likely to affect the scope of the meritocratic consensus in the short- and medium-term.

6.1 Funding Allocation

The shift toward higher funding at more selective institutions has leveled out or even become less pronounced since about 2010 (Figure A-12). The between-institution associations between average SAT score and both net tuition revenue and private funding have continued their slow growth in recent years, in part likely driven by rising international undergraduate and graduate enrollment at more-selective institutions (Bound et al., 2021). However, the relationship between average SAT score and government appropriations has become meaningfully negative over the past two decades. This has been almost wholly driven by increasing funding from student grant programs to low-SAT institutions, not by changes in states' block appropriations. State financial aid programs like New York's Excelsior Scholarship and New Jersey's Tuition Aid Grant have played an important role in this nascent trend.[16] Recent federal actions targeting the finances of high-revenue private universities, such as the expanded university endowment tax in the One Big Beautiful Bill Act of 2025, could compound this trend, though whether it continues will depend on federal and state legislative priorities.

6.2 College Admissions

The early 2020s brought two of the most significant changes to university admissions policies in over 50 years. First, after about 25 percent of private universities adopted "test-optional" admissions policies between the 1990s and 2010s, the 2020 Covid-19 pandemic then forced many other institutions to drop their test requirements. The share of "test-mandatory" institutions fell to just 8 percent in 2023 – comparable to its level in the 1930s. Some universities went test-optional, meaning that applicants could decide

[16]The Appendix includes a number of figures (A-12 to A-14) that offer robustness checks and additional perspectives on the issues discussed here.

whether to submit standardized test scores, while others (including all public universities in California) have gone test-blind, prohibiting applicants from submitting scores. The Ivy League schools largely reinstated their testing requirements between 2024 and 2026, but as of mid-2026, most top-50 private and flagship public institutions had not.

The effect of test-optional and test-blind policies on student stratification by pre-college academic achievement is unknown – not least because tests are the most common measure of cross-university differences in academic achievement – but their widespread adoption may herald meaningful changes in academic achievement's centrality in university admissions. Schools that implemented test-optional admissions prior to 2016 saw the stabilization of previously-declining application rates and a 1-2 percentage point increase in their enrollment shares of underrepresented minorities (Bennett, 2022).

Second, the US Supreme Court decision in *Students for Fair Admissions (SFFA) v. Harvard* (600 US 181 [2023]) prohibited race-based affirmative action in college admissions. The year following the Court decision, first-year enrollment of underrepresented minorities at top-50 private universities declined from 22 to 18 percent. However, an enrollment cascade of students into less-selective institutions (Conrad and Sharpe, 1996) led 2024 enrollment of first-year underrepresented minority students at flagship public universities to rise by 3 percent. The downstream effects of the *SFFA* decision – and related actions by the Trump administration to obligate greater 'meritocratic' selection in university admissions – are still evolving.

6.3 Within-Institution Stratification

While enrollment stratification by levels of pre-college academic achievement has been largely stable for decades and inequity in resource allocations between institutions with higher and lower test scores has stopped growing, recent decades have seen the increasing allocation of greater resources to higher-achieving students *within* institutions.

As one example, college students in the early 1960s at nearly all institutions could declare any available major in which they could earn passing grades. In recent decades, however, many institutions (and especially public universities) have begun rationing access to high-demand majors, often using admissions processes that resemble those used

for college admissions. Among lucrative majors like engineering and business, over 25 percent of enrollment-weighted programs at the "R1" research-intensive public universities have minimum grade-point-average requirements in introductory courses of above 2.3 (higher than the passing grade of 2.0) and another 20 percent implement admissions policies mirroring holistic review (based on grades, essays, and other materials). Between 2010 and 2024, the share of computer-science programs at R1 public universities – often the institution's highest-value major – that restricted access on the basis of academic preparation rose from 10 to 40 percent (Bleemer and Mehta, 2022a). 'Soft' restrictions like relatively low grade distributions in courses related to science, technology, engineering, and mathematics (STEM) have also become more prevalent (for example, Butcher, McEwan and Weerapanda, 2014).

These policies have contributed to substantial between-major stratification by academic achievement, with higher-achieving students allocated to higher-value majors. This in turn manifests as swiftly growing gaps between the average relative value of college majors earned by more- or less-advantaged students (Bleemer and Mehta, 2022a; Bleemer and Quincy, 2025a). As with allocations across institutions, there is little evidence that within-institution "meritocratic" allocation of majors is efficient, with the available evidence instead suggesting that providing lucrative major access to *lower*-achievement students would result in greater human capital production (Bleemer and Mehta, 2022b).

The growing prevalence of "honors colleges" – separate programs within large universities that provide additional resources to selected high-achieving students that are not available to the broader student body – is another recent avenue by which resources are preferentially allotted to higher-performing students within institutions. Between 1958 and 2021, the number of research-intensive universities with honors colleges rose from three to about 200, with growth accelerating over time. Honors colleges generally provide dedicated curricula in small classrooms, special advising services, and additional student amenities, typically without any additional cost for the student (Cognard-Black and Smith, 2023). While there is little evidence on the long-run effects of honors college participation, studies of 'gifted education' programs at the K-12 level suggest that expanding program access to lower-performing students would provide those students substantial long-run value (Card, Chyn and Giuliano, 2024; Lavy and Goldstein, 2022).

7 Conclusion

Colleges and universities in the United States are highly stratified by students' pre-college academic achievement, with the highest-achieving students attending colleges that spend dramatically more on instruction than those that lower-achieving students attend. Neither of these is a necessary feature of a higher education market; universities need not be as stratified by academic achievement as they are in the United States, nor do resources need to be so inequitably distributed. There is little economic evidence justifying the status quo on efficiency grounds, and substantial room for skepticism. Whether this feature of the American higher education system persists depends on whether universities – and their primary owners, the American public – wish it to.

References

Abramitzky, Ran, Jennifer Kowalski, Santiago Pérez, and Joseph Price. 2024. "The GI Bill, Standardized Testing, and Socioeconomic Origins of the US Educational Elite Over a Century." NBER Working Paper, 33164.

Akhtari, Mitra, Natalie Bau, and Jean-William Laliberte. 2024. "Affirmative Action and Precollege Human Capital." American Economic Journal: Applied Economics, 16(1): 1–32.

Arcidiacono, Peter. 2018. "Expert Report of Peter S. Arcidiacono." United States District Court for the District of Massachusetts Expert Report No. 14-cv-14176-ADB, Boston, MA. Filed in *Students for Fair Admissions, Inc. v. Harvard* (D. Mass.).

Arcidiacono, Peter, and Michael Lovenheim. 2016. "Affirmative action and the quality–fit trade-off." Journal of Economic Literature, 54(1): 3–51.

Arcidiacono, Peter, Josh Kinsler, and Tyler Ransom. 2022. "Legacy and athlete preferences at Harvard." Journal of Labor Economics, 40(1): 133–156.

Barahona, Nano, Caue Dobbin, and Sebastian Otero. 2023. "Affirmative Action in Centralized College Admissions Systems." Manuscript.

Bastedo, Michael, Nicholas Bowman, Kristen Glasener, and Jandi Kelly. 2018. "What are We Talking About When We Talk About Holistic Review? Selective College Admissions and its Effects on Low-SES Students." The Journal of Higher Education, 89(5): 782–805.

Baum, Sandy, and Eban Goodstein. 2005. "Gender imbalance in college applications: Does it lead to a preference for men in the admissions process?" Economics of Education Review, 24(6): 665–675.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John van Reenen. 2019. "Who Becomes an Inventor in America? The Importance of Exposure to Innovation." Quarterly Journal of Economics, 134(2): 647–713.

Benabou, Roland. 2000. "Meritocracy, Redistribution, and the Size of the Pie." In Meritocracy and Economic Inequality. , ed. Kenneth Arrow, Samuel Bowles and Steven Durlauf, 317–339. Princeton, NJ:Princeton University Press.

Bender, William J. 1952. "Admission Policy Memo." In Admissions 1957-1958: Report, General Harvard Admissions Collection (HUC 957.3), Harvard University Archives, Cambridge, MA.

Bennett, Christopher T. 2022. "Untested Admissions: Examining Changes in Application Behaviors and Student Demographics under Test-Optional Policies." American Educational Research Journal, 59(1): 180–216.

Betts, Julian. 2011. "The Economics of Tracking in Education." Handbook of the Economics of Education, 3: 341–381.

Bielby, Rob, Julie Posselt, Ozan Jaquette, and Michael Bastedo. 2014. "Why are Women Underrepresented in Elite Colleges and Universities? A Non-Linear Decomposition Analysis." Research in Higher Education, 55: 735–760.

Black, Sandra, Jeffrey Denning, and Jesse Rothstein. 2023. "Winners and Losers? The Effect of Gaining and Losing Access to Selective Colleges on Education and Labor Market Outcomes." American Economic Journal: Applied Economics, 15(1): 26–67.

Bleemer, Zachary. 2018. "The UC ClioMetric History Project and Formatted Optical Character Recognition." Center for Studies in Higher Education Research Paper, 3(18).

Bleemer, Zachary. 2021. "Top Percent Policies and the Return to Postsecondary Selectivity." CSHE Research Paper, 21(1).

Bleemer, Zachary. 2022. "Affirmative Action, Mismatch, and Economic Mobility after California's Proposition 209." Quarterly Journal of Economics, 137(1): 115–160.

Bleemer, Zachary. 2023. "Affirmative Action and its Race-Neutral Alternatives." Journal of Public Economics, 220: 104839.

Bleemer, Zachary, and Aashish Mehta. 2022a. "College Major Restrictions and Student Stratification." NBER Working Paper, 33269.

Bleemer, Zachary, and Aashish Mehta. 2022b. "Will Economics Make You Rich? A Regression Discontinuity Analysis of the Return to College Major." American Economic Journal: Applied Economics, 14(2): 1–22.

Bleemer, Zachary, and John Friedman. 2026. "Standardized Tests in Holistic University Admissions." Manuscript.

Bleemer, Zachary, and Sarah Quincy. 2025a. "Changes in the College Mobility Pipeline Since 1900." NBER Working Paper, 33797.

Bleemer, Zachary, and Sarah Quincy. 2025b. "Junior Colleges and Human Capital." Manuscript.

Bleemer, Zachary, Mukul Kumar, Aashish Mehta, Chris Muellerleile, and Christopher Newfield. 2023. Metrics that Matter: Counting What's Really Important to College Students. Baltimore, MD:Johns Hopkins University Press.

BOARS. 2012. Comprehensive Review in Admissions at the University of California: An Update. Oakland, CA:University of California.

Bound, John, and Sarah Turner. 2007. "Cohort crowding: How resources affect collegiate attainment." Journal of Public Economics, 91(5-6): 877–899.

Bound, John, Brad Hershbein, and Bridget Terry Long. 2009. "Playing the admissions game: Student reactions to increasing college competition." Journal of Economic Perspectives, 23(4): 119–146.

Bound, John, Breno Braga, Gaurav Khanna, and Sarah Turner. 2021. "The Globalization of Postsecondary Education: The Role of International Students in the US Higher Education System." Journal of Economic Perspectives, 35(1): 163–184.

Bound, John, Michael Lovenheim, and Sarah Turner. 2010. "Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources." American Economic Journal: Applied Economics, 2(3): 129–157.

Butcher, Kristen, Patrick McEwan, and Akila Weerapanda. 2014. "The Effects of an Anti-Grade-Inflation Policy at Wellesley College." Journal of Economic Perspectives, 28(3): 189–204.

Card, David, Eric Chyn, and Laura Giuliano. 2024. "Can Gifted Education Help Higher-Ability Boys from Disadvantaged Backgrounds?" NBER Working Paper, 33282.

Castilla, Emilio J., and Ethan J. Poskanzer. 2022. "Through the Front Door: Why Do Organizations (Still) Prefer Legacy Applicants?" American Sociological Review, 87(5): 782–826.

CCHE. 1966. Annual Report on Faculty Salaries, Benefits, and Recruitment. Sacramento and San Francisco:Coordinating Council for Higher Education.

Chetty, Raj, David J Deming, and John N Friedman. 2026. "Diversifying society's leaders? The determinants and causal effects of admission to highly selective private colleges." The Quarterly Journal of Economics, 141(1): 51–145.

Chetty, Raj, Gaia Dossi, Matt Smith, John van Reenen, Owen Zidar, and Erick Zwick. 2025. "Creating New Businesses in America: The Determinants of and Returns to Entrepreneurship." Manuscript.

Chetty, Raj, John Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2020. "Income Segregation and Intergenerational Mobility Across Colleges in the United States." Quarterly Journal of Economics, 135(3): 1567–1633.

Chingos, Matthew, and Kristin Blagg. 2017. Do Poor Kids Get Their Fair Share of School Funding? Washington, DC:Urban Institute.

Clothier, Robert. 1946. "Letter to Hon. Mattie S. Doremus." In Records of the Robert C. Clothier Administration, 1925-1952 (RG 04/A16), Box 1, Folder 3, Rutgers University Archives, New Jersey.

Coate, Stephen, and Glenn C Loury. 1993. "Will affirmative-action policies eliminate negative stereotypes?" The American Economic Review, 1220–1240.

Cognard-Black, Andrew, and Patricia Smith. 2023. "Characteristics of the 21st Century Honors College." In Honors Colleges in the 21st Century. , ed. Richard Badenhousen, 23–79. Lincoln, NE:National Collegiate Honors Council.

Cohodes, Sarah, and Joshua Goodman. 2014. "Merit Aid, College Quality, and College Completion: Massachusetts' Adams Scholarship as an In-Kind Subsidy." American Economic Journal: Applied Economics, 6(4): 251–285.

Committee of Ten. 1892. Report of the Committee of Ten on Secondary School Studies. New York:American Book Company on behalf of the National Educational Association.

Conrad, Cecilia A., and Rhonda V. Sharpe. 1996. "The impact of the California Civil Rights Initiative (CCRI) on university and professional school admissions and the implications for the California Economy." The Review of Black Political Economy, 25: 13–59.

Coons, Arthur G., Thomas C. Holy, Arthur D. Browne, Dean E. McHenry, Howard A. Campion, Henry T. Tyler, Glenn S. Dumke, and Robert J. Wert. 1960. A Master Plan for Higher Education in California, 1960-1975. Sacramento, CA:California State Department of Education.

Cooper, William John. 1931. Biennial Survey of Education, 1928-1930, Volume II. Washington, DC:US Government Printing Office.

Cunha, Flavio, and James Heckman. 2007. "The technology of skill formation." American Economic Review, 97(2): 31–47.

Dale, Stacy, and Alan Krueger. 2002. "Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables." The Quarterly Journal of Economics, 117(4): 1491–1527.

Dale, Stacy, and Alan Krueger. 2014. "Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data." Journal of Human Resources, 49(2): 323–358.

Deming, David, and Christopher Walters. 2014. "The Impact of Price Caps and Spending Cuts on U.S. Postsecondary Attainment." NBER Working Papers, 23736.

Dillon, Eleanor, and Jeffrey Smith. 2017. "Determinants of the Match between Student Ability and College Quality." Journal of Labor Economics, 35(1): 45–66.

Dillon, Eleanor Wiske, and Jeffrey Andrew Smith. 2020. "The Consequences of Academic Match between Students and Colleges." Journal of Human Resources, 55(3): 767–808.

Ehrenberg, Ronald. 2012. "American Higher Education in Transition." Journal of Economic Perspectives, 26(1): 193–216.

Espeland, Wendy, and Michael Sauder. 2007. "Rankings and Reactivity: How Public Measures Recreate Social Worlds." American Journal of Sociology, 113(1): 1–40.

Freedman, Leora. 2023. The Employees of the California State University, 2023. Long Beach, CA:California State University Systemwide Human Resources.

Fryer, Roland G., Glenn C. Loury, and Tolga Yuret. 2008. "An Economic Analysis of Color-Blind Affirmative Action." The Journal of Law, Economics, and Organization, 24(2): 319–355.

Goldin, Claudia, and Lawrence Katz. 1999. "The Shaping of Higher Education: The Formative Years in the United States, 1890 to 1940." Journal of Economic Perspectives, 13(1): 37–62.

Goldin, Claudia, Lawrence Katz, and Ilyana Kuziemko. 2006. "The Homecoming

of American College Women: The Reversal of the College Gender Gap." Journal of Economic Perspectives, 20(4): 133–156.

Grodsky, Eric, John Warren, and Erika Felts. 2008. "Testing and Social Stratification in American Education." Annual Review of Sociology, 34: 385–404.

Grosz, Michel. 2021. "Admissions Policies, Cohort Composition, and Academic Success: Evidence from California." Journal of Human Resources.

Hansen, John D, and Justin Reich. 2015. "Democratizing education? Examining access and usage patterns in massive open online courses." Science, 350(6265): 1245–1248.

Hendricks, Lutz, Christopher Herrington, and Todd Schoellman. 2021. "College Quality and Attendance Patterns: A Long-Run View." American Economic Journal: Macroeconomics, 13(1): 184–215.

Hinckley, William. 1941. Handbook of College Entrance Requirements. Vol. 13, Washington DC:US Department of Education Bulletin.

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

Holmström, Bengt. 1979. "Moral hazard and observability." The Bell journal of economics, 74–91.

Hoxby, Caroline M. 2009. "The changing selectivity of American colleges." Journal of Economic perspectives, 23(4): 95–118.

Hurwitz, Michael. 2011. "The impact of legacy status on undergraduate admissions at elite colleges and universities." Economics of Education Review, 30(3): 480–492.

Jackson, C. Kirabo, and Claire Mackevicius. 2024. "What Impacts Can We Expect from School Spending Policy? Evidence from Evaluations in the United States." American Economic Journal: Applied Economics, 16(1): 412–446.

Johnson, Rucker C, and C Kirabo Jackson. 2019. "Reducing inequality through dynamic complementarity: Evidence from Head Start and public school spending." American Economic Journal: Economic Policy, 11(4): 310–349.

Kane, Thomas J, and Cecilia Elena Rouse. 1999. "The community college: Educating students at the margin between college and work." Journal of economic Perspectives, 13(1): 63–84.

Kane, Thomas J, Peter R Orszag, and Emil Apostolov. 2005. "Higher education appropriations and public universities: Role of Medicaid and the business cycle." In Brookings-Wharton Papers on Urban Affairs. Vol. 2005, , ed. Gary Burtless and Janet Pack, 99–146. Washington, DC:Brookings Institution Press.

Kapor, Adam. 2024. "Transparency and percent plans." NBER Working Paper, 32372.

Karabel, Jerome. 1989. Freshman Admissions at Berkeley: A Policy for the 1990s and Beyond. Berkeley, CA:Committee on Admissions and Enrollment at the University of

California, Berkeley.

Karabel, Jerome. 2005. The Chosen: The Hidden History of Admission and Exclusion at Harvard, Yale, and Princeton. New York:Houghton Mifflin Harcourt.

Kofoed, Michael, Lucas Gebhart, Dallas Gilmore, and Ryan Moschitto. 2024. "Zooming to Class? Experimental Evidence on College Students' Online Learning during COVID-19." American Economic Review: Insights, 6(3): 324–340.

Krueger, Alan B. 1999. "Experimental estimates of education production functions." The Quarterly Journal of Economics, 114(2): 497–532.

Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach. 2018. "School finance reform and the distribution of student achievement." American Economic Journal: Applied Economics, 10(2): 1–26.

Lavy, Victor, and Yoav Goldstein. 2022. "Gifted Children Programs' Short and Long-Term Impact: Higher Education, Earnings, and the Knowledge Economy." NBER Working Paper, 29779.

Lemann, Nicholas. 2000. The Big Test: The Secret History of the American Meritocracy. New York:Macmillan.

Lemann, Nicholas. 2024. Higher Admissions: The Rise, Decline, and Return of Standardized Testing. Princeton, NJ:Princeton University Press.

Loury, Linda Datcher, and David Garman. 1995. "College Selectivity and Earnings." Journal of Labor Economics, 13(2): 289–308.

Mare, Robert D. 2014. Holistic review in freshman admissions at the University of California-Los Angeles, 2009-11 update. Los Angeles:Committee on Undergraduate Admissions and Relations with Schools, University of California, Los Angeles.

McVay, Ruth. 1969. Undergraduate Enrollment by Ethnic Group in Federally Funded Institutions of Higher Education. Washington, DC:U.S. Department of Health, Education, and Welfare Office of Civil Rights.

Meder, Jr, Albert. 1939. "Letter to Robert C. Clothier." In Records of the Robert C. Clothier Administration, 1925-1952 (RG 04/A16), Box 77, Folder 15, Rutgers University Archives, New Jersey.

Miller, Lois. 2025. "Switching Schools: Effects of College Transfers. EdWorkingPaper No. 25-1159." Annenberg Institute for School Reform at Brown University.

Mountjoy, Jack. 2022. "Community colleges and upward mobility." American Economic Review, 112(8): 2580–2630.

Mountjoy, Jack. 2024. "Marginal Returns to Public Universities." NBER Working Paper, 32296.

Mountjoy, Jack, and Brent Hickman. 2021. "The Returns to College(s): Relative Value-Added and Match Effects in Higher Education." NBER Working Paper, 29276.

NCES. 2021. Digest of Education Statistics 2019 (NCES 2021-009). Washington,

D.C.:National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Reber, Sarah, Rina Nagashima, and Gabriela Goodman. 2023. Admissions at most colleges will be unaffected by Supreme Court ruling on affirmative action. Washington, DC:Brookings Institution.

Reeves, Richard V. 2022. Of Boys and Men: Why the Modern Male Is Struggling, Why It Matters, and What to Do About It. Bloomsbury Publishing USA.

Roemer, John. 1998. Equality of opportunity. Cambridge, Mass:Harvard University Press.

Rothstein, Jesse, and Albert Yoon. 2008. "Mismatch in Law School." NBER Working Paper, 14275.

Rothstein, Jesse, and Diane Whitmore Schanzenbach. 2022. "Does money still matter? Attainment and earnings effects of post-1990 school finance reforms." Journal of Labor Economics, 40(S1): S141–S178.

Rothstein, Jesse M. 2006. "Good principals or good peers? Parental valuation of school characteristics, Tiebout equilibrium, and the incentive effects of competition among jurisdictions." American Economic Review, 96(4): 1333–1350.

Sallee, James, Alexandra Resch, and Paul Courant. 2008. "On the Optimal Allocation of Students and Resources in a System of Higher Education." The B.E. Journal of Economic Analysis & Policy, 8(1).

Scrivener, Susan, and Erin Coghlan. 2011. "Opening Doors to Student Success: A Synthesis of Findings from an Evaluation at Six Community Colleges. Policy Brief." MDRC.

Scrivener, Susan, Dan Bloom, Allen LeBlanc, Christina Paxson, Cecilia Elena Rouse, and Colleen Sommo. 2008. "Opening Doors. A Good Start: Two-Year Effects of a Freshmen Learning Community Program at Kingsborough Community College." MDRC.

Scrivener, Susan, Michael J Weiss, Alyssa Ratledge, Timothy Rudd, Colleen Sommo, and Hannah Fresques. 2015. "Doubling graduation rates: Three-year effects of CUNY's Accelerated Study in Associate Programs (ASAP) for developmental education students." Scrivener, Susan, Michael J. Weiss, Alyssa Ratledge, Timothy Rudd, Colleen Sommo, and Hannah Fresques, Doubling Graduation Rates: Three-Year Effects of CUNY's Accelerated Study in Associate Programs (ASAP) for Developmental Education Students. New York: MDRC.

Spindt, Herman, Ralph Prator, J. Edward Sander, Basil Peterson, and Grant Jensen. 1961. Selection and Retention of Students in California's Institutions of Higher Education. Sacramento, CA:California State Department of Education.

Stanley, Marcus. 2003. "College education and the midcentury GI Bills." The Quarterly Journal of Economics, 118(2): 671–708.

Summers, Anita A, and Barbara L Wolfe. 1977. "Do schools make a difference?" The American Economic Review, 67(4): 639–652.

Terrin, Eder, and Moris Triventi. 2023. "The Effect of School Tracking on Student Achievement and Inequality: A Meta-Analysis." Review of Educational Research, 93(2): 236–274.

The Carnegie Foundation for the Advancement of Teaching. 2000. The Carnegie Classification of Institutions of Higher Education, 2000 Edition. Menlo Park, CA:The Carnegie Foundation for the Advancement of Teaching.

Thelin, John. 2010. "The Attrition Tradition in American Higher Education: Connecting Past and Present." AEI Future of American Education Project, 1.

Truffa, Francesca, and Ashley Wong. 2025. "Undergraduate Gender Diversity and the Direction of Scientific Research." American Economic Review, 115(7): 2414–2448.

University of California. 2023. Compensation at the University of California. Oakland, CA:The Regents of the University of California.

U.S. Department of Education. 2025. "Equity in Athletics Data Analysis." Web-based data collection cutting tool.

Veysey, Laurence. 1965. The Emergence of the American University. Chicago:University of Chicago Press.

Walton, Gregory, Steven Spencer, and Sam Erman. 2013. "Affirmative Meritocracy." Social Issues and Policy Review, 7(1): 1–35.

Westrick, Paul A., Jessica P. Marini, Linda Young, Helen Ng, Doron Shmueli, and Emily J. Shaw. 2019. Validity of the SAT for Predicting First-Year Grades and Retention to the Second Year. New York, NY:College Board.

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

Zitzewitz, Eric. 2025. "The Academic Consequences of Affirmative Action Bans Combined with Diversity Targets." NBER Working Paper, 33563.

Online Appendix

The Meritocratic Consensus and Stratification in Higher Education

Zachary Bleemer and Jesse Rothstein

August 2026

Table of Contents

A Appendix A: Data Appendix

This project combines annual university-level statistics from the US Biennual Surveys of Education, the College Blue Books, the Higher Education General Information Survey (HEGIS), the College Board Annual Survey of Colleges, the Integrated Postsecondary Education Data System (IPEDS), and the College Mobility Report Cards. This data appendix discusses data availability and construction. All academic years refer to their start year; e.g. 2019-2020 is referred to as 2019.

A.1 US Biennial Survey of Education

The United States Office of Education published the Biennial Survey of Education biennially from 1916 until 1958, replacing the earlier (but similarly-detailed) annual Report of the Commissioner of Education. The biennials included annual enrollment by institution until the mid-1940s, but only included institution-level revenue and expenditure data between 1928 and 1940. We restrict attention to the 1928-1930 biennial and digitize the enrollment, expenditure, and revenue statistical tables using ExtractTable in October 2025, hand-correcting a very small number of errors.[17] Per-student instructional expenditures (referred to as "Resident Instruction" expenditures) and total revenues (which is the difference between "Total receipts, exclusive of additions to endowment" and both "private benefactions for increase of plant" and "state and city for increase of plant") are CPI-adjusted to 2025 and defined as the quotients with the total number of students (including preparatory and graduate students). Four-year institutions are defined as those that awarded at least 10 "first degrees" in that year. All reported statistics are winsorized at the top and bottom 5 percent to avoid statistical noise from misreporting and the role of individual institutions.

We winsorize all series 5 percent to reduce the influence of any individual institution. The primary impact of this is to reduce the influence of the City University of New York, which was a significant outlier among public four-year institutions in 1929 in both enrollment (73 percent larger than the next-largest institution) and per-student instructional expenditures (32 percent lower than the next-lowest institution). Both likely reflect high part-time enrollment, but the 1929 data do not permit calculation of enrollment measured in full-time equivalent students.

A.2 College Blue Books

The *College Blue Books* were statistical records originally collected by Huber William Hurt, a professor of education at Columbia University. They include detailed institution-level information on universities across the United States. We employ digitized versions of the Blue Books from the

[17] We use the "Statistics of Universities and Colleges" chapter of volume 2, Tables 7, 10, and 21.

years 1923, 1928, 1933, 1939, 1947, and 1962, available through HathiTrust with hand-recorded data provided by Bleemer and Quincy (2025a).

We impute undergraduate enrollment as the product between total enrollment (including graduate students) and the share of undergraduate degrees among all degrees (weighting Masters by 0.5 relative to undergraduate or doctoral degrees). Institutions are marked as requiring standardized tests in admissions if at least 50 percent of students are admitted at least in part by examination (1923-1947) or if the school reports requiring a standardized exam in admissions (1962). Junior colleges are classified as two-year institutions and are omitted; public institutions are those controlled by the federal, state, city, or other local government.

A.3 US Office of Civil Rights Reports

The US Office of Civil Rights published biannual reports on the ethnic composition of full-time undergraduate at each US college and university between 1968 and 1984. These are the only known comprehensive data on college ethnic composition prior to 1976. We digitized the main Table 1 of the 1968 volume (McVay, 1969) using ExtractTable in October 2025, hand-correcting a very small number of errors. Enrollments are reported overall and for the following racial groups: "Negro," "American Indian," "Oriental," and "Spanish Surnamed." We take the first, third, and fourth of these as the number of Black, Asian, and Hispanic enrollees at each institution, and divide by the total to get enrollment shares at each school.

A.4 HEGIS

A subset of the complete federal 1966-1986 HEGIS database has been digitized and made available as fixed-width files through ICPSR, series 30. We construct financial and enrollment series for 1969, 1970, 1972, 1975, 1978, and 1981. Instructional and student services expenditures are directly reported. Research expenditures before 1975 are the sum of "Sponsored Research" and "Other Separately Budgeted Research"; research expenditures are directly reported starting in 1975. "Academic Support" (like the library), "Student Services" (like the Registrar and athletics), and "Institutional Support" (like general management and HR), where the latter two are combined in 1969-1970. Total non-hospital revenue is difference between "Total current funds revenue" and "sales and services of hospitals". University revenues are not reported in a manner that permits reconstruction of the categories used in this paper because government tuition grants included in student tuition payments. All expenditure and revenue measures are CPI-adjusted to 2025.

We collect three enrollment measures for each university: first-time first-year enrollment, full-time undergraduate enrollment, and total FTE enrollment (which is full time enrollment plus one-third of part-time enrollment). The former is used for all cross-university statistical weights; the

middle is used for ethnicity shares of the full-time freshman class in 1976 (and is also collected for Black, Hispanic, and Asian students); and the last is used for all measurement of per-student expenditures and revenues.

Universities are linked to IPEDS by a FICE-to-UnitID dictionary available in the replication package. Public, private, and two-year institutions are identified using the link to IPEDS (privileging the 2002 status).

A.5 College Board Annual Survey of Colleges

The College Board has conducted annual surveys of all US colleges and universities since 1983; data are available in fixed-width files for license from College Board starting in 1985. We construction admissions and enrollment series for 1985, 1988, 1993, 1996, and 1998. Test-required universities are identified as those which report requiring any standardized test in admissions. Public and private universities are directly observed; universities are linked to IPEDS by a FICE-to-UnitID dictionary available in the replication package. The 25th and 75th percentile of SAT math and reading scores is observed for most test-required universities; we define the average test score as the sum of the mean of the 25th and 75 percentiles of each component. An additional 10-20 percent of institutions only report the 25th and 75th percentile of the school's composite ACT score; we predict those schools' average SAT score using the two ACT components separately in each year ($R^2 \approx 0.8$) and construct predicted SAT scores for those institutions.

We directly measure first-time first-year enrollment and use it to weight cross-university averages.

A.6 IPEDS

The complete federal university-level IPEDS database is available for 1980 and 1984-2023 at the website of the National Center for Education Statistics; we construct annual series for each year. Public, private, for-profit, 2000 Carnegie-classification "R1", and two-year institutions are directly observed; we fix institutions' categories in 2002. Tables A-3 and A-4 identify the top-50 private and state flagship universities by IPEDS UnitID.

The 25th and 75th percentile of SAT math and reading scores is observed for most test-required universities annually since 2001; we define the average test score as the sum of the mean of the 25th and 75 percentiles of each component. An additional 10-20 percent of institutions only report the 25th and 75th percentile of the school's composite ACT score; we predict those schools' average SAT score using the two ACT components separately in each year ($R^2 \approx 0.8$) and construct predicted SAT scores for those institutions. The admissions rate is directly observed for each institution in 2019.

We collect three enrollment measures for each university: first-time first-year enrollment, full-time undergraduate enrollment, and total FTE enrollment (which is full time enrollment plus one-third of part-time enrollment). The former is used for all cross-university statistical weights; the middle is used for ethnicity shares of the full-time freshman class in 1976 (and is also collected for Black, Hispanic, and Asian students); and the last is used for all measurement of per-student expenditures and revenues.

The student-to-faculty ratio is the ratio of total FTE enrollment to full-time faculty employment, including both tenure-track and term faculty. The six-year graduation rate is the share of first-time full-time undergraduates in that year who have earned a degree within six years of enrollment, excluding students who leave the university due to death, permanent disability, military enlistment, Peace Corps enlistment, or church mission, following the definition in IPEDS.

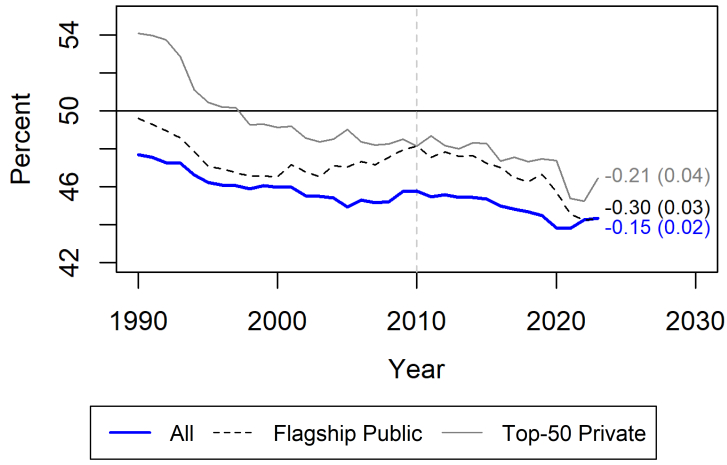
We take care to construct consistent measures of university revenues before and after the 1993 FASB Statement 117 (which changed accounting standards for non-profit private universities) and the 1999 GASB Statement 35 (which changed accounting standards for public universities). "Gross Tuition" is the sum of reported tuition payments collected by the institution and – in years that they are separated out – the portion of governmental and institutional student grants that are applied to tuition and fees. "Net Tuition" is the difference between gross tuition and total governmental and institution student grants (including those which subsidize auxiliary enterprises like housing), resulting in the total amount of money paid by students for tuition (understanding housing subsidies as negative tuition payments). Net tuition is not observed prior to 1988 because student aid receipts are not observed.

"Government Appropriations and Grants" include federal, state, and local block grants *as well as* federal Pell and non-Pell student aid grants and state and local grants (like the CalGrant or the Excelsior Scholarship), resulting in the complete student-oriented funding provided to universities by government. "Government Contracts" includes federal, state, and local grants and contract funding provided for research purposes, which subtracts government student aid grants from the IPEDS fields. Government appropriations and contracts are not observed prior to 1986.

"Private Funding" is the sum of private gifts, grants, and contracts, contributions from affiliated entities (like fundraising foundations), and all revenue from university endowment funds. Private funding is not observed prior to 2002.

"Instructional Expenditures," "Research Expenditures," and "Academic Support Expenditures" are directly observed. "Administrative Expenditures" are the sum of "Academic Support" (like the library), "Student Services" (like the Registrar and athletics), and "Institutional Support" (like general management and HR). We omit expenditures on public service (like community education programs), auxiliary enterprises (like dormitories and cafeterias), hospitals, and other independent operations from our main expenditure categories. "Total Non-Hospital Revenues" are the differ-

Figure BB-1: Male Enrollment Among Freshmen at Top-50 Private Universities



Note: This figure shows that top private universities have seen slightly less of a trend toward majority-female enrollment than state flagship universities since 2010 ($p = 0.10$). The annual share of freshman students who are male at all four-year universities, state flagship institutions, and all top-50 private institutions. Numbers on the plot report the average annual change in the male share between 2010 and 2023 (with robust standard error). The difference in mean slopes between state flagships and top-50 privates is 0.091, with an approximate standard error of 0.055. Source: IPEDS.

ence between total revenues and revenues from hospitals. All revenue and expenditure measures are CPI-adjusted to 2025.

A.7 Parental Income

We employ parental income data produced by Chetty et al. (2020), who in turn constructed parental income quintile shares by cohort and institution using data from the Internal Revenue Service. Parental income rank is measured for the 1980-1982 or 1989-1991 birth cohort by students' modal institution. Institutions are assigned (by necessity) to the lowest tier of all institutions sharing their IRS tax ID.

B Appendix B: Male Affirmative Action

Bielby et al. (2014) provide evidence that male college applicants were no more likely to attend a highly-selective university than academically-similar female applicants in the mid-2000s. However, the continued decline in the male share of all four-year college-goers – to 44.3 percent in 2023 – provided increasing an incentive for colleges with an interest in approximate gender balance to provide admission preference to male students. Figure BB-1 shows the annual share of male first-year students at flagship public universities – where gender-specific admissions prefer-

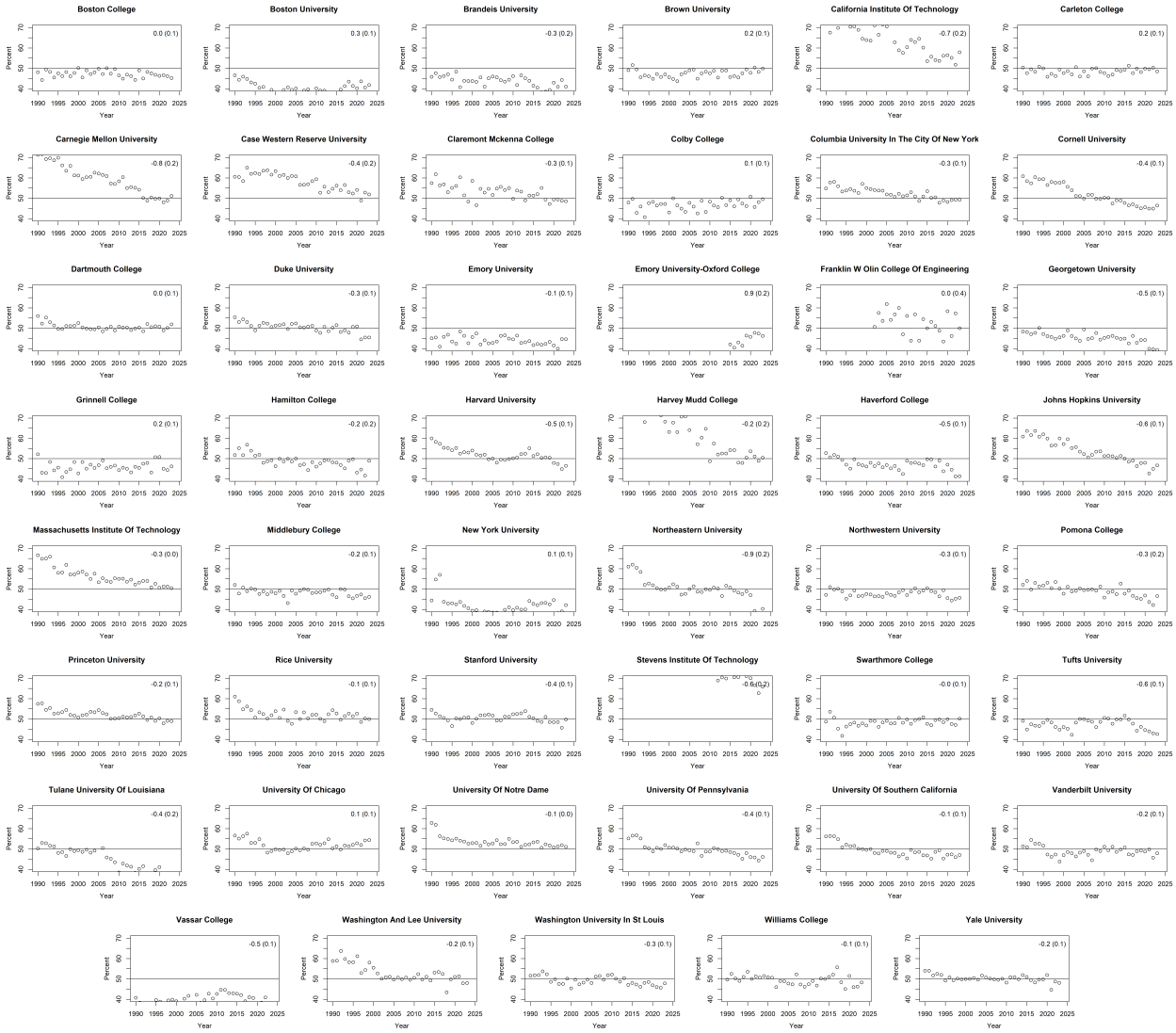
ences are believed to have been weak since the University of Georgia's were found to violate Title IX in *Johnson v. Board of Regents* in 2000 – and at top-50 private universities. Top-50 private universities had disproportionate male enrollment until 2010, when flagship state and top-50 private universities each enrolled about 48 percent male students. Since 2010, flagship public universities have become about 0.30 percentage points more female each year, whereas top-50 private universities have become about 0.21 percentage points more female each year. Treating flagship public enrollment as reflecting the characteristics of the study body eligible for top-50 private enrollment, this suggests that male affirmative action could explain as much as 1 to 2 percentage points – 0.09 percentage points per year over 14 years – of augmented male enrollment at top-50 private universities.

Figure BB-2 visualizes the male shares of first-year students at each of the top-50 private universities, excluding three all-female institutions. Some universities, like Brandeis University, have student gender compositions that closely mirror the average flagship public university. Others, like Dartmouth University and Swarthmore College, have gender compositions that are closely reflect those of institutions that impose a 50-50 gender quota, which has likely required a steadily-growing preference for male applicants since 2010 in order to maintain. Yet other schools – see Tulane and Northeastern University – appear to have implemented admissions policies that maintained an approximate gender quota for many years before policy changes led to increased female enrollment. These various patterns confirm the possibility that a subset of Top-50 universities implemented admissions policies in the 2010s that aggregate up to result in a small increase in male enrollment at top-50 private universities in recent years.

We conclude that while male affirmative action may operate at some highly-selective private universities, its enrollment effects even at those schools are on the order of a few percentage points, and there is little reason to believe that male affirmative action is a first-order determinant of university allocation in the US.

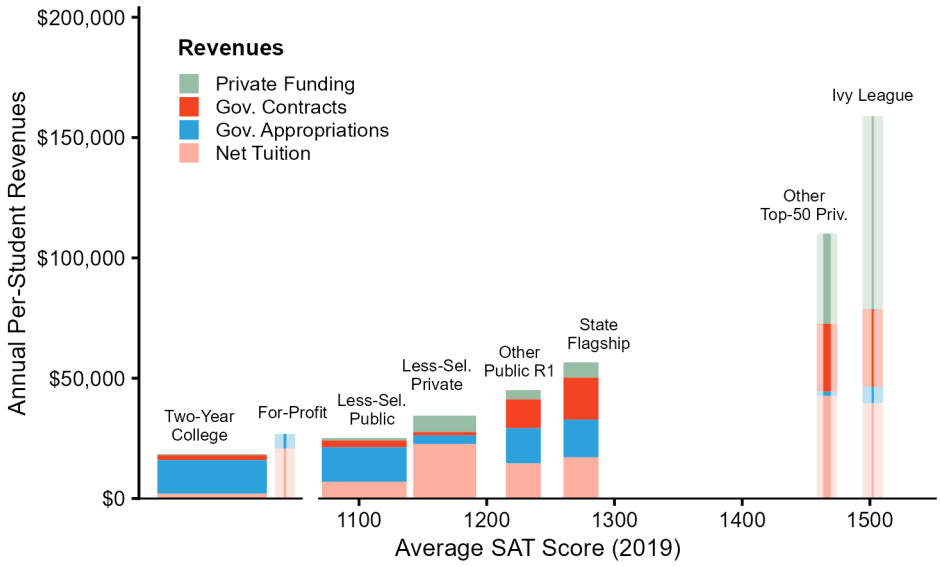
A Appendix Figures and Tables

Figure BB-2: Male Enrollment Among Freshmen at Top-50 Private Universities



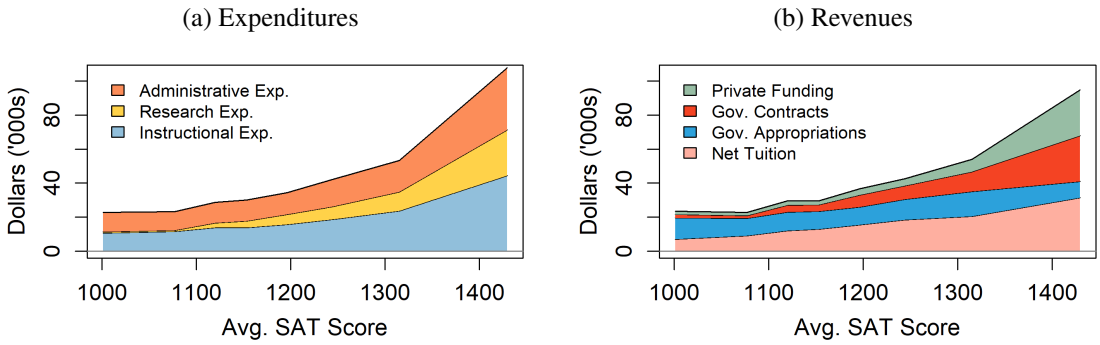
Note: The annual share of freshman students who are male at each of the top-50 private institutions. Numbers on the plot report the average annual change in the male share between 2010 and 2023 (with robust standard error). Source: IPEDS.

Figure A-1: Universities' 2019 Revenues by Tier



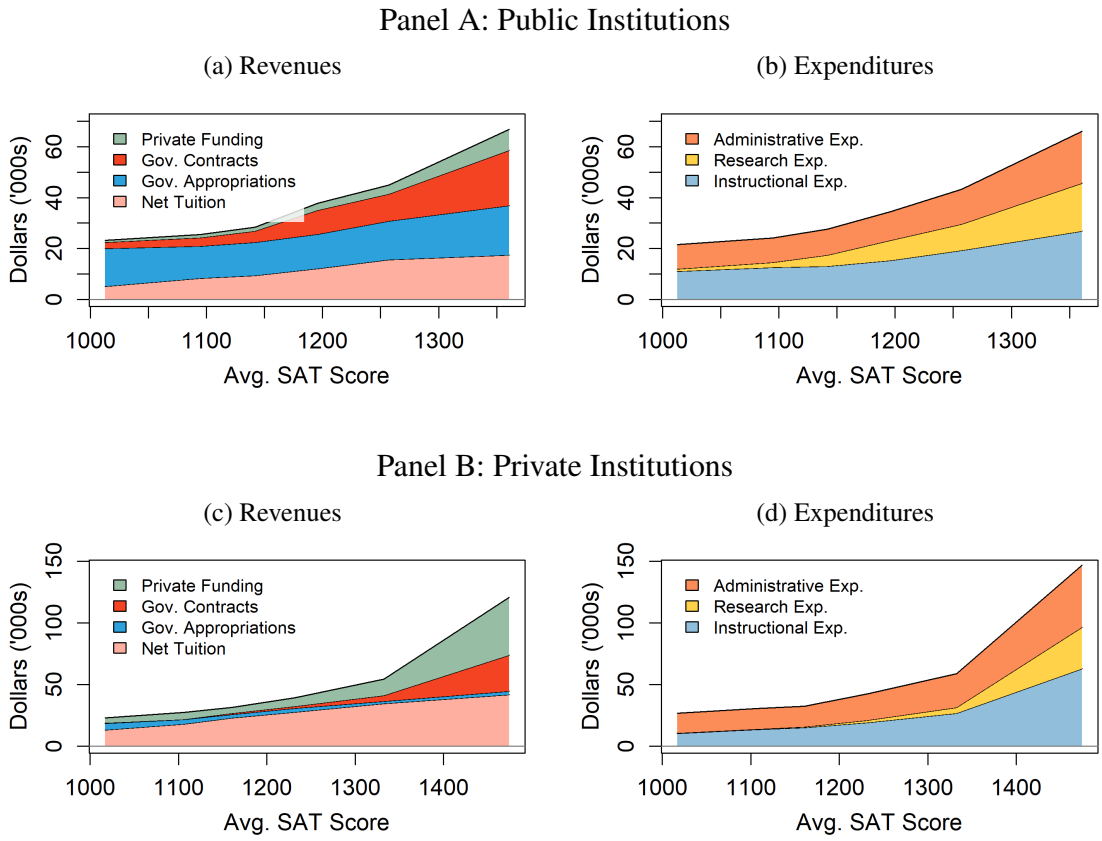
Note: This figure shows that the larger expenditures at selective institutions are financed by substantially higher net tuition funding (gross tuition payments minus gross financial aid), private funding (including alumni donations, affiliated groups like booster clubs, and endowment returns), and government research contracts. This figure shows average 2019 revenues per full-time equivalent student by tier of institution and source of revenue. "Net Tuition" is the difference between gross tuition and total governmental and institution student grants (including those which subsidize auxiliary enterprises like housing). "Government Appropriations and Grants" include federal, state, and local block grants *as well as* federal Pell and non-Pell student aid grants and state and local grants (like the CalGrant or the Excelsior Scholarship). "Government Contracts" includes federal, state, and local grants and contract funding provided for research purposes. "Private Funding" is the sum of private gifts, grants, and contracts, contributions from affiliated entities (like fundraising foundations), and all revenue from university endowment funds. Schools' revenues may not exactly equal expenditures due to cross-subsidization from hospitals and auxiliary enterprises (like dormitories and dining halls) and intertemporal substitution. Students are defined as full-time (undergraduate and postgraduate) students plus one-third of part-time students. All dollar amounts are CPI-adjusted to 2025. Source: IPEDS.

Figure A-2: Universities' 2019 Per-Student Revenues and Expenditures by Average SAT Score



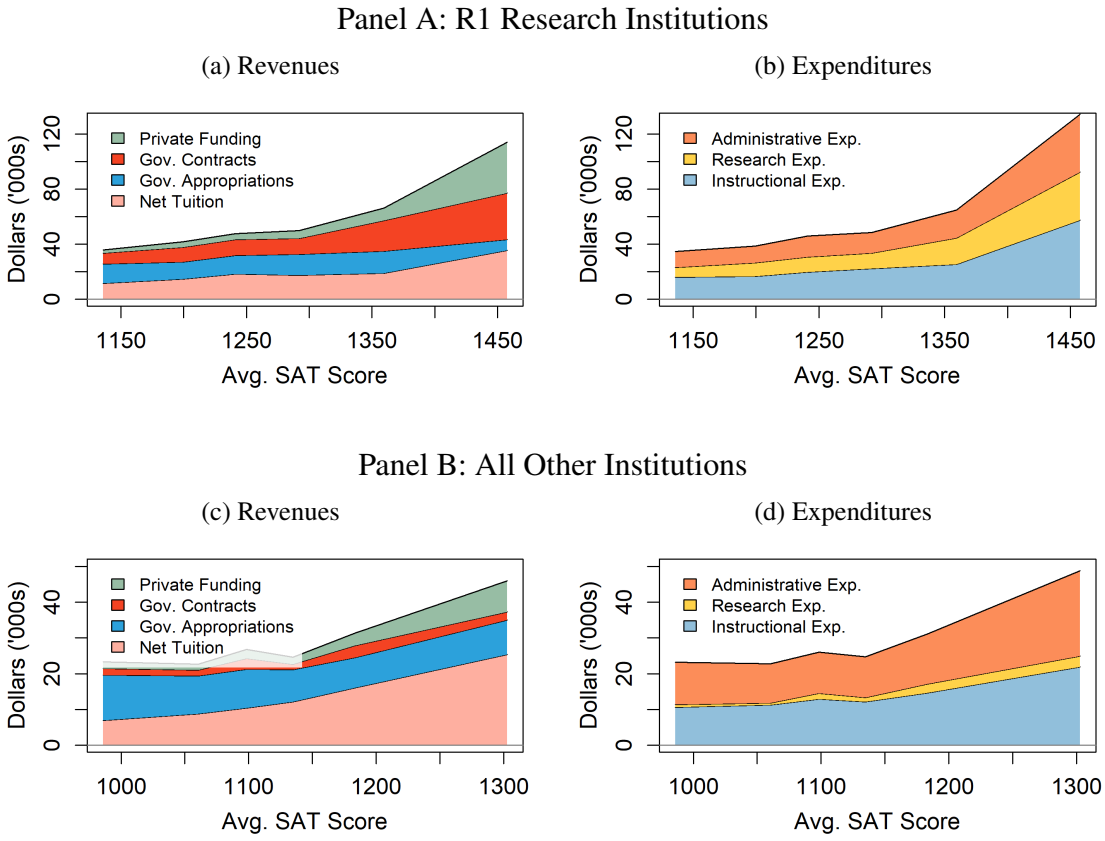
Note: This figure organizes universities by SAT quantile instead of tier, showing that revenues and expenditures sharply rise by average SAT score. The 1,162 four-year institutions with reported test scores, enrollments, revenues, and expenditures are divided into eight quantile bins by their average 2019 SAT scores; this figure shows mean expenditures and revenues per FTE pupil by SAT bin. See notes to Figure 1 and A-1 for definitions of expenditure and revenue categories. Institutional average SAT score is the sum of the averaged 25th and 75th percentile of SAT math and verbal exam scores among enrolled freshman students, with equating for ACT schools as discussed in Appendix A. All dollar amounts are CPI-adjusted to 2025. Schools' revenues may not exactly equal expenditures due to cross-subsidization from hospitals and auxiliary enterprises (like dormitories and dining halls) and intertemporal substitution. Source: IPEDS.

Figure A-3: Universities' 2019 Per-Student Revenues and Expenditures by Average SAT Score, **by Public/Private**



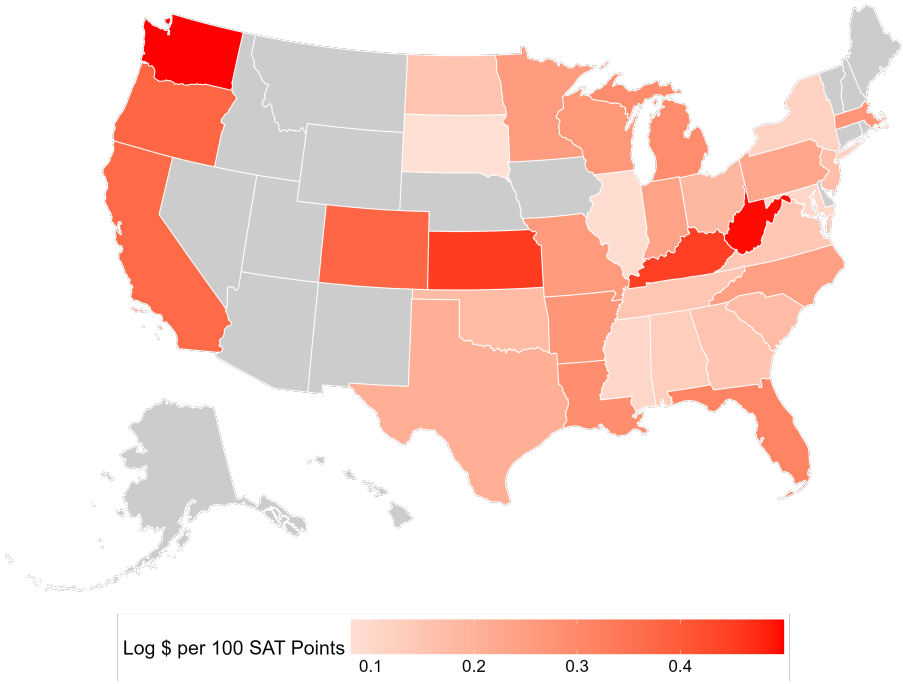
Note: This figure shows reasonably similar revenue and expenditure patterns by institutions' average SAT scores between public and private institutions, though the differences are larger at private universities and the higher per-student revenues of high-SAT public universities are driven relatively more by research-oriented government contracts and less by private funding. Average 2019 per-student revenues and expenditures in primary categories by eight quantile bins of institutions' average 2019 SAT score, among the 434 public and 728 private institutions with reported test scores, enrollments, revenues, and expenditures. 'Net Tuition' is gross tuition payments minus gross financial aid (from both government and institutional sources, including non-tuition aid). 'Government Appropriations and Grants' include both direct appropriations and all governmental financial aid programs (like Pell and state grant programs). 'Government Contracts' include all other government grants and contracts, largely for research purposes. 'Private Funding' includes private gifts, grants, and contracts, contributions from affiliated entities like booster clubs, and investment returns. 'Instructional expenditures' include expenditures for general academic instruction and departmental research. 'Instructional Expenditures' include expenditures for general academic instruction and departmental research. 'Research Expenditures' include all expenditures for research commissioned by an external organization or an internal organizational unit. 'Administrative Expenditures' include "Academic Support" (like the library), "Student Services" (like the Registrar and athletics), and "Institutional Support" (like general management and HR). FTE students are defined as full-time (undergraduate and postgraduate) students plus one-third of part-time students. Average SAT score is the sum of the averaged 25th and 75th percentile of SAT math and verbal exam scores among enrolled freshman students. Private institutions include all non-public institutions. All dollar amounts are CPI-adjusted to 2025. Source: IPEDS.

Figure A-4: Universities' 2019 Per-Student Revenues and Expenditures by Average SAT Score, **by Research Status**



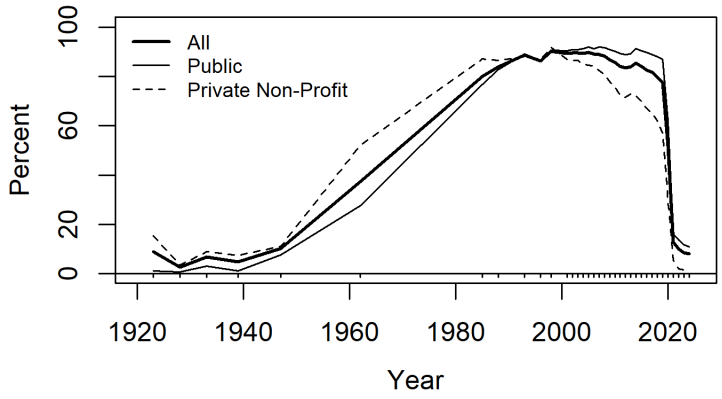
Note: This figure shows that the strong relationships between institutions' average SAT scores and both overall revenues/expenditures and instructional expenditures are present for both research-oriented R1 and non-research-oriented universities, suggesting that high-testing universities' research mission is not the sole driver of stratification across higher- and lower-testing universities. Average 2019 per-student revenues and expenditures in primary categories by eight quantile bins of institutions' average 2019 SAT score, among the 141 institutions classified by the Carnegie Foundation as "R1: Doctoral/Research Universities–Extensive" in 2000, and 1,021 non-R1 institutions with reported test scores, enrollments, revenues, and expenditures. 'Net Tuition' is gross tuition payments minus gross financial aid (from both government and institutional sources, including non-tuition aid). 'Government Appropriations and Grants' include both direct appropriations and all governmental financial aid programs (like Pell and state grant programs). 'Government Contracts' include all other government grants and contracts, largely for research purposes. 'Private Funding' includes private gifts, grants, and contracts, contributions from affiliated entities like booster clubs, and investment returns. 'Instructional Expenditures' include expenditures for general academic instruction and departmental research. 'Research Expenditures' include all expenditures for research commissioned by an external organization or an internal organizational unit. 'Administrative Expenditures' include "Academic Support" (like the library), "Student Services" (like the Registrar and athletics), and "Institutional Support" (like general management and HR). FTE students are defined as full-time (undergraduate and postgraduate) students plus one-third of part-time students. Average SAT score is the sum of the averaged 25th and 75th percentile of SAT math and verbal exam scores among enrolled freshman students. All dollar amounts are CPI-adjusted to 2025. Source: IPEDS.

Figure A-5: State-Level Relationship Between Instructional Expenditures and SAT



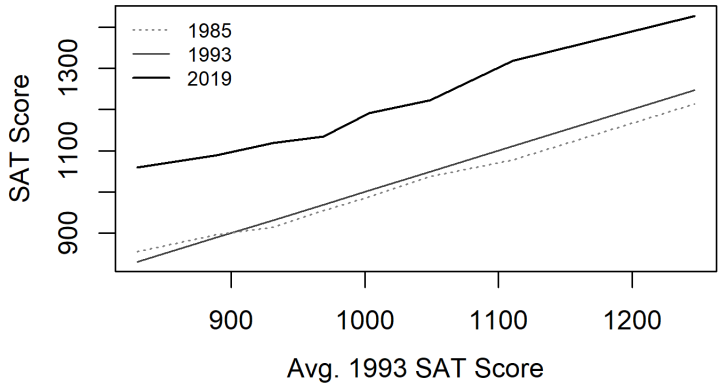
Note: This figure shows substantial heterogeneity across US states in the degree of instructional expenditure stratification between higher- and lower-testing public universities, with lower stratification in the Southeast (10-20 percent higher per-student expenditures per 100 average SAT points) and higher stratification in the West (30-50 percent per 100 SAT points). The state-level OLS coefficient from a regression of 2019 log per-student instructional expenditures on average SAT score, across enrollment-weighted four-year public institutions in the state. States with fewer than 5 public institutions with observed test scores are omitted. SAT scores are predicted within-year for institutions that only report ACT scores ($R^2 \approx 0.8$). The national relationship is 0.24 log dollars per 100 SAT points. Source: IPEDS.

Figure A-6: Share of Four-Year Institutions Requiring Tests for Admissions



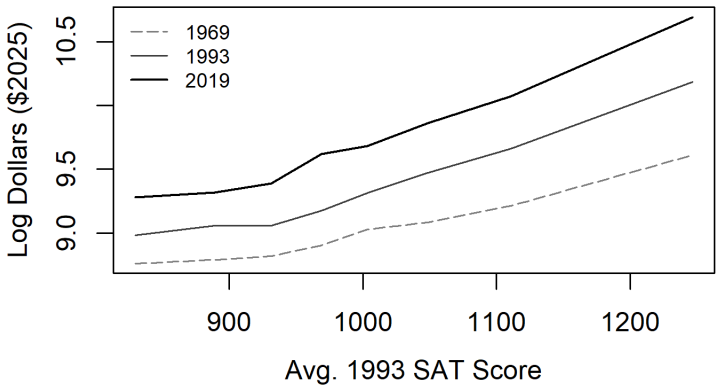
Note: This figure shows that the share of universities requiring standardized tests in admissions rose from about 10 to about 90 percent between the 1940s and the 1980s, before precipitously falling during the Covid-19 pandemic. This figure shows the enrollment-weighted share of four-year colleges and universities that require standardized tests for admission, overall and by public/private status. Statistics are available in years demarcated by ticks in the x-axis. Prior to 1960, colleges and universities are counted as requiring tests if at least 50 percent of students are admitted 'by exam' or by 'both exam and certificate.' Enrollment is defined as total enrollment prior to 1970 and total FTE enrollment thereafter. Source: College Blue Books, College Board Annual Survey of Colleges, and IPEDS.

Figure A-7: Universities' Average **SAT Score** in Different Years by Average 1993 SAT Score



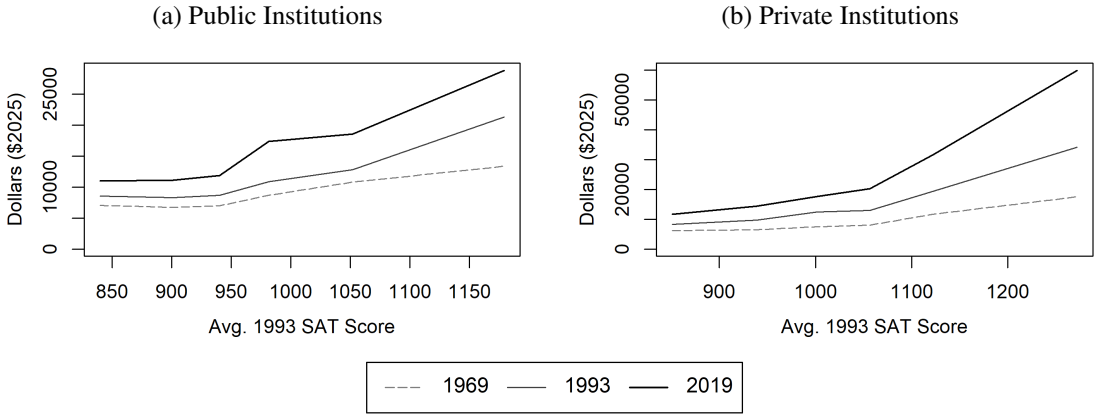
Note: This figure shows that the relative ranking of colleges and universities by average SAT score has remained largely unchanged, with very minimal mean reversion, since the mid-1980s, though average SAT scores rose considerably following College Board's 1995 'recentering'. Average freshmen-enrollment-weighted SAT score by eight quantile bins of institutions' average 1993 SAT score, among the 666 four-year institutions with reported 1985, 1993, and 2019 test scores and enrollments. FTE students are defined as full-time (undergraduate and postgraduate) students plus one-third of part-time students. Average SAT score is the sum of the averaged 25th and 75th percentile of SAT math and verbal exam scores among enrolled freshman students; for the 195 institutions where only ACT scores are reported, we impute average SAT scores by a cross-institution linear regression of average SAT score on the interaction between the 25th and 75th ACT score percentiles ($R^2 = 0.83$). Source: HEGIS, IPEDS, and the College Board Annual Survey of Colleges.

Figure A-8: Universities' Average **Log** Instructional Expenditures Per Student by Average 1993 SAT Score



Note: This figure shows that the growth in stratification of per-student instructional expenditures by institutions' average SAT score holds in percent terms as well as real dollar terms (Figure 4). Average freshmen-enrollment-weighted log instructional expenditures per FTE student by eight quantile bins of institutions' average 1993 SAT score, among the 157 public and 384 private four-year institutions with reported 1993 test scores and 1969, 1993, and 2019 enrollments and instructional expenditures. Instructional expenditures include expenditures for general academic instruction and departmental research and are CPI-adjusted to 2025. FTE students are defined as full-time (undergraduate and postgraduate) students plus one-third of part-time students. Average SAT score is the sum of the averaged 25th and 75th percentile of SAT math and verbal exam scores among enrolled freshman students; for the 165 institutions where only ACT scores are reported, we impute average SAT scores by a cross-institution linear regression of average SAT score on the interaction between the 25th and 75th ACT score percentiles ($R^2 = 0.83$). Source: HEGIS, IPEDS, and the College Board Annual Survey of Colleges.

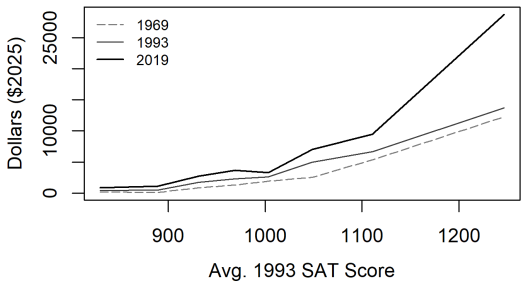
Figure A-9: Universities' Per-Student Instructional Expenditures by Average 1993 SAT Score **By Control**



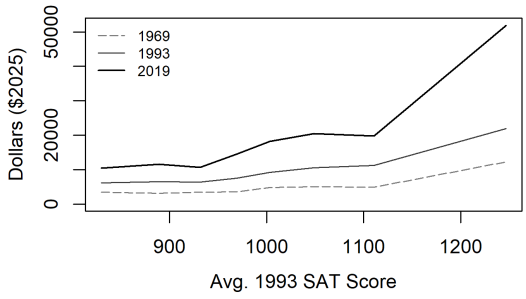
Note: This figure shows that the growth in stratification of per-student instructional expenditures by institutions' average SAT score has occurred at both public and private institutions since 1969, though to a greater degree at private institutions. Average freshmen-enrollment-weighted instructional expenditures per FTE student by eight quantile bins of institutions' average 1993 SAT score and by whether the institution is public or private, among the 559 (169 public and 390 private) four-year institutions with reported 1993 test scores and 1969, 1993, and 2019 enrollments and instructional expenditures. Instructional expenditures include expenditures for general academic instruction and departmental research and are CPI-adjusted to 2025. FTE students are defined as full-time (undergraduate and post-graduate) students plus one-third of part-time students. Average SAT score is the sum of the averaged 25th and 75th percentile of SAT math and verbal exam scores among enrolled freshman students; for the 165 institutions where only ACT scores are reported, we impute average SAT scores by a cross-institution linear regression of average SAT score on the interaction between the 25th and 75th ACT score percentiles ($R^2 = 0.83$). Private institutions include all non-public institutions. Source: HEGIS, IPEDS, and the College Board Annual Survey of Colleges.

Figure A-10: Universities' Average Other Per-Student Expenditures by Average 1993 SAT Score

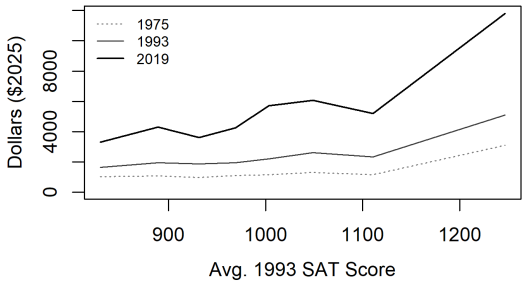
(a) Research Expenditures



(b) All Administrative Expenditures

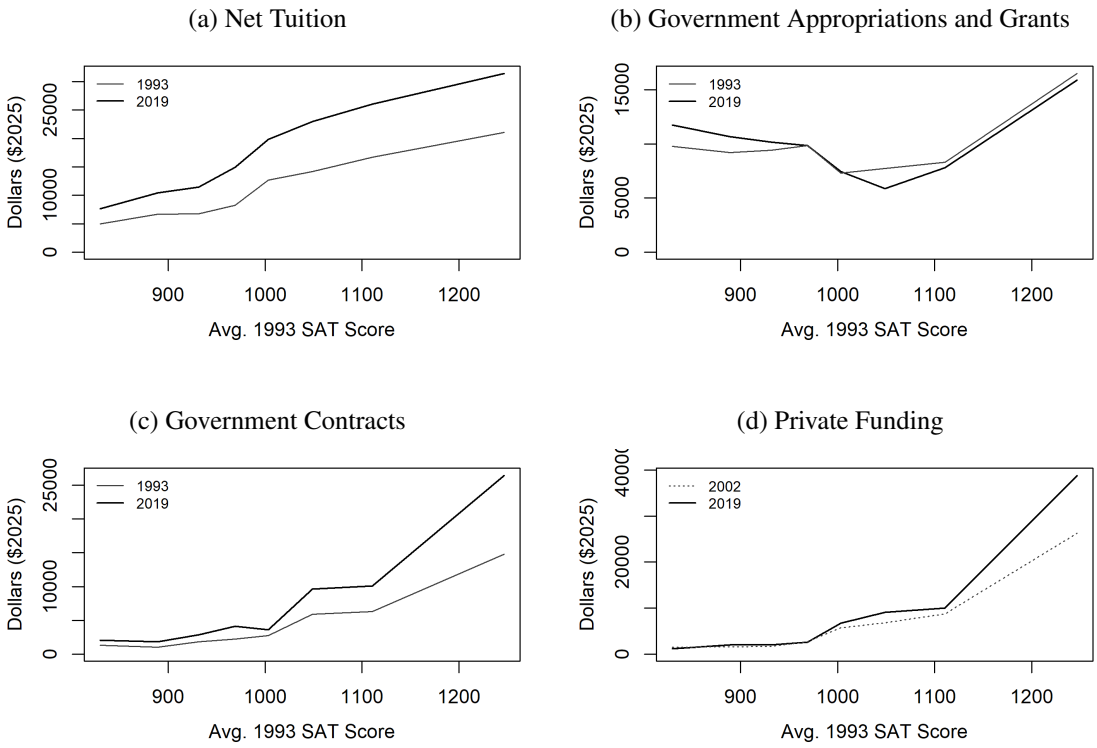


(c) Student Services Admin. Expenditures



Note: This figure shows that per-student research and administrative expenditures have also stratified by institutions' average SAT score since 1969, especially at the top quantile of institutions and largely since 1993. Average freshmen-enrollment-weighted expenditures per FTE student by eight quantile bins of institutions' average 1993 SAT score, among the 559 four-year institutions with reported 1993 test scores and 1969 (or 1975 for student services), 1993, and 2019 enrollments and expenditures. 'Research Expenditures' include all expenditures for research commissioned by an external organization or an internal organizational unit. 'Administrative Expenditures' include "Academic Support" (like the library), "Student Services" (like the Registrar and athletics), and "Institutional Support" (like general management and HR). 'Student Service Expenditures' include only administrative expenditures that support student development and well-being ("Student Services"), including admissions, registrar, financial aid, counseling, and health services. All expenditures are CPI-adjusted to 2025. FTE students are defined as full-time (undergraduate and post-graduate) students plus one-third of part-time students. Average SAT score is the sum of the averaged 25th and 75th percentile of SAT math and verbal exam scores among enrolled freshman students; for the 165 institutions where only ACT scores are reported, we impute average SAT scores by a cross-institution linear regression of average SAT score on the interaction between the 25th and 75th ACT score percentiles ($R^2 = 0.83$). Source: HEGIS, IPEDS, and the College Board Annual Survey of Colleges.

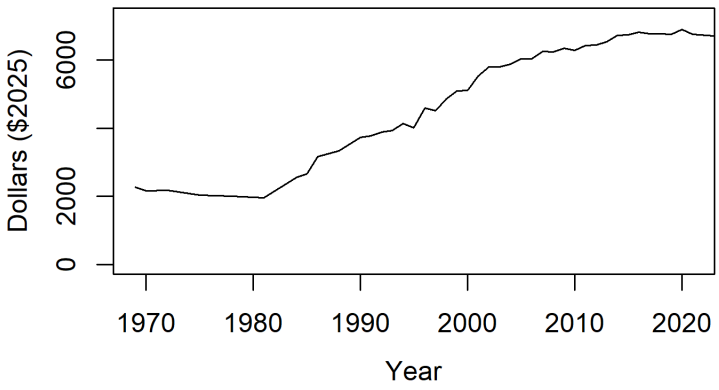
Figure A-11: Universities' Average Revenues Per Student by Source and Average 1993 SAT Score



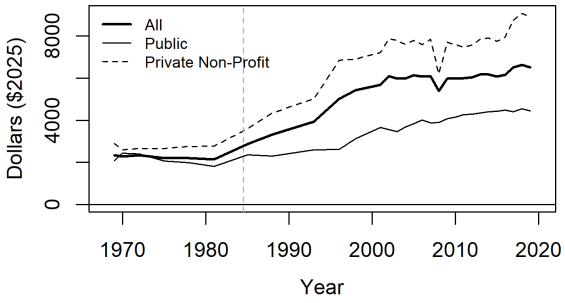
Note: This figure shows that the relative tuition, government research funding, and private funding received by higher-testing universities has risen since the mid-1990s, whereas government appropriations and grants have become slightly more evenly distributed over that period. Average freshmen-enrollment-weighted source-specific revenues per FTE student by eight quantile bins of institutions' average 1993 SAT score, among the 559 four-year institutions with reported 1993 test scores and 1969, 1993, and 2019 enrollments and revenues. Due to changes in revenue reporting, no revenue data are available prior to 1988. 'Net Tuition' is gross tuition payments minus gross financial aid (from both government and institutional sources, including non-tuition aid). 'Government Appropriations and Grants' include both direct appropriations and all governmental financial aid programs (like Pell and state grant programs). 'Government Contracts' include all other government grants and contracts, largely for research purposes. 'Private Funding' includes private gifts, grants, and contracts, contributions from affiliated entities like booster clubs, and investment returns. FTE students are defined as full-time (undergraduate and postgraduate) students plus one-third of part-time students. Average SAT score is the sum of the averaged 25th and 75th percentile of SAT math and verbal exam scores among enrolled freshman students; for the 165 institutions where only ACT scores are reported, we impute average SAT scores by a cross-institution linear regression of average SAT score on the interaction between the 25th and 75th ACT score percentiles ($R^2 = 0.83$). Source: IPEDS and the College Board Annual Survey of Colleges.

Figure A-12: Annual Association Between Per-Student Instructional Exp. and Average SAT Scores

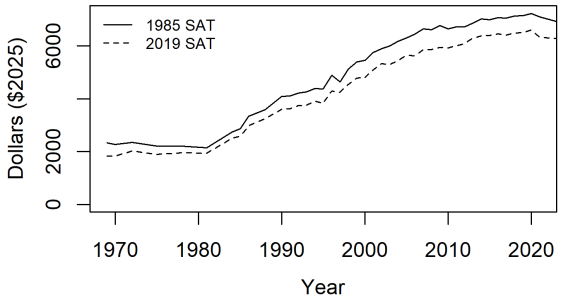
(a) SAT Fixed in 1993



(b) Contemporaneous SAT

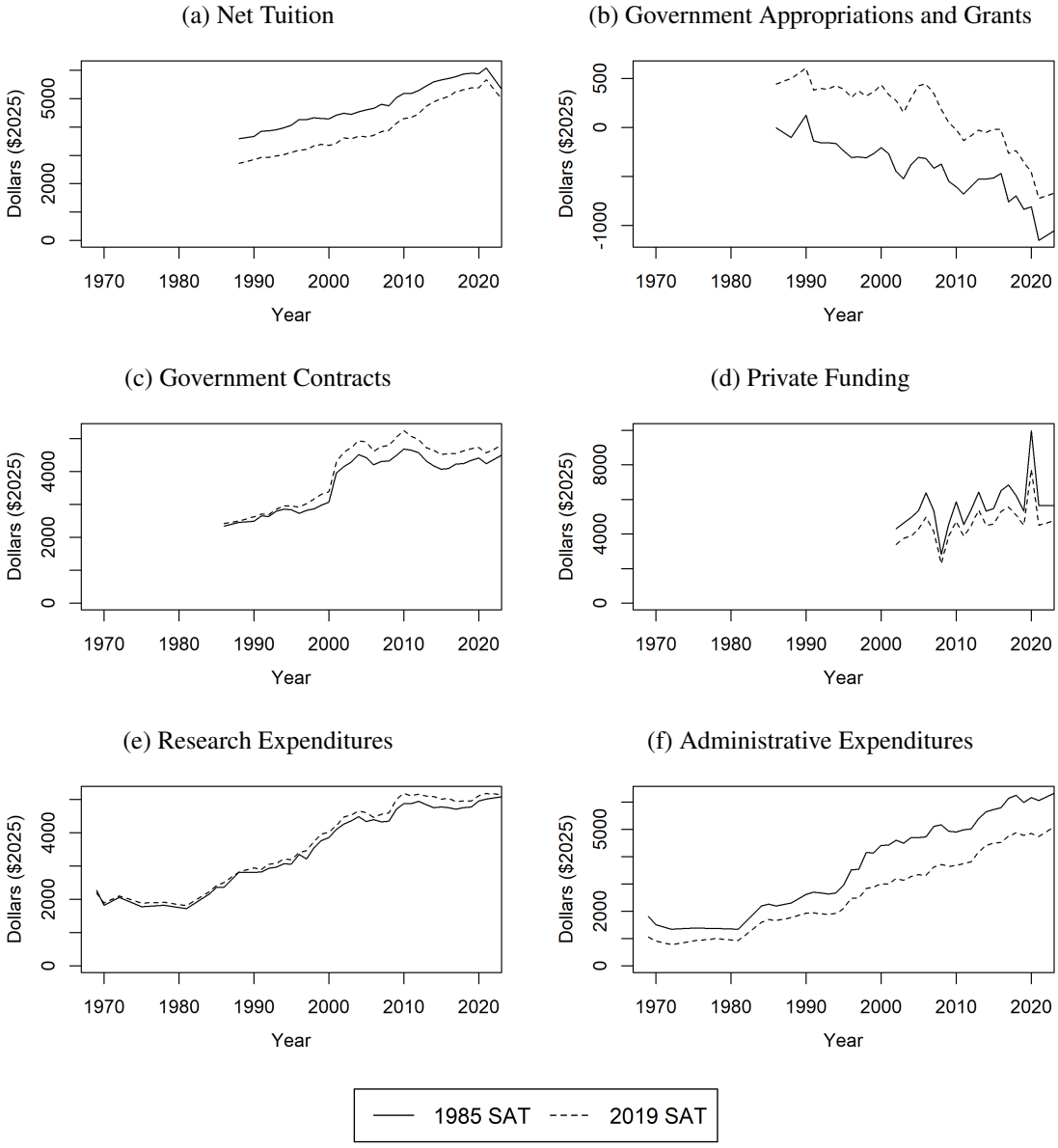


(c) SAT in Other Fixed Years



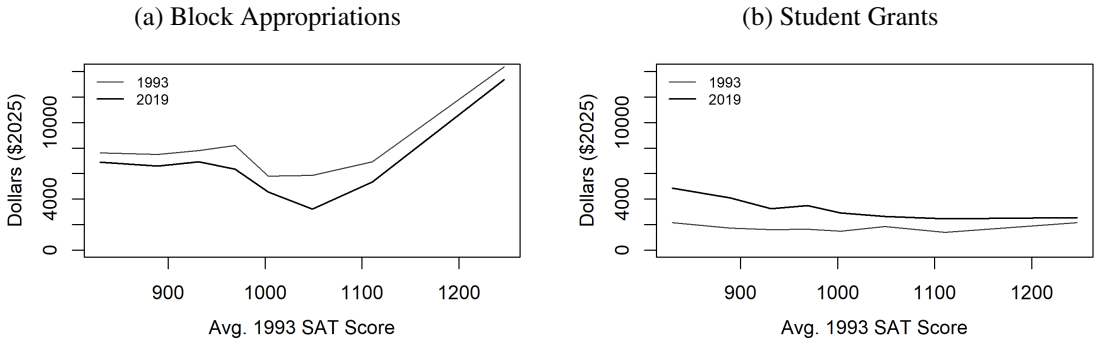
Note: This figure presents the annual student-weighted cross-institution relationship of per-student instructional expenditures per 100 additional points of 1993 SAT score, showing that the relationship considerably strengthens between 1980 and 2010 before flattening in recent years. It then recapitulates the pattern using contemporaneous SAT scores (after 1985) overall and by public-private control as well as using 1985 or 2019 SAT scores, showing a similar trend overall (though the leveling-out may have occurred slightly before 2010) and in each sector (though the overall association is consistently much weaker in the public sector). Annual freshman-enrollment-weighted OLS relationship between average instructional expenditures per FTE student (in 2025 dollars) and either contemporaneous average SAT score divided by 100, across four-year institutions overall and by public/private control, or average 1985, 1993, or 2019 SAT score divided by 100. For example, in 2019, a 100 point increase in universities' average SAT score was associated with higher per-student instructional expenditures of about $6,300. 'Instructional expenditures' include expenditures for general academic instruction and departmental research. FTE students are defined as full-time (undergraduate and postgraduate) students plus one-third of part-time students. Average SAT score is the sum of the averaged 25th and 75th percentile of SAT math and verbal exam scores among enrolled freshman students; for institutions where only ACT scores are reported, we impute average SAT scores by a cross-institution linear regression of average SAT score on the interaction between the 25th and 75th ACT score percentiles (R^2 is between 0.77 and 0.83). Prior to 1985 (the dotted line), SAT scores are not observed and are replaced by 1985 scores. All dollar amounts are CPI-adjusted to 2025 and annually winsorized at 1 percent. Source: HEGIS, IPEDS, and the College Board Annual Survey of Colleges.

Figure A-13: Annual Association Between Per-Student Revenues and Expenditures and Average SAT Score



Note: This figure shows that while the association between SAT scores and both per-student net tuition and private funding continued to rise in the 2010s, an accelerated decline in the SAT association with per-student government appropriations and grants may help to explain the leveling-out of the SAT association with per-student instructional expenditures in that period. Annual freshman-enrollment-weighted OLS association between per-student revenues or expenditures per FTE student and average 1985 or 2019 SAT score across four-year institutions. 'Government Appropriations and Grants' include both direct appropriations and all governmental financial aid programs (like Pell and state grant programs). 'Government Contracts' include all other government grants and contracts, largely for research purposes. 'Private Funding' includes private gifts, grants, and contracts, contributions from affiliated entities like booster clubs, and investment returns. 'Research Expenditures' include all expenditures for research commissioned by an external organization or an internal organizational unit. 'Administrative Expenditures' include "Academic Support" (like the library), "Student Services" (like the Registrar and athletics), and "Institutional Support" (like general management and HR). FTE students are defined as full-time (undergraduate and postgraduate) students plus one-third of part-time students. FTE students are defined as full-time (undergraduate and postgraduate) students plus one-third of part-time students. Average SAT score is the sum of the averaged 25th and 75th percentile of SAT math and verbal exam scores among enrolled freshman students; for institutions where only ACT scores are reported, we impute average SAT scores by a cross-institution linear regression of average SAT score on the interaction between the 25th and 75th ACT score percentiles (R^2 is between 0.77 and 0.83). All dollar amounts are CPI-adjusted to 2025 and annually winsorized at 1 percent. Source: HEGIS, IPEDS, and the College Board Annual Survey of Colleges.

Figure A-14: Universities' Average Government Revenues Per Student by Average 1993 SAT Score



Note: This figure shows that the reason government appropriations have become less correlated across universities with average SAT scores since the 1990s is that growth in state financial aid programs (like the Excelsior Scholarship and CalGrant) has provided disproportionate funding to lower-testing institutions. Average freshmen-enrollment-weighted source-specific revenues per FTE student by eight quantile bins of institutions' average 1993 SAT score, among the 559 four-year institutions with reported 1993 test scores and 1969, 1993, and 2019 enrollments and revenues. 'Block Appropriations' include direct appropriations from federal, state, and local governments to colleges and universities; 'Student Grants' include Pell grants and all other federal, state, and local financial aid programs that subsidize student tuition. The sum of these two equals total 'Government Appropriations and Grants' as defined elsewhere. FTE students are defined as full-time (undergraduate and postgraduate) students plus one-third of part-time students. Average SAT score is the sum of the averaged 25th and 75th percentile of SAT math and verbal exam scores among enrolled freshman students; for the 165 institutions where only ACT scores are reported, we impute average SAT scores by a cross-institution linear regression of average SAT score on the interaction between the 25th and 75th ACT score percentiles ($R^2 = 0.83$). Source: IPEDS and the College Board Annual Survey of Colleges.

Table A-1: Changes In Cross-University Per-Student Total Current Revenues Over Time

	1929	1969	1984	1996	2008	2019
Panel A: Average Total Non-Hospital Revenues Per Student ($2025)						
All Four-Year Institutions	10,500	24,300	32,700	40,500	48,800	57,800
Ivy League	20,700	89,200	107,100	144,300	229,500	264,800
State Flagship	12,500	38,900	42,600	56,900	73,400	83,900
Public Non-Flagship	10,500	18,800	29,500	34,100	39,900	45,200
Ivy vs. Flagship Gap (%)	65	129	151	154	213	215
Flagship vs. Non-Flagship Gap[1] (%)	9	54	51	61	70	69
Panel B: Coefficient of Variation of Per-Student Total Non-Hospital Current Revenues (%)						
Public Institutions	46	45	56	58	56	54
All Four-Year Institutions	58	51	52	59	59	59

Note: This table mirrors Table 2 but replaces per-student total current revenues with per-student instructional expenditures, showing similar trends overall and by tier over the past 100 years. Panel A shows the enrollment-weighted average total current revenues per FTE student at all, Ivy League, state flagship, and public non-flagship four-year institutions for six selected years. Panel B shows the across-institution coefficient of variation (the standard deviation divided by the mean) for the per-student total current revenues of public four-year institutions and all four-year institutions, weighted by enrollment. All dollars are reported in CPI-adjusted 2025 dollars. Total revenues exclude revenues from hospitals, revenues earmarked for capital expenditures, and direct endowment contributions. All series are winsorized at 5 percent to limit the contribution of any single university. Enrollment is measured as total enrollment in 1929 and total FTE enrollment in other years (counting part-time enrollees as 1/3). State flagship public institutions are defined in Table A-4.

[1]The flagship/non-flagship gap is measured via an OLS regression of log expenditures on a flagship indicator, including state fixed effects to isolate within-state comparisons.

Source: Biennial Survey of Education, 1928-1930 (Cooper, 1931), HEGIS, and IPEDS.

Table A-2: Racial Composition by Tier Over Time

	1968	1976	1984	1996	2008	2019
Panel A: Percent of Students Black						
All	6.0†	9.0	8.7	10.3	11.5	10.7
Ivy League	2.7	6.0	5.7	6.1	7.3	7.2
Other Top-50 Private	2.5	5.5	4.5	5.4	5.7	5.8
State Flagship	1.6	3.6	4.1	5.3	5.4	5.2
Panel B: Percent of Students Hispanic						
All	1.8†	2.5	3.2	5.9	7.9	15.1
Ivy League	0.6	2.0	3.3	5.6	7.2	12.2
Other Top-50 Private	0.7	2.4	3.2	6.1	7.1	12.1
State Flagship	1.0	1.6	2.3	4.4	5.8	10.2
Panel C: Percent of Students Asian						
All	1.1†	1.6	3.0	5.8	6.4	7.5
Ivy League	1.0	2.7	6.9	15.9	16.2	19.6
Other Top-50 Private	1.5	3.0	5.9	14.3	15.2	17.2
State Flagship	1.3	2.8	4.2	8.2	9.1	10.3

Note: This table presents numerical statistics mirroring those shown in Figure 3. Share of full-time undergraduate students who are Black, Hispanic, or Asian at all four-year universities and by university tier. Tables A-3 and A-4 identify the top-50 private and state flagship universities. †'All four-year universities' includes some two-year institutions (those that do not have 'junior' or 'community' in their name) in 1968.
Source: McVay (1969), HEGIS, and IPEDS.

Table A-3: Top-50 Private Institutions

State	University	UnitID	State	University	UnitID
CA	California Institute of Technology	110404	MA	Williams College	168342
CA	Claremont Mckenna College	112260	MD	Johns Hopkins University	162928
CA	Harvey Mudd College	115409	ME	Colby College	161086
CA	Pomona College	121345	MN	Carleton College	173258
CA	Scripps College	123165	MO	Washington University in St Louis	179867
CA	Stanford University	243744	NC	Duke University	198419
CA	University of Southern California	123961	NH	Dartmouth College*	182670
CT	Yale University*	130794	NJ	Princeton University*	186131
DC	Georgetown University	131496	NJ	Stevens Institute of Technology	186867
GA	Emory University	139658	NY	Barnard College	189097
GA	Emory University	487092	NY	Columbia University*	190150
IA	Grinnell College	153384	NY	Cornell University*	190415
IL	Northwestern University	147767	NY	Hamilton College	191515
IL	University of Chicago	144050	NY	New York University	193900
IN	University of Notre Dame	152080	NY	Vassar College	197133
LA	Tulane University of Louisiana	160755	OH	Case Western Reserve University	201645
MA	Boston College	164924	PA	Carnegie Mellon University	211440
MA	Boston University	164988	PA	Haverford College	212911
MA	Brandeis University	165015	PA	Swarthmore College	216287
MA	Franklin W Olin College of Engineering	441982	PA	University of Pennsylvania*	215062
MA	Harvard University*	166027	RI	Brown University*	217156
MA	Massachusetts Institute of Technology	166683	TN	Vanderbilt University	221999
MA	Northeastern University	167358	TX	Rice University	227757
MA	Tufts University	168148	VA	Washington And Lee University	234207
MA	Wellesley College	168218	VT	Middlebury College	230959

Note: This table lists the 50 private institutions with the highest summed 25th percentile SAT math and reading scores in 2019 (with the cutoff at 1340). Asterisks denote the eight Ivy League institutions; the other colleges of Cornell (190433) are also considered part of the Ivy League. Two public universities – the Universities of Michigan and Virginia – also have scores above this cutoff but are omitted from this tier.

Table A-4: State Flagship Universities

State	University	UnitID	State	University	UnitID
AK	University of Alaska, Fairbanks*	102614	MT	University of Montana, Missoula*	180489
AL	University of Alabama	100751	NC	University of North Carolina, Chapel Hill	199120
AR	University of Arkansas	106397	ND	University of North Dakota*	200280
AZ	University of Arizona	104179	NE	University of Nebraska, Lincoln	181464
CA	University of California, Berkeley	110635	NH	University of New Hampshire	183044
CO	University of Colorado, Boulder	126614	NJ	Rutgers University, New Brunswick	186380
CT	University of Connecticut	129020	NM	University of New Mexico	187985
DE	University of Delaware	130943	NV	University of Nevada, Reno	182290
FL	University of Florida	134130	NY	SUNY Stony Brook	196097
GA	University of Georgia	139959	OH	Ohio State University	204796
HI	University of Hawaii, Manoa	141574	OK	University of Oklahoma, Norman	207500
IA	University of Iowa	153658	OR	University of Oregon	209551
ID	University of Idaho	142285	PA	Pennsylvania State University	214777
IL	University of Illinois, Urbana-Champaign	145637	RI	University of Rhode Island	217484
IN	Indiana University, Bloomington	151351	SC	University of South Carolina, Columbia	218663
KS	University of Kansas	155317	SD	University of South Dakota*	219471
KY	University of Kentucky	157085	TN	University of Tennessee	221759
LA	Louisiana State University	159391	TX	University of Texas, Austin	228778
MA	University of Massachusetts, Amherst	166629	UT	University of Utah	230764
MD	University of Maryland, College Park	163286	VA	University of Virginia	234076
ME	University of Maine	161253	VT	University of Vermont	231174
MI	University of Michigan, Ann Arbor	170976	WA	University of Washington	236948
MN	University of Minnesota, Twin Cities	174066	WI	University of Wisconsin, Madison	240444
MO	University of Missouri, Columbia	178396	WV	West Virginia University	238032
MS	University of Mississippi	176017	WY	University of Wyoming	240727

Note: This table lists the designated flagship institution of each US state, following the College Board designation. SUNY Stony Brook was founded in 1957; the University at Buffalo is omitted as a second flagship institution in New York. Rutgers University was a private institution prior to 1945, but remains designated as the flagship in earlier years. Universities marked by asterisks are designated as R2 institutions; all others are R1.