The Meritocratic Consensus and Stratification in Higher Education

Zachary Bleemer and Jesse Rothstein*

Introduction

Alice and Bob are high school seniors. Alice has a 1300 SAT score (the 85th percentile among all test-takers), a correspondingly high GPA, and leadership roles in several extracurricular activities. She comes from an upper-middle-class family with two college-educated parents. She has carefully planned her college application strategy with the help of a college counselor. Bob's SAT score is 1050, just above the national median, and his GPA is also lower. Neither of his parents attended college and he is unfamiliar with the college application process. Not surprisingly, Alice is a more competitive applicant. She is admitted to many selective schools and chooses to attend her state's flagship public university, or maybe a prestigious, research-intensive private university. Bob attends his local state college, which admits all students deemed academically college-ready.

Alice's future classmates will have strong academic backgrounds like hers, while Bob's peers will be socioeconomically diverse but not especially distinguished academically. Moreover, Bob's school is much less well funded than Alice's. Per-student expenditures at Alice's school are likely between $50,000 (if public) and $140,000 (if private), while Bob's school spends only $20,000. This in part reflects a much larger research footprint at Alice's school, but even per-student instructional expenditures are between 2 and 5 times higher for Alice. Alice will pay higher tuition than Bob, though not nearly to the same degree.

Not surprisingly, Alice is much more likely than Bob to graduate from college and to achieve a high-earning career. This is partly because of Alice's talents and her family's pre-college investments in her human capital, but the better resources of her university also play an important role.

The modern higher education ecosystem came into being in the period following World War II. The rise of federal scientific research funding meant that some institutions became much more research intensive, with dramatic increases in funding and scientific activity. Simultaneously, to accommodate a perceived need to send many more students to postsecondary education than had attended previously, states created a large number of teaching-focused public colleges and universities. This created the need to answer a question: Which students should attend which institutions? In this paper, we review what we call the "meritocratic consensus" in postsecondary education: the idea that students with greater pre-college academic achievement should attend better-funded and higher-quality colleges and universities. This idea has been the organizing principle of U.S. higher education since the middle of the 20th century, but it has been largely taken for granted by economists and others. How stratified should institutions be, and what are the ben-

* Bleemer: Princeton University, Princeton NJ, bleemer@princeton.edu. Rothstein: University of California, Berkeley, Berkeley, CA, rothstein@berkeley.edu. Thanks to Sarah Turner for helpful comments and Dylan Black, Ivy Chen, and Shinrea Su for excellent research assistance.

efits and costs of allocating high-testing students to high-resource universities? How has the meritocratic consensus evolved in recent years, and how is it likely to change in the decades to come?

Section 1 describes relevant aspects of the U.S. higher education system. Colleges and universities in the United States are highly heterogeneous, ranging from small, teaching-focused institutions that in many ways resemble high schools to city-sized behemoths with tens of thousands of students. We highlight two important dimensions along which institutions vary: greater and lesser selectivity and higher vs. lower per-student educational funding. We show that students are highly stratified across both of these dimensions: students from more advantaged backgrounds and students with stronger pre-college academic achievement attend distinctly different institutions, with much higher instructional expenditures, than do students from less advantaged backgrounds and those with weaker achievement.

Many discussions of higher education focus on the Ivy League and other highly selective private colleges. Stratification is extreme at those institutions: the most selective schools have large endowments and very high per-student expenditures. But the 50 highest-scoring private universities enroll only 4 percent of first-year university students. Public institutions are much more numerically important: research-intensive public universities enroll one-third of students who attend four-year colleges and universities. These institutions are also stratified, relative both to each other and to less research-intensive public colleges that spend less and enroll lower-scoring students. Endowments are less important here, but the more selective institutions benefit from higher tuition, larger state institutional subsidies, and more spillovers from research to instruction.

Section 2 provides a history of the "meritocratic consensus" in US higher education. A century ago, the American higher education sector was much smaller. It was far less segregated by academic achievement and exhibited much less cross-university variation in per-student revenues and expenditures. In the mid-20th century, however, higher education enrollment grew dramatically. Demand for a small set of colleges and universities grew well beyond their capacity. Rather than setting market-clearing prices, these institutions implemented non-market admissions policies to select their enrollees. Early proponents of meritocratic selection, like Harvard's James Conant, advocated for segregating students by academic achievement.[1] As late as 1960, however, this was relevant only at a small number of institutions; most institutions remained effectively open-access. Moreover, the growth in stratification by prior achievement was not yet accompanied by growing variation in expenditures.

This began to change in the 1960s. California's 1960 Master Plan for Higher Education was an influential marker. Public institutions stratified into tiers – flagship research universities at the top, two-year institutions at the bottom – and they adopted private universities' achievement-based selection criteria to ration access to the schools in the upper ranks. By 1980, when the new sector of community colleges had grown to accommodate a large number of high school graduates who would not previously have been thought college-ready, educational resources and quality were as aligned with students' academic achievement across public institutions as in the much smaller private sector. This stratification then intensified, in both public and private institutions, from the 1980s until the early 2010s.

Section 3 discusses the role of non-academic criteria in admissions. Contemporary universities have many competing objectives in choosing who to admit, including preferences for athletic talent, family con-

[1] Conant saw standardized tests as measuring student 'aptitude' rather than 'achievement.' The distinction between the two has always been fraught and highly debated. When first developed, the SAT was named the "Standardized Aptitude Test." It was later renamed the "Standardized Achievement Test" and then simply the "SAT."

nections, potential donations, and various dimensions of diversity. These objectives tend to cut against the admission of the most academically accomplished students by traditional metrics, but are often couched in the language of identifying students' true 'merit.' In any event, they are quantitatively quite small; relatively few students are shifted toward higher-expenditure institutions by policies like athletic 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 in any event is no longer practiced following the 2023 *Students for Fair Admissions (SFFA) v. Harvard* decision.

Section 4 discusses potential economic justifications for a highly stratified university market structure. Under what circumstances would it be desirable to spend more money on the instruction of high-achieving students than on that of students with worse high school achievement? We discuss several possible models. One is based on a view of the production function where student achievement and college resources are q-complements: students with stronger prior achievement may benefit more from smaller classes and from access to cutting-edge research. There is surprisingly little evidence in support of this view; instead, most empirical scholarship suggests the exact opposite, that achievement and resources are *substitutes* in production. A second possibility is that the social welfare function is highly convex in student outcomes, such that there is 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 productively invest in cognitive skill development. We review the limited empirical evidence on each of these three hypotheses. We conclude that there is considerable room for future research in evaluating the efficiency of the meritocratic consensus.

Section 5 discusses 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, the relationship has leveled out since 2010, due in part to growth of state financial aid programs that effectively direct public dollars toward the institutions that lower-income students attend. Recent federal actions targeting high-expenditure institutions – including threats to university research funding and the recent increase in the private university endowment tax – may compound this trend. On the side of student allocation, the rising prevalence of 'test-optional' and 'test-free' admissions policies suggest that university priorities may be shifting away from an admissions regime dominated by preferences over academic achievement. This may in part be a consequence of the 2023 *SFFA* decision, which removed a valued tool used by universities to simultaneously achieve their achievement and diversity objectives.[2] However, pressure from the federal government and supporters of 'meritocratic' allocation may stem this trend.[3]

The most striking trend in the stratification we discuss in recent years, however, is its movement *inside* colleges and universities, particularly at public institutions. These institutions have increasingly created

[2] An influential model from Chan and Eyster (2003) suggests that colleges prohibited from the consideration of race will reduce their reliance on pre-college academic achievement as a way of maintaining racial diversity, and there is some indication that this has happened in states that banned affirmative action earlier (e.g. Antonovics and Backes, 2014). Geiser (2017) argues that affirmative action allows admissions offices to offset racial gaps in SAT scores and that, as a result, the case against using these scores in admissions is strengthened by the *SFFA* decision.

[3] Recent demands that the Trump Administration has made of elite universities, like Harvard and Columbia, have included requirements that admissions be "merit-based" and that the institutions report admissions rates broken down by standardized test scores. A February 2025 "Dear Colleague" letter from the U.S. Department of Education asserts that it would be unlawful for an institution to "eliminate standardized testing...to increase racial diversity."

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	67,100	37,300	79,600	10.0	96.1
Other Top-50 Private	3.3	15.2	1466	80.0	59,600	31,500	40,700	10.4	92.0
State Flagship	15.2	64.9	1275	71.7	21,500	14,600	14,400	15.9	78.7
Other Public R1	15.2	63.9	1232	68.4	19,300	11,400	14,000	19.5	72.5
Less-Sel. Private	27.4	67.1	1168	68.5	16,000	1,300	19,300	22.9	64.6
Less-Sel. Public	36.9	73.1	1105	62.3	11,700	1,700	9,500	23.6	54.6
For-Profit	1.2	70.6	-	44.3	7,600	0	15,500	141.5	40.3

Note: Four-year colleges and universities are divided into tiers: the eight Ivy League institutions, the other 42 'Top-50 Private' institutions with at least a 25th percentile 1340 SAT score in 2019 (see Table A-2), the 50 state flagship universities (see Table A-3), the 56 non-flagship R1 public universities, other four-year private and public universities, and all for-profit 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. '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 admissions, registrar, financial aid, counseling, and health services, as well as other categories (e.g., building maintenance) not directly related to student well-being. 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 institution, and institutions are assigned (by necessity) to the lowest tier of all institutions sharing their IRS tax ID (Chetty et al., 2020). All other statistics are measured in 2019.
Source: IPEDS and Chetty et al. (2020).

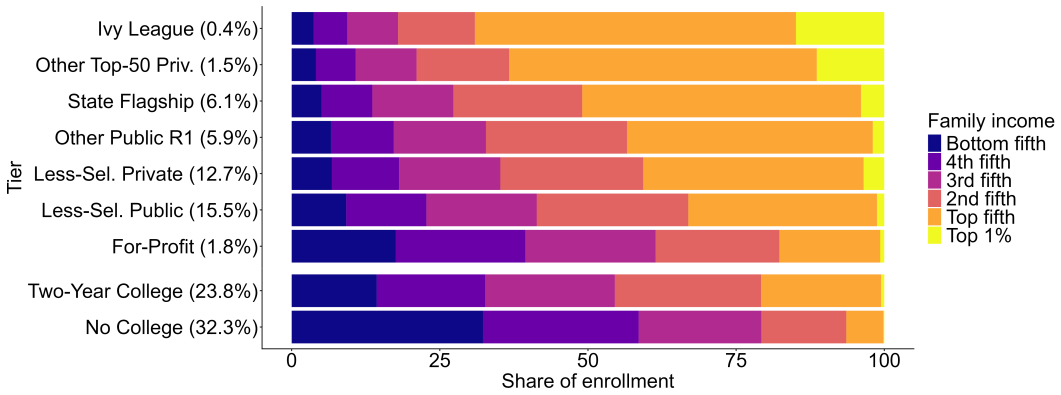
"honors colleges" and other programs designed to separate and provide additional resources to their highest-achieving students. Novel restrictions on access to popular college majors have also contributed. More than 40 percent of lucrative majors (like engineering and business) at public research universities have adopted supplemental achievement-based admissions policies over the past 50 years, stratifying students even after they've been admitted to the same university and generating equity gaps – and perhaps efficiency losses – that mirror those across universities.

Background

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% have a four-year Bachelor's degree or more (NCES, 2021).

Students receive these credentials from institutions that differ widely in their educational quality. Table 1 shows summary statistics for seven tiers of four-year colleges and universities, distinguished by their control (public vs. private), research expenditures, and selectivity. Many discussions of American higher education focus on highly-selective private institutions like the eight Ivy League universities, but those schools enroll less than 1 percent of full-time first-year American college students. The 50 private universities with the highest average SAT scores together enroll only about 4 percent of four-year college students. More

Figure 1. 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 (2,848,000; NCES, 2021); the 'no college' share is the residual. Data come from Chetty et al. (2020); parental income is measured in IRS tax records when the child was 15-19 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). For parental income statistics, institutions are assigned (by necessity) to the lowest tier of all institutions sharing their IRS tax ID; for example, the four University of Colorado campuses are reported under the same tax ID. Source: IPEDS and Chetty et al. (2020).

quantitatively important are the large, research-intensive public "flagship" universities like the University of California, Berkeley, the University of Wisconsin, and the University of Virginia. These institutions are also selective – though less so than the top-50 private colleges – and enroll about 15 percent of students, with another 15 percent enrolled at non-flagship research-intensive public universities like Utah State University and the University of Pittsburgh.[4] Almost 40 percent of students attend less-selective public institutions, with the remaining quarter attending less-selective private institutions or the very small for-profit sector.[5]

Figure 1 shows the family income distribution of enrollment across these eight tiers of US higher education, with rows showing two-year college students and non-college-goers for comparison. The more selective the college, the higher-income are its students: More than half of students at the top 50 private and state flagship universities, and roughly two-thirds of students at Ivy League institutions, come from top-quintile families, while 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% of the income distribution.

These groups of colleges differ not just in their selectivity and their student populations, but also in their funding. Figure 2 shows average instructional, research, and administrative expenditures per pupil at institutions in each tier. It is impossible to cleanly delineate different functions, but these estimates, which follow university reporting guidelines, provide a reasonable approximation.[6] There are enormous

[4]Research-intensive "R-1" institutions are designated following The Carnegie Foundation for the Advancement of Teaching (2000).

[5]The for-profit share of all first-time first-year college enrollment (including two-year schools) was 2.6 percent in 2019, down from its peak of nearly 10 percent in the late 2000s.

[6]Colleges' self-reported instructional expenditures include "compensation for academic instruction, occupational and vocational in-

differences here: the most selective schools spend many multiples of what is spent at less-selective and two-year colleges, overall and in dedicated instructional expenditures.

These college tiers combine several dimensions of institutional differences: control, research expenditures, and selectivity. Figure 3 organizes colleges into eight quantile bins by a single metric: the average 2019 SAT score of their students. The more selective institutions have much higher revenues and expenditures: those with the highest SAT scores have instructional expenditures four times larger than the average, and total expenditures that are almost five times larger. These larger expenditures are financed by substantially higher tuition funding, private funding (like alumni donations), and government research contracts; government educational appropriations are roughly evenly distributed across more- and less-selective institutions.[7] This pattern is similar among public and private universities (Figure A-1) and in both research-intensive and non-research universities (Figure A-2). At public universities, for example, per-student instructional expenditures was 24 percent higher for each 100 average SAT points in 2019.[8]

Does it matter that the colleges that enroll higher-testing (and generally higher-income) students also have the highest revenues and expenditures? Empirical evidence indicates that it does. While there is some disagreement across settings and research designs about the relationship between university selectivity and value-added (Dale and Krueger, 2002; Hoekstra, 2009; Cohodes and Goodman, 2014; Bleemer, 2022), there is broader agreement that universities with higher per-student instructional expenditures provide greater wage value-added to their students (Chetty, Deming and Friedman, 2023; Bleemer, 2021; Mountjoy and Hickman, 2021).[9] Deming and Walters (2014) provide suggestive evidence that increasing university expenditures improves students' educational outcomes, paralleling growing evidence from studies of K-12 education documenting positive effects of large expenditure increases on student test scores and longer-run outcomes (Jackson and Mackevicius, 2024; Lafortune, Rothstein and Schanzenbach, 2018; Rothstein and Schanzenbach, 2022). If this result is correct, the patterns seen in Table 1 and Figure 3 indicate that the sorting of the most successful high school students to the highest-resource institutions yields an important relative advantage for those students. Whether that is desirable depends on the degree of complementarity between resources and student type in the educational production function and on the convexity of the social welfare function. We discuss each of these below.

struction, community education, preparatory and adult basic education, and remedial and tutorial instruction conducted by the teaching faculty for the institution's students." In practice, the full salaries of teaching professors are allocated to this category, even those who also conduct research. Whether this results in overstated instructional expenditures depends on the degree to which research time should be considered a compensating amenity for teachers and on teaching-research complementarity, but in any case such overstatement is unlikely to account for high-expenditure institutions' full advantage. Even if the state flagship and other public R1 instructional expenditures are overstated by a factor of two, this would still put them well above the two-year colleges. Administrative expenditures include academic support (like the library) and student services (like the Registrar), so arguably contribute to the instructional function. We omit expenditures on public service (like community education programs), auxiliary enterprises (like dormitories and cafeterias), hospitals, and other independent operations.

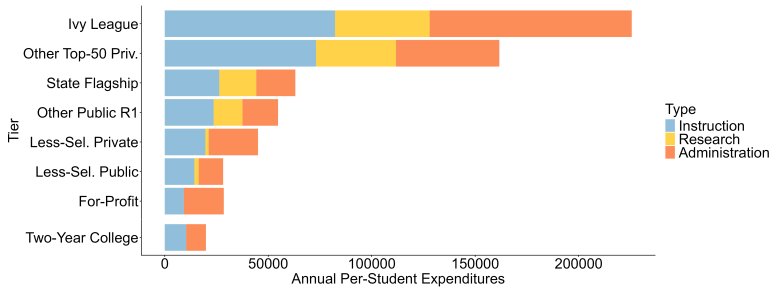
[7] Government appropriations include both direct appropriations and governmental financial aid programs (like Pell and CalGrant). Private funding includes private gifts and endowment returns. See Appendix A for details on data construction.

[8] States vary dramatically in the degree to which spending is aligned with funding across public institutions. 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 (see Figure A-3).

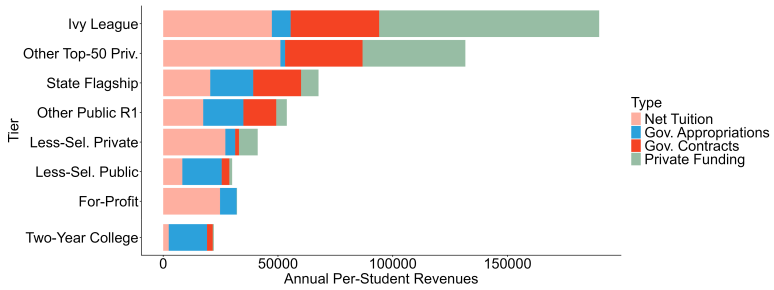
[9] Of particular note is Figure 5 in Mountjoy and Hickman (2021), which reproduces the Dale and Krueger (2002) finding of no relationship between value-added and selectivity but identifies a positive relationship with institutional resources. Several studies show that decreased public university revenues lead to worse quality and lower graduation rates (Kane, Orszag and Apostolov, 2005; Bound and Turner, 2007; Bound, Lovenheim and Turner, 2010) .

Figure 2. Universities' 2019 Revenues and Expenditures by Tier

(a) Expenditures per student



(b) Revenues per student



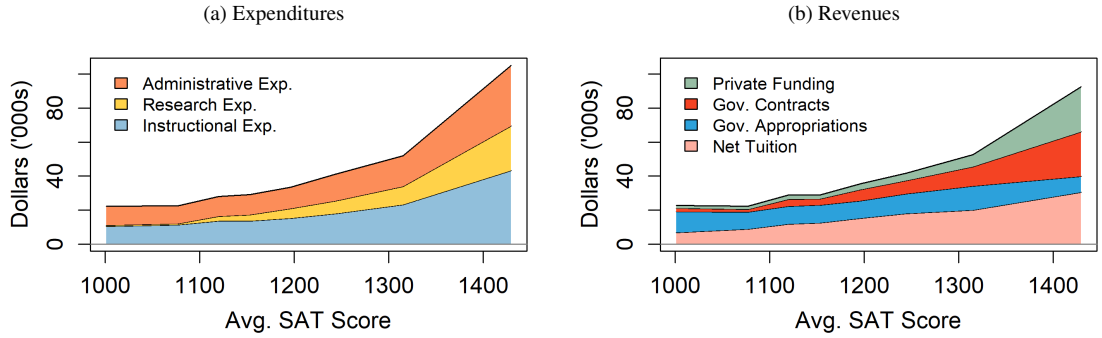
Note: This figure shows average 2019 expenditures and revenues per full-time equivalent student by tier of institution and type of expenditure or source of revenue. 'Instructional Expenditures' include expenditures for general academic instruction and departmental research. 'Research Expenditures' include 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). On the revenue side, '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. All dollar amounts are CPI-adjusted to 2024. 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.

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, as the sector grew from a small niche of the population to one that serves more than two-thirds of all young people and has become essential to labor market success.

In discussing long-run changes in the American higher education system, it is not always possible to have perfectly comparable measures of university characteristics. We generally rely on enrolled students' average SAT scores as measures of institutional selectivity; Figure A-4 shows that the SAT rankings of institutions have been largely stable since at least the mid-1980s, and we sometimes use the post-1980

Figure 3. Universities' 2019 Per-Student Revenues and Expenditures by Average SAT Score



Note: 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 2 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 2024. 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.

SAT as a proxy for hard-to-measure earlier selectivity.[10] Our main measures of the resources available for students are per-student instructional expenditures and total current (non-hospital) revenues, which have been collected in a roughly comparable way by the National Center for Education Statistics since the 1920s. Neither of these perfectly capture instructional quality, in part because there is no clear way to separate out the instructional, research, and other functions of a university, but we believe that they are good enough to capture broad trends. We discuss the consequences of potential mismeasurement below. Appendix A discusses our different data sources in greater detail.

The pre-war landscape: 1900-1945

At the beginning of the twentieth century, the college-going rate was between 5 and 10 percent (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 Department of Education's 1913 bulletin on "College Entrance Requirements" was wholly dedicated to between-institution variation in the number of units required in each academic subject (Kingsley, 1913).[11] Some universities additionally "administer[ed]

[10] Universities generally measure students' academic achievement by a combination of standardized test scores and measures of high school performance like their grade point average (GPA). We characterize university selectivity by standardized test scores just because of poor data availability regarding GPAs. Some schools provide average ACT scores (an alternative standardized test) instead of SAT scores; we convert these to SAT scores using an annual concordance (see Appendix A).

[11] The influential 1892 Committee of Ten – which included Harvard University president Charles Eliot and US Commissioner of Education 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 state's high school curriculum to enforce minimum pre-college academic preparation.

Table 2—Distribution of Instructional Expenditures Per Student Over Time

	1929	1969	1984	1996	2008	2019
Panel A: Average Instructional Expenditures per Student ($2024)						
All Four-Year Institutions	4,000	8,200	10,400	13,000	14,400	16,400
Ivy League	8,200	22,500	28,700	43,100	75,200	73,000
State Flagship	4,900	10,500	12,600	15,300	19,700	21,400
Public Non-Flagship	3,700	7,500	10,000	11,100	12,300	13,800
Ivy vs. Flagship Gap (%)	68	115	128	181	281	242
Flagship vs. Non-Flagship Gap[1] (%)	21	35	36	43	54	50
Panel B: Coefficient of Variation of Per-Student Instructional Expenditures (%)						
Public Institutions	38	30	46	53	50	49
All Four-Year Institutions	53	42	50	59	60	60

Note: Panel A shows the enrollment-weighted average instructional expenditures 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 instructional expenditures of public four-year institutions and all four-year institutions, weighted by enrollment. All dollars are reported in CPI-adjusted 2024 dollars. 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-3.

[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.

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 (private) universities in 1923 were $1,350 ($3,060) in 2024 dollars (Bleemer and Quincy, 2025a).[12]

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, and the variation was not strongly correlated with selectivity. Table 2 shows how cross-university variation in per-student institutional expenditures has risen over the past 100 years.[13] At the 111 public institutions that awarded Bachelor's degrees in 1929, the enrollment-weighted standard deviation of per-student instructional expenditures was about

[12]Universities maintained academic standards by evaluating classroom performance rather than standardized test scores; universities dismissed many low-performing students following their first or second year, often resulting in graduation rates below 50 percent at even the most prestigious institutions like Harvard and Brown (Thelin, 2010).

[13]The tiers included here are necessarily fewer than above to permit comparability over time.

$1,600 (in 2024 dollars), or 39 percent of the contemporaneous mean. 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" spent 68 more than the flagship publics. Gaps in total per-student revenues were slightly smaller in magnitude (Table A-1).[14]

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), in part due to federal financial aid from the GI Bill (Stanley, 2003). This shock in student demand forced supply-constrained universities to quickly 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.

Rutgers, The State University of New Jersey, is a paradigmatic example. Rutgers was an early adopter of standardized testing, but used test scores only to determine whether students satisfied the university's minimum academic qualifications. 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 (Albert Meder, 1939). In the Fall of 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).

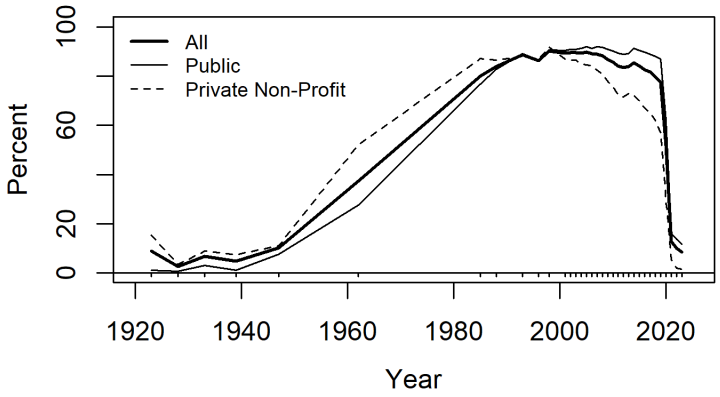
While many universities' 'meritocratic' admissions policies were initially intended as temporary responses to a short-run demand shock, the continued growth in college demand motivated their permanent codification. One popular blueprint was the 1960 California Master Plan for Higher Education. California was growing rapidly in this period, and the state was opening new college campuses at a very fast rate.[15] The Master Plan laid out a three-tiered higher education system – the research universities of the University of California (UC) at the top; the teaching-oriented, bachelor's-only California State University (CSU) campuses in the middle; and two-year community colleges below – and suggested using applicants' pre-college academic achievement to track students between the segments. The stated objective of admissions was to minimize dropout, and admissions criteria were valued for their predictive "validity" in identifying students who were likely to successfully complete their studies: "The more capable high school students should have the greater freedom of choice of collegiate institution, and selection procedures should give preference to the more able ... [to] predict success in the state colleges" (Spindt et al., 1961).[16] The authors of the Master

[14] All series are winsorized at 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; 1929 data do not permit calculation of FTE enrollment.

[15] 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. Eight new California State University campuses were also opened in this period.

[16] Three studies on the predictive validity of College Board standardized tests were published in 1960. By 1970, over 30 scholarly

Figure 4. Share of Four-Year Institutions Requiring Tests for Admissions



Note: 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.

Plan further recognized that rising demand among qualified students for certain institutions – especially UC Berkeley and UCLA – would soon outstrip those campuses' enrollment capacities, and suggested two remedies: (1) new 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) diverting 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 UC and CSU campuses (Coons et al., 1960).

Figure 4 shows that the use of standardized tests in admissions 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 (Hoxby, 2009a); nearly all private institutions required applicants to submit test scores by the early 1980s.[17] Public institutions were slower to adopt testing, with a large private-public gap opening in the 1950s. But public colleges and universities caught up in the late 1980s, and by 1990 nearly all required tests.

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.[18] Deviations from that assortative matching "overwhelmingly result from choices

publications per year were evaluating College Board's 'predictive validity' (as indexed by Google Scholar); by 2010, this number had risen above 150 (e.g. Westrick et al., 2019).

[17] 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.

[18] 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 strong peers (Hendricks, Herrington and Schoellman, 2021;

Table 3—Racial Composition by Tier Over Time

	1968	1976	1984	1996	2008	2019	2024
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: Share of full-time undergraduate students who are Black, Hispanic, or Asian at all four-year universities and by university tier. Tables A-2 and A-3 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.

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. Schools that wanted to admit more Black students but that relied on standardized tests for admissions faced a problem, however: Due in large part to the legacy of segregation and discrimination in pre-college education, very few Black students met 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. Table 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

Rothstein, 2006).

lower base) at flagship public universities.[19] 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 them further below.

Growing resource gaps: 1970-2019

The cross-university test score distribution had largely stabilized by 1970, with a mean test score gap between the bottom and top university testing deciles of between 2 and 2.5 national standard deviations.[20] But over the decades since then, long-standing funding gaps between universities grew sharply. The coefficient of variation (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 (see Table 1). Flagship public universities were spending about 35 percent more on per-student 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 (Freedman, 2023; University of California, 2023).[21]

Figure 5 shows average instructional expenditures per student by institutional average SAT score in 1969, 1993, and 2019. To focus on changes in resources at the institution level, institution SAT scores are held fixed here at their 1993 level.[22] In 1969, per-student instructional expenditures – which largely go toward labor costs for teaching faculty and assistants – 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.[23] 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).

[19] 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). Table 3 shows that Asian students were already over-represented at selective universities in 1968 (when the Asian population share was 0.8%) and are substantially overrepresented at those schools today, reflecting not an affirmative action boost but their high average pre-college academic achievement.

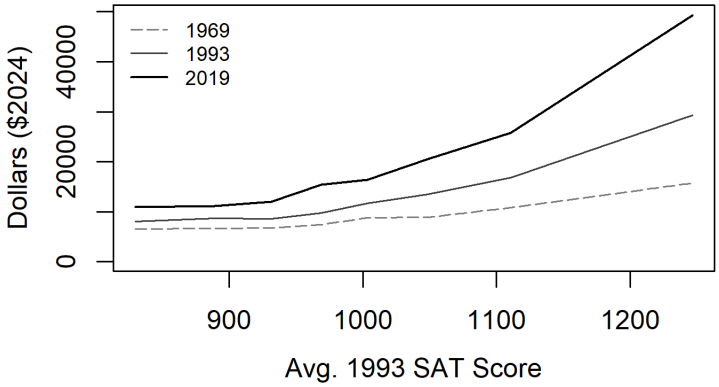
[20] See Figure A-4 regarding test score stabilization, which aligns with the post-1970 stabilization shown in Figure 1 of Hoxby (2009b). Bleemer and Quincy (2025a) show that students have been similarly stratified across higher- and lower-value institutions (as measured by fixed wage value-added) by parental income since the 1930s. Abramitzky et al. (2024) show that the share of elite college students from high-income families declined between 1940 and 1980, but that this had reverted by 2000. Alon and Tienda (2008) find that selective universities shifted some admissions 'weight' from high school rank to test scores in the 1980s.

[21] In dollar terms, CSU and UC mean wages were $11,450 and $14,070, respectively, in 1960, and $125,200 and $279,800 in 2023. The respective salaries for assistant professors were $7,360 vs. $7,550 (2.5 percent gap) in 1960 and $91,100 vs. $141,800 (56 percent) in 2023. CSU professors' wages have risen only slightly faster than inflation in this period; $11,450 is $118,000 in 2023 dollars, while $7,360 is $75,800.

[22] Because test score quantiles are largely unchanged over time (Figure A-4), this figure is qualitatively unchanged if SAT scores are measured contemporaneously. Figure A-5 shows the same trends in log dollars. Even on a log scale, gaps have widened over time.

[23] See Appendix Figures A-6 and A-7.

Figure 5. Instructional Expenditures Per Student by Institutions' Average SAT Score in 1993



Note: Average freshmen-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. Instructional expenditures include expenditures for general academic instruction and departmental research and are CPI-adjusted to 2024. 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 6 presents another view of the same trend, showing how the slopes seen in Figure 5 have changed over time. The relationship between student pre-college achievement and school resources grew steadily stronger between 1980 and the mid-2000s, with greater growth at private institutions, and was more than three times as strong in 2010 as it was in 1980.

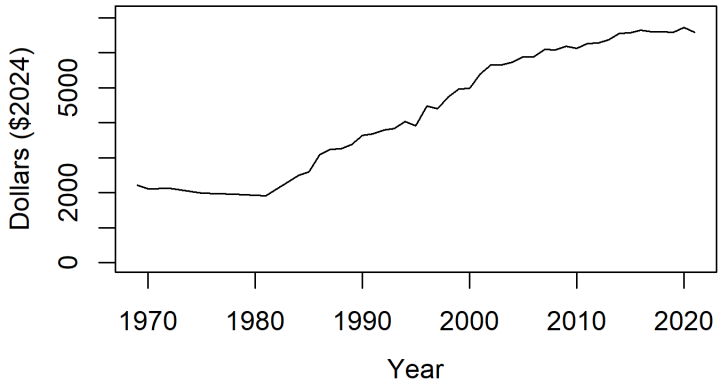
An examination of the revenue changes that financed the $10,000 increase in the gap in per-student instructional expenditures between top- and bottom-test-quartile institutions from 1993 to 2019 reveals evidence of the growing gap's causes.[24] The high-scoring institutions saw relative increases in 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 increased those schools' instructional expenditures.[25]

The stratification of instructional resources by standardized test score would look even more extreme if two-year colleges were included in these trends. 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)

[24] See Figure A-8. The revenue changes discussed here add up to more than the change in instructional expenditures because the gaps in research, student services, and administrative expenditures also grew over this period; see Figure A-7.

[25] Private funding statistics use a 2002 baseline instead of 1993, since data are not available for earlier years; see Appendix A.

Figure 6. Annual Association Between Per-Student Instructional Expenditures and Average SAT Scores



Note: This figure shows the coefficients from annual OLS regressions of per-student instructional expenditures on institutions' average SAT score (divided by 100). Regressions are weighted by freshman enrollment. Sample consists of four-year institutions, with average SAT scores held fixed at their 1993 level. For example, a 100 point increase in universities' average 1993 SAT score was associated with higher 2019 per-student instructional expenditures of about $6,600. '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). All dollar amounts are CPI-adjusted to 2024 and annually winsorized at 1 percent. Source: HEGIS, IPEDS, and the College Board Annual Survey of Colleges.

– to about half of four-year enrollment in the 2010s. 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 than that of typical public institutions (Mountjoy, 2022). Their negligible tuition costs and open-access admissions policies 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).[26]

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 we refer to as the "meritocratic consensus."

[26] A small share of high-demand community college programs implement 'meritocratic' admissions policies (Grosz, 2021). 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).

Deviations from "meritocratic" admissions

> "Admission requirements often reflect more realistically the objectives of the college than does its statement of aims and purposes."
>
> *Ralph Prator, President, San Fernando Valley State College, 1961*

While we've told a story of rising stratification of students by academic achievement and growing funding disparities across colleges and universities, some university interests and trends do not align with the rise of this "meritocratic consensus." Priorities like student diversity, successful football teams, and strong connections with their alumni and donors may conflict with admitting the students with the strongest academic backgrounds. Universities often finesse this by including all of these dimensions in an expanded understanding of "merit," 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.

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. 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.[27]

The tension between intercollegiate athletes and meritocratic admissions arises because 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.[28] Athletic recruitment also affects universities' ethnic composition, in perhaps surprising ways: Arcidiacono, Kinsler and Ransom (2022) show that Harvard athletes are almost twice as likely to be white and twice as likely to be from the top 1 percent of the parental income distribution as 'typical' enrollees (see also Chetty, Deming and Friedman, 2023).

This tension is mitigated in two ways. First, the NCAA and other athletic conferences set minimum requirements for athletes' pre-college academic achievement and standards for college GPAs that scale with the institution's own grading policies, requiring that athletes' academic preparation is not *too* different from that of their peers at even Division I institutions. Famously, however, some college sports powerhouses routinely evade these rules, for example through courses designed to boost athletes' grades. Second, and more importantly, the tension is limited by the small size of the athlete pool in the larger pool of selective institutions. While athletes make up substantial shares of enrollment at some very small schools, overall only about 7 percent of college students are intercollegiate athletes (in 2023; this is up from 5 percent in 2004). Only 2 percent are NCAA Division I athletes eligible for preferential tuition pricing and other benefits. This is the same as the athlete share of students at R1 public universities; at top-50 private universities, the share is 9 percent.[29] As a result, athletics contributes very little to the overall allocation of students to US

[27] Institution-level athletics data are available from the Equity in Athletics Data Analysis platform of the US Department of Education.
[28] Calculation by the authors (Bleemer, 2018).
[29] Many of the latter schools are quite small: while 20 are NCAA Division I, another 22 are Division III.

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.

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. These policies are justified as benefits for institutional stakeholders and incentives for financial support for the institution (Castilla and Poskanzer, 2022).

Arcidiacono, Kinsler and Ransom (2022) show that the children of alumni, donors, and faculty were 5.2, 6.9, and 4.1 times more likely, respectively, to be admitted to Harvard in 2014 than other academically-similar applicants. Fully 14 percent of admitted students were the children of alumni (Arcidiacono, 2018). Chetty, Deming and Friedman (2023) report that, due to preferences for the children of alumni (known commonly as "legacies"), 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.[30] But legacy admits are still quite high scoring in traditional academic terms; while admitted legacies may have lower test scores than the marginal non-legacy applicant, the gap is not enormous. And, like athletes, legacies are relatively limited in their overall impact, as legacy preferences are strongly concentrated at a few mostly-private institutions. Only 9 percent of enrollment-weighted R1 public universities employ such preferences (Hurwitz, 2011).

Affirmative Action

As discussed above, 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 prominently justified by the claim that these students' underrepresentation reflects mismeasurement of their underlying 'merit' by standardized tests and other measures of pre-college academic achievement, which may not take into account the students' poorer pre-college academic opportunity and circumstances (e.g. Walton, Spencer and Erman, 2013).[31] If intelligence is reasonably uniformly distributed across the population but traditional academic achievement measures are biased against students from underrepresented groups, then preferences for these groups could improve the university's ability to identify and admit the most intelligent applicants (Geiser, 2017).[32]

[30] The top 1 percent share is further increased by 1-2 percentage points due to athletic preferences.

[31] Other universities justify diversity policies as contributing to the educational content of the institution (Bleemer and Jaynes, 2025) or to social justice goals.

[32] As discussed in the next section, the absence of consistent evidence (or even theoretical justification) on the complementarity between university quality/resources and *either* measured academic achievement or innate 'intelligence' limits scholars' ability to arbitrate this issue on efficiency grounds.

The most prominent example of diversity-based admissions preferences is race-based affirmative action. Pre-college academic achievement sharply differs 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).[33] In fact, 8 percent of those colleges' freshmen were Black. In 2023, 76 (47) percent of highly-selective (selective) universities reported providing race-based admissions preferences (Reber, Nagashima and Goodman, 2023).[34]

Race-based affirmative action provided large admissions advantages to applicants from underrepresented racial minorities (URM). One study found that in the 1990s URM applicants were twice as likely to be admitted to UC Berkeley and UCLA as academically-comparable non-URM applicants (Bleemer, 2023), and the URM students who enrolled at those campuses had about 1 standard deviation lower academic achievement (Arcidiacono, Aucejo and Hotz, 2016).[35] However, 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*). As a result, race-based affirmative action no longer plays a role in the allocation of students to colleges and universities.

Some universities have also used affirmative action to manage their gender distributions, though the impact is quantitatively too small to play a significant role in the aggregate allocation of students to universities. 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% female at most highly-selective colleges (except those focused on certain STEM fields). Baum and Goodstein (2005) find evidence of male admissions preferences in the late 1990s at small liberal arts colleges, but by the mid-2000s the net effect of gender-based preferences on selective university enrollment was very small (Bielby et al., 2014).[36] Appendix B provides suggestive evidence that admissions quotas and preferences for male students have become more impactful in recent years, increasing male enrollment at top-50 private universities by perhaps 1-2 percentage points since 2010.

Holistic Review and Alternative Metrics

Nearly all colleges and universities consider quantitative measures of academic achievement beyond standardized tests, including students' high school grade point average and many other achievement metrics. Some less-selective colleges use formulas or cutoffs for admissions (Zimmerman, 2014; Mountjoy, 2024;

[33] Krueger, Rothstein and Turner (2006) estimate that in 2000, less than 1.5% as many Black students as white students scored above 1400.

[34] Reber, Nagashima and Goodman (2023) define highly-selective (selective) universities as those which deny admission to at least 85 (50) percent of applicants.

[35] Not all of this admissions advantage was due to affirmative action; some was due to differences in harder-to-measure qualifications that are rewarded by holistic review (see the next subsection). In the absence of affirmative action, the URM advantage would have been around 40 percent, not 100 percent (Bleemer, 2023).

[36] The highly-selective private universities that first admitted women in the 1960s subsequently implemented admissions quotas that limited female enrollment, presaging male admissions advantages that have lasted for decades in some cases (Malkiel, 2016). Data from 2014-2018 show that Harvard University provided small admissions advantages to female applicants relative to academically-similar male applicants, about one-tenth the size of its preference for Black applicants (Arcidiacono, 2018).

Miller, 2025), but nearly all highly-selective public and private universities conduct admissions by "holistic review" (Bastedo et al., 2018): rather than admitting students using a fixed-weight ranking of students by academic criteria (e.g., a minimum SAT threshold), universities provide evaluative flexibility to 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).

Selective public universities have had two reasons to adopt holistic review in admissions.[37] One is that traditional measures like standardized tests and high school grades are thought to miss key aspects of academic achievement that can be better captured by subjective assessments of course-taking patterns, essays, and extracurricular activities. Another is related to the discussion of affirmative action above: If traditionally underrepresented students score lower on standardized tests than their true "merit," either because the tests are biased or because the students have not had adequate educational opportunity, then one might be able to do a better job of capturing their "merit" through an application review that takes account of the student's context.[38]

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 is based primarily on an assessment of academic achievement.[39] Its net enrollment effect is generally to increase relative enrollment among two groups of students: (1) students with strong interpersonal and non-cognitive skills and (2) students with poor pre-college academic opportunity (Bastedo et al., 2018). Selective public universities that implement holistic review generally increase their URM enrollment by 1-2 percentage points (Bleemer, 2023), though 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.[40] This overstates the effect on selection by academic achievement, however, as in many cases holistic review merely replaces students with standardized test scores with others whose applications show other indications of achievement (e.g., difficult course schedules).

Another alternative admissions policy is more formulaic, but also downweights test scores. Five states' public university systems have adopted "top percent" policies, which provide an admissions advantage to the highest-ranking (by high school GPA) graduates from every eligible high school, regardless of their test scores.[41] The net effect of these policies is to provide admissions advantages to high-GPA low-SAT students

[37] Holistic review was introduced in the 1920s at elite private universities, which used it to admit traditional upper-class students rather than high-testing (generally Jewish) applicants (Karabel, 2005).

[38] This movement was given momentum by Supreme Court jurisprudence that prohibited explicit racial quotas (*Regents of the University of California v. Bakke*, 1978) or point-based admissions procedures in providing race-based admissions advantages (*Gratz v. Bollinger*, 2003) but that endorsed taking account of students' background in more flexible ways. It received further momentum in states where affirmative action was banned.

[39] A study of holistic review at UCLA found that the largest weights were put on grades, standardized tests, college preparatory coursework, and AP test scores, but that factors like extracurricular activities, community involvement, out-of-school work, and socioeconomic hardship were also rewarded (Mare, 2014). Arcidiacono (2018) similarly shows that the academic index and admissions readers' academic ratings were the most significant determinants of admission to Harvard in the 2010s.

[40] While holistic review reallocates admissions weight from test scores to essays and other components that are relatively more predictive of parental income and race (Alvero et al., 2021), universities' revealed preferences over those other components (e.g. for the essay topics chosen by lower-income students) generally leads to relatively increased enrollment among disadvantaged students (Bleemer, Portier and Rothstein, 2026).

[41] The five states are Texas, California, Florida, Georgia, and New York. These policies have a precedent in "highest seventh plan" implemented by Harvard and other private Massachusetts colleges in the 1920s (Hinckley, 1941). The recent plans were adopted in part as race-neutral alternatives to affirmative action, because in the presence of secondary school segregation, within-school rankings

from low-performing high schools, since top students at high-performing high schools generally have high test scores and would be admitted under traditional rules (Black, Denning and Rothstein, 2023). As a result of high school stratification by race and class, implemented top percent policies tend to slightly increase URM enrollment (1-2 percentage points; Kapor, 2024; Bleemer, 2021) and enrollment from families with below-median income at participating selective universities (as much as 1 percentage point; Bleemer, 2023).

As this discussion makes clear, it is too simple to imagine that US college admissions are conducted solely or primarily on the basis of standardized test scores. Some students are admitted for their athletic ability or personal connections despite weak academic achievement, while others get boosts for their social context aimed at increasing diversity. But most college and university admissions decisions, while not formulaic, are based on demonstrated academic achievement, particularly after the end of affirmative action.

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? While the current instructional allocation represents a market outcome – institutions set their admissions rules and prices and students set their application strategies – the higher education market does not resemble the perfect competition ideal. Institutions do not set prices to market-clearing rates but instead make admissions decisions based on non-price factors; nearly all institutions are public or not-for-profit, and cannot be assumed to be profit maximizing; output is jointly produced by the college and the student and depends on both student effort and ability; and externalities – also known in this context as peer effects – are a potentially important component of the production process. As a result, no welfare theorem ensures that the market outcome is efficient.[42]

We examine the allocative efficiency of US higher education by considering higher education 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 earnings maximization the only goal, or are there other forms of learning that should be pursued? – and assume an agreed-upon, unidimensional output HC. Each university u produces this output by combining institutional inputs such as instructors, Z_u, with student characteristics, X_u:

$$HC_u = f(X_u, Z_u).$$

Here, f is the production function, which could be quite general. In particular, the dependence of output on student characteristics X_u allows for a range of types of peer effects, where the outcomes of one student

yield a more diverse class than between-school competition.

[42] Earlier studies of intergenerational income transmission (e.g. Loury, 1981; Becker and Tomes, 1986; Cameron and Taber, 2004) focused on imperfect capital markets as explaining differential educational access by family income, but as institutional resources have grown more unequal and admission has gotten more competitive, non-market university allocation by admissions to highly resourced institutions based on pre-college academic achievement provides a more direct friction precluding market-generated efficiency in recent years.

depend on the characteristics of his or her classmates. An allocation improves efficiency if $\sum_u W(HC_u)$ rises for the same aggregate set of students and institutional inputs, where W assigns welfare to the newly-produced HC_u.

For simplicity, we consider the case where X_u indexes students' average pre-college academic achievement and Z_u indexes universities' average per-student instructional expenditures. As a result, the "meritocratic consensus" in higher education comprises the resource and student allocation processes that have resulted in both substantial variation in X_u and Z_u and a strong correlation between X_u and Z_u.[43]

Notice that the meritocratic consensus is formally very different from meritocracy in labor markets, which requires that high-merit workers (measured by either innate ability or labor market productivity) are rewarded by high wages (e.g. Loury, 1981). This enhances efficiency by directly incentivizing workers' production and is generally produced by competitive market forces. Wage dispersion creates rewards for high effort and ensures that the most able workers wind up in the jobs where their ability is most valuable. In education markets, on the other hand, institutional resources are less a reward for high-achieving students than an investment in those students' human capital production. They are distributed not by competitive markets, but by policy and institutional decisions in an environment that is quite far from the perfect competition ideal.[44] Efficiency of this allocation cannot be assumed, and must be demonstrated based on claims about the production process or social welfare function.

We discuss three plausible theories about why the type of stratification that we have called the meritocratic consensus in higher education might improve efficiency.

q-Complementarity

The first theory is that Z and X are q-complements in the production of human capital: increases in X may increase the marginal return to Z (and vice versa). Formally, this corresponds to $\frac{\partial^2 f}{\partial X \partial Z} > 0$. This is a plausible model: 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 up-to-date lab facilities. If so, it makes sense to segregate those high-achieving students from others and to direct substantial resources in their direction (Sallee, Resch and Courant, 2008).[45]

There is another plausible view of the production process, in which X and Z are substitutes rather than complements: that is, $\frac{\partial^2 f}{\partial X \partial Z} < 0$. Perhaps weaker students benefit more from individualized attention, out-of-classroom advising, tutoring, and student support services, while stronger students can learn effectively with

[43] One initial question is whether it is efficient for X_u to vary so much; in many countries (like Canada and Germany) institutions are much less stratified on student achievement, 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 X_u increases inequality but has little effect on efficiency (Betts, 2011; Terrin and Triventi, 2023).

[44] Among other deviations, it is nearly impossible for a new college entrant to assemble the resources or reputation needed to compete with established elite institutions.

[45] Complementarity creates an equity-efficiency tradeoff: it is efficient to direct more resources to the highest-X students, but because high-X 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 *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 Z to the schools that high-ability, disadvantaged students attend, which might be low-X schools if X is sufficiently correlated with family background. See related discussion in Roemer (1998).

only a lecture and a textbook. Elementary and secondary education funding policies are often based on an assumption like this about the production process; see, for example, "weighted student funding" formulas that direct more resources to schools serving disadvantaged students (e.g. Chingos and Blagg, 2017). If student aptitude or preparedness is a 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 practical terms, this would mean more generous funding for community colleges and open access institutions, and lower funding – at least for instruction – at state flagships.[46] However, while there are some policies that aim to provide compensatory funding – e.g., Pell grants for low-income students – there is little sign that policymakers are interested in generating net negative correlations between Z_u and X_u.

Researchers have investigated the complementarity or substitutability of higher education production in several distinct ways. Many 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. While an early series of observational and structural studies of race-based affirmative action – which admits lower-testing URM students to more-selective universities – provided mixed evidence that the policy 'mismatched' students (e.g. 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) 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). Importantly, however, these studies generally do not distinguish between peer groups and resources as determinants of college value-added; Dillon and Smith (2020), for example, construct an index of college quality combining peer strength and expenditures.[47] Because they do not separately identify $\frac{\partial^2 f}{\partial X \partial Z}$ and $\frac{\partial^2 f}{\partial X^2}$, they thus offer little direct evidence about the efficiency benefits or costs of matching resources to student preparedness.

There is a bit more evidence on the topic in the elementary and secondary context. Models of dynamic complementarity ("skill begets skill") in 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, it is not clear whether the production function in elementary and secondary school has the same basic shape as that of postsecondary education.

Perhaps the most direct postsecondary evidence comes from specific program evaluations. There are

[46] If students value resources highly in their college choices, this might not be sustainable as an equilibrium: under such a funding system, the most prepared students would most want to enroll in community colleges, which would be forced to institute a rationing mechanism that could wind up reproducing the current hierarchy. However, resources are not the only institutional characteristic that students care about; strongly prepared students also prefer to attend school with other well-prepared students. If this peer-group concern is strong enough, an allocation in which the most selective schools have the lowest funding could be sustained as an equilibrium (Rothstein, 2006).

[47] Sander (2004) is an exception: He emphasizes the importance of class rank as a mechanism for mismatch effects, implicitly assuming that peer strength rather than resources is the relevant contextual measure.

several examples of resource-intensive programs that yield large benefits for college students with low prior achievement. The ASAP program provides financial, academic, and personal support to low-achievement students at the City University of New York (CUNY) and significantly raises student success (Scrivener et al., 2015). This follows earlier evidence from the evaluation of the Opening Doors demonstration project (Scrivener et al., 2008; Scrivener and Coghlan, 2011). This is again consistent with substitutability of resources for student preparedness, though it is difficult to compare the benefits of these programs to the potential benefits of devoting similar resources to high-prior-achievement students.[48]

A large "predictive validity" literature provides evidence that standardized test scores are highly predictive of academic performance in college (e.g. University of California STTF, 2020; Friedman et al., 2025; Joyce et al., 2026). While this evidence could imply that $\frac{\partial f}{\partial X} > 0$ if high-performing college students derive relatively greater educational value from college attendance – though there is some evidence to the contrary (Bleemer and Mehta, 2022b) – there is no obvious relationship between $\frac{\partial f}{\partial X}$ and the efficiency of merit-based stratification. As a result, the predictive validity of standardized tests is largely orthogonal to their usefulness in efficiently allocating students between institutions.

Convex Social Welfare

A second model in which the observed patterns might be efficient is one in which the social welfare function W puts very high weight on the human capital of the highest achievers. These top achievers may be responsible for a disproportionate share of scientific breakthroughs, entrepreneurial successes, and artistic masterpieces. It is thus plausible that the social welfare function is convex in HC: $W(HC)$, with $W' > 0$ and $W'' > 0$). In this case, even if production does not exhibit complementarity, social welfare may.[49] It would then be optimal to concentrate the best students with the highest resources, not because they receive the highest HC benefit from these resources but because any HC increment that they receive is enormously valuable to society.

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 (X_i) rather than a consequence of school-constructed HC. The relationship between entrepreneurial success and SAT scores, on the other hand, is positive but slightly *concave* (Chetty et al., 2025).

[48] Another piece of evidence on this point comes from Massively Open Online Courses (MOOCs). These attracted a great deal of attention in the 2010s as a potentially much less resource intensive way to deliver postsecondary instruction. Evidence indicated, however, that this model was less effective for students with low prior achievement (Hansen and Reich, 2015; Kofoed et al., 2024).

[49] Note that

$$\frac{\partial^2 W(f(X,Z))}{\partial X \partial Z} = W''(f)\frac{\partial f}{\partial X}\frac{\partial f}{\partial Z} + W'(f)\frac{\partial^2 f}{\partial X \partial Z}.$$

If $W''(HC)$ is sufficiently large, the positive contribution of the first term could outweigh any negative contribution from substitutability in the second term.

High School Incentivization

A third model in which it might be efficient to direct greater resources to colleges enrolling students with higher pre-college academic achievement is one that focuses on 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 (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. The incentive effects of large rewards for high school achievement may induce sufficient effort in high school to justify even inefficient allocations in college, but there is little available evidence on the potential magnitude of these incentive effects.

Overall, 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 that, at current margins, overall learning would be increased by diverting some resources to the schools that enroll students with lower prior achievement.

The Future of the Meritocratic Consensus

Students do not need to be as stratified by test score across postsecondary institutions as they currently are in the United States, nor do the institutions that enroll high-testing students need to be so relatively well-funded. This section discusses recent developments that may expand or contract the scope of the meritocratic consensus in future years.

Funding Allocation

Figure 6 shows that the enrollment-weighted slope of per-student expenditures (in 2024 $'s) with respect to institutional average student SAT score (as measured in either 1985 or 2019) grew almost linearly from about $2,000 in 1980 to about $6,000 in 2010. The relationship then flattened, remaining at about $6,500 in 2021.[50] The between-institution relationship between research expenditures and average SAT score has also flattened since 2010, though the administrative expenditures relationship continues to grow (Figure A-10).

Expenditures must align with revenues. 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 meaningfully declined over the past two decades from slightly positive to meaningfully negative (Figure A-10). The creation and expansion of state financial aid programs like New York's Excelsior Scholarship

[50] Using universities' contemporaneous average SAT scores suggests a similar trend through 2019, after which SAT scores are poorly observed (Figure A-9).

and New Jersey's Tuition Aid Grant have played an important role in this nascent trend.[51] Recent federal actions targeting the finances of high-revenue private universities – like 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.

College Admissions

The early 2020s brought two of the most significant changes to universities admissions policies in over 50 years. The first regarded the use of standardized tests in admissions. Figure 4 shows that about 25 percent of private universities steadily adopted 'test-optional' admissions policies between the 1990s and 2010s, permitting applicants the choice of whether to submit standardized test scores as part of their application. The 2020 Covid-19 pandemic, however, decreased the share of 'test-mandatory' institutions to levels comparable to the 1930s (8 percent in 2023). Five years later, only a handful of top-50 private and flagship public institutions have reinstituted their testing requirements. Some universities have gone further; public universities in California have adopted 'test-free' admissions policies that prohibit applicants from submitting standardized test scores. The second significant change to university admissions policies in the early 2020s was the nationwide prohibition of race-based affirmative action resulting from the *SFFA* Supreme Court decision.

The adoption of test-optional and test-free policies has several motivations beyond the proximate 2020 Covid-19 disruptions to test administration. First, competition for qualified students incentivized some institutions to attract more applications from students who had been previously discouraged from applying by either the logistical and financial cost of test-taking or by their own judgment of their low test score. Second, the large test score gap between URM and non-URM students motivated omission of test scores from admissions decisions in order to increase URM enrollment as a form of race-neutral affirmative action. Indeed, Bennett (2022) shows that 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 URM enrollment shares. This may have motivated institutions to maintain their Covid-era test-optional policies, particularly after the *SFFA* decision cut off other options for addressing URM test score gaps (Geiser, 2017). Other potential motivations include solving a principal-agent problem in which university administrators could not control the weight that admissions officers gave to applicants' test scores and minimizing the political pressure that arises from large ethnic test score gaps by not measuring them (Dessein, Frankel and Karthik, 2025). 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.

As discussed above, preferences regarding racial diversity comprised the largest wedge between US higher education and 'meritocratic' student allocation for several decades, and the *SFFA* decision generated considerable discussion among selective universities regarding their admissions objectives. As of 2025, the

[51] Figure A-11 shows that the decline in the correlation between government funding and institutions' average SAT scores has been almost wholly driven by increasing funding from student grant programs to low-SAT institutions, not by changes in states' block appropriations.

downstream effects of the decision – and related actions by the Trump administration to obligate greater 'meritocratic' selection in university admissions – on student stratification are still evolving.[52]

Within-Institution Stratification

While institutional stratification has been stable for decades and inequity in resource allocations has stopped growing, there is one area where the allocation of greater resources to higher-achieving students has continued: *within* institutions. College majors offer a paradigmatic example. In the 1960s, students 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 R1 public universities have minimum GPA 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 R1 public universities' computer science programs – often the institutions' highest-value major – that restricted access on the basis of academic preparation rose from 10 to 40 percent (Bleemer and Mehta, 2022a), and 'soft' restrictions like relatively low grade distributions in STEM courses have become more prevalent (e.g. Butcher, McEwan and Weerapanda, 2014). Restricted majors tend to be in STEM fields and have relatively high instructional costs (Altonji and Zimmerman, 2019), providing one motivation for the restrictions' implementation.

These policies have generated substantial between-major stratification by academic achievement – with higher-achieving students allocated to higher-value majors – which in turn manifest as swiftly growing gaps between the average relative value of college majors earned by URM and non-URM students and by lower- and higher-income students (Bleemer and Mehta, 2022a; Bleemer and Quincy, 2025a). As with university admissions, there is little evidence that 'meritocratic' major allocation 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 within-institution 'honors colleges' is another recent avenue by which resources are allotted to higher-performing students. Between 1958 and 2021, the number of R1 and R2 universities with honors colleges rose from 3 to about 200, with growth accelerating over time. Honors colleges generally provide dedicated curricula in small classrooms, special advising services, and additional student amenities, all typically provided without any additional cost for the student (Cognard-Black and Smith, 2023). Their growth in recent years seems to be motivated (at least in part) by competition to attract high-achieving students who would otherwise enroll elsewhere, either because these students provide valuable peer effects to other students or, more likely, because these students contribute to the institution's prestige and ranking. While there is little evidence on the long-run effects of honors college participation, Card, Chyn and Giuliano (2024) and Lavy and Goldstein (2022) show that the additional resources provided

[52]The year following the *SFFA* decision, freshman URM enrollment at top-50 private universities declined 18 percent (from 22 to 18 percent freshmen), though this followed years of increasing Black and Hispanic enrollment; 2024 Black enrollment was 10 percent lower than its 2000-2023 average. An enrollment cascade of URM students into less-selective institutions (Conrad and Sharpe, 1996) led 2024 freshman URM enrollment at flagship public universities to rise 3 percent.

by parallel 'gifted education' programs in K-12 education would provide substantial long-run value even to lower-performing students excluded from the programs.

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.

Albert Meder, Jr. 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.

Alon, Sigal, and Marta Tienda. 2008. "Diversity, Opportunity, and the Shifting Meritocracy in Higher Education." American Sociological Review, 72(8): 487–511.

Altonji, Joseph G., and Seth D. Zimmerman. 2019. "The Costs of and Net Returns to College Major." In Productivity in Higher Education. , ed. Caroline M. Hoxby and Kevin Stange, 133–176. Chicago, IL:University of Chicago Press.

Alvero, AJ, Sonia Giebel, Ben Gebre-Medhin, Anthony Antonio, Mitchell Stevens, and Benjamin Domingue. 2021. "Essay Content and Style Are Strongly Related to Household Income and SAT Scores: Evidence from 60,000 Undergraduate Applications." Science Advanced, 7(42).

Antonovics, Kate, and Ben Backes. 2014. "The Effect of Banning Affirmative Action on College Admissions Policies and Student Quality." Journal of Human Resources, 49(2): 295–322.

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, Esteban Aucejo, and V. Joseph Hotz. 2016. "University Differences in the Graduation of Minorities in STEM Fields: Evidence from California." American Economic Review, 106(3): 525–562.

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.

Becker, Gary, and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." Journal of Labor Economics, 4(3): S1–S39.

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 Gerald Jaynes. 2025. "Affirmative Action and Racial Integration." 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, Gabriel Portier, and Jesse Rothstein. 2026. "College Essays and Parental Income." 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.

Cameron, Stephen, and Christopher Taber. 2004. "Estimation of Educational Borrowing Constraints Using Returns to Schooling." Journal of Political Economy, 112(1): 132–182.

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.

Chan, Jimmy, and Erik Eyster. 2003. "Does banning affirmative action lower college student quality?" American Economic Review, 93(3): 858–872.

Chetty, Raj, David Deming, and John Friedman. 2023. "Diversifying Society's Leaders? The Determinants and

Causal Effects of Admission to Highly Selective Private Colleges." NBER Working Paper, 31492.

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.

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.

Dessein, Wouter, Alex Frankel, and Navin Karthik. 2025. "Test-Optional Admissions." American Economic Review, 115(9): 3130–3170.

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.

Friedman, John, Bruce Sacerdote, Douglas O. Staiger, and Michele Tine. 2025. "Standardized Test Scores and Academic Performance at Ivy-Plus Colleges." American Economic Association Papers and Proceedings, 115: 676–681.

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.

Geiser, Saul. 2017. "Norm-Referenced Tests and Race-Blind Admissions: The Case for Eliminating the SAT and ACT at the University of California." CSHE Research and Occasional Paper Series, 17(15).

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.

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

Hoxby, Caroline M. 2009b. "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.

Joyce, Theodore, Mina Afrouzi Khosroshahi, Sarah Truelsch, Kerstin Gentsch, and Kyle Du. 2026. "Standardized Test Scores and Academic Performance at a Public University System." NBER Working Paper, 34975.

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." 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.

Kingsley, Clarence. 1913. College Entrance Requirements. Vol. 7, Washington, DC:US Department of Education Bulletin.

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.

Krueger, Alan, Jesse Rothstein, and Sarah Turner. 2006. "Race, income, and college in 25 years: Evaluating Justice O'Connor's conjecture." American Law and Economics Review, 8(2): 282–311.

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, Glenn. 1981. "Intergenerational Transfers and the Distribution of Earnings." Econometrica, 49(4): 843–867.

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

Malkiel, Nancy Weiss. 2016. "Keep the Damned Women Out": The Struggle for Coeducation. Princeton, NJ:Princeton University Press.

Mare, Robert D. 2014. "Holistic review in freshman admissions at the University of California-Los Angeles, 2009-11 update."

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.

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).

Sander, Richard H. 2004. "A systemic analysis of affirmative action in American law schools." Stan. L. Rev., 57: 367.

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."

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.

University of California STTF. 2020. Report of the Academic Council Special Committee on Standardized Testing (STTF). Oakland, California:University of California Academic Senate.

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

November 2025

Table of Contents

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.

A1. 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.[53] 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 2024 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.

A2. 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 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.

A3. 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.

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

A4. 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 2024.

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).

A5. 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.

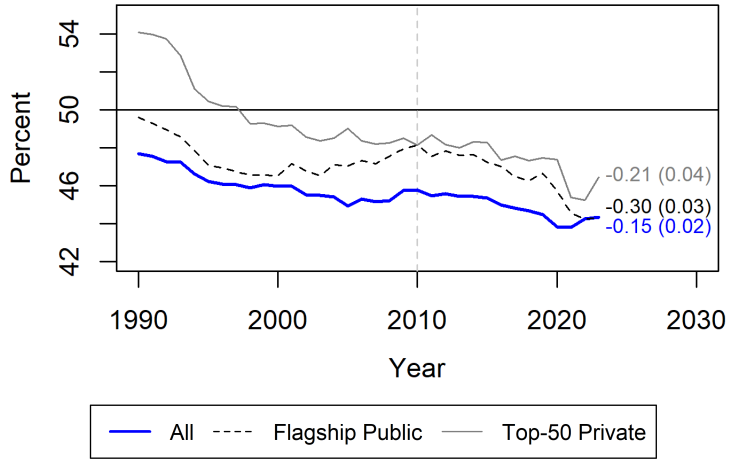
A6. IPEDS

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-2 and A-3 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

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.

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.

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.

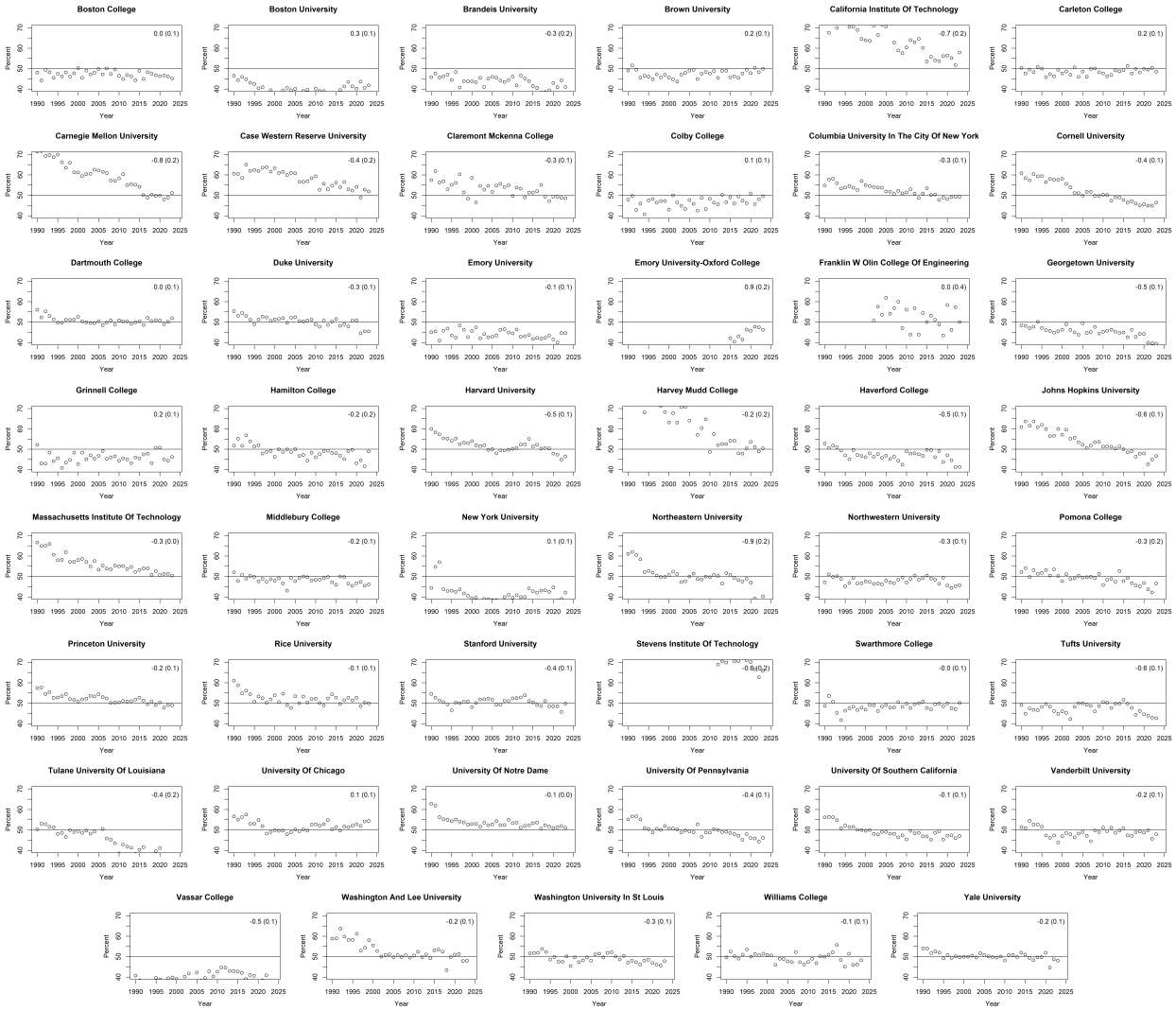
"Private Funding" is the sum of private gifts, grants, and contracts, contributions from affiliated entities (like fundraising foundations), and all revenue from university endowment funds. Private funding is not observed prior to 2002.

"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). "Total Non-Hospital Revenues" are the difference between total revenues and revenues from hospitals. All revenue and expenditure measures are CPI-adjusted to 2024.

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

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.

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 preferences 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.

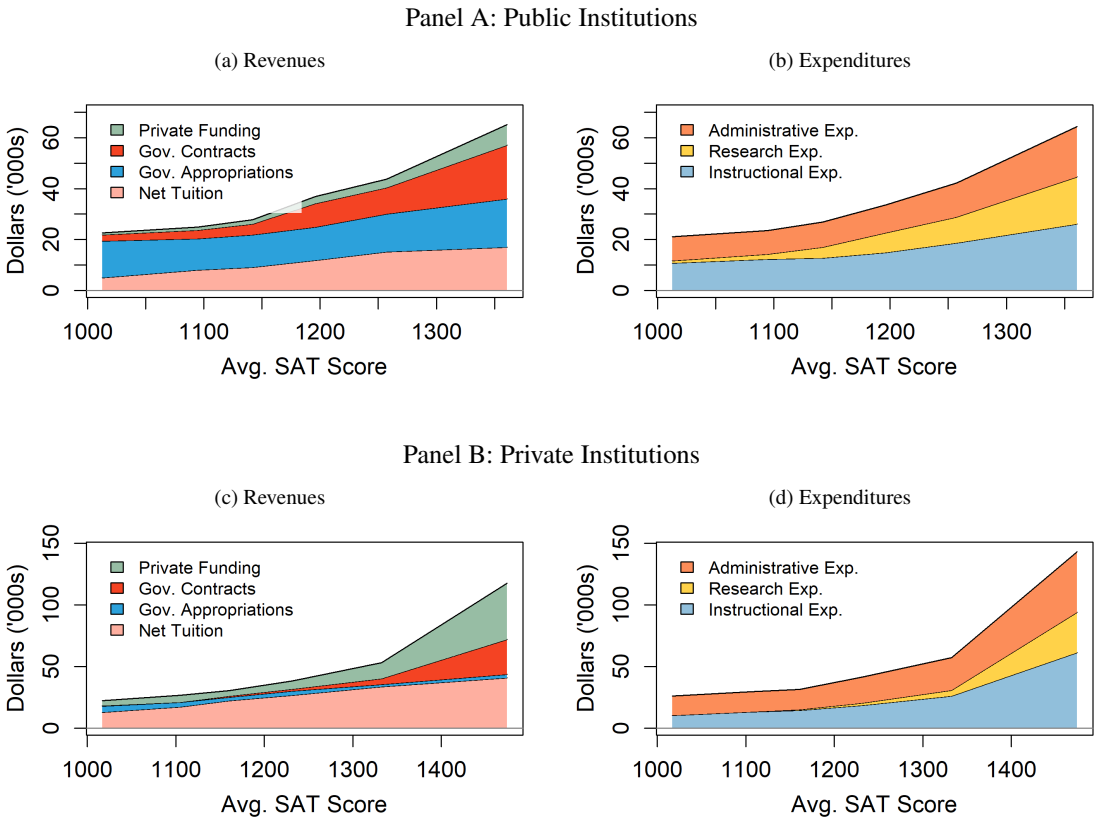
Figure BB-2 visualizes the male shares of first-year students at each of the top-50 private universities, excluding three all-female institutions. Some universities, like Brandeis University, have student gender compositions that closely mirror the average flagship public university. Others, like Dartmouth 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.

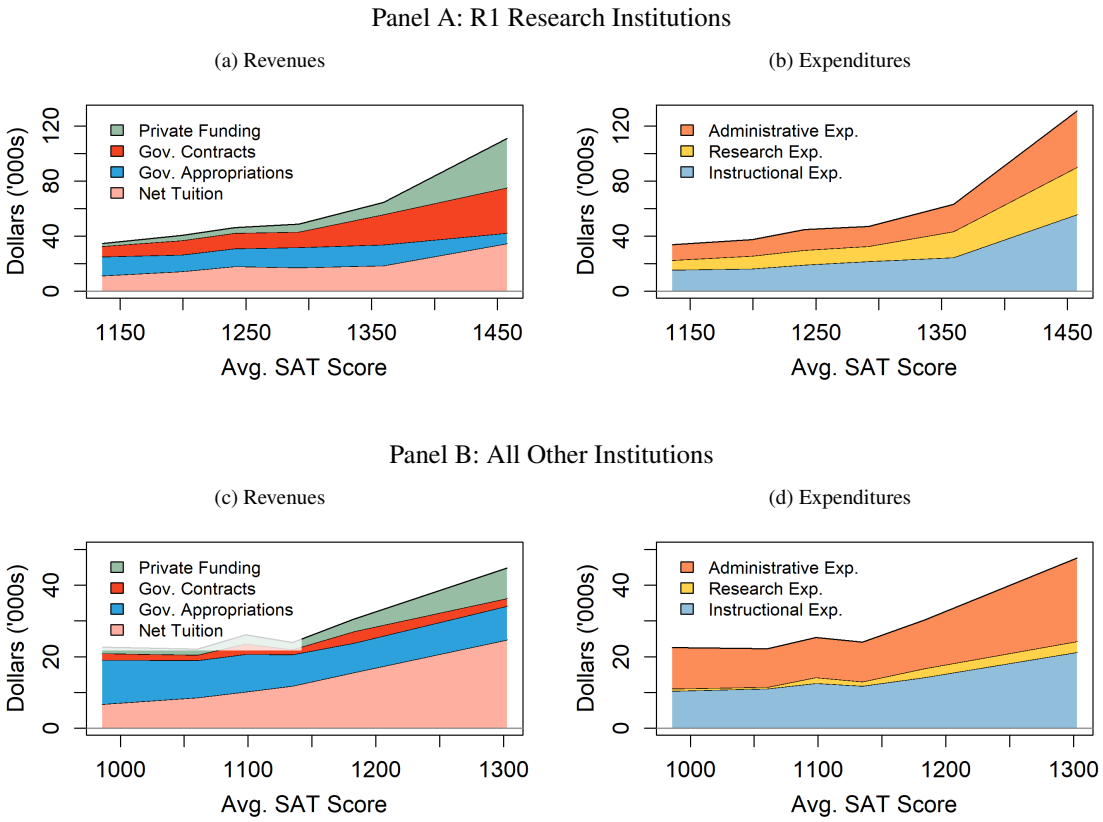
APPENDIX FIGURES AND TABLES

Figure A-1. Universities' 2019 Per-Student Revenues and Expenditures by Average SAT Score, by Public/Private



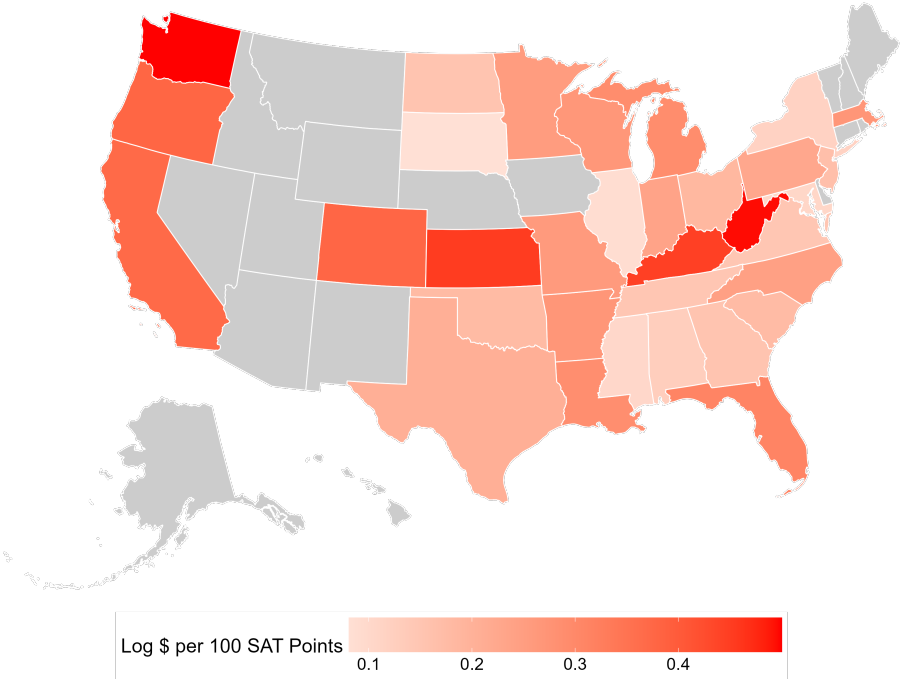
Note: This figure shows reasonably similar revenue and expenditure patterns by institutions' average SAT scores between public and private institutions, though the differences are larger at private universities and the higher per-student revenues of high-SAT public universities are driven relatively 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 2024. Source: IPEDS.

Figure A-2. Universities' 2019 Per-Student Revenues and Expenditures by Average SAT Score, by Research Status



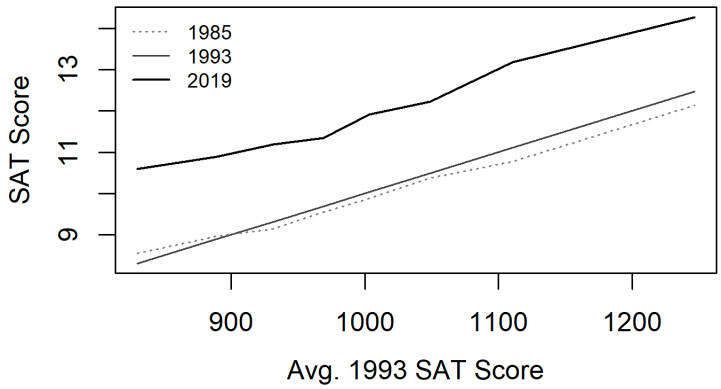
Note: This figure shows that the strong relationships between institutions' average SAT scores and both overall revenues/expenditures and instructional expenditures are present for both research-oriented R1 and non-research-oriented universities, suggesting that high-testing universities' research mission 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 2024. Source: IPEDS.

Figure A-3. State-Level Relationship Between Instructional Expenditures and SAT



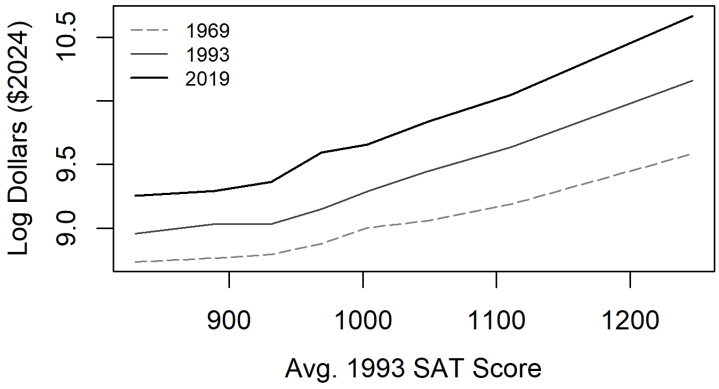
Note: This figure shows substantial heterogeneity across US states in the degree of instructional expenditure stratification between higher- and lower-testing public universities, with lower stratification in the Southeast (10-20 percent higher per-student expenditures per 100 average SAT points) and 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-4. Universities' Average SAT Score in Different Years by Average 1993 SAT Score



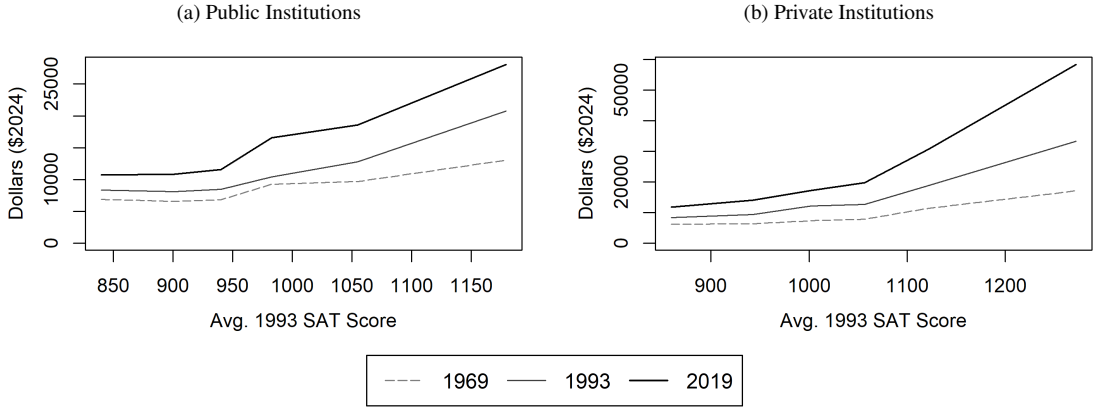
Note: This figure shows that the relative ranking of colleges and universities by average SAT score has remained largely unchanged, with very minimal mean reversion, since the mid-1980s, though average SAT scores rose considerably following College Board's 1995 'recentering'. Average freshmen-enrollment-weighted 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-5. 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 5). 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 2024. 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.

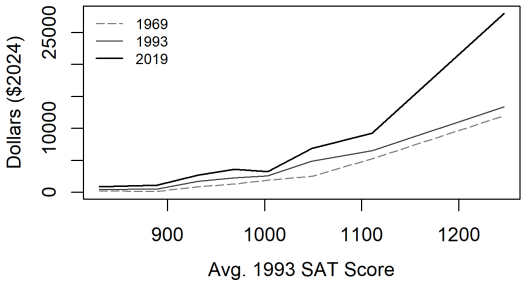
Figure A-6. Universities' Per-Student Instructional Expenditures by Average 1993 SAT Score By Control



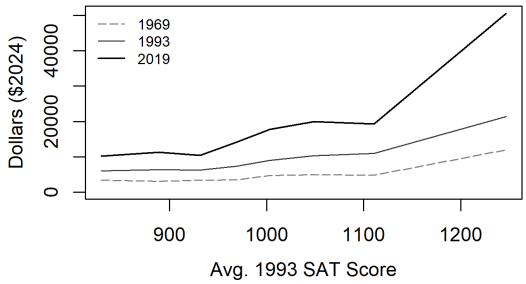
Note: This figure shows that the growth in stratification of per-student instructional expenditures by institutions' average SAT score has occurred at both public and private institutions since 1969, though to a greater degree at private institutions. Average freshmen-enrollment-weighted instructional 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 2024. 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 ($R^2 = 0.83$). Private institutions include all non-public institutions. Source: HEGIS, IPEDS, and the College Board Annual Survey of Colleges.

Figure A-7. Universities' Average Other Per-Student Expenditures by Average 1993 SAT Score

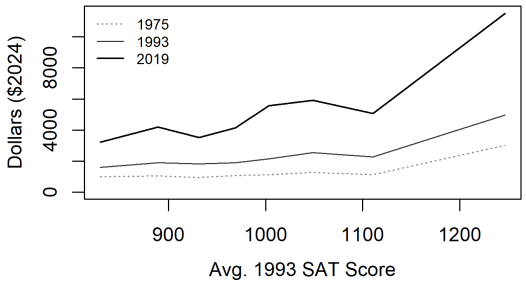
(a) Research Expenditures



(b) All Administrative Expenditures

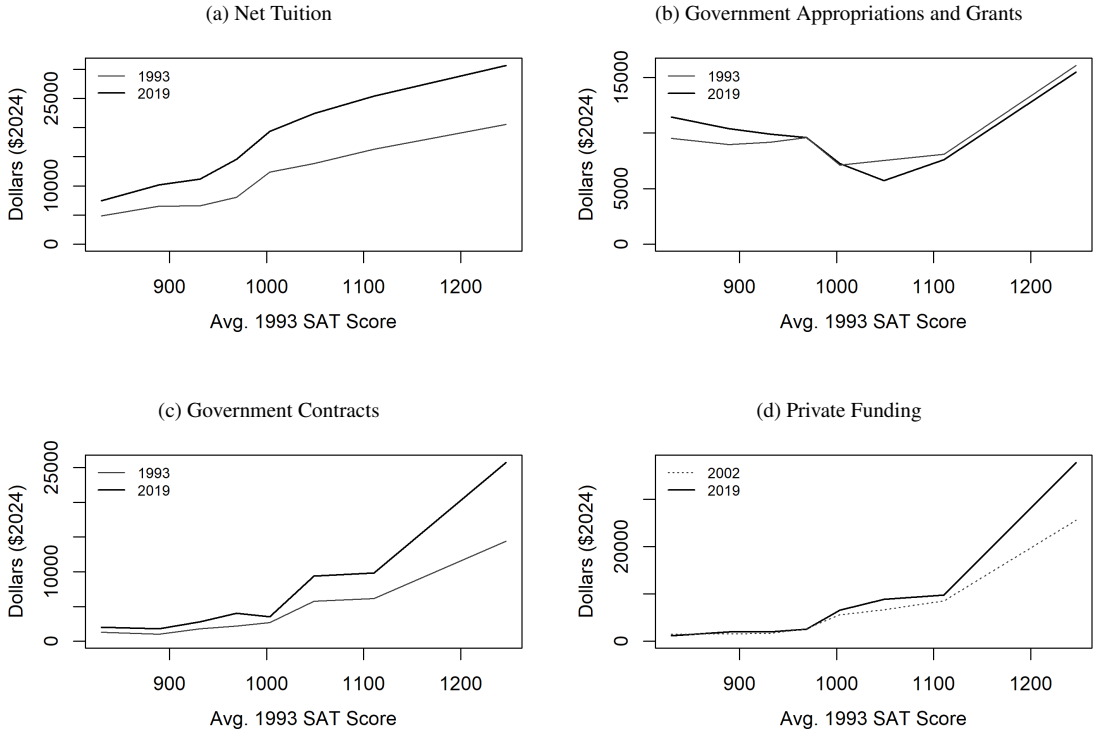


(c) Student Services Admin. Expenditures



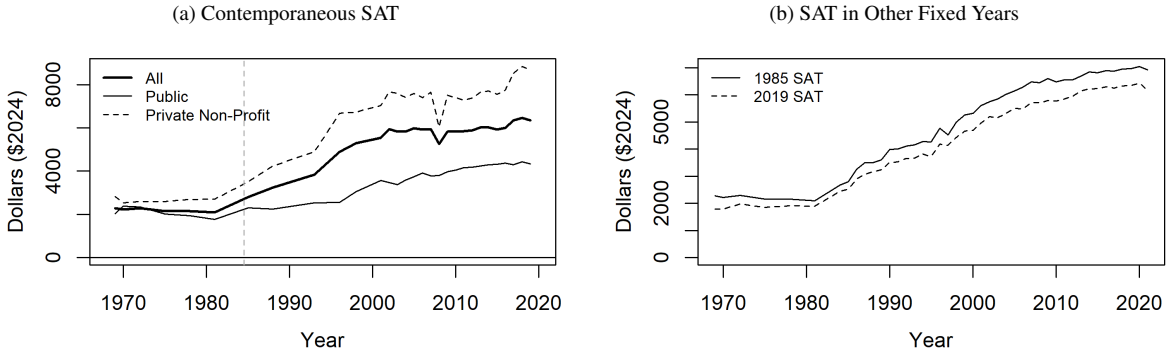
Note: This figure shows that per-student research and administrative expenditures have also stratified by institutions' average SAT score since 1969, especially at the top quantile of institutions and largely since 1993. Average freshmen-enrollment-weighted expenditures per FTE student by eight quantile 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 2024. 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-8. Universities' Average Revenues Per Student by Source and Average 1993 SAT Score



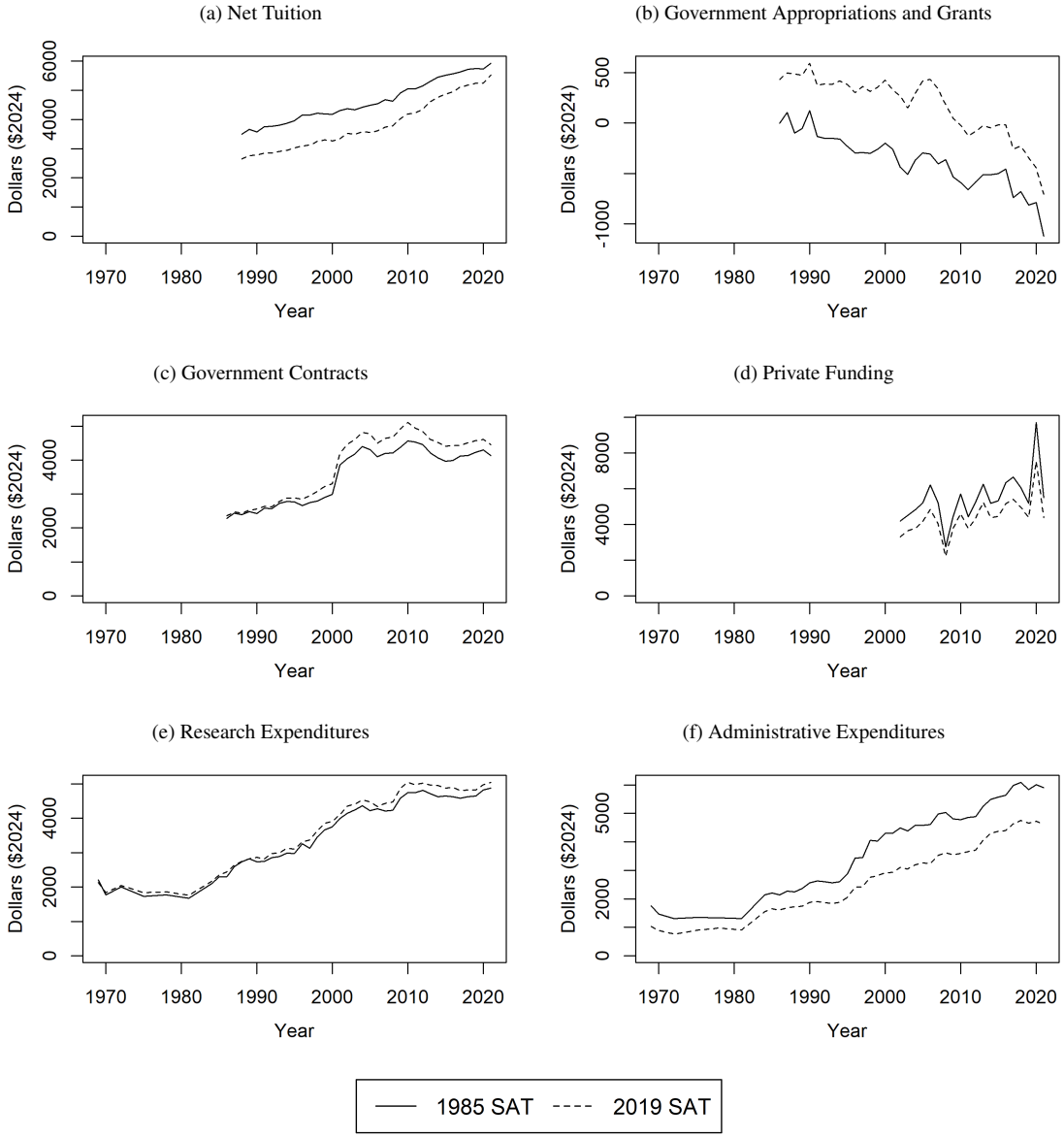
Note: This figure shows that the relative tuition, government research funding, and private funding received by higher-testing universities has risen since the mid-1990s, whereas government appropriations and grants have become slightly more evenly distributed over that period. Average freshmen-enrollment-weighted 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-9. Annual Association Between Per-Student Instructional Exp. and Average SAT Scores



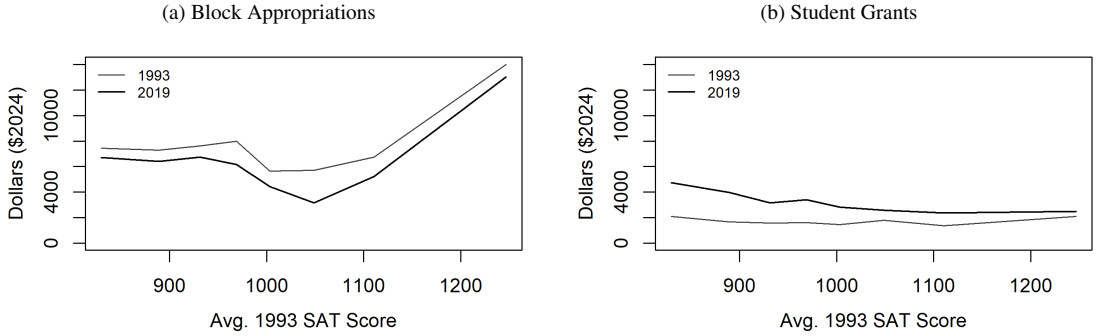
Note: This figure recapitulates the pattern shown in Figure 6 using contemporaneous SAT scores (after 1985) overall and by public-private control as well as using 2019 SAT scores, showing a similar trend overall (though the leveling-out may have occurred slightly before 2010) and in each sector (though 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 2024 dollars) and either contemporaneous average SAT score divided by 100, across four-year institutions overall and by public/private control, or average 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 2024 and annually winsorized at 1 percent. Source: HEGIS, IPEDS, and the College Board Annual Survey of Colleges.

Figure A-10. 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 2024 and annually winsorized at 1 percent. Source: HEGIS, IPEDS, and the College Board Annual Survey of Colleges.

Figure A-11. 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 ($2024)						
All Four-Year Institutions	10,300	23,700	31,900	39,500	47,500	56,300
Ivy League	20,200	86,900	104,400	140,600	223,600	258,000
State Flagship	12,200	37,900	41,500	55,400	71,500	81,800
Public Non-Flagship	10,300	18,300	28,500	32,900	38,700	43,900
Ivy vs. Flagship Gap (%)	65	129	151	154	213	215
Flagship vs. Non-Flagship Gap[1] (%)	9	54	51	62	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: 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 2024 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-3.

[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—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-3—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.