STANDARDIZED TESTING TASK FORCE (STTF)
Eddie Comeaux and Henry Sánchez, Co-Chairs

Assembly of the Academic Senate
1111 Franklin Street, 12th Floor
Oakland, CA 94607-5200

January 27, 2020

KUM-KUM BHAVNANI, CHAIR, ACADEMIC COUNCIL

Dear Kum-Kum:

Please find attached the final report of the Academic Council's Standardized Testing Task Force (STTF).

In January 2019, at the request of President Napolitano, Senate Chair Robert May empaneled the STTF to examine the University's current use of standardized testing for admission and consider whether the University and its students are best served by UC's current testing practices, a modification of current practices, another testing approach, or the elimination of testing. Chair May asked the STTF to develop a set of actionable recommendations to BOARS and the Academic Council.

The STTF and its Writing Subcommittee consulted dozens of studies concerning standardized tests, their predictive value, and their impact on access and diversity. It met with the national testing agencies, critics of standardized testing, State education leaders, UC campus admissions officers, UCOP institutional researchers, BOARS, and other UC-based and non-UC content experts. It examined the effectiveness of UC's current comprehensive review and "9 x 9" admissions policies, and considered the potential benefits and drawbacks of eliminating the testing requirement or moving to a "test optional" policy. It also considered potential alternative assessment tools, as well as alternative tests that might be more effective in meeting UC's goals around academic success and diversity.

The STTF engaged in complex and principled conversations about how to define and measure academic preparedness without contributing to preexisting disparities in educational opportunity along lines of race and class. It assessed how to weigh the goal of accurate prediction of UC performance against other, potentially competing goals, such as educating future leaders, representing the diversity of California, promoting socioeconomic mobility, and countering inequality. It investigated concerns about the potential correlation between race and SES status, on the one hand, and standardized test scores, on the other, as well as the impact of school segregation and resource inequality in California schools on UC applicants. These discussions revealed the complexity of the task of selecting students who should be admitted to UC, consistent with the goals of ensuring preparedness, pursuing diversity, and supporting opportunity, access, and fairness. The STTF and its subcommittee requested many additional analyses, including regressions and simulations of the effects of various possible reforms, to assist in the development of recommendations.

The resulting report and recommendations, grounded in evidence-based research and UC values, reflects the high quality of work one expects of UC faculty. The STTF co-chairs would like to recognize the hard work and dedication exhibited by the entire Task Force over the past year, and in particular the efforts of the STTF Writing Subcommittee (Chair Jonathan Glater, Julian Betts, Li Cai, Patricia Gandara, Andrea Hasenstaub, and George Johnson), whose extraordinary and sustained dedication made this report possible. The Subcommittee conducted original analyses of new data, endured long and intense conversations about the analyses and recommendations—including over the winter holiday season, on weekends, while on family vacation, and after hours—and relentlessly kept at the task of writing to meet accelerated deadlines of completing the report.

We also would like to thank Robert May for empaneling an outstanding group of UC faculty, graduate student representative, and Academic Senate staff. Finally, we thank you, Kum-Kum, for your leadership in guiding the Task Force through a set of challenging deadlines. We believe the attached report and appendix represents a living document to help frame actionable next steps to advance the University of California missions of access, opportunity, diversity, inclusion, and fairness.

Sincerely,

Henry Sánchez,	Eddie Comeaux,
STTF Co-Chair	STTF Co-Chair

UNIVERSITY OF CALIFORNIA Academic Senate

Report of the UC Academic Council Standardized Testing Task Force (STTF)

January 2020
Systemwide Academic Senate
University of California

Table of Contents

Executive Summary

In July 2018, President Napolitano wrote to 2017–18 Academic Senate Chair Shane White asking the Academic Senate to examine the current use of standardized testing for admission to the University of California (UC or the University); review the testing principles developed in 2002[1] by the Board of Admissions and Relations with Schools (BOARS) and revised by BOARS in 2010[2]; and determine whether any changes in policies on use of test scores are needed.

In early 2019, 2018–19 Senate Chair Robert May empaneled an eighteen-member Standardized Testing Task Force (STTF) to consider whether the University and its students are best served by UC's current testing practices, a modification of current practices, another testing approach, or the elimination of testing. Chair May asked the STTF to develop a set of actionable recommendations to the Academic Council; to approach its work analytically, without prejudice or presupposition; and to consult with experts from a broad range of perspectives. The STTF met 12 times between February 2019 and January 2020, and empaneled a six-member subcommittee to draft specific recommendations.

Chair May's charge to the STTF identified the following questions:

- *How well do UC's current standardized testing practices assess entering high school students for UC readiness?*
- *How well do UC current standardized testing practices predict student success in the context of its comprehensive review process?*
- *Do standardized testing assessments fairly promote diversity and opportunity for students applying to UC?*
- *Does UC's use of standardized tests increase or contract the eligibility pool compared to two other possibilities: 1) de-weighting standardized tests; or 2) eliminating the testing requirement?*
- *Should UC testing practices be improved, changed, or eliminated?*

By way of background, the University admits students through a two-stage process: first, a determination of eligibility for the University overall, then selection by a specific campus using comprehensive review. Test scores are used to establish eligibility for some (but not all) students and are also used in admissions decisions. Some campuses also use test scores for purposes other than admissions, such as awarding of scholarships, placement in classes, identification of students who might benefit from extra support, and admissions to honors programs. Standardized test scores tend to exhibit differences along lines of race and class, with students who belong to many of the demographic groups historically excluded from educational (and other) opportunities on average receiving lower scores.

How well do UC's current standardized testing practices assess entering high school students for UC readiness? How well do UC current standardized testing practices predict student success in the context of its comprehensive review process?

The STTF found that standardized test scores aid in predicting important aspects of student success, including undergraduate grade point average (UGPA), retention, and completion. At UC, test scores are currently better predictors of first-year GPA than high school grade point average (HSGPA), and about as good at predicting first-year retention, UGPA, and graduation.[3] For students within any given (HSGPA) band, higher standardized test scores correlate with a higher freshman UGPA, a higher graduation UGPA,

[1] https://senate.universityofcalifornia.edu/_files/committees/boars/admissionstests.pdf
[2] https://senate.universityofcalifornia.edu/_files/reports/hp2mgy_boars-testing_010609.pdf
[3] UCOP, Relationship of the SAT/ACT to College Performance at the University of California (2019).

and higher likelihood of graduating within either four years (for transfers) or seven years (for freshmen). Further, the amount of variance in student outcomes explained by test scores has increased since 2007, while variance explained by high school grades has decreased, although altogether does not exceed 26%. Test scores are predictive for all demographic groups and disciplines, even after controlling for HSGPA. In fact, test scores are better predictors of success for students who are Underrepresented Minority students (URMs), who are first-generation, or whose families are low-income: that is, test scores explain more of the variance in UGPA and completion rates for students in these groups. One consequence of dropping test scores would be increased reliance on HSGPA in admissions. The STTF found that California high schools vary greatly in grading standards, and that grade inflation is part of why the predictive power of HSGPA has decreased since the last UC study.

Do standardized testing assessments fairly promote diversity and opportunity for students applying to UC? Does UC's use of standardized tests increase or contract the eligibility pool compared to two other possibilities: 1) de-weighting standardized tests; or 2) eliminating the testing requirement?

The racial and ethnic makeup of the population admitted to the University differs significantly from that graduating from California high schools. In 2019, only 37% of the California resident students in the admitted freshman class, and 26% of all admitted students, were URMs, defined as Latino, African-American, and Native American students. Approximately 59% of California high school graduates were URMs. Further, the 37% (or 26%) are not distributed evenly across campuses; at Berkeley, for example, only 26% of admitted California freshmen (and only 18% of the admitted students overall) were URMs. The difference in fraction of URMs between the high school student population and the undergraduate population is a matter of concern. The Task Force was concerned about the extent to which standardized tests may contribute to this underrepresentation of URMs at UC.

Analysis of admissions results by the Task Force concluded that UC takes into account students' contexts when evaluating test scores. Applicants from less advantaged demographic groups are admitted at higher rates for any given test score as a result of comprehensive review, which is a process that evaluates applicants' academic achievements in light of the opportunities available to them and takes into consideration the capacity each student demonstrates to contribute to the intellectual life of the campus. The point was not what the STTF expected before we commenced our data analysis. Task Force members took note of the average differences in test scores among groups and expected to find that test score differences explain differences in admission rates. That is not what we found. Instead, the STTF found that UC admissions practices compensated well for the observed differences in average test scores among demographic groups. This likely reflects UC's use of comprehensive review, as well as UC's practice of referencing each student's performance to the context of their school.

The Task Force found that test scores appear to receive less weight in admission than high school grades: for example, for every campus, the admit rate is much higher for applicants with high HSGPA and low SATs than for those with high SATs and low HSGPAs. On average, disadvantaged and URM students applied to UC with HSGPAs significantly lower than the HSGPAs of more advantaged groups, and the UC admissions process did not appear to compensate for HSGPA as it does for test scores.

The Task Force found that of the 22,613 students guaranteed admission through the statewide index (the admissions pathway in which test scores can compensate for lower HSGPA) but not through Eligibility in the Local Context (ELC), which only considers high school grades, about 25% were members of underrepresented minority groups, and 47% were low-income or first generation students. These students would not have been guaranteed admission on the basis of their grades alone.

The Task Force examined the pipeline for high school seniors for the class of 2018 to the admitted pool of freshmen that year. There was a 22-percentage point gap between the share of URMs in the grade 12 class and the pool of California resident students admitted by UC. The Task Force concluded that multiple factors contribute to this underrepresentation at UC. Roughly 75% of the opportunity gap arises from factors rooted in systemic racial and class inequalities that precede admission: lower high school graduation rates for URMs, lower rates of completion of the A-G courses required by UC and CSU, and lower application rates. The most significant contributor was lack of eligibility as a result of failure to complete all required A-G courses with a C or better. Roughly 25% of underrepresentation was due to UC admissions decisions overall. Test scores play a role in those decisions, and thus account for some of that 25%, even if they are not the primary barrier to admission. It is also unknown to what extent low scores on the standardized tests deter some students from even submitting applications.

In sum, mean differences in standardized test scores between different demographic groups are often very large, and many of the ways these tests could be used in admissions would certainly produce strong disparate impacts between groups. However, UC weights test scores less strongly than GPA, and comprehensive review appears to help compensate for group differences in test scores. The distributions of test scores among applicants are very different by group, but the distributions of test scores among admitted students are also very different by group, and in almost exactly the identical way. The Task Force did not find evidence that UC's use of test scores played a major role in worsening the effects of disparities already present among applicants and did find evidence that UC's admissions process helped to make up for the potential adverse effect of score differences between groups.

Yet this is not to conclude that consideration of test scores does not adversely affect URM applicants. If standardized test scores must be compensated in order to achieve the entering class sought by UC, that is reason to question whether it is necessary to use the tests at all, and/or whether it is possible to design an alternative instrument that does not require such compensation. UC admissions practices do not fully make up for disparities that persist along lines of race and class. Whether these disparities arise from test scores, GPA, or others among the 14 factors that comprise comprehensive review at UC, the outcome of UC admissions processes is that many of the populations historically excluded from opportunity are still underrepresented by wide margins. Some members of the Task Force emphasized UC's responsibility to assist disadvantaged and URM students who attend schools with lesser resources than those attended by students from affluent families, and they worried that continuing to use tests would help preserve the status quo. These members contended that the University has an obligation to interrupt perpetuation of inequality, especially when the state played a historic role in creating it, and must do more to serve the state's aspiring college students more equitably.

Should UC testing practices be improved, changed or eliminated?

In reaching its recommendations, the Task Force weighed both the importance of admitting students most likely to succeed at UC and the importance of enrolling classes representative of California's diversity. STTF members had varying perspectives on the relative importance of these goals. Members of the Task Force were also aware of the public scrutiny of admissions practices generally and consideration of standardized test scores in particular. The Task Force carefully considered the relative risks and benefits of possible steps that UC could take to ensure that the population of admitted students reflects the diversity of the state.

The STTF produced the following consensus recommendations, offered below with brief explanations:

1) Review and update components of the statewide index used to identify UC-eligible students. While the University *admits* students to campuses based on 14 factors, we base *systemwide eligibility* on only two — HSGPA and standardized test scores. The Task Force recommends that BOARS consider whether, among the 14 factors, there are additional, quantitative features that could be included in the statewide eligibility index.

2) Expand the ELC pathway, which admits students in the top 9% of each school based on HSGPA alone, to admit more than the top 9%, so as to increase access and representation, while leaving the statewide index formula unchanged. This would result in growth in the student body, but would avoid likely opposition that would result if students who would have been admitted under the current statewide eligibility rules are excluded as a result of shrinking the statewide pool to accommodate growth in the ELC pool. In addition, allowing growth in the undergraduate population that would come from expanding ELC would send a signal to Californians about the high value that UC places on becoming more representative of the state over time.

3) Undertake further analysis of the admissions process to identify those points in the process and the factors at work at each point that contribute to disproportionate declines in representation of students who belong to populations that have historically been excluded from higher education opportunity. This research matters because the Task Force's preliminary analysis indicates that disparities along lines of race, ethnicity, and socioeconomic status (SES) evident in the undergraduate population of the University are a function of multiple factors, and that the SAT and ACT are smaller contributors. For example, disparities in access to and completion of A-G courses account for a disproportionate lack of UC eligibility for students who are members of underrepresented groups. Differences in A-G completion rates across schools also reduce the yield of students through ELC. This study may inform new reforms to the admissions process to achieve greater equity in access to UC.

4) Study and expand student academic support services. A multi-pronged effort to study, fine-tune and expand the system of academic and socio-emotional supports that UC provides to its undergraduate students would begin by identifying and reporting on existing programs, then proceed to evaluate them. UC would then be in position to design and evaluate new support programs implementing best practices, as needs are identified. The Task Force believes this effort will promote University efforts to improve important measures of student success, including undergraduate GPA and likelihood of graduation, for students who are members of groups historically excluded from and underrepresented on college campuses, and who might benefit from support to excel in UC classes.

5) Obtain the necessary item-level data to perform an updated analysis of potential bias in current standardized tests. With the most recent operational student-level data from College Board and ACT in the form of raw item responses, as well as necessary applicant and school background information as they are available, UC can attempt to provide a definitive answer to the question of whether the tests and the items making up the tests are biased against members of particular subgroups of test-takers.

6) Develop a new assessment system that will be continuously accessible to students and that will assess a broader array of student learning and capabilities than any of the currently available tests. This new assessment system should provide more detailed feedback to schools and to students, allow students to learn the item types through multiple opportunities to practice, and align programs in Recommendation 4 to be targeted to students' learning needs. The new assessment may enable UC to admit classes of students more representative of the diversity of the state. A new system should allow the measurement of student progress over time and development of a broader range of skills and attributes relevant to an applicant's success at UC, reflecting the principle that the University "should select those students who demonstrate a strong likelihood that they will persist to graduation." In particular, based on the research reviewed, the Task Force believes that UC will be able to articulate a set of skills and attributes that a) can be measured in a uniform and quantitative way, b) are likely to be predictive of success at UC, and c) could potentially show smaller disparities than current measures along the lines of race, ethnicity, and SES. The tests should continue to evaluate math and reading achievement due to their foundational importance in college and career success. The formats

of the tasks on the new assessments can and should also take full advantage of technological advances to let the aspiring applicants exercise a wider range of knowledge, skills, and reasoning than multiple-choice questions do.

The timeline for development of a new suite of assessments would put widespread implementation nine years out. Members of the Task Force differed on the question of whether to recommend that UC cease consideration of standardized test scores sooner — in all likelihood *before* availability of the replacement suite of assessments. The differences reflected uncertainty within the Task Force, given the data available to us, over the relative risks and benefits of possible, additional steps that UC should take to ensure that the population of admitted students reflects the diversity of the state.

The Task Force considered, but does not recommend, the following possibilities.

7) The Task Force does not recommend that UC make standardized tests optional for applicants at this time. UC should conduct additional research on the impact of going "test optional" before deciding whether and how to implement such a policy. STTF had pragmatic concerns about how campuses would evaluate and compare applicants who submit standardized test scores relative to applicants who do not; whether and how campuses would impute, explicitly or implicitly, test scores to applicants; and ethical concerns about how to treat students in the two groups fairly. The current UC admissions practice of putting each applicant's test scores into context by comparing them to all applicants from the same school, thus allowing readers to identify students who performed exceptionally well given available opportunities, could no longer be used if students could choose whether or not to submit their test scores. Finally, given the state of research and UC's size and unique system of "in the context" review, it was not clear that going "test-optional" would increase diversity of incoming classes, and the likely impact on preparation of incoming classes was not clear.

8) The Task Force does not recommend adopting the Smarter Balanced (SBAC) Assessment in lieu of currently used standardized tests. The Task Force was deeply concerned with a wide range of risks related to the adoption of the SBAC to meet UC admissions requirements, including (but not limited to) test security, item exposure control, and item bank size, as well as inconsistent implementation of the SBAC across states, ambiguity regarding instructional validity, and the likelihood that this change would reduce the utility of SBAC for its main purpose.

I. Introduction

The University of California ("UC" or "the University") faces a crisis of legitimacy around its undergraduate admissions processes. On one hand, some believe that the prior academic achievements of entering undergraduates affect UC's institutional excellence and that students' scores on standardized admissions tests – the SAT and ACT – reflect academic achievement. Consequently, for this group, abandoning consideration of test scores in admissions risks compromising the legitimacy of the University's claim to excellence. On the other hand, some believe that continued consideration of standardized test scores undermines institutional legitimacy from a different perspective: that by using these tests, on which students who belong to historically excluded groups receive lower scores, the University is complicit in perpetuating existing, longstanding inequities in educational opportunity.

The debate over the admissions process consequently implicates additional debates over what it means to be an excellent institution – i.e., are admitted students' standardized test scores critical to excellence – and what it means to be a public institution. We believe that the path to reconciliation rests on development of a definition of excellence that enables the University both to assert academic preeminence and to use its powerful role to interrupt inequality of educational opportunity along lines of race, class, and ethnicity. We believe that the role of a public university is not simply to identify smart students but also to develop informed citizens who will contribute to the common good.

The Organic Act of 1868 establishing the University of California included a set of principles to guide the institution. Among these was a definition of whom it would serve:

> "[It] shall be the duty of the regents, according to population, to so apportion the representation of students, when necessary, that all portions of the State shall enjoy equal privilege therein." (Organic Act, 1868, Sec. 14)

To some extent UC has been grappling with this principle ever since. Its founders could not have imagined how the state and the University would change over time or how "all portions of the state" might come to be re-defined. In the nearly 60 years since the Master Plan for Higher Education[4] was established, limiting eligibility for UC to the top 12.5% of California high school graduates, the state's population has changed dramatically. In 1960, California's population was overwhelmingly white and English speaking and immigration was at an historic low point. Income inequality was also at a low point and California enjoyed a reputation for having some of the nation's best schools. California had invested heavily in both its K-12 and higher education systems. The College Board's SAT, and later the ACT, were incorporated into UC's admissions requirements in part to determine the eligible pool of students and in part to standardize and establish a common metric for all applicants. However, as the California population has become increasingly diverse, the University has come under greater pressure to ensure that the students it admits reflect the state's high school graduates with respect to race, ethnicity, and socio-economic status (SES). These student characteristics were not considered in the language of the 1960 Master Plan.

In 2017, the most recent year for which we have data, 61% of California high school graduates were underrepresented minorities (URM), a category that UC defines as students who are African-American, Native American, and Latino) while just 31% of freshman enrollees to UC were URMs.[5] At Berkeley, fewer than 18% of freshman enrollees were URMs, and at UCLA, fewer than 28% were. URM students

[4] https://www.ucop.edu/acadinit/mastplan/MasterPlan1960.pdf
[5] https://www.universityofcalifornia.edu/infocenter/admissions-residency-and-ethnicity

are enrolled disproportionately at the University's less competitive campuses. These disparities across UC campuses have raised questions about fairness among some critics.

There are reasons unrelated to standardized tests that some URM students are not eligible to attend UC, and other hurdles to admission contribute to their disproportionate exclusion. But the tests play a role in eligibility and admissions decisions. It is ironic that the SAT was developed during the early part of the last century for the purpose of identifying intellectually talented male students without ties to the nation's elite families, or who resided outside the closed circle of prep schools in the Northeast.[6] This allowed elite colleges to enroll a somewhat more diverse student body. Today, some critics assert that use of the test scores has the opposite effect.

A recent nationwide scandal that involved wealthy families paying for fraudulent test scores on standardized admission tests brought further scrutiny of the use of these tests in admissions. It also brought an outcry against the tests from many quarters; some University Regents, state legislators, community groups and public interest lawyers have demanded an end to their use.[7] Several UC leaders have weighed in on the issue. Amidst these concerns, as well as concern over excessive testing in K-12 schools and recognition that admissions tests measure only a fraction of the skills and knowledge students are likely to require in college and later life, a growing number of colleges and universities has abandoned the SAT and ACT test requirements or gone test-optional. The University of Chicago's decision in 2018 to make the SAT and ACT tests optional has been particularly noteworthy as it, like UC, is very selective. It is also noteworthy, however, that the debates over the use of standardized tests in undergraduate admissions is longstanding and dates back to the mid-1960s.[8]

In July 2018, UC President Janet Napolitano asked the Academic Senate to undertake a study to "examine the current use of standardized testing for UC undergraduate admission; review the testing principles developed in 2002 and revised in 2010; and determine whether any changes in admission testing policies or practices are necessary to ensure that the University continues to use standardized tests in the appropriate way." At the September 2018 meeting of the University of California Regents, Robert May, then chair of the Academic Senate, accepted the request and in January 2019 laid out the charge to the Standardized Testing Task Force (STTF). The charge included "examin[ing] how well our current standardized testing is working, how well it predicts student success in the context of our comprehensive review process for admissions; could it be improved, changed or eliminated; and is it fair and unbiased for the diversity of students applying to UC?"[9]

The work of the STTF is guided in part by existing University policies. In May 1988, the UC Regents affirmed its position on undergraduate admissions in Policy 2102, which has remained in force to the present. It includes the following text:

> The entrance requirements established by the University follow the guidelines set forth in the Master Plan, which requires that the top one-eighth of the state's high school graduates, as well as those transfer students who have successfully completed specified college work, be eligible for admission to the University of California. These requirements are designed to ensure that all eligible students are adequately prepared for University-level work.

---

[6] Lemann, N. (1999). The Big Test. The Secret History of the American Meritocracy. New York: Farrar, Straus and Giroux

[7] Watanabe, T. (2019). *Put to the test, GPA and SATs may not tell all*, Los Angeles Times (December 22, 2019), at A1.

[8] Douglass, J.A. (2004). The Conditions of Admission. Access, Equity and the Social Contract of Public Universities. Palo Alto: Stanford University Press.

[9] https://senate.universityofcalifornia.edu/_files/committees/sttf/standardized-testing-tf-charge.pdf

> Mindful of its mission as a public institution, the University of California has an historic commitment to provide places within the University for all eligible applicants who are residents of California. The University seeks to enroll, on each of its campuses, a student body that, beyond meeting the University's eligibility requirements, demonstrates high academic achievement or exceptional personal talent, and that encompasses the broad diversity of cultural, racial, geographic, and socioeconomic backgrounds characteristic of California.
>
> Because applicant pools differ among the campuses of the University, each campus shall establish procedures for the selection of applicants to be admitted from its pool of eligible candidates. Such procedures shall be consistent with the principles stated above and with other applicable University policies.[10]

The loss of affirmative action in UC admissions policy, first in 1995 by a vote of the Regents on Special Policy 1 (SP-1), and then in 1996 by a state-wide referendum,[11] dealt a blow to the University's ability to admit a diverse class from which UC has never fully recovered.[12] An important tool for diversity was removed, which put greater pressure on the system to find new ways to support diversity in admissions.

This is the third time in almost two decades that the University of California has undertaken a review of its admissions procedures to ensure fairness to all applicants. The most recent studies were undertaken by the Board of Admissions and Relations with Schools (BOARS) in 2002 and in 2010. In 2002, BOARS concluded unanimously that "admissions tests serve a useful purpose in helping both to determine UC eligibility and to select applicants for admission to campuses that cannot accommodate all UC-eligible applicants. While the grade point average (GPA) in UC-approved college preparatory courses taken in high school is clearly the best indicator of likely success, standardized tests can provide additional information that is helpful in illuminating the high school record …" (BOARS, 2002, executive summary, p. 2).[13] In this report, BOARS also articulated a set of principles to be applied to all admissions *tests.* BOARS wrote:

1. Admissions tests will be used at the University of California
    - To assess academic preparation and achievement of UC applicants;
    - To predict success at UC beyond that predicted by high school GPA;
    - To aid in establishing UC eligibility; and
    - To aid in selecting students for admission at individual UC campuses.

2. The desired properties of admissions tests to be used for these purposes include the following:
    - An admissions test should be a reliable measurement that provides uniform assessment and should be fair across demographic groups.
    - An admissions test should measure levels of mastery of content in UC-approved high school preparatory coursework and should provide information to students, parents, and educators enabling them to identify academic strengths and weaknesses.
    - An admissions test should be demonstrably useful in predicting student success at UC and provide information beyond that which is contained in other parts of the application. (It is recognized that predictors of success are currently limited, and generally only include first-

[10] https://regents.universityofcalifornia.edu/governance/policies/2102.html
[11] Proposition 209.
[12] Kidder, W. & Gándara, P. (2015). Two Decades After the Affirmative Action Ban: Evaluating the University of California's Race-Neutral Efforts. Princeton, NJ: Educational Testing Service.
[13] https://senate.universityofcalifornia.edu/_files/committees/boars/testsummary.pdf

year college GPA and graduation rate. As this field advances, better predictors should be identified and used in validating admissions tests.)

- An admissions test should be useful in a way that justifies its social and monetary costs.

BOARS also recommended that, as a matter of principle, the faculty regularly review UC's admissions testing policy and practices to ensure that tests are being used in a way that is consistent with these principles and desired properties of admissions tests. In 2003, the UC Regents accepted the BOARS recommendations, and College Board and ACT revamped their tests to better align with high school subject matter.

In 2010, BOARS reaffirmed the testing principles and recommended that the Regents remove the provisional status of the SAT-R (the revised SAT designed to better align with high school subject matter). BOARS also noted that it planned "to continue to study and monitor the use of tests in UC admissions, revise the testing principles, *and determine whether alternative paths are possible to diminish reliance on tests in the future"* [italics added].[14] Thus in 2010, the door was left open to consider other admissions practices and procedures.

Beginning in 1998, not long after the passage of SP-1, the Berkeley campus instituted a form of comprehensive review that evolved over time, and that the other campuses came to adopt in varying forms.

In 2001, BOARS developed a set of principles for Comprehensive Review, which all campuses were expected to adhere to, each in its own autonomous way. These procedures are very relevant to the examination of fairness of the UC admissions processes as they are conducted at each campus. The UC Regents define comprehensive review as: "the process by which students applying to UC campuses are evaluated for admission using multiple measures of achievement and promise while considering the context in which each student has demonstrated academic accomplishment."[15] In designing campus procedures, campus admissions committees should adhere to the following guiding principles:

1. The admissions process honors academic achievement and accords priority to students of high academic accomplishment. At the same time, merit should be assessed in terms of the full range of an applicant's academic and personal achievements and likely contribution to the campus community, viewed in the context of the opportunities and challenges that the applicant has faced.
2. Campus admissions procedures should involve a comprehensive review of applications using a broad variety of factors to select an entering class.[16]
3. No fixed proportion of applicants should be admitted based solely on a narrow set of criteria.
4. Campus policies should reflect continued commitment to the goal of enrolling classes that exhibit academic excellence as well as diversity of talents and abilities, personal experience, and backgrounds.
5. Faculty on individual campuses should be given flexibility to create admission policies and practices that, while consistent with University-wide criteria and policies, are also sensitive to local campus values and academic priorities.
6. The admission process should select students of whom the campus will be proud, and who give

[14]Admissions Tests and UC Principles for Admissions Testing, available at https://senate.universityofcalifornia.edu/_files/reports/hp2mgy_boars-testing_010609.pdf

[15] https://regents.universityofcalifornia.edu/governance/policies/2104.html

[16] Comprehensive review includes 14 factors that are to be considered in all applications, including GPA and standardized tests scores, but also special talents and achievements, proficiency in other languages and academic achievement in the context of disadvantage, among others. The full list can be found at https://admission.universityofcalifornia.edu/how-to-apply/applying-as-a-freshman/how-applications-are-reviewed.html.

evidence that they will use their education to make contributions to the intellectual, cultural, social, and political life of the State and the Nation.

7. The admissions process should select those students who demonstrate a strong likelihood that they will persist to graduation.
8. Campus selection policies should ensure that no applicant will be denied admission without a comprehensive review of his or her file.[17]

Because differences in educational opportunity vary substantially by race and class among California's students, the 2010 Recommendations on Comprehensive Review state that a comprehensive review involving both quantitative evaluation as well as evaluation of non-quantified criteria, should be undertaken for each applicant who meets the minimum criteria for UC eligibility.[18] Since 2010, incremental changes have increased the fraction of students admitted through this "Entitled to Review" (ETR) pathway as well as the fraction of students admitted through "Eligibility in the Local Context" (ELC), under which students whose grades put them in the top 9% in each California public high school are deemed eligible[19] based on GPA in a specific set of courses regardless of admissions test scores.

Although each review of admissions procedures has resulted in some changes, standardized tests have remained a part of the freshman admissions evaluation since they were instituted. The comprehensive review is the latest major reform to the admissions process, but it still includes consideration of standardized test scores, which some campuses weigh heavily and to varying degrees. The charge given to the STTF to review current admissions procedures necessarily begs the question of the goal of a University of California education: Is the goal to maintain UC's reputation for excellence, which some contend depends on admitting students who have shown academic achievement in the form of high test scores that correlate with greater likelihood of receiving higher grades and graduating? Or is it to admit a class representative of the public that UC was created to serve, thereby acknowledging some of the vast inequities in California society? Currently, the University is under increasing pressure to ensure that the student body more closely represents the state population and includes a broader range of students, especially at the flagship campuses. Standardized tests may identify some number of students overlooked by other metrics, but they also operate to exclude some number of students.

The choices before the STTF were to recommend changes to admissions policy that either eliminated or altered the way the tests are used in the process of admitting students or to maintain the status quo, while possibly tweaking the system once again at the margins. On the one hand, BOARS and the Academic Senate have reviewed the use of standardized tests several times before and decided repeatedly to retain them in some form. These decisions are defensible in that the standardized tests do have predictive power above and beyond high school GPA (HSGPA), they provide a common metric across students from very different schools and communities, and as admissions officers told the STTF, "We know how to do this." Any change would require doing something that UC does not yet know how to do and could conceivably result in greater inequities and worse student outcomes, particularly if changes were mandated immediately and without adequate consultation with admissions officers on the various campuses. On the other hand, if UC chooses to resist (or appears to resist) the pressures to change the current admissions procedures, outside political forces might compel change, and mandated reforms likely would be less informed, less effective, and more prone to produce unexpected and undesired outcomes than reforms

[17] University of California, Guidelines for Implementation of University Policy on Undergraduate Admissions, p. 2, at https://www.ucop.edu/student-affairs/_files/guidelines-for-implementation-of-undergraduate-admissions--rev-7-2019.pdf.

[18] For example, students must have met the A-G requirement. This and other requirements are specified in Regents Policy 2103, https://regents.universityofcalifornia.edu/governance/policies/2103.html

[19] That is to say, not admitted to a particular campus, but eligible for admission to one of the UC campuses.

produced within the University. Such reforms might also miss the opportunity to place greater emphasis on holistic review than is now practiced, especially on some campuses.

The Task Force also considered that if UC chooses not to revise how it uses standardized test scores in admissions decisions, the University could be passing up an opportunity to address the extent to which public higher education in the state should weigh more heavily in admissions the non-quantitative, non-academic achievements reflected principally in comprehensive review factor 11.[20]

University admissions pursues multiple goals, including developing academic excellence, advancing equity in access to higher education opportunity, and producing graduates who will serve the campus and later the larger society well in whatever capacity, with an eye to the greater good. These goals will not be met by strictly prioritizing students who have received the highest grades and scores on standardized tests at the time they apply to college. As expressed in Regents Policy 2106:

> The University of California is committed to excellence and equity in education for all of California's students to secure the social well-being and economic prosperity of the individual and the State. The University affirms that a fundamental part of its mission is to engage in efforts to promote the academic achievement and success of all students, including students who, because they are educationally disadvantaged and underrepresented, therefore need additional assistance.[21]

It was with this understanding that the STTF reviewed the options before it.

---

[20] Factor 11 is defined as follows: "Special talents, achievements and awards in a particular field, such as visual and performing arts, communication or athletic endeavors; special skills, such as demonstrated written and oral proficiency in other languages; special interests, such as intensive study and exploration of other cultures; experiences that demonstrate unusual promise for leadership, such as significant community service or significant participation in student government; or other significant experiences or achievements that demonstrate the student's promise for contributing to the intellectual vitality of a campus." https://senate.universityofcalifornia.edu/_files/committees/boars/documents/guidelines-implementation-of-ug-admission-rev-7-2019.pdf

[21] https://regents.universityofcalifornia.edu/governance/policies/2106.html

II. How UC Uses Standardized Tests in Admissions

The University of California ("UC" or "the University") admits students through a two-stage admissions process: first, a determination of eligibility for the UC system overall, and then selection by a specific campus using comprehensive review. Test scores are used to establish eligibility for some (but not all) students. Test scores are used in admissions decisions. Test scores are also used at some campuses for purposes beyond campus admissions, such as part of scholarship decisions, for placement, to identify students who would benefit from extra support, and in honors program admissions. These uses of tests are described below, beginning with a review of the two-stage UC admissions process and the uses of tests in admissions.

Eligibility versus selection. UC campuses consider 14 factors in shaping their freshman classes by selecting among qualified applicants. These "comprehensive review" factors include quantifiable features of academic records such as high school GPA (HSGPA) and standardized test scores, the number of A-G and honors classes the student completed before the end of the junior year, and how many A-G and honors courses the student is taking in the senior year. The University also considers factors related to background, opportunities, and socioeconomic status (such as first-generation, family income, and academic opportunities in their local school), as well as more difficult to quantify factors such as a student's trajectory of improvement and the strength of a student's extracurricular activities. However, UC eligibility is based on only two of these factors: 1) GPA in a specific set of sophomore and junior year courses, and 2) in some cases, test scores. Eligible students are guaranteed admission to UC (i.e., guaranteed to be admitted to at least one campus). The concept of eligibility applies only to in-state public and private high school students, not to out-of-state, international, or homeschooled students.

Eligibility. Applicants may be eligible for UC admission in either "the local context" or the "statewide context." Any student who is UC eligible is guaranteed admission to at least one UC campus, although not necessarily their top choice. The top 9% of students in each participating high school in California are deemed "eligible in the local context" (ELC). This determination is made based on HSGPA in a specific set of courses, without considering admissions test scores. In 2019, roughly 32% of California freshman applicants to UC had earned eligibility for admission based on eligibility in the local context. The rationale for ELC was that "Using superior scholarship identified in every high school in the state as a basis for eligibility will foster equal opportunity to attend UC for students from all parts of California, regardless of their socio-economic background or the richness of educational opportunities available to them."[22] In addition, the top 9% of students statewide are deemed "statewide eligible" based on an index that combines HSGPA with a core test score. The index allows a student's strength in one measure to compensate for weakness in the other. Index eligibility allows UC to admit academically strong students who are not in the top 9% of their high school class.

Individual UC campuses may also enroll students who are neither index eligible nor ELC (i.e. Entitled to Review [ETR]), provided they have completed a specific complement of A-G courses with a 3.0 GPA and submit a SAT or ACT score but there is no guarantee of admission as there is with students meeting the statewide index or ELC. There is no minimum test score requirement for ETR students. In addition, one-third of the seats in each junior class are reserved for transfer students from California community colleges, who are admitted based on the community college courses taken and grades received and are not required to take a core admissions test.

---

[22] Regents item: Establishment of UC Freshman Eligibility in the Local Context (ELC), Proposed Changes in Academic Course Requirements for UC Freshman Eligibility (VPA Requirement), Proposed Changes in GPA Calculation for UC Freshman Eligibility. Available at:
https://regents.universityofcalifornia.edu/minutes/1999/edpol399.pdf

Selection through comprehensive review. UC considers itself to be "one system with many campuses." Each undergraduate campus selects its own incoming class: that is, each campus is independently responsible for issuing a set of offers that will produce an enrolled freshman pool that 1) matches overall enrollment to physical and instructional capacity 2) balances enrollment across majors 3) is academically qualified, such that most students will re-enroll after the freshman year and ultimately complete degrees 4) admits out-of-state/international students who are more qualified, on average, than in-state students 5) "demonstrates high academic achievement or exceptional personal talent, and encompasses the broad diversity of backgrounds characteristic of California".[23]

While each campus has a different strategy for achieving these (sometimes conflicting) goals, **all UC campuses select students for admission by conducting a comprehensive review of all application information, including standardized test scores**. It is here useful to review the principles for Comprehensive Review developed by BOARS in 2001.[24] BOARS defines comprehensive review as: *the process by which students applying to UC campuses are evaluated for admission using multiple measures of achievement and promise while considering the context in which each student has demonstrated academic accomplishment.* The comprehensive review criteria established by BOARS include both quantitative and non-quantitative criteria. As noted in the Introduction, BOARS has identified the following principles of comprehensive review, noting that all campuses are expected to adhere to these principles, with each campus given leeway to adapt as best suits local priorities and circumstances:

> *1) The admissions process honors academic achievement and accords priority to students of high academic accomplishment. At the same time, merit should be assessed in terms of the full range of an applicant's academic and personal achievements and likely contribution to the campus community, viewed in the context of the opportunities and challenges that the applicant has faced.*
>
> *2) Campus admissions procedures should involve a comprehensive review of applications using a broad variety of factors to select an entering class.*
>
> *3) No fixed proportion of applicants should be admitted based solely on a narrow set of criteria.*
>
> *4) Campus policies should reflect continued commitment to the goal of enrolling classes that exhibit academic excellence as well as diversity of talents and abilities, personal experience, and backgrounds.*
>
> *5) Faculty on individual campuses should be given flexibility to create admission policies and practices that, while consistent with University-wide criteria and policies, are also sensitive to local campus values and academic priorities.*
>
> *6) The admission process should select students of whom the campus will be proud, and who give evidence that they will use their education to make contributions to the intellectual, cultural, social, and political life of the state and the nation.*
>
> *7) The admissions process should select those students who demonstrate a strong likelihood that they will persist to graduation.*

[23] https://regents.universityofcalifornia.edu/regmeet/may01/re28new.pdf
[24] Guidelines for Implementation of University Policy on Undergraduate Admissions, Guiding Principles For Comprehensive Review available at https://senate.universityofcalifornia.edu/_files/committees/boars/CRGUIDELINES.pdf

8) Campus selection policies should ensure that no applicant will be denied admission without a comprehensive review of his or her file.

Each campus implements comprehensive review in different ways. Most use holistic review, in which no applications are filtered out based only on standardized test scores or only on HSGPA, no fixed weights are assigned to any of the 14 factors, and applications are read by trained readers who see – in addition to raw GPAs, test scores, number of honors courses taken, and so on – a rich array of information that places these raw numbers "in the context" in which the achievements were obtained. For instance, GPAs are shown both as raw GPAs and as percentiles with respect to the local high school pool of applicants to UC as a whole and to the specific campus, in addition to the percentile of all applicants to UC as a whole and to the specific campus. Number of honors courses taken is shown next to the number of courses available at the school. Test scores are shown next to the percentile the test score represents with respect to test-takers in the local high school class and with respect to the applicant pool overall.

Two campuses incorporate a fixed-weight system in which applications are scored prior to comprehensive review: Riverside uses a fixed weight system in which test scores are ~41% of the total weight (GPA represents ~50% total weight), and Santa Barbara uses a data-driven complex system in which tests ultimately comprise 20-25% of the total admissions score (depending on major). Even at score-using campuses, applications are read as well as scored. Admissions decisions are ultimately made based on a review of all the information available in the applicant's file.

Students who are eligible for UC admission, but are not selected for admission by one of the campuses to which they applied, are offered admission to another campus/major combination (generally Merced) through the "referral pool."

Scholarships. Some, but not all, campuses consider test scores in assigning scholarships. Davis uses an "Academic Index Score," which includes test scores, to assign the Regents' Scholarship, Wakeman Scholarship, and other scholarship funds. Irvine includes test scores in the holistic review that determines assignment of Regents' Scholarships and Chancellors' Excellence Scholarships. San Diego uses test scores for assigning scholarships in a similar way. Overall, most campuses that use test scores for scholarship assignment use them holistically, in the context of the entire application, rather than by assigning a fixed weight to the test score.

Placement. All campuses allow the University's Entry Level Writing Requirement to be met by standardized test scores. Other uses of test scores for placement vary on a campus by campus basis. For instance, at Irvine, introductory hard-science courses allow students to enroll with high enough test scores, as an alternative to completing prerequisite coursework or locally-developed placement examinations. Los Angeles' minimum graduation requirement for quantitative reasoning can be satisfied using test scores. San Diego allows using test scores for math placement, as an alternative to taking a math placement exam.

Other. UC has an "admission by examination" mechanism, to allow students from nontraditional backgrounds (such as homeschooled students) who are unable to meet A-G requirements to be admitted solely based on standardized test scores. This mechanism is rarely used and amounts to less than 1% of admissions. Some campuses, including Merced and Riverside, require minimum test scores for students who are admitted "by exception" (i.e., without completing the required A-G classes with a 3.0 GPA). Some campuses use test scores to identify students who may struggle without additional academic support and provide support to those students – for instance, San Diego uses test scores to identify candidates for the Triton Summer Success Program. Some campuses use test scores to identify candidates for honors programs (e.g. San Diego's Revelle program).

III. Rationale for Current Use of Standardized Tests

1) Recent history of standardized testing at the University of California ("UC" or "the University").

Before discussing why UC is considering changing its use of tests, it is important to review the reasons why UC currently uses admissions tests at all. When BOARS studied this issue in 2001–2002, they noted three "*different, but related assumptions regarding the value of tests in making admissions decisions:*

1. *Admissions tests are a valid measure of student preparation and/or promise that have been proven statistically to add to an institution's ability to predict student success beyond the predictive information that high school grades alone provide.*
2. *Admissions tests provide a standardized measure of preparation that is independent of the variability among grading patterns inevitably present when reviewing the records of students from thousands of high schools across the country.*
3. *Admissions tests can identify as-yet undeveloped talent in students who for a variety of reasons may not have worked to their full potential in high school, but who will nonetheless excel in college."*[25]

This 2002 BOARS report noted each of these assumptions was in principle testable against UC admissions and college success data, that such tests had not been adequately done in the past, and that examination of this data was essential for rational decision making in the future.[26] BOARS concluded that only the first two factors were well-justified by available data and should form the rationale for continued use of standardized tests at UC. The report identified four policy issues that should be considered in regular reviews of the use of tests:

1. *"The relative merits of tests that measure achievement versus those that purport to measure aptitude.*
2. *The messages that our test requirement sends to students and their families.*
3. *The degree to which admissions tests should be related to the curriculum UC applicants are expected to study in high school.*
4. *The burden that our test requirements place on students and schools."*

These policy issues, combined with the previously identified value of admissions tests as a standardized measure that increases predictability of student success, led to the following principles for use of standardized tests at UC:

1) *Admissions tests will be used at the University of California:*
   - *to assess academic preparation and achievement of UC applicants;*
   - *to predict success at UC beyond that predicted by high school GPA;*
   - *to aid in establishing UC eligibility; and*

---

[25] The Use of Admissions Tests by the University of California, available at: https://senate.universityofcalifornia.edu/_files/committees/boars/admissionstests.pdf

[26] The University of California has been publicly debating this issue for at least one hundred years. "As suggested above, most arguments concerning the efficiency of college entrance examinations as a test of fitness to succeed in college are based merely upon mere opinions; there have been few careful studies to determine the facts." "A suggested simplification in the present method of admitting freshmen to the university," W. Scott Thomas, The University of California Chronicle, 1914.

- *to aid in selecting students for admission at individual UC campuses.*

2) *The desired properties of admissions tests to be used for these purposes include the following.*
    - *An admissions test should be a reliable measurement that provides uniform assessment and should be fair across demographic groups.*
    - *An admissions test should measure levels of mastery of content in UC-approved high school preparatory coursework and should provide information to students, parents, and educators enabling them to identify academic strengths and weaknesses.*
    - *An admissions test should be demonstrably useful in predicting student success at UC and provide information beyond that which is contained in other parts of the application. (It is recognized that predictors of success are currently limited, and generally only include first-year college GPA and graduation rate. As this field advances, better predictors should be identified and used in validating admissions tests.)*
    - *An admissions test should be useful in a way that justifies its social and monetary costs.*

BOARS also recommended that "as a matter of principle, the faculty regularly review UC's admissions testing policy and practices to ensure that tests are being used in a way that is consistent with these principles and desired properties of admissions tests."

When BOARS addressed the issue again over three years of meetings in 2007–2009, it conducted additional statistical analyses focused on outcomes and predictability. BOARS affirmed that, at that time, the UC use of tests was generally compatible with the UC principles on use of tests, but revised the testing principles to include "*to increase the options available for applicants to demonstrate their preparedness for college*."

2) Does UC's current use of tests meet UC/BOARS criteria?

This section of the report focuses on whether the University's current use of tests is compatible with the principles for use of tests at UC as they are currently formulated (i.e., as they have been inherited from previous generations of BOARS reports). In subsequent sections of the report, we will examine whether UC's principles on testing should be modified, and more broadly, whether continued adherence to the California Master Plan for Higher Education and the ongoing prioritization of college preparation and academic merit should be revised in favor of an approach to college admissions more focused on demographic equity and the righting of historic wrongs.

This section is divided into five subsections that answer the following questions:

A. Are admissions test scores currently useful to predict student success beyond the predictive information in high school grades alone?
B. Does UC appropriately weight tests in admissions?
C. Does UC's use of standardized test scores inappropriately exclude members of certain groups?
D. Do standardized test scores counterbalance the variability among grading patterns at different high schools?
E. Additional rationales implicit in UC's current use of tests.

Much of the data presented here comes from the UCOP Department of Institutional Research and Academic Planning (IRAP) study entitled "Relationship of the SAT/ACT to College Performance at the University of California, January 2020," attached as Appendix I to this report.

A. Admissions tests add substantially to UC's ability to predict student success beyond the predictive information in high school grades alone.

A.1. What does it mean to succeed in college, and why is predictive validity important? The definition of college success can be broader and more nuanced than is commonly appreciated. Some aspects of success are nearly universally agreed on: for instance, there is broad consensus that success in college requires graduating with a credential rather than leaving college partway through with debt and no degree. The relative importance of other metrics of success is debated. While in the past BOARS has generally agreed that college grades are a valuable measure of student success, the specific importance of freshman GPA as an outcome metric is arguable[27]. It can be argued that as long as students complete their degrees, their grades are largely irrelevant to their success in life after they have graduated. It can also be argued that good grades are important for students who will attend graduate or professional school, and since UC is the primary research arm of the state's tripartite college and university system, it is reasonable to ensure that our undergraduate students are well prepared for post-baccalaureate activity. Many other aspects of success are difficult to measure or the University does not keep systematic records on them (e.g., participation in campus research opportunities or leadership).

In order to ensure public accountability, UC is obligated to make data-driven admissions decisions and report these decisions to the public, as it does in the annual Accountability Report.[28] Thus, despite debate on what exactly qualifies as success in college and consensus that the definition of success could include more metrics than are currently used, admissions accountability must focus on measurable (and measured) aspects of college success. The measurable outcomes that UC and BOARS typically focus on include 1) retention to second year, 2) first-year GPA, 3) graduation (whether on-time or not), and 4) GPA at graduation. The use of these metrics to determine whether admissions criteria are valid and appropriate goes back at least to the Master Plan in 1960[29]; their use has been regularly re-affirmed since then, including in the 2002 and 2010 BOARS studies[30] of the use of standardized tests at UC.

The primary conclusions of the IRAP report and the STTF's studies on the usefulness of admissions test scores in identifying successful students by these four metrics include the following.

1) Overall, both grades and admissions test scores are moderate predictors of college GPA at UC. The predictive power of test scores has gone up, and the predictive power of high school grades has gone down, since the 2010 BOARS study of this issue. At present, test scores are a slightly better predictor of freshman grades than high school grades are. Both grades and scores are

---

[27] https://senate.universityofcalifornia.edu/files/committees/boars/admissionstests.pdf

[28] https://accountability.universityofcalifornia.edu

[29] From the Master Plan: "*The Technical Committee suggested as a guiding principle that admission requirements are valid for any one college if, first, they serve to qualify for admission those applicants whose educational purposes are properly met by the college and whose abilities and training indicate probable scholastic success in the college and, secondly, they serve to eliminate applicants not meeting these requirements. The Technical Committee suggested the following measures of validity:*

1. *Scholastic success in the first semester or year*
2. *Continuance in college*
3. *Rate of dismissal for poor scholastic performance*
4. *Comparative standing on objective tests*

The Technical Committee regards scholastic success as the best single measure of validity. The Survey Team agrees, but prefers the use of several criteria in combination." (Emphasis added.)

[30] https://senate.universityofcalifornia.edu/committees/boars/reports.html

stronger predictors of early outcomes (freshman retention and GPA) than of longer-term outcomes (eventual graduation and graduation GPA).

2) Test scores contribute a statistically significant increment of prediction when added to regression analyses that already include high school grades as predictors. This improved prediction can translate to fairly large differences in predicted student outcomes (e.g. fourfold changes in non-retention rates, even for students with similar high school GPAs [HSGPAs], depending on test scores).
3) Test scores contribute significant predictive power across all income levels, ethnic groups, across both first-generation and non-first-generation students, and across all campuses and majors. There are meaningful differences between groups in the predictive improvement conferred by test scores.
4) Although there are large test score gaps between applicants from different demographic groups, UC does not use test scores in a way that prevents low-scoring students from disadvantaged groups from being admitted to UC as long as their applications show academic achievement or promise in other ways. Some students admitted with weak test scores through comprehensive review do perform well at UC. However, the lowest-scoring students typically do not perform well at UC.

A.2. What does a 25% [15%, 10%] correlation mean for student outcomes? Later chapters of this report will discuss whether a given increment in predictive power is worth considering in admissions decisions. Here, we provide three guides for interpreting correlation coefficients.[31] A detailed explanation of how correlations are computed and interpreted is attached as Appendix IV.

1) In statistical studies of human behavior,
   a. a correlation above 0.5, corresponding to R^2 or share of variance explained of above 0.25, is considered high (the R^2 is the square of the correlation);
   b. R^2 / share variance explained of roughly 0.1 to 0.25 is considered moderate;
   c. R^2/ share variance explained of below 0.1 is considered low.[32]
2) Given the very wide range of scores and grades presented by applicants for admission, even a low correlation coefficient can reveal large differences between the expected success levels for students in different grade or score bands. "The SAT 'explaining' less than 10% variance given HSGPA may seem trivial, but the difference between a 16% success rate and a 73% success rate <u>for students with similar high school records</u>, but different SAT scores, appears less trivial."[33] (Underlines are added by the STTF for emphasis.)
3) The practical usefulness of a given strength of correlation depends on the fraction of applicants who can be selected.[34] At UC overall, and particularly at certain campuses, space constraints, financial constraints, and the requirements of the Master Plan allow admitting only a small fraction of applicants. "For nearly a century it has been understood…that the smaller the selection ratio, the greater the payoff a particular correlation will yield."[35] When the fraction of applicants who can be admitted is small,

[31] Correlation coefficients are used to measure/summarize the strength of the association between two variables
[32] Zwick; Cohen 1998.
[33] Bridgeman, Burton, and Pollack 2008. Predicting Grades in College Courses: A Comparison of Multiple Regression and Percent Succeeding Approaches. Bridgeman, Brent; Pollack, Judith; Burton, Nancy. Journal of College Admission, n199 p19-25 Spr 2008. Available at https://files.eric.ed.gov/fulltext/EJ829428.pdf
[34] Taylor and Russell 1939. Taylor, H. C., & Russell, J. T. (1939). The relationship of validity coefficients to the practical effectiveness of tests in selection: discussion and tables. Journal of Applied Psychology, 23(5), 565–578.
[35] Zwick, "Who Gets In Zwick, R. (2017). Who gets in? Strategies for fair and effective college admissions. Cambridge, MA: Harvard University Press

even a weakly correlated predictor can produce a substantial increase in the proportion of students who will perform at a given level (e.g., maintaining a minimum GPA or graduating).

A.3. Predictive validity: College success measures correlate with admissions test scores. At UC, each of the primary success measures is clearly correlated with, and predicted by, scores on standard college admissions tests. Extensive tables and analyses are shown in the IRAP study included as **Appendix I**. A table of regression slopes and squared correlation coefficients for admissions years 2010-2012 is shown as **Table 3A-1**. Additional correlations are shown in **Appendix II** (Tables 1-4). In data from 2010-2012, 21% of variance in first-year GPA was predicted by SAT score. This number may sound small, but it translates into large differences in student outcomes. For instance, among students admitted with SAT scores below 700, 35% left UC after only one year and only 50% graduated within seven years; among students admitted with SAT scores above 1400, only 3% were not retained past freshman year and 92% graduated within seven years. Outcomes data are shown in **Figure 3A-1**.

HSGPA compared to SAT as a predictor of UC freshman grades across demographic groups, 2010-2012

Category	N	GPA-only models		SAT-only models	
		Slope	R^2	Slope	R^2
All	90504	0.39	0.16	0.46	0.21
<$30K	20178	0.34	0.12	0.41	0.17
$30-$60K	19232	0.36	0.13	0.42	0.18
$60-$120K	19307	0.36	0.13	0.39	0.15
$120K+	21339	0.41	0.17	0.35	0.12
Asian	37899	0.38	0.15	0.42	0.18
Black	3652	0.29	0.09	0.37	0.14
Hispanic	22522	0.33	0.11	0.39	0.15
White	23046	0.38	0.14	0.34	0.11
Not 1st Gen	47626	0.39	0.15	0.37	0.14
First Gen	40405	0.34	0.12	0.4	0.16

Table 3A-1. High school GPA (HSGPA) compared to SAT as predictors of freshman grades at UC, for all students and for different demographic groups, 2010-2012. SAT is calculated as SATRW + SATRM. Both SAT and HSGPA are normalized (z-scored) prior to ordinary least-squares regression.

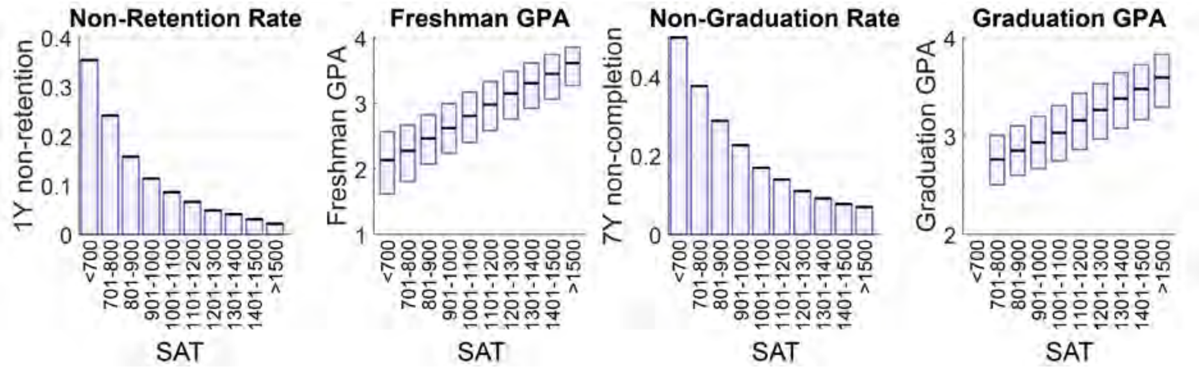


Figure 3A-1 Percent of freshmen who fail to return the following fall, average first year UC GPA, percent of students who fail to earn a degree within seven years, and average UC GPA at graduation by SAT bands. SAT = SATRM+SATRR. All panels but Graduation GPA contain data from 2010-2012 admits. Graduation GPA was only available for 2010 admits. Retention/graduation: Lines represent mean rate of non-retention or non-graduation. GPA: Lines represent medians and ranges shown are interquartile intervals.

A.4. Predictive validity: Standardized test scores provide information beyond what is available in HSGPA. Contrary to the narrative that is commonly expressed, the information available in test scores is not redundant with the information available in grades. Even among students with similar high school academic records, we still see a substantial relationship between test scores and college retention, grades, and graduation. For any given high school GPA, a student admitted with a low SAT score is between two and five times more likely to drop out after one year, and up to three times less likely to complete their degree compared to a student with a high score (**Figure 3A-2**).

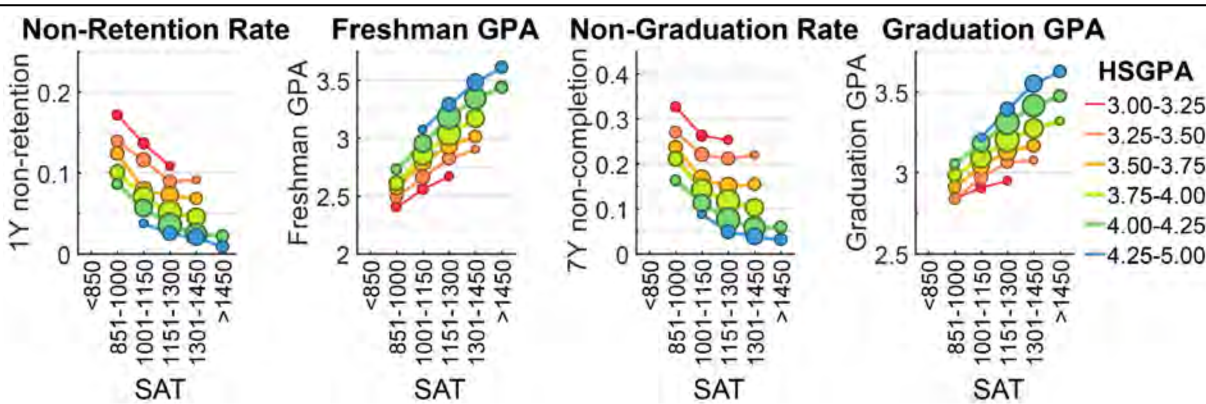


Figure 3A-2 Percent of freshmen who fail to return the following fall, average first year UC GPA, percent of students who fail to earn a degree within seven years, and average UC GPA at graduation by SAT and HSGPA bands. Dot area is proportional to number of students in each SAT and HSGPA band. SAT scores are summed SATRM + SATRW from the 2010-2012 application pools, except for the plots of graduation GPA which are for 2010 only. HSGPAs are weighted. Color indicates HSGPA bands

HSGPA <u>plus</u> SAT to predict UC freshman grades across demographic groups, 2010-2012

Category	**N**	**Combined models**		
		GPA slope	**SAT slope**	**R^2**
All	90504	0.25	0.36	0.26
<$30K	20178	0.23	0.33	0.22
$30-$60K	19232	0.23	0.34	0.23
$60-$120K	19307	0.25	0.29	0.20
$120K+	21339	0.32	0.21	0.20
Asian	37899	0.26	0.32	0.24
Black	3652	0.19	0.30	0.17
Hispanic	22522	0.22	0.31	0.20
White	23046	0.29	0.22	0.18
Not 1st Gen	47626	0.29	0.26	0.21
First Gen	40405	0.23	0.32	0.21

Table 3A-2. High school GPA (HSGPA) combined with SAT to predict freshman grades at UC for different demographic groups, 2010-2012. SAT is calculated as SATRR + SATRM. Both SAT and HSGPA are normalized (z-scored) prior to ordinary least-squares regression.

These results are derived from UC admissions data in the 2010–2012 cohort. We chose this cohort because it is the most recent cohort for which seven-year graduation rates are available. Some STTF members felt that overreliance on four-year graduation rates might be unnecessarily discriminatory against low-income students, who are more likely to have to interrupt their studies to take care of family members or earn money to support themselves[36]. The graduation GPA data comes only from the 2010 cohort because 2011-2012 graduation GPAs were not yet available from UCOP at the time of this study.

These results are broadly consistent with studies from other public university systems and nationwide studies as well as with UC historical data.[37] The major difference is that in current UC data, test scores explain a greater fraction of outcome variance than at universities nationwide. Indeed, the variance in outcomes explained by test scores has increased since 2007, while the variance explained by grades has decreased. Analysis in the IRAP report (attached as Appendix I) concludes that this is likely due to the way UC uses grades and test scores in admissions, which has changed over this time: the Regents increased the minimum HSGPA required for UC freshman eligibility from 2.8 to 3.0, and rapid growth of the qualified applicant pool has increased selectivity (and thus reduced the fraction of students admitted with low HSGPAs) at all campuses. The fraction of students admitted through the Eligibility in the Local

[36] Soria, Krista M.; Weiner, Brad; and Lu, Elissa C. (2014) "Financial Decisions among Undergraduate Students from Low-Income and Working-Class Social Class Backgrounds," Journal of Student Financial Aid: Vol. 44: Iss. 1

[37] Reviewed in "Eight Myths about Standardized Admissions Testing," Paul Sackett and Nathan Kuncel, book chapter in "Measuring Success: Testing, Grades, and the Future of College Admissions", ed. Buckley et al 2018.

Context (ELC) program (which admits students with high grades without regard to their test scores) has increased. At the same time, "the distribution of HSGPAs was increasingly compressed as a result of high schools' grade inflation,"[38] "students have tended towards taking larger courses with more formal test-based examinations,"[39] and "high-HSGPA students are increasingly enrolled in low-grade courses like Engineering and Economics."[40]

We also note that because the data supplied to STTF by the UC Office of the President (UCOP) only included years for which seven-year graduation data were available, the STTF was unable to directly examine outcomes of students admitted after the ELC expansion. This highlights the relationship between two closely related issues: what do we believe are the most important goals of the admissions process, and how often UC should consider non-incremental changes to its admissions strategy? Following the last BOARS report on admissions tests in 2010, UC made significant changes to its admissions procedures in 2012 (expanding the ELC program and adding a new "Entitled to Review" [ETR] category guaranteeing that a student who submits a complete application and has a 3.0 GPA in A-G courses will not be denied admission without a comprehensive review). We do not yet know how the students admitted through these expanded access programs have fared in the long term. In addition, the new iteration of the SAT[41] is only a few years old; only one-year outcomes are available for students admitted using the new SAT and we do not know how those students will fare in long-term metrics. There is a tradeoff between prioritizing long-term outcome measures and being able to rapidly analyze the consequences of changes in policy. One reason why BOARS analyses have often focused on freshman-year grades is that they are assessable rapidly, and allow BOARS to justify more rapid changes in policy than could be supported based on longer-term outcome metrics.[42] Caution is warranted in making another large change to UC admissions procedures before we understand the consequences of the last large changes to UC admissions procedures.

A.5. Predictive validity: Standardized test scores are predictive of outcomes for all demographic groups. Many on STTF were concerned that the predictive power of test scores might differ for members of different demographic groups, and that therefore members of certain groups typically underrepresented at UC would be disadvantaged by mandatory consideration of test scores in their application. This issue was noted in the BOARS 2002 report on testing, which stated that *"It is well known that admissions tests*

---

[38] IRAP report, page 44.
[39] IRAP report, page 45.
[40] IRAP report, page 45.
[41] STTF reviewed the College Board's preliminary report, showing generally similar relationships between scores, demographics, and outcomes for the new version of the SAT as for prior SAT versions. STTF was unable to College Board representatives back after our initial meeting with them to go over their statistical report in more detail, due to the shortened turnaround for this report.
[42] The 2002 BOARS report noted the same issue but chose to prioritize rapid turnaround over evaluation of longer-term outcomes: "It should be noted here that historically the measure of "student success" most commonly used in validity studies of admissions tests—and the one on which the College Board and Educational Testing Service base statistical studies of SAT I and SAT II scores—has been the freshman-year GPA. This is because intervening variables (including differing academic programs pursued by individual students and the intellectual and personal development students experience during their college years) serve to weaken substantially the predictive relationship between pre-college factors like high school GPA and test scores and eventual college GPA. Predictive validity studies justify use of freshman-year GPA by citing its correlation with GPA at graduation. Nonetheless, faculty and admissions officers involved in determining eligibility and selection criteria and processes point out that maximizing freshman year GPA is at best a relatively minor goal of the admissions evaluation process. While highly cognizant of the weaknesses involved in using freshman GPA, BOARS could not identify an alternative and therefore followed the generally accepted practice of using freshman-year GPA as the outcome variable indicating success in college." https://senate.universityofcalifornia.edu/files/committees/boars/admissionstests.pdf

of all types — along with high school grades and other indicators of academic achievement — are strongly correlated with family income. This does not reflect bias in the tests, but rather the inescapable fact that schools in California — like those throughout the country — vary widely in available resources and students from poor families are more likely to attend schools with fewer resources. The members of BOARS are well aware that they cannot eliminate this level of "disparate impact" admissions tests have on students from socio-economically disadvantaged circumstances."[43]

IRAP and STTF also performed predictive validity analyses that disaggregated UC students according to demographic factors including parental education, family income, and racial/ethnic identity. **Figures 3A-3, 3A-4, and 3A-5; Tables 3A-1 and 3A-2** above, and the IRAP report show the findings. The IRAP study showed that models combining test scores with GPA explained 14-22% of the variance in UC grades for all ethnic groups, and 18-20% for all incomes (see **Appendix C** of IRAP report). As described above, this may or may not sound like a large amount of variance, but it corresponds to fairly significant changes in outcomes between students with high and low test scores, even for students with equivalent high school GPA and similar demographic profiles. The STTF study determined the increment in predictive power attributable to adding test scores to HSGPA and the relative statistical power of the two measures for the set of students enrolling in the years 2010-2012.

Perhaps counterintuitively, we found that test scores were better predictors of outcomes for underrepresented groups than for majority groups.

For instance, for the wealthiest incoming freshmen (those from families earning above $120,000 per year – the top quartile of UC's income distribution), HSGPA-only models explained about 17% of variance in freshman GPA (**Figure 3A-3**; **Table 3A-1** above; **Appendix II Table 2**). Adding SAT to a HSGPA-only model produced about a 20% improvement in predictive power (**Table 3A-2)**. Further, the predictive power of SAT systematically increases as student income decreases: for students in the lowest income quartile (<$30K household income), HSGPA-only models explain about 12% of variance in freshman GPA, while adding SAT to HSGPA improves predictability to 20%, an 80% increase. Similarly, for students in the highest income quartile, HSGPA-only models explained about 16% of the variance in graduation GPA, while SAT-only models explained only 10% of the variance for outcomes among these high-income students. Models in which SAT and HSGPA were combined accounted for 19% of the variance, only a 17% improvement over models containing HSGPA alone. Yet for students in the lowest income quartile, adding SAT to a HSGPA-only model produces a 60% improvement in predictability of final GPA.

Similarly, for non-first-generation students, a HSGPA-only model explains about 15% of variance in freshman GPA (**Table 3A-1)** and adding SAT to HSGPA increased explained variance to 21%, about a 40% increase (**Figure 3A-4**; **Table 3A-2**). But for first-generation students, a HSGPA-only model only explained 12% of variance in freshman GPA (**Table 3A-1**), and combining SAT and HSGPA increased explained variance by 75% over HSGPA alone (**Table 3A-2**). For non-first-generation students, a HSGPA-only model explained 16% of graduation GPA variance, and adding SAT to HSGPA increased explained variance to 19%, about a 20% increase. Yet for first-generation students, a HSGPA-only model explained 11% of variance, and combining SAT and HSGPA increased explained variance by 55% above GPA alone (**Figure 3A-4**; **Appendix II Tables 3 and 4**). In other words, the increase in predictive power was greatest for the student population less well represented at UC.

[43] https://senate.universityofcalifornia.edu/_files/committees/boars/admissionstests.pdf

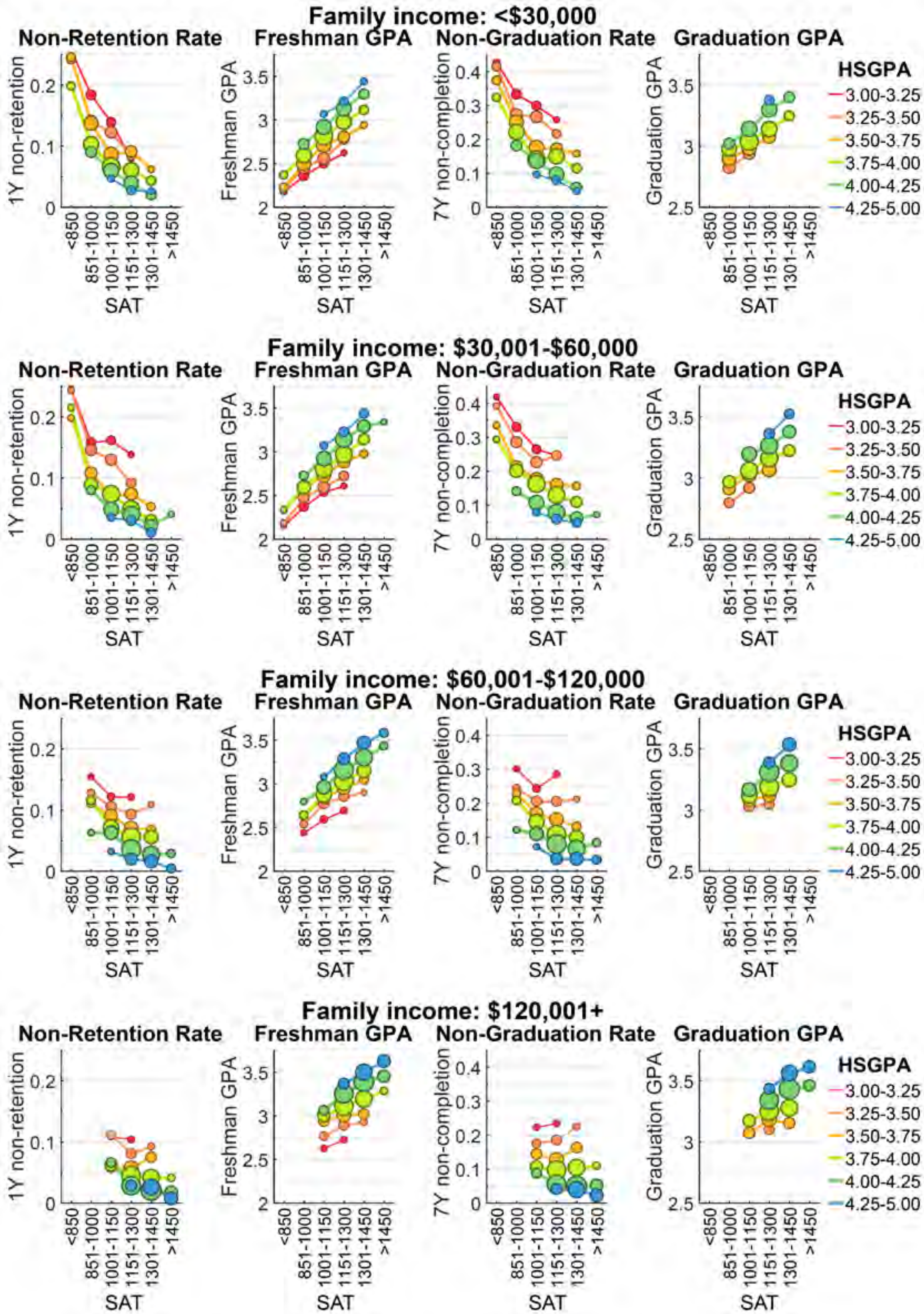


Figure 3A-3 Mean percent of freshmen who fail to return the following fall, average first-year UC GPA, percent of students who fail to earn a degree within seven years, and average UC GPA at graduation by SAT and HSGPA band for students of varying family incomes. Dot area is proportional to number of students in each SAT and HSGPA band (normalized by number of students in each demographic group). SAT scores are summed SATRM + SATRW from the 2010–2012 application pools, except for the plots of graduation GPA which are for 2010 only. HSGPAs are weighted. Color indicates HSGPA bands.

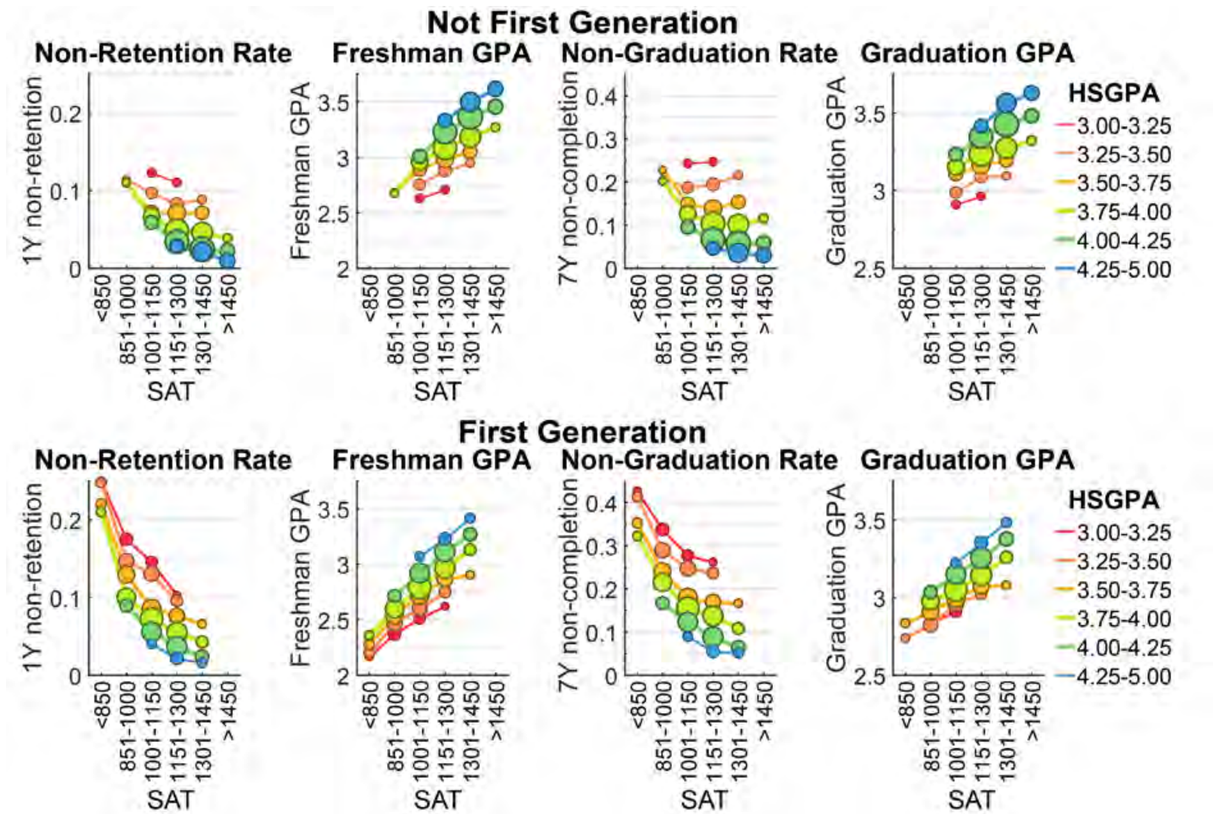


Figure 3A-4. Mean percent of freshmen who fail to return the following fall, average first-year UC GPA, percent of students who fail to earn a degree within seven years, and average UC GPA at graduation by SAT and HSGPA band for students who self-reported as having (top) or lacking (bottom) at least one parent who attended college. Dot area is proportional to number of students in each SAT and HSGPA band (normalized by number of students in each demographic group). SAT scores are summed SATRM + SATRW from the 2010–2012 application pools, except for the plots of graduation GPA, which are for 2010 only. HSGPAs are weighted. Color indicates HSGPA bands.

We observe the same phenomenon among different racial and ethnic groups. When added to a statistical model relating first-year GPA to HSGPA, test scores increase predictability of freshman GPA by 29% for Caucasian students, by 60% for Asian students, by 77% for Hispanic students, and by 96% for African-American students. Standardized test scores increase predictability of graduation GPA by 22% for Caucasian students, by 39% for Asian students, by 61% for Hispanic students, and by 56% for African-American students (**Figure 3A-5**; **Appendix II Table 4**).

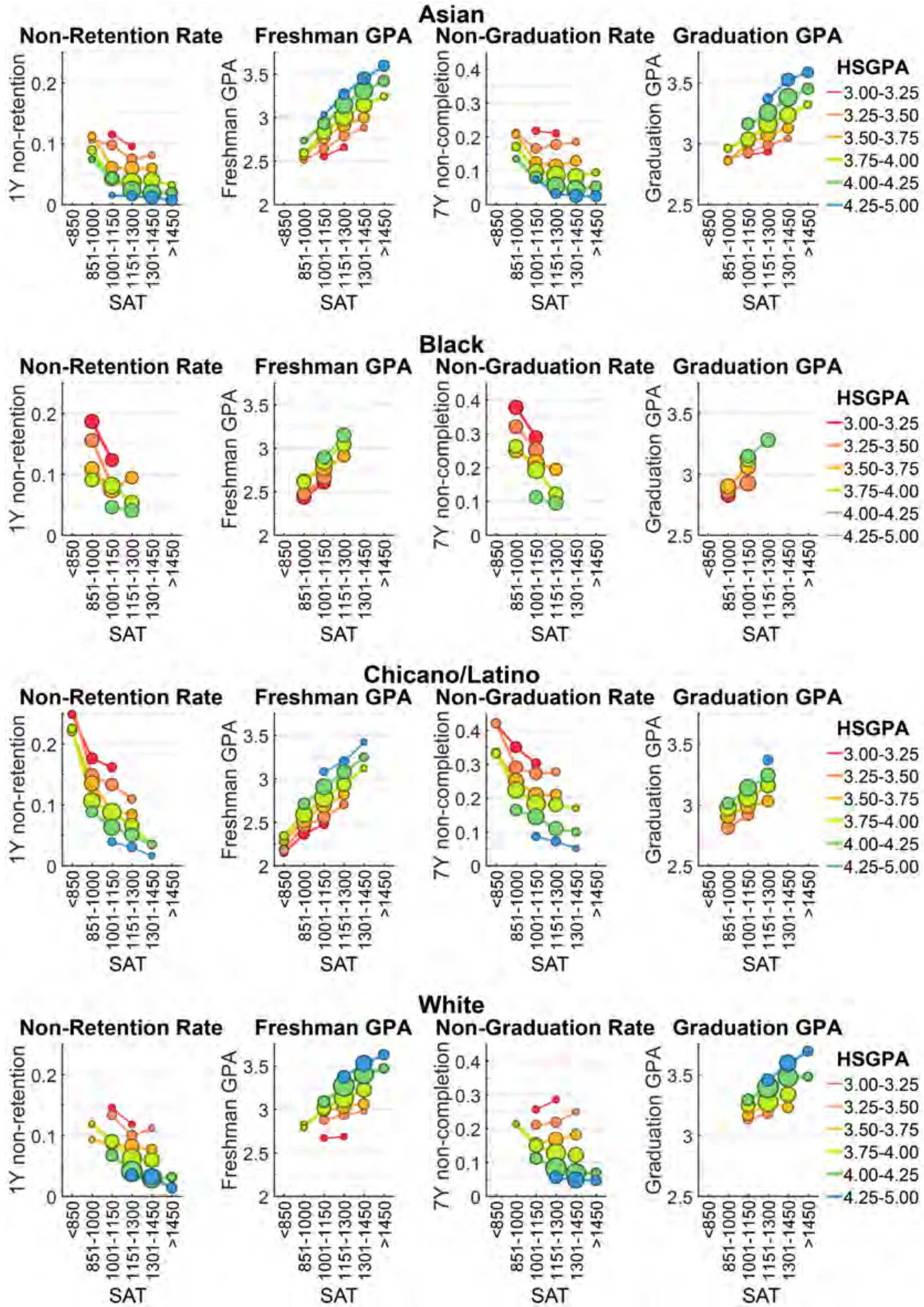


Figure 3A-5. Mean percent of freshmen who fail to return the following fall, average first-year UC GPA, percent of students who fail to earn a degree within seven years, and average UC GPA at graduation by SAT and HSGPA band for students self-reporting as belonging to different racial/ethnic groups. Dot area is proportional to number of students in each SAT and HSGPA band (normalized by number of students in each demographic group). SAT scores are summed SATRM + SATRW from the 2010–2012 application pools, except for the plots of graduation GPA, which are for 2010 only. HSGPAs are weighted. Color indicates HSGPA bands.

Similarly, the utility of test scores for identifying students who were likely to persist past the first year, or students who were likely to graduate – or conversely, for identifying students who are at great risk for non-persistence or non-graduation without additional support – was greater for members of underrepresented groups (**Figures 3A-3 to 3A-5**, columns 1 and 3).

Again, the results in the universitywide predictive validity study are based on UC admissions data from 2010–2018, while the figures are based on UC admissions and outcomes data from 2010–2012. Both sets of results are consistent with the broad literature on the topic, including other public university systems, other selective institutions, and nationwide[44].

A-6. Predictive validity: Performance prediction is strongest in quantitative disciplines. IRAP also studied the predictive power of grades and test scores on college grades of students in different majors. This analysis disaggregated freshman GPA at the course level, which allowed for a more precise examination of the effects of test scores and HSGPA in explaining student performance in different academic disciplines. IRAP identified first-year courses taught in four broad areas of study – humanities, social sciences, natural sciences, and engineering – then modeled the student outcomes in those four courses separately in terms of tests and HSGPA. This analysis considered the effects of each of the three components of the SAT that were taken at the time, mathematics, reading and writing, separately.

The results are summarized in Table 6 of the IRAP report (see Appendix I). The first four models consider each of the four academic indicators (HSGPA, SAT Reading, SAT Math and SAT Writing) independently. In every case, the percent of within-course variation explained by these indicators in isolation was least for HSGPA, sometimes dramatically so. For the natural sciences, for example, the GPA only model had an R^2 of 4.7%, while the SAT Math only model had an R^2 of 13.5%. As might be expected, the SAT reading and writing components are most explanatory variables for the humanities and social sciences, while the SAT math is most explanatory for the natural sciences and engineering.

Combining the indicators increased the amount of variance explained substantially above what is available using only HSGPA. For the humanities, the four-indicator model (Model 6 in Table 6 of the IRAP report), has an R^2 that is 2.7 times that of the HSGPA-only model. For engineering, R^2 for the full model is 5.2 times that of the HSGPA-only model. Clearly, the tests are providing a great deal of information on prospective student performance that is not available in high school grades alone.

We note that the results of this subsection do not speak to the question of institutional values. That is, the statistical power of tests in predicting outcomes does not address whether UC **ought** to consider expected college success when deciding which students to admit. These results address the empirical question of whether admissions test scores can assist in identifying students likely to succeed at UC. They do not answer the moral and practical questions of whether we ought to use that information to select for admission students whose records indicate that they will likely persist to graduation and do well in their classes, whether we ought to use the information to target aid towards students who will struggle without additional assistance, or whether we ought to ignore the information entirely. Policy implications of these findings are discussed and debated in the following sections.

[44] Reviewed in "Eight Myths about Standardized Admissions Testing", Paul Sackett and Nathan Kuncel, book chapter in "Measuring Success: Testing, Grades, and the Future of College Admissions", ed. Buckley et al 2018.

B. UC does not appear to violate BOARS principles of comprehensive review by prioritizing standardized test scores in admissions. UC admissions place more value on the high school record.

B.1. It is very rare that otherwise qualified students could be blocked from UC based only on test scores. Prior BOARS studies of this issue investigated whether students with strong grades could be blocked from UC admission based on a low test score, and conversely, whether students with high test scores were being admitted to UC despite weak grades. The current STTF repeated this analysis on the 2018 dataset to look for evidence that UC inappropriately uses test scores to block otherwise strong students from admission. Test scores and grades are generally moderately correlated (R^2 of 28%). This is a strong enough correlation that it is unusual for students to have "discrepant" grades and scores – that is, to have grades more than one standard deviation above the mean but scores more than one standard deviation below the mean, or vice versa. Of the ~18,000 students with SATs more than one standard deviation below average, only ~300 of these had grades more than one standard deviation above the average, and 84% of those students were admitted. In other words, of the 120,000+ California resident applicants to UC for 2018, fewer than 50 students with strong grades and low scores were denied admission. (We do not know whether these students were denied admission based on test scores or based on other weaknesses in their applications; this analysis simply tells us that no more than 50 students fall into this category.) Conversely, of the 18,000 students with HSGPAs more than one standard deviation below the mean, about 400 had test scores more than one standard deviation above the mean, and only 41% of these were admitted. This indicates that, consistent with BOARS policy, UC admissions is at most very rarely blocking otherwise qualified students based on scores, and generally appears to reward high grades more than high test scores.

B.2. UC appears to consider grades more than test scores in admissions decisions. To examine this point more fully, we looked at the admit rates by campus as well as by the system as a whole. In this analysis, each of the academic indicators is normalized by computing the number of standard deviations that a particular indicator is away from the campus- or system-level mean and dividing that difference by the relevant standard deviation. If the original data follows a normal or Gaussian distribution, the normalized data follows standard normal distribution with a mean of zero and a standard deviation of 1. That way one can compare the influence of HSGPA and SAT with each other using the same scale. The tables below show the percent of applicants admitted along each of four divisions of the academic indicators: High (one standard deviation or more above the mean), Mid-High (between the mean and one standard deviation above), Mid-Low (between the mean and one standard deviation below), and Low (one standard deviation or more below the mean). Given the importance of HSGPA and SAT in admissions, it is natural to suppose that an applicant who is in the "high" category in both indicators is much more likely to be admitted than a student who is "low" in both. Indeed, this is the case.

Table 3B-1 shows the admit rates for the system as a whole. For clarity, we discuss this table in some detail. Tables for a few campuses are also included in the body of the report. Tables for all campuses may be found in **Appendix III**.

Percent of Applicants Admitted to at Least One Campus					
	High SAT	Mid-High SAT	Mid-Low SAT	Low SAT	Net
High HSGPA	95%	95%	94%	84%	95%
Mid-High HSGPA	85%	75%	73%	64%	75%
Mid-Low HSGPA	65%	51%	43%	28%	43%
Low HSGPA	41%	25%	17%	9%	16%
Net	86%	68%	52%	27%	58%

Table 3B-1. Admission rate for the UC system for California resident applicants who submitted an SAT score for fall 2018. High (one standard deviation or more above the mean), Mid-High (between the mean and one standard deviation above), Mid-Low (between the mean and one standard deviation below), and Low (one standard deviation or more below the mean). Colors correspond to the fraction of applicants admitted in each group. The large differences between rows indicate that admission probability changes very much depending on HSGPA. The smaller differences between columns indicate that admission probability is less dependent on SAT.

The top row shows the admit rate for applicants who all reported high HSGPAs, divided into cells according to their SAT scores. As discussed above, among students with high HSGPAs, even students in the lowest SAT range had an 84% admit rate. The leftmost column has results for applicants with high SAT scores divided across the HSGPA ranges. Again, as discussed above, for this group of applicants, the admit rate was only 41% for students in the lowest HSGPA range, less than half that for applicants with high grades and low test scores. Nearly 27% of applicants with low SAT were admitted compared with about 16% for those with low HSGPA. The clear implication of this is that **as a system, UC values high school grades more than test scores in making admission decisions**.

Campus-level results can be even more extreme, depending upon the selectivity of the campus. **Table 3B-2** shows results for Berkeley, one of the most selective campuses (16% overall admit rate), **Table 3B-3** shows results for Santa Barbara (29% admit rate), and **Table 3B-4** shows results for Merced (68% admit rate).

Percent of Applicants Admitted to Berkeley					
	High SAT	Mid-High SAT	Mid-Low SAT	Low SAT	Net
High HSGPA	54%	42%	43%	30%	47%
Mid-High HSGPA	36%	20%	18%	14%	23%
Mid-Low HSGPA		2%	2%	2%	2%
Low HSGPA					
Net	36%	17%	11%	4%	16%

Table 3B-2. Admission rate for California resident applicants to Berkeley who submitted an SAT score for fall 2018. High (one standard deviation or more above the mean), Mid-High (between the mean and one standard deviation above), Mid-Low (between the mean and one standard deviation below), and Low (one standard deviation or more below the mean). Colors correspond to the fraction of applicants admitted in each group. The large differences between rows indicate that admission probability changes very much depending on HSGPA. The smaller differences between columns indicate that admission probability is less dependent on SAT. To preserve student privacy, cells with fewer than 30 students are blanked.

Percent of Applicants Admitted to Santa Barbara

	High SAT	Mid-High SAT	Mid-Low SAT	Low SAT	Net
High HSGPA	90%	74%	69%	42%	80%
Mid-High HSGPA	55%	32%	31%	17%	35%
Mid-Low HSGPA	11%	7%	8%	2%	6%
Low HSGPA					1%
Net	63%	31%	19%	5%	29%

Table 3B-3. Admission rate for California resident applicants to Santa Barbara who submitted an SAT score for fall 2018. High (one standard deviation or more above the mean), Mid-High (between the mean and one standard deviation above), Mid-Low (between the mean and one standard deviation below), and Low (one standard deviation or more below the mean). Colors correspond to the fraction of applicants admitted in each group. The large differences between rows indicate that admission probability changes very much depending on HSGPA. The smaller differences across columns indicate that admission probability is less dependent on SAT. To preserve student privacy, cells with fewer than 30 students are blanked.

Percent of Applicants Admitted to Merced

	High SAT	Mid-High SAT	Mid-Low SAT	Low SAT	Net
High HSGPA	99%	99%	98%	91%	99%
Mid-High HSGPA	98%	97%	90%	57%	90%
Mid-Low HSGPA	83%	74%	62%	19%	59%
Low HSGPA	16%	15%	10%		9%
Net	91%	82%	66%	25%	68%

Table 3B-4. Admission rate for California resident applicants to Merced who submitted an SAT score for fall 2018. High (one standard deviation or more above the mean), Mid-High (between the mean and one standard deviation above), Mid-Low (between the mean and one standard deviation below), and Low (one standard deviation or more below the mean). Colors correspond to the fraction of applicants admitted in each group. The large differences between rows indicate that admission probability changes very much depending on HSGPA. The smaller differences across columns indicate that admission probability is less dependent on SAT. To preserve student privacy, cells with fewer than 30 students are blanked.

For every campus, the admit rate was much higher for applicants with high HSGPA and low SAT than for those with high SAT and low HSGPA, offering campus-level validation for the suggestion that grades matter more than tests in admission decisions at UC. At the highly selective or moderately selective campuses (admit rates less than 35%), **the admit rate for applicants with high GPA and low SAT was at least ten times that of applicants with high SAT and low GPA.**

These results show that UC campuses use test scores to help select from the large number of applicants with high HSGPAs, but that test scores by themselves do not block students from UC admission, and that test scores are generally less important in the admissions process than grades.

B-3. Under the current eligibility system, tests do identify otherwise ineligible applicants who come from relatively disadvantaged backgrounds. Admission tests find talented students who do not stand out in terms of high school grades alone. The SAT was originally designed to identify students who might not have had high HSGPA or attended elite or well-resourced high schools but who showed academic promise. The SAT and ACT appear to play that role for many students at the UC today. Many students who do not qualify for ELC, (i.e., being in the top 9% of their classes based on HSGPA), become eligible to attend UC through statewide eligibility. Unlike ELC, statewide eligibility, which is also referred to as Index Eligibility, uses the SAT as well as GPA.

Moreover, this population of students does not consist solely of affluent students in high- performing competitive high schools. On the contrary, in the 2018 applicant pool, 22,613 students were guaranteed admission to UC through the statewide eligibility index but not ELC, meaning that their SAT score made them eligible. Of these students for whom the SAT score was decisive in guaranteeing them admission to UC, 4,931 were low-income and 5,704 were first generation college. Turning to race/ethnicity, 5,609 were under-represented Minorities (URMs), comprised of 4,442 Latino students, 999 African-American students, and 168 Native American students.

To put these numbers into context, **Figure 3B-1** shows, for the students in each of these disadvantaged groups who were guaranteed UC admission, the percentages who earned their spots because of the SAT. That is, they were statewide index eligible and not eligible through ELC, the latter of which uses only GPA ranking within each high school.

The figure shows that about one quarter of all low-income, first-generation college and URM students who were guaranteed admission earned this guarantee through the statewide eligibility path, but *not* through ELC. This means that these students had relatively high SAT scores but their GPA was not high enough to guarantee them admission through ELC. This is a substantial proportion of these disadvantaged groups of students who would not have been guaranteed admission had they not had high SAT scores.

The three rightmost bars in the figure divide the URM category into the three ethnic/racial groups that together constitute under-represented minorities. Just under a quarter of Latino students earned their guarantee of admission through their SAT score. The percentage of African-American and Native American students who gained admission through Statewide Eligibility but not ELC is markedly higher at 40% and 47% respectively.

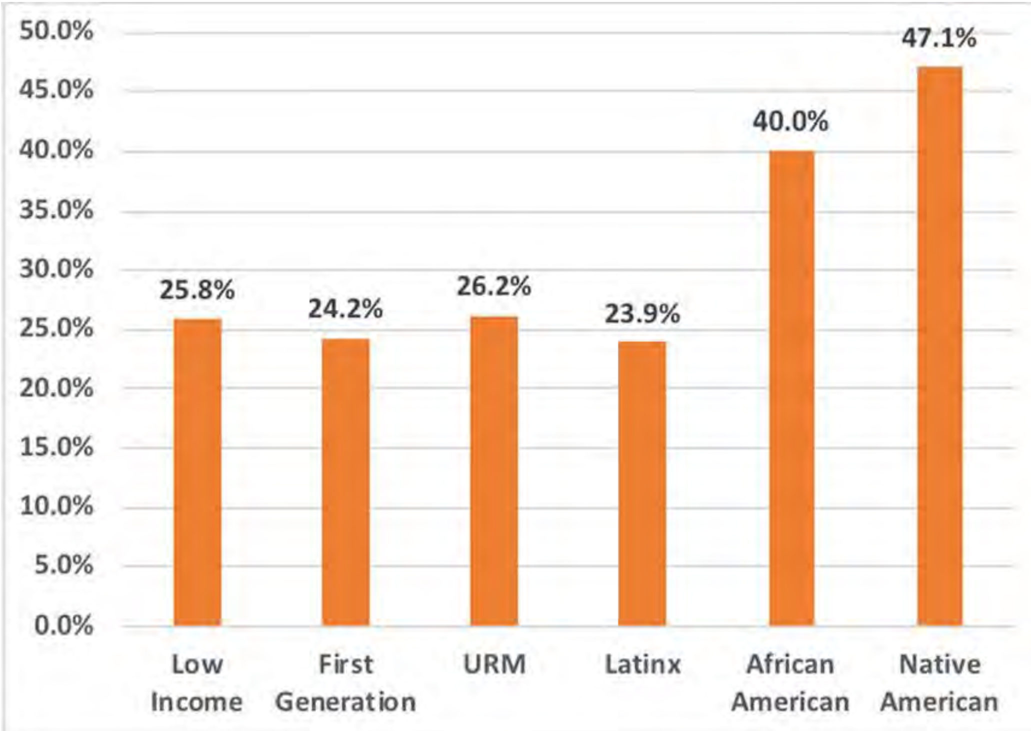


Figure 3B-1. The percentage of students in disadvantaged groups who were guaranteed admission in 2018 solely due to the statewide eligibility index. The denominator for these calculations is the total number of students in the given group who were eligible through: i) Statewide index only, ii) Eligibility in the Local Context (ELC) Only, and iii) Both Statewide Index and ELC. These three pools do not represent all admits, because students are also admitted through ETR and by exception. To give some perspective, in 2018, about 17,000 out of 71,000 admits were admitted through means other than the three guaranteed admission categories.

To sum up, the SAT allows many disadvantaged students to gain guarantees of admission to UC. As a share of all students in disadvantaged groups who are guaranteed admission to UC, the percentages who earn this guarantee due to their SAT scores range from a low of 24% for Latino students to highs of 40% and 47% for African-Americans and Native Americans. The original intent of the SAT was to identify students who came from outside relatively privileged circles who might have the potential to succeed in university. This original intent is clearly being realized at UC.

B-4. Some students admitted with weaker standardized test scores through comprehensive review do perform well at UC. However, the lowest-scoring students typically do not perform well at UC. The STTF has considered whether UC should increase admissions through the ELC pathway (in which the highest scoring students at each high school are admitted irrespective of scores) or through the ETR pathway (in which all students with a 3.0 in A-G courses have their application comprehensively reviewed at the campuses to which they apply) in order to increase enrollment of underrepresented groups. The question arises of whether admissions test scores should still be required for those applicants. **Figures 3B-2** and **3B-3** show side-by-side the relationship between SAT score and outcomes for examples of groups underrepresented based on family income (left), ethnicity (center), and family educational history (right). **Figure 3B-2** shows rate of retention to the sophomore year, and **Figure 3B-3** shows the graduation rate. These data suggest that in some ways, admissions test scores are a helpful and informative predictor of whether the students of diverse backgrounds targeted by ELC/ETR will succeed at UC. These data indicate that whether a students' test score is high, fairly high, or moderately high has little effect on their odds of graduating UC. However, students admitted with the lowest test scores show a strikingly lower probability of graduating (top row). This pattern remains evident when stratifying applicants based on high school GPA (bottom row).

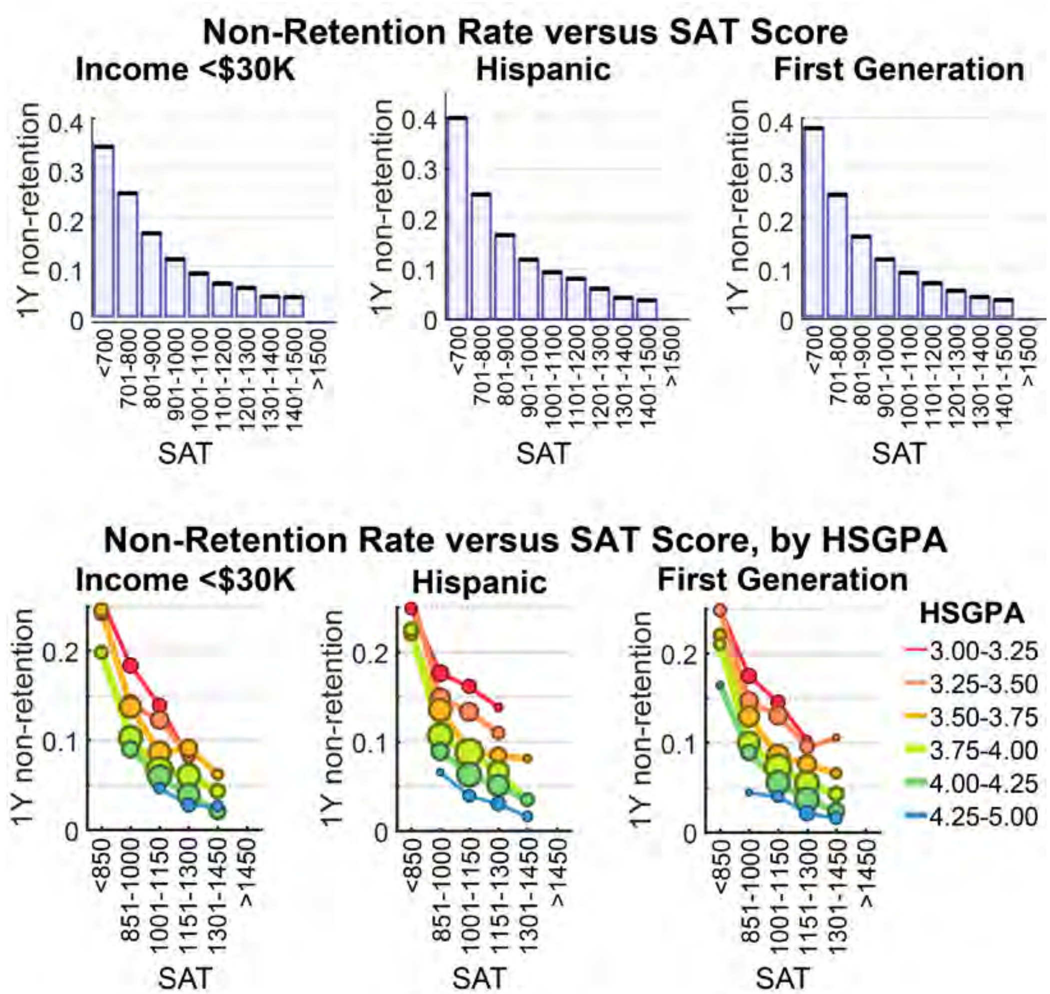


Figure 3B-2. Students with poor test scores are at especially high risk of leaving UC within one year. Top row: SAT scores are summed SATRM + SATRW from the 2010-2012 admissions pool. Lines represent mean rate of non-retention. Bottom row: Dot area is proportional to number of students in each SAT and HSGPA band (normalized by number of students in each demographic group). HSGPAs are weighted. Color indicates HSGPA bands. Bottom row panels are the same data as in Figures 3A-3, 3A-4, and 3A-5.

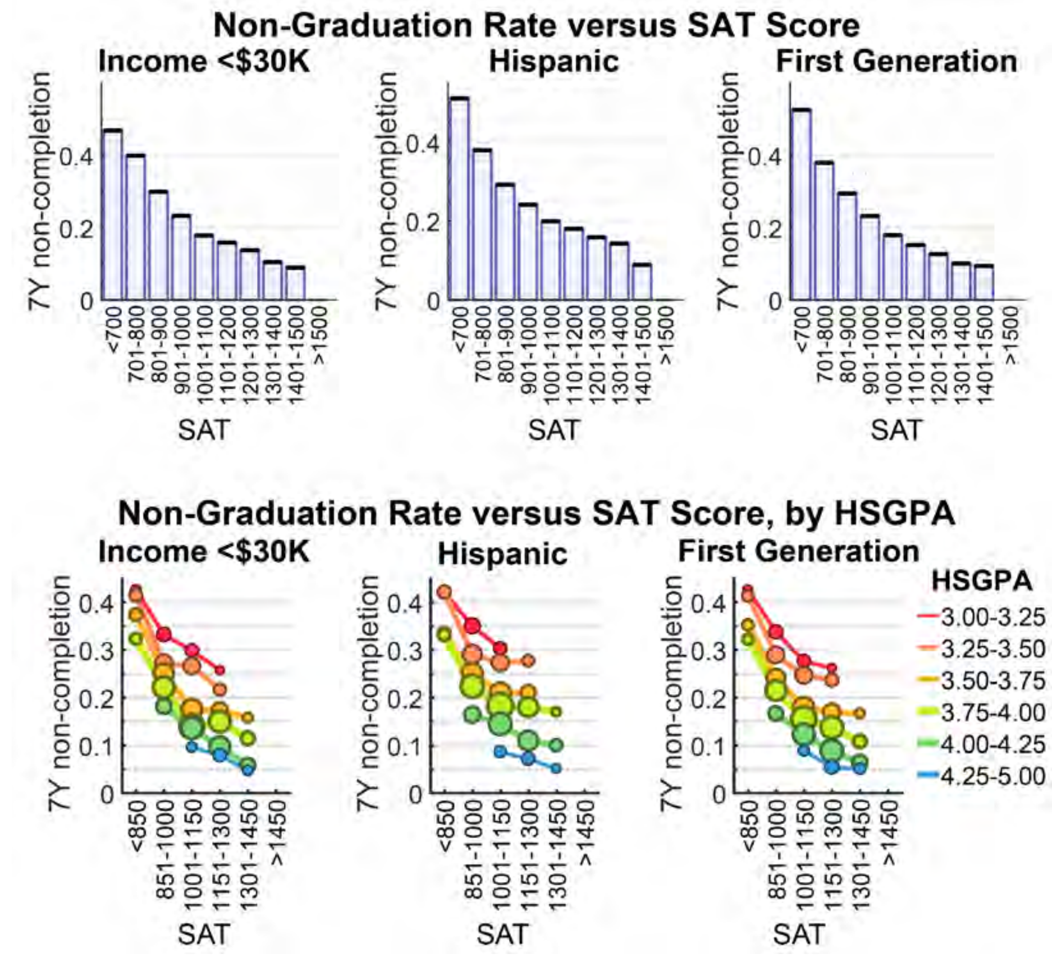


Figure 3B-3. Students with poor test scores are at especially high risk of leaving UC with no degree. Top row: SAT scores are summed SATRM + SATRW from the 2010-2012 admissions pool. Lines represent mean rate of non-retention. Bottom row: Dot area is proportional to number of students in each SAT and HSGPA band (normalized by number of students in each demographic group). HSGPAs are weighted. Color indicates HSGPA bands. Bottom row panels are the same data as in Figures 3A-3, 3A-4, and 3A-5.

This does not necessarily mean UC should not admit those students. One of the goals of UC admissions is to expand college access by providing opportunities to promising students who have not had equal opportunities to manifest high academic achievement. This goal has been previously endorsed by BOARS in its statements on Comprehensive Review[45]. However, these results show that

[45]https://senate.universityofcalifornia.edu/_files/committees/boars/BOARSREPORTCOMPREHENSIVEREVIEW2012.pdf "One measure of excellence relates to the traditional academic indicators, GPA, standardized test scores, numbers of A-G and honors courses, discussed in the prior section, but an important test of comprehensive review is whether selection processes fulfill the mission of a great public university: to provide access to diverse quarters of the state. It is critical to this mission for admission policies and practices to balance increasing selectivity while acknowledging persistent inequalities in California schools and other opportunity deficiencies that impede the ability of talented students to meet their potential. The extent to which UC is fulfilling this mission can be assessed by examining systemwide and campus-specific outcomes using a range of demographic indicators, and by placing the academic indicators discussed in the previous section in context."

1) Discussions about expanding ELC, and particularly about increasing admission through the ETR pool, must be made with full understanding that test scores do provide predictive power about UC outcomes.
2) In the absence of other changes to support for vulnerable students, modifications to UC admissions processes to admit higher numbers of students with low test scores (as the STTF is recommending) are likely to decrease retention, decrease GPA, increase time-to-degree, and decrease graduation rates. This is the basis for the Task Force recommendation that campuses examine whether their supports for vulnerable students are adequate.

C. UC does not appear to use standardized test scores in a way that amplifies racial disparities or that prevents low-scoring students from disadvantaged groups from being admitted to UC.

Mean differences in test scores between members of various racial, ethnic, and socioeconomic groups can be very large (hundreds of points on the SAT, one standard deviation or even more). Prior BOARS committees, UC governing bodies, and outside advocacy and lobbying groups have at times expressed concern that this may needlessly block talented members of certain groups from UC access: in other words, that there are applicants who are otherwise prepared, would do well at UC, and are prevented from attending UC based on a single test score. It is clear that many of the possible ways of using test scores in admissions (for instance, hard score cutoffs or fixed-weight systems) could produce this result. UC's policy on comprehensive review is in part designed to address this concern, identify such applicants, and ensure that they are offered UC admission irrespective of their test scores. "Guiding Principles for Comprehensive Review" in the BOARS document entitled "Guidelines for Implementation of University Policy on Undergraduate Admissions"[46] defines comprehensive review as "The process by which students applying to UC campuses are evaluated for admission using multiple measures of achievement and promise while considering the context in which each student has demonstrated academic accomplishment" and includes the provision that: "No fixed proportion of applicants should be admitted based solely on a narrow set of criteria."

The current Task Force shares the concern that large differences in test scores across demographic groups could lead to large disparities in admissions rates across these groups for students who were otherwise quite similar. STTF members undertook a series of analyses aimed at determining how often this happens and how well our comprehensive review system functions to promote access.

C-1. How UC campuses use standardized test scores in selection. Here, it is useful to review how most UC campuses use test scores in selection. UCOP is responsible for producing contextual data – aimed at ensuring that an applicant's academic accomplishments are evaluated relative to peers in their high school. To do this, UCOP annually generates percentile tables for all high schools that have at least 20 UC applicants across the past three years. Percentiles are provided within the applicant pool and within a particular school, by individual UC campus and systemwide for standardized test scores as well as for high school GPA, number of A-G courses, and honors/AP courses. Normalizing academic achievements and test scores to each high school allows strong students to stand out even if the school is not a particularly high performing or well- resourced school. Centralizing this data production at UCOP provides systematization, fairness, and economies of scale, ensuring all campuses have access to the information that would support comprehensive review.

[46] https://senate.universityofcalifornia.edu/_files/committees/boars/documents/guidelines-implementation-of-ug-admission-rev-7-2019.pdf

C-2. A note on test optional. It is important to note that this system works as well as it does because UCOP receives both test scores and grades for all the applicants to any UC campus from a given high school. Because UCOP receives scores from so many of the students at each school, they can supply the campus admissions officers with scores normalized by high school, thus letting the readers judge whether a student performed exceptionally well in the local context. A switch away from mandatory submission of test scores to a "test-optional" regime in which students choose whether or not to take a test/submit a score would remove UCOP's ability to normalize scores by school and thus to compensate for school to school variability in educational quality.

C-3. Family income and first-generation status play a strong role in admissions. UC does not use hard score cutoffs. UC admits members of different groups with widely varying test scores. It is well known that students in disadvantaged groups tend, on average, to have lower HSGPAs and test scores than students without such disadvantage. The UC application asks students to report, among many other things, their annual family income and whether they would be the first member of their immediate family to graduate from a four-year institution (first-generation status). Table **3C-1** presents the differences in average HSGPA and SAT for three groups: low-income vs. not low-income; first-generation vs. not first-generation; and applicants who are both low-income and first-generation vs. those who are neither. These group average differences are substantial, especially for those applicants who are both low-income and first-generation[47].

	Income	**First Generation**	**Both**
Average Difference in HSGPA	0.18	0.20	0.25
Average Difference in SAT	194	216	226

Table 3C-1. Differences in average HSGPA and SAT between various demographic groups, for applicants to any UC campus for fall 2018. Breakdown of this result for individual campuses is in Appendix V, Tables 1-3, which show generally similar trends at the highly selective campuses, and indeed at almost all campuses, with the exception of Riverside and Santa Cruz

Examination of the admission of applicants for fall 2018 show clearly that family income and first-generation status play a strong role in UC admissions. **Figure 3C-1** and **Figure 3C-2** show the percentage of applicants who were admitted to any campus by bands of raw (not normalized) SAT scores and HSGPAs. Focusing on the vertical (% admitted) difference between the disadvantaged and not disadvantaged groups, the results show that applicants who were either low-income, first-generation, or especially both were much more likely to have been admitted within nearly all bands of SAT and HSGPA.

[47] The difference of average SAT scores of 226 points between those who are both low-income and first-generation is about 1.3 standard deviations below the average of the non-disadvantaged group.

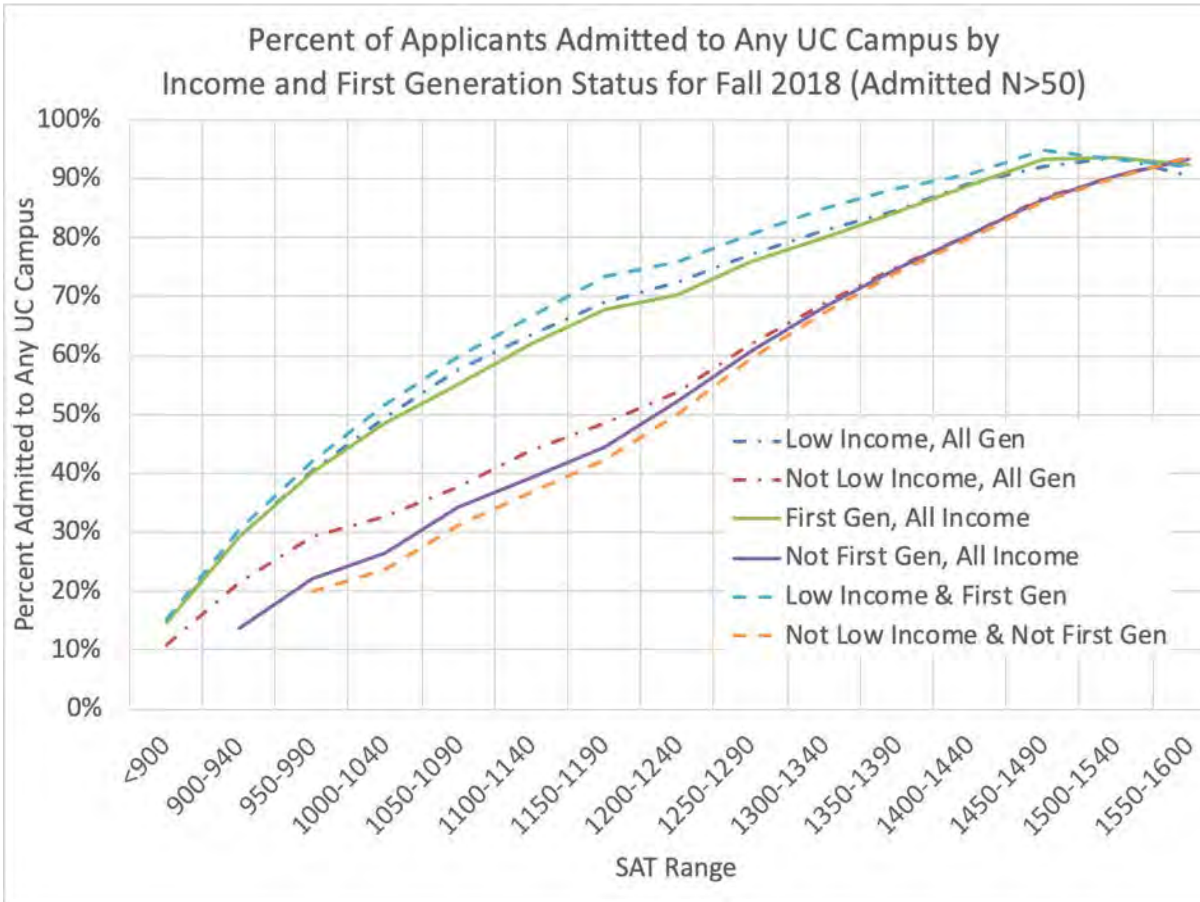


Figure 3C-1. Admit rate for fall 2018 comparing applicants in various demographic groups by SAT. Breakdown of this result for individual campuses is in Appendix V, Figures 11-30 D, which show generally similar trends at the highly selective campuses, and indeed at almost all campuses, with the exception of Riverside and Santa Cruz.

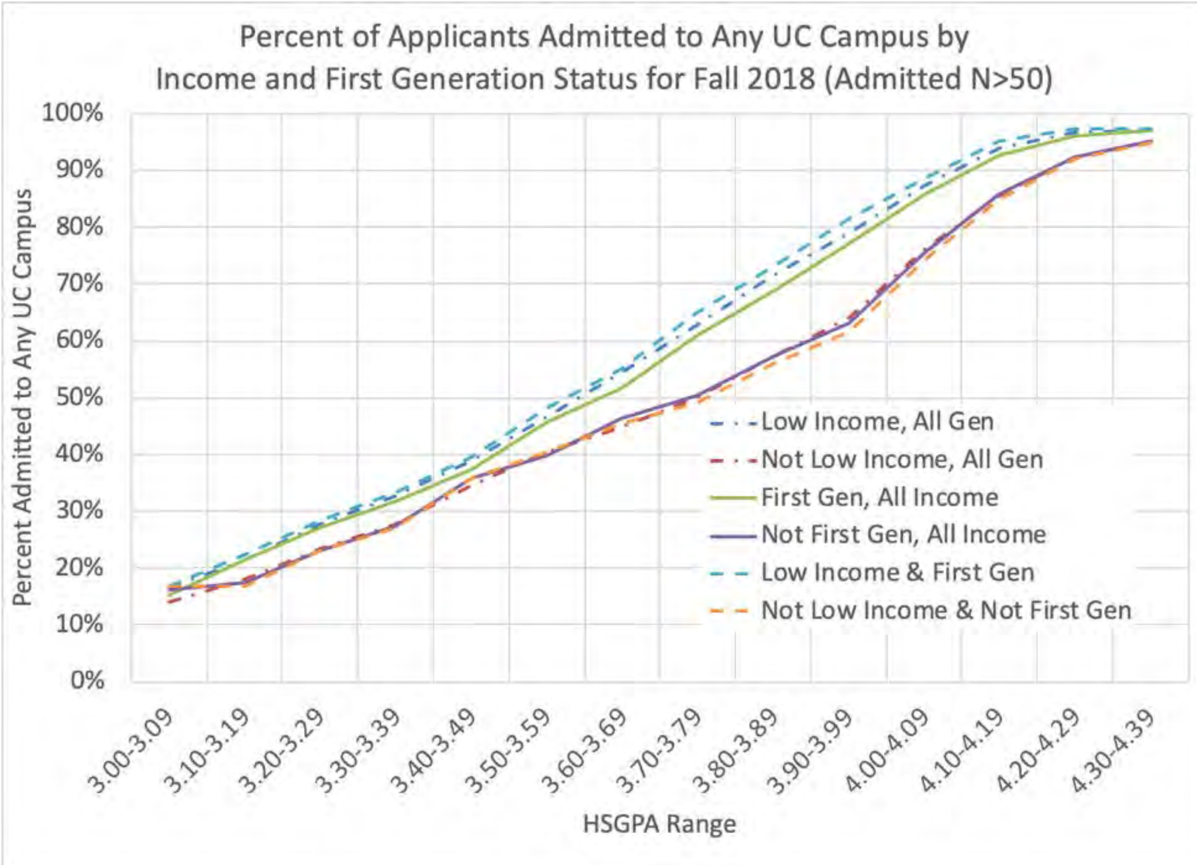


Figure 3C-2. Admit rate for fall 2018 comparing applicants in various demographic groups by HSGPA. Breakdown of this result for individual campuses is in Appendix V, Figures 11-30 B, which show generally similar trends at the highly selective campuses, and indeed at almost all campuses, with the exception of Riverside and Santa Cruz.

If campuses did nothing to account for the fact that students are members of these groups whose academic indicators are likely to be lower than the entire applicant pool, their admission rates would be very low. However, consistent with university policy on comprehensive review,[48] UC does appear to use an applicant's non-academic information in making admission decisions.

Race, ethnicity and gender are not, of course, considered in admissions, but other demographic factors such as family income and first-generation status are. BOARS has identified 14 groups of factors that should be used in some combination during comprehensive review of applications. One of these states:

> *13. Academic accomplishments in light of the applicant's life experiences and special circumstances. These experiences and circumstances may include, but are not limited to, disabilities, low family income, first generation to attend college, need to work, disadvantaged social or educational environment, difficult personal and family situations or circumstances, refugee status, or veteran status.*[49]

[48] https://senate.universityofcalifornia.edu/_files/reports/hp2mgy_boars-testing_010609.pdf

[49] https://senate.universityofcalifornia.edu/_files/committees/boars/documents/guidelines-implementation-of-ug-admission-rev-7-2019.pdf

	Racial/Ethnic Group					
	Latino	White	Asian	African-American	American Indian	Unknown
% of group who are low income	61%	17%	32%	51%	28%	32%
% of group who are first-generation	78%	19%	33%	50%	34%	30%
% of low income total	57%	9%	24%	8%	~0%	2%
% of first-generation total	61%	9%	21%	6%	~0%	2%

Table 3C-2. Fraction of applicants receiving low-income and first-generation preferences.

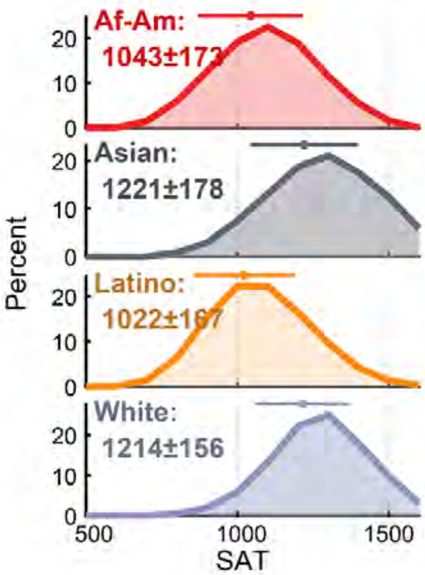
This preference for low-income and first-generation students benefits all applicants in these categories, so it is relevant to consider what percentage of applicants in various racial/ethnic groups are affected. This is shown in **Table 3C-2**. The groups with the highest in-group percentages for both low-income and first-generation are Latino and African-American applicants. Within these groups, 61% of Latino applicants and 51% of African-American applicants are identified as being low-income, while 78% of Latino applicants and 50% of African-American applicants are identified as being first-generation. Thus, these groups should benefit substantially from the preference given to low-income and first-generation status, even though race and ethnicity are not explicitly considered.

C-4. UC's admissions process does not appear to increase racial/ethnic disparities by giving undue weight to admissions test scores. STTF members were keenly aware of the large differences between mean test scores among different racial and ethnic groups. The STTF was concerned that the unique challenges faced by members of certain minority groups may produce disparities in admissions outcomes, mediated at least in part by differences in test scores, which may not be fully compensated for by considering parental income, educational history, or the other comprehensive review factors UC is legally allowed to consider in admissions. This section expands upon the work presented in the previous subsection to assess the extent to which racial and ethnic disparities are ameliorated by use of the 14 factors in comprehensive review.

Table 3C-3 shows the mean SAT scores of applicants by group. The rightmost column displays gaps between the given group and a comparison group, as stated in the column headers. **Figure 3C-3** shows the distribution of SAT scores of applicants by racial/ethnic group. On average, applicants from different racial/ethnic groups do have significantly different test scores, with gaps in average test scores as large as 234 points between Asian and Latino applicants.

Because of these gaps, STTF was concerned that UC might be using the SAT in ways that disparately affects particular racial/ethnic groups. STTF regarded this as a plausible and sensible claim that an intelligent layperson concerned about diversity would be likely to make. In the absence of persuasive evidence that this did not happen, the STTF was prepared to recommend significant restrictions to UC's use of tests. Indeed, elsewhere in this document the Task Force has marshalled the main arguments for dropping admissions tests, and that section writes stirringly about differences by group in SAT scores such as those shown in **Table 3C-3**, which many consider to be evidence that the SAT has a disparate impact on admissions by racial and socioeconomic groups.

Example test score differentials by applicant group
2018 (left) and 2019 (right)



Race/Ethnicity	Mean SAT	Gap Vs. Asian
African-American	1166	190
Native American	1259	97
Asian	1356	0
Latino	1122	234
Pacific Islander	1216	140
White	1338	18
Unknown	1356	0

Figure 3C-3. SAT Scores of 2018 Applicants by Racial/Ethnic Group. Breakdown of this result for individual campuses is in Appendix V, Figures 1-10 Cab, which show generally similar trends at the highly selective campuses, and indeed at almost all campuses, with the exception of Riverside and Santa Cruz.

Table 3C-3. Mean SAT Scores of 2019 Applicants by Racial/Ethnic Group. Source: Mean SAT scores provided by UCOP. (Note: The 2019 SAT scores are from the new version of the SAT, on which students tend to receive higher numerical scores than on the old SAT.) Breakdown of this result for individual campuses is in Appendix V, Tables 4-6, which show generally similar trends at the highly selective campuses, and indeed at almost all campuses, with the exception of Riverside and Santa Cruz.

We did not find evidence that UC's use of test scores worsened disparities that are already present among applicants. Rather, we found that, even without explicitly considering race or ethnicity in admissions, UC's use of test scores along with the other factors incorporated into the comprehensive review of each application compensated for the majority of the score differences between groups. If the SAT mattered significantly for admissions, we would expect to see that the admission rate is very low up to some SAT cutoff, and then rise steeply. This would greatly reduce the admission rate for the groups in **Table 3C-3** with lower mean test scores. We did not observe this to be the case.

Table 3C-4 repeats the columns in **Table 3C-3**, but also adds a column showing the mean SAT score of those admitted to UC for each group. **Table 3C-4** also provides information for other student groupings, including grouping based on family income and whether either parent has graduated from college. The column on the left shows large differences in mean SAT scores between low-income and non-low-income applicants, and between first-generation and non-first-generation families. For instance, the gap in mean SAT scores between applicants who are first generation and those who are not is 216 points, a gap almost as large as that between Latino and Asian applicants.

2019 Admissions	Applicants		Admits	
Race/Ethnicity	Mean SAT	Gap Vs. Asian	Mean SAT	Gap Vs. Asian
African-American	1166	190	1230	156
Native American	1259	97	1308	78
Asian	1356	0	1386	0
Latino	1122	234	1172	214
Pacific Islander	1216	140	1269	117
White	1338	18	1368	18
Unknown	1356	0	1406	-20
First-Generation College		Vs. Not First Generation		Vs. Not First Generation
First-Generation	1145	216	1201	192
Not First-Generation	1361	0	1393	0
Unknown	1286	75	1341	52
Low-Income Status		Vs. Not Low Income		Vs. Not Low Income
Low-Income	1143	194	1200	171
Not Low-Income	1337	0	1371	0
Unknown	1348	-11	1392	-21

Table 3C-4. Mean SAT Scores of 2019 Applicants and Admits by Racial/Ethnic Group. Source: Mean SAT scores provided by UCOP. (Note: The 2019 SAT scores are from the new version of the SAT, on which students tend to receive higher numerical scores than on the old SAT.) Lines above distributions show means and standard deviations. Means and standard deviations are also printed on each panel. Breakdown of this result for individual campuses is in Appendix V, Tables 4-6, which show generally similar trends at the highly selective campuses, and indeed at almost all campuses, with the exception of Riverside and Santa Cruz.

The rightmost column in **Table 3C-4** shows the gap in test scores between those admitted in each group and in a selected group (that with the highest mean test score). If the SAT was the predominant admissions factor, and UC ignored an applicant's context in its comprehensive review, these gaps in mean test scores among *admits* should be much smaller than among *applicants*. The gaps remain large – approximately the same size as the gaps in the test scores of applicants. For instance, the Latino/Asian test score gap among applicants, 234 points, is almost exactly preserved among the admitted students, with a gap of 214 points. *The implication is that comprehensive review effectively renorms SAT scores through its consideration of the broader set of 14 factors.*

This important finding, that admissions procedures appear to compare students' SAT scores within socioeconomic groupings, but not across groupings, is reinforced by using graphs to look at the entire distributions of tests scores of applicants and accepted students. **Figure 3C-4** shows SAT scores of students of different racial/ethnic groups accepted to UC overlaid on scores of applicants to UC. Counter to the belief implicit in the claim that the SAT plays too big of a role, and that UC pays less attention to other factors, the mean SAT scores of those *admitted* varies dramatically by group. This level of inter-group difference would be implausible if the SAT played such a predominant role in admissions as is often thought. Just as **Table 3C-4** suggests, the SAT scores seem to be renormed to take into account students' personal context.

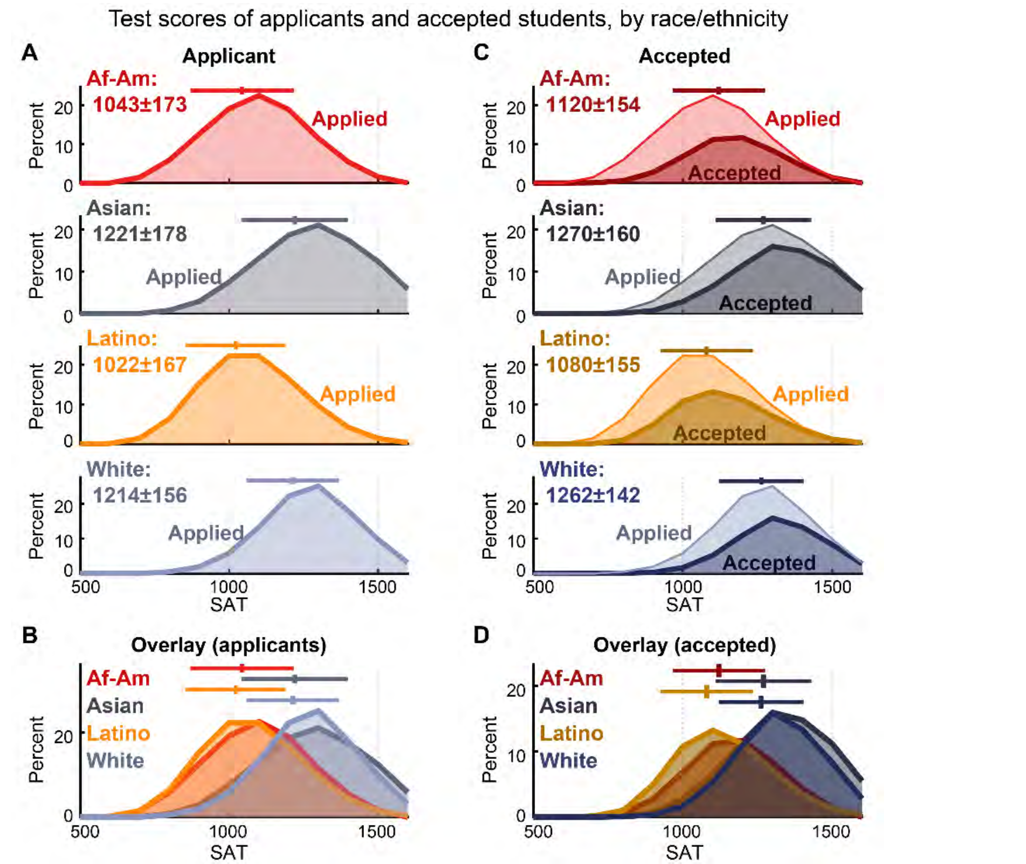


Figure 3C-4. Distributions of SAT scores among applicants and among admits, by racial/ethnic group. The figure shows the distribution of SAT scores for applicants (left, and grayed-out on right for ease of comparison) and admits (right) by racial/ethnic group. Data represent admission decisions for students applying for admission in fall 2018. Lines above distributions show means and standard deviations. Means and standard deviations are also printed on each panel. Breakdown of this result for individual campuses is in Appendix V, Figures 1-10 C, which show generally similar trends at the highly selective campuses, and indeed at almost all campuses, with the exception of Riverside and Santa Cruz.

The STTF considered whether this implied that the use of tests in admissions was small enough that test scores could be dropped without serious consequences. Yet STTF considered one goal should be to identify students who are the most likely to be prepared to take advantage of UC's rigorous educational opportunities. As shown in Section II-A of this chapter, the SAT is highly predictive of a broad range of student outcomes at UC, within, not merely across, specific demographic groups. Further, the admissions procedures do take into account differences in test scores within group. The rationale for use of test scores is that it increases the likelihood that students in the admitted pool will graduate from UC.

Figure 3C-5 (left column, top row) shows even more clearly that UC comprehensive review policy effectively compensates for most group differences in the SAT scores, by plotting the admission rate of each racial/ethnic group by SAT score. It shows that less advantaged racial/ethnic groups are admitted at higher rates for any given SAT score. This finding is inconsistent with claims that UC considered SAT in a way which fails to take into account disadvantage. Indeed, to some readers, the differences in admission rates by group, holding constant SAT scores may seem surprisingly large. To give one striking example: for students with SAT scores of 1000, admission rates are about 30% for whites but roughly 50% for Latino students. How high do SAT scores need to be for whites to have the same 50% admission rate as Latino students with scores of 1000 do? The answer is that whites do not reach admission rates of 50% until they reach an SAT score of 1200. As **Table 3C-4** shows, this gap almost exactly matches the gap in mean test scores among different groups of applicants. The fact that it is socioeconomic factors that are at work here is clear from the fact that on average Asian students benefit more than white students due to the fact that the percentage of Asians who are low-income and/or first-generation students is roughly twice that of whites. (See **Table 3C-2**.)

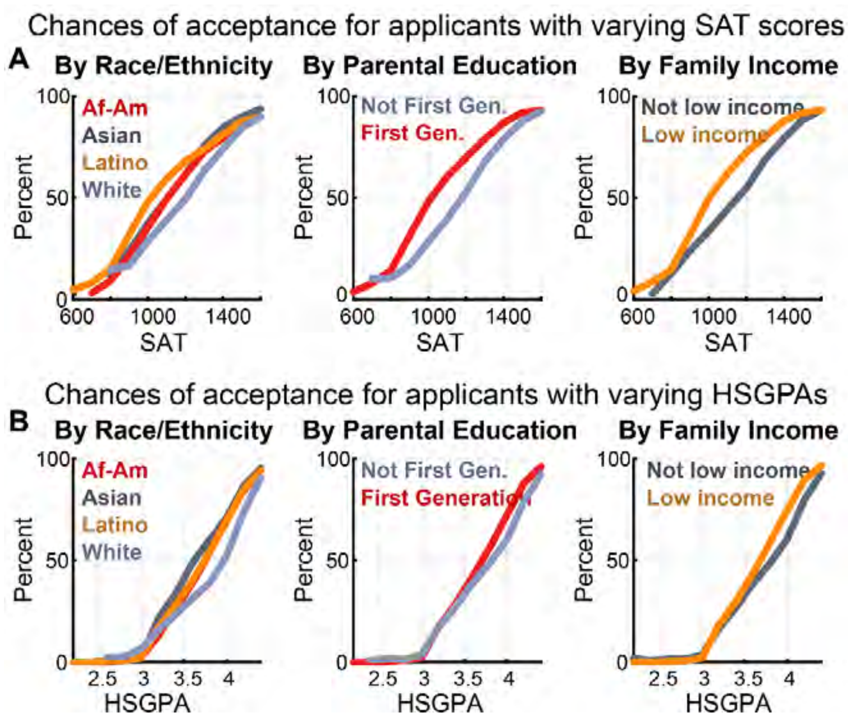


Figure 3C-5. Admission rates by racial/ethnic group, parental education level, and family income level for given SAT scores or HSGPA. Data represent admission decisions for students applying for admission in fall 2018. Breakdown of this result for individual campuses is in Appendix V, Figures 1-30 B and D, which show generally similar trends at the highly selective campuses, and indeed at almost all campuses, with the exception of Riverside and Santa Cruz.

Figure 3C-5 (center column, top row) shows corresponding admission rates based on whether applicants are the first in their family to graduate from college (similar to **Figure 3C-1**). It shows large advantages in admission rates for first-generation students compared to not-first-generation students. Students whose parents have not graduated from college are admitted at rates 15 to 20 percentage points higher for all levels of SAT scores except the very lowest range, roughly below 900 (where both groups are equally likely to be admitted), and those approaching a perfect 1600 score (who are likely to be admitted no matter their family educational history).

Figure 3C-5 (right column, top row) shows admission rates for each SAT score based on whether the family is low-income or not (similar to **Figure 3C-1**). Again, the relatively disadvantaged group has substantially larger admission rates for students with SAT scores between about 900 and 1500.

Further confirming that the comprehensive review process has the practical effect of compensating for the variations in SAT scores across groups, **Figure 3C-4** above shows for each racial/ethnic group, and **Figure 3C-6** below shows for each income and educational group, the distribution of applicants' test scores (left) and the distribution of admits' test scores (right). Although there is considerable overlap in the admits' test scores across groups, the figure clearly shows that campuses are taking into account disparities in educational opportunities and socioeconomic status of the various groups.

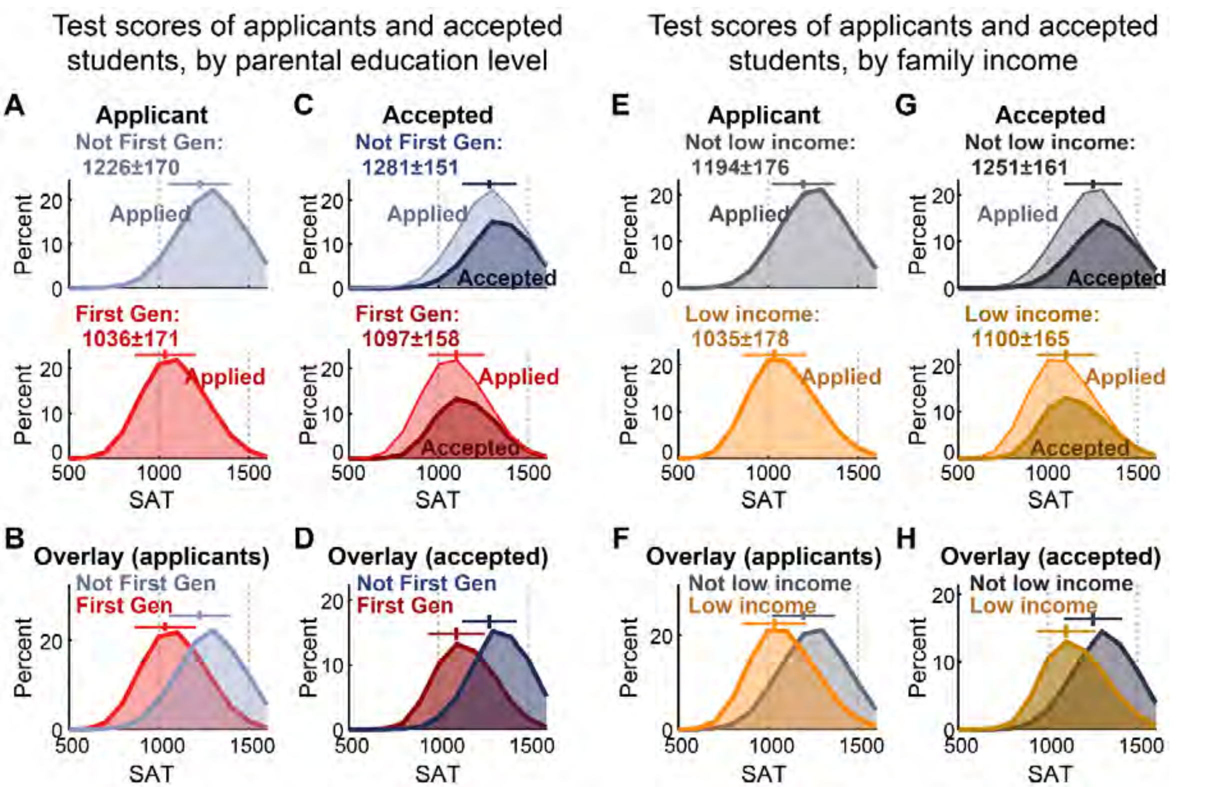


Figure 3C-6. Distributions of SAT scores among applicants and among admits by parental education (left) and family income level (right). Data represent admission decisions for students applying for admission in fall 2018. Distributions of test scores for applicants are shown in paler colors, alone (left columns) and underlaid beneath the distributions of test scores for admitted students (darker colors, right columns). Lines above distributions show means and standard deviations. Means and standard deviations are also printed on each panel. Breakdown of this result for individual campuses is in Appendix V, Figures 11-30 A and C, which show generally similar trends at the highly selective campuses, and indeed at almost all campuses, with the exception of Riverside and Santa Cruz.

C-5. The overall admissions procedures at UC make a smaller than expected contribution to racial and ethnic differences in admissions rates. Failure to complete minimum admissions requirements is a substantially larger contributor than factors differentiating among eligible students. UC has a unique two-stage selection process: determination of systemwide eligibility, followed by campus- and major-specific selection leading to admissions, in which campuses choose among students meeting the minimum admissions requirements. Strategies for addressing underrepresentation at UC will be very different depending on whether the disparities arise from failure to meet minimum UC requirements, or whether they arise from differences in test scores, HSGPAs, curricular rigor, or other factors causing campuses to choose among eligible students at different rates. The STTF was particularly concerned that the requirement that students even take a standardized test might be a major contributor to the racial/ethnic differences in admissions rate we observed.

Surprisingly, the STTF's preliminary study of the factors limiting minority enrollment found that much of the difference between the racial makeup of UC's resident undergraduate population and the state's high school population depends on other factors that limit eligibility, particularly UC/CSU's requirement that applicants complete a broad college preparatory curriculum (A-G) with grades of C or higher. These data suggest that a move to drop testing and instead to rely on HSGPA alone for eligibility, and the remaining 12 factors for comprehensive review, would not do as much to increase diversity as might be expected, because these other factors play such large roles in reducing the enrollment of Latino, African-American, and Native American students at UC.

First, consider overall patterns when these three groups are combined. In the 2017-18 school year, the URM groups constituted 59.1% of high school seniors in California, but only 37.0% of UC admits. The gap of 22.1% is large by any standard.

What factors lead to this opportunity gap? Students need to complete a series of three actions before they can be considered for admission to a UC campus. At a minimum, 1) they must have graduated from high school. In order to be eligible for UC admission, 2) they must also have satisfied the A-G requirements. Having graduated with the requisite courses, 3) they must then apply, and finally 4) an admission decision is made.

Figure 3C-7 shows the contributions of these four factors. The three pre-admission factors explain much of the disparity. The single biggest contributing factor is the lower shares of URM students who graduate having completed the A-G requirements. URM students are also less likely to graduate from high school than other students, accounting for just over 2 percentage points of the 22 percentage point gap. Differences between URM students and others in their rates of application and rates of admission account for about 5 percentage points each.

Figure 3C-8 depicts this information differently by showing the percentage contributions to the admissions gap of each of these factors. The figure shows that the single largest contributing factor is the lower than average completion rate by URM of A-G college preparatory coursework in high school. This factor alone accounts for 40% of the gap between the shares of URM students in high school and in the admitted freshmen at UC. Lower rates of application to UC and acceptance to UC conditional upon having applied each explain about 25% of the gap. Finally, lower high school graduation rates explain about 10% of the gap.

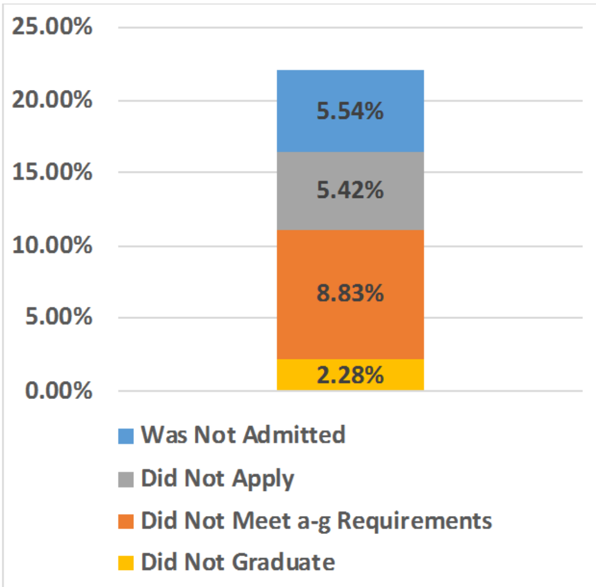


Figure 3C-7 The contributions of various factors to the gap between underrepresented minority shares of high school seniors and UC in-state admits, 2017-2018. Notes: See notes to **Table 3-C-5**.

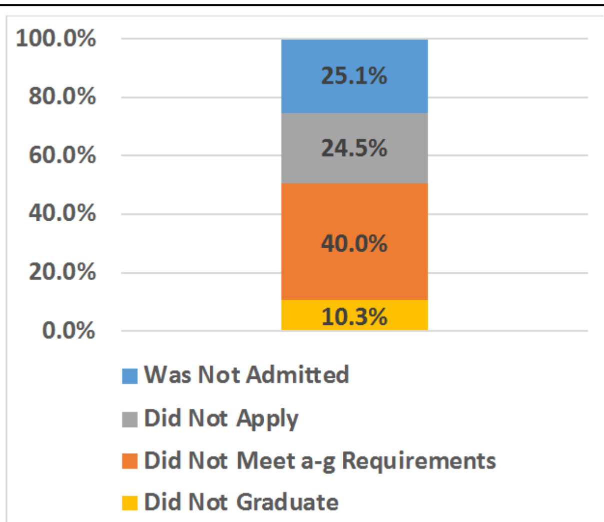


Figure 3C-8 The percentage contributions of various factors to the gap between underrepresented minority shares of high School seniors and UC in-state admits, 2017-2018 (numbers add to 100%). See notes to **Table 3C-5.**

Table 3C-5 below shows the data underlying the above figures, while also providing details by ethnic/racial group. The first three columns show for California the racial makeup in 2017–18 of high school seniors, high school graduates, and graduates meeting the UC/CSU's A-G requirements with grades of C or higher, all obtained from the California Department of Education.[50] This last piece of information is crucial, because students must complete the A-G requirements to be eligible to attend UC. The rightmost columns add information on the racial mix of applicants to UC and admits to UC for fall 2018.[51] [52] The rows for the three groups that UC considers as underrepresented minorities are highlighted in yellow. The final row combines the three URM groups. (It is these data that appear in **Figures 3C-7** and **3C-8**.)

Table 3C-6 calculates the differences between various columns in **Table 3C-5** to show how the racial mix changes at each step as we compare high school seniors to graduates, graduates to graduates who complete the A-G requirements and so on. The first column shows the difference between each group's percentage in the high school senior population and the percentage in UC's admit population. At the bottom of the first column, we see that overall there is a 22.1 percentage point gap between the share of under-represented minorities in the high school class of 2018 and the students admitted to UC in 2018. This crucial first column also shows that Latino students are under-represented among UC admits by 18.2%, while the corresponding percentages are 1.85% for African-Americans and 0.28% for Native Americans. Because these three groups vary in their shares of the high school population, it makes sense to compare these gaps to their overall shares in the high school class of 2018, shown in the first column. Relative to their share of the high school population, Latino students are under-represented by about a third, compared to just over half for African-Americans and two-thirds for Native Americans.

[50] For source see notes to the table.

[51] For source see notes to the table.

[52] It is important to note that the way UC presents most data by race and ethnicity does not match the federal standard. The data in the above table and the preceding graphs in this subsection all uses the federal ethnicity/race standard. The difference is that the UCOP uses the following hierarchical order to determine ethnic/racial classification for an applicant when more than one ethnicity is selected: African-American, Chicano, Latino, American Indian, Asian, Pacific Islander, White, Missing/Unknown. This means that when an applicant selects African-American and Latino, UC puts this applicant in the African-American category. The federal standard takes a different approach in which choice of Hispanic/Latino ethnicity by a person trumps any choice of race that person might make. Federal guidelines can be found at http://data.ucop.edu/support-training/support-resources-files/Business%20Rules%20-%20UC%20and%20IPEDS%20Ethnicities.pdf. For UC applicants and admits, the biggest difference this makes is that UC counts students with race of black and ethnicity of Hispanic/Latino as black students. Under the federal definition, these students would be listed as Hispanic/Latino. Elsewhere in this report, calculations we made were based on a database of individual students that use the UC definition which under the federal definition will overcount blacks and undercount Hispanic/Latino students. The differences in the racial mix that result from the UCOP system are small in terms of percentage points, but can be large in terms of percentages (proportions) for the smaller groups such as African-Americans. To give one example, in 3C-5 below, we calculate using federal definitions that African-American, Native American and Latino students account for 2.8%, 0.1% and 37.0% of UC admits. If we had used the UCOP racial/ethnic rules, the respective numbers would have been 4.8%, 0.5% and 33.0% respectively.

Race/Ethnicity	Cohort Students (Enrolled in Expected Graduation Cohort)	Regular HS Diploma Graduates	Graduates Meeting a-g Require-ments	UC CA Resident Applicants	UC CA Resident Admits
Hispanic or Latino	52.29%	50.82%	43.31%	38.72%	34.12%
Asian *	12.24%	13.79%	20.03%	26.46%	31.77%
White	24.32%	25.51%	27.87%	22.67%	22.48%
African American	6.18%	5.46%	4.33%	3.69%	2.78%
American Indian or Alaska Nat	0.62%	0.53%	0.34%	0.15%	0.13%
Pacific Islander	0.52%	0.51%	0.44%	0.44%	0.35%
Two or More Races	2.73%	2.77%	3.10%	5.90%	6.05%
Not Reported	1.10%	0.61%	0.58%	1.97%	2.32%
Underrepresented Minorities	59.09%	56.82%	47.98%	42.57%	37.03%

Table 3C-5 The Racial/Ethnic Mix of Students in High School and in Applications Data for the UC, 2017-2018 Notes: * Asian group includes Filipinos. Data from the first three columns from Dataquest downloaded 11/2019 at https://dq.cde.ca.gov/dataquest/dqcensus/CohRate.aspx?cds=00&agglevel=state&year=2017-18. Data from the final two columns downloaded 11/19 (CA Residents only) from https://www.universityofcalifornia.edu/infocenter/admissions-residency-and-ethnicity. The first three columns show a different racial/ethnic breakdown than the last two columns because the federal government and UC divide students by racial/ethnic identity along different lines.

Race/Ethnicity	Gap between Cohort Student Percentage and UC Resident Admit Percentage	Of this Gap, Amount Due to Differences in Graduation Rates and Rates of Meeting UC Requirements	Amount that Can Be Explained by Difference between Application and Admit Rates	Percentage of Difference Explained by Admissions
Hispanic or Latino	18.17%	8.98%	4.60%	25.34%
Asian *	-19.53%	-7.79%	-5.31%	27.19%
White	1.84%	-3.55%	0.18%	9.98%
African American	3.40%	1.85%	0.92%	26.90%
American Indian or Alaska Native	0.49%	0.28%	0.02%	4.58%
Pacific Islander	0.17%	0.08%	0.09%	50.95%
Two or More Races	-3.32%	-0.37%	-0.15%	4.60%
Not Reported	-1.22%	0.52%	-0.35%	28.72%
	0.00%	0.00%	0.00%	0.00%
Underrepresented Minorities	22.06%	11.11%	5.54%	25.12%

Table 3C-6 Decomposition of the Gap between the Racial/Ethnic Mix of Students in High School and UC Admitted Students, 2017-2018. See notes to **Table 3C-5**. The rightmost column calculates the third column as a percentage of the first column.

The columns further right in **Table 3C-6** break down the percentage gaps into gaps due to various factors. The second column compares the racial mix of high school seniors to the racial mix of high school graduates who complete the A-G requirements. The calculations show that 11.1 percentage points of the 22.1 percentage point gap in the share of under-represented minorities between high school seniors and UC admits can be accounted for by the lower than average rates at which under-represented minorities graduate having completed A-G requirements. The column furthest to the right shows the extent to which the gap between applications and admissions can explain the overall gaps between each group's shares of the high school class of 2018 and the students admitted to UC. For Latino and African-American students, about one-quarter of the gap can be explained by admission decisions. For Native Americans, about 5% can be explained by admission decisions.

To obtain a better sense of the relative contributions of all of the factors, **Figure 3C-9** shows the breakdown on the relative contributions of the four factors, ranging from differences in high school graduation rates to differences in admission rates to UC. The first bar repeats the overall results for all three URM groups combined, while the bars to the right show the results by specific racial/ethnic group.

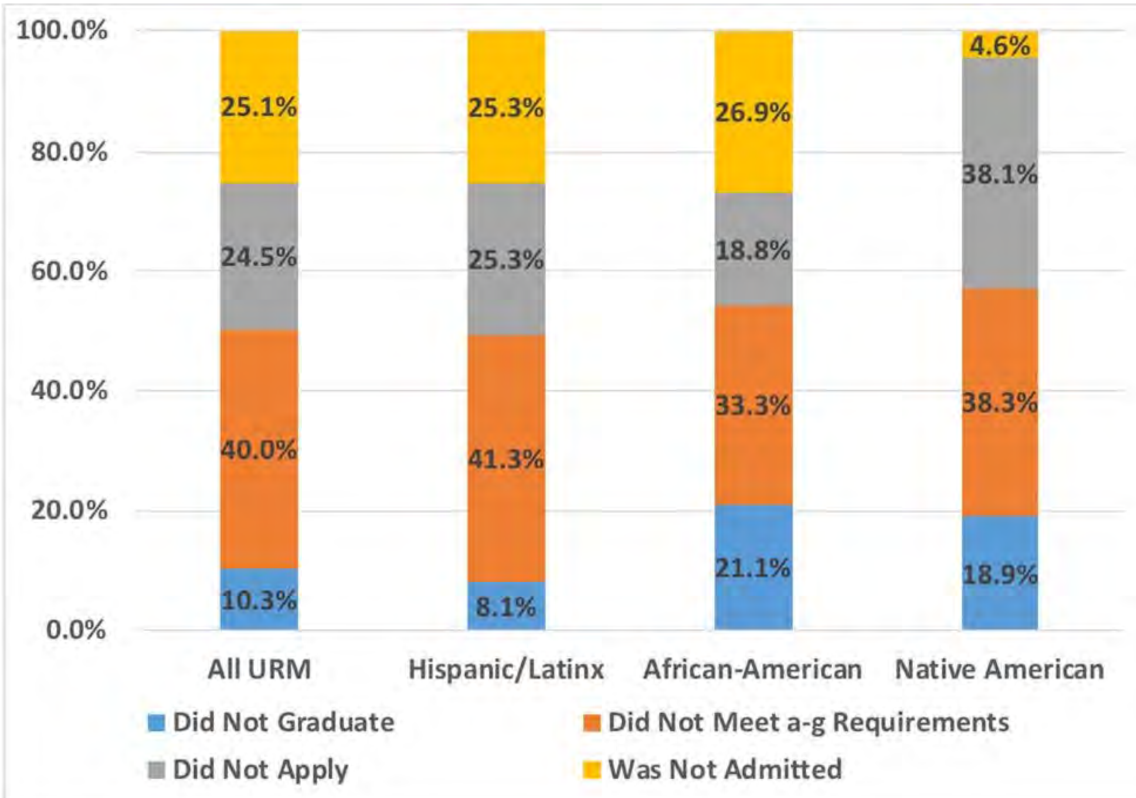


Figure 3C-9 The Percentage Contributions of Various Factors to the Gap between Each Underrepresented Minority's Share of High School Seniors and UC In-State Admits, 2017-2018 (Numbers add to 100%). Notes: See notes to **Table 3C-5.**

For all groups, about one half of the gap between the given group's share in the high school class of 2018 and the admit pool comes from lower than average high school graduation rates and lower A-G completion rates. These pre-UC-application factors were especially important for Native American and African-American students. Lower than average UC application rates for these groups also contributed, and were particularly important for Native Americans, for whom this factor explained 38.1% of the disparity. This factor was least important for African-American students.

Turning to the item of key interest, the contribution to the disparity made by lower than average UC acceptance rates ranged from a low of 4.6% of the total gap for Native Americans, to 25.3% for African-Americans and 26.9% for Hispanic/Latino students.

The conclusion from this analysis is that the comprehensive review admissions procedures at UC, of which standardized test scores and HSGPA are two components, is not the main determinant of the gap between the racial mix of high school seniors and the racial mix of UC admits.

C-6. HSGPA also shows strong racial disparities. Much of the gap between the proportions of URM high school seniors and the representation of URMs among UC admit pools reflect disparities in graduating from high school or completing the A-G coursework. Nonetheless, a smaller fraction of underrepresented minority students are admitted to UC than graduate with A-G courses completed. We examined whether factors other than differences in test scores account for any of this further drop-off.

One example of the possibility that other elements among the 14 admissions criteria may contribute to racial/ethnic gaps is the distribution of HSGPA by racial/ethnic group. Each row of **Figure 3C-10** below shows, for one of the four racial/ethnic groups most represented at UC, the distribution of HSGPA among applicants on the left, and admits on the right (with the applicant distribution shown in the paler colors on the right as well). The figure shows large differences in the distribution of HSGPA among applicants, with African-American and Hispanic students having much lower HSGPAs on average. The bottom row of the figure overlays the various racial/ethnic groups, which makes the disparities in HSGPA among groups even clearer. The right hand side of the figure shows the distribution of HSGPA among admits. UC appears to admit very few of the students with low GPAs. Differences in HSGPA distributions are much lower among admitted students than among applicants.

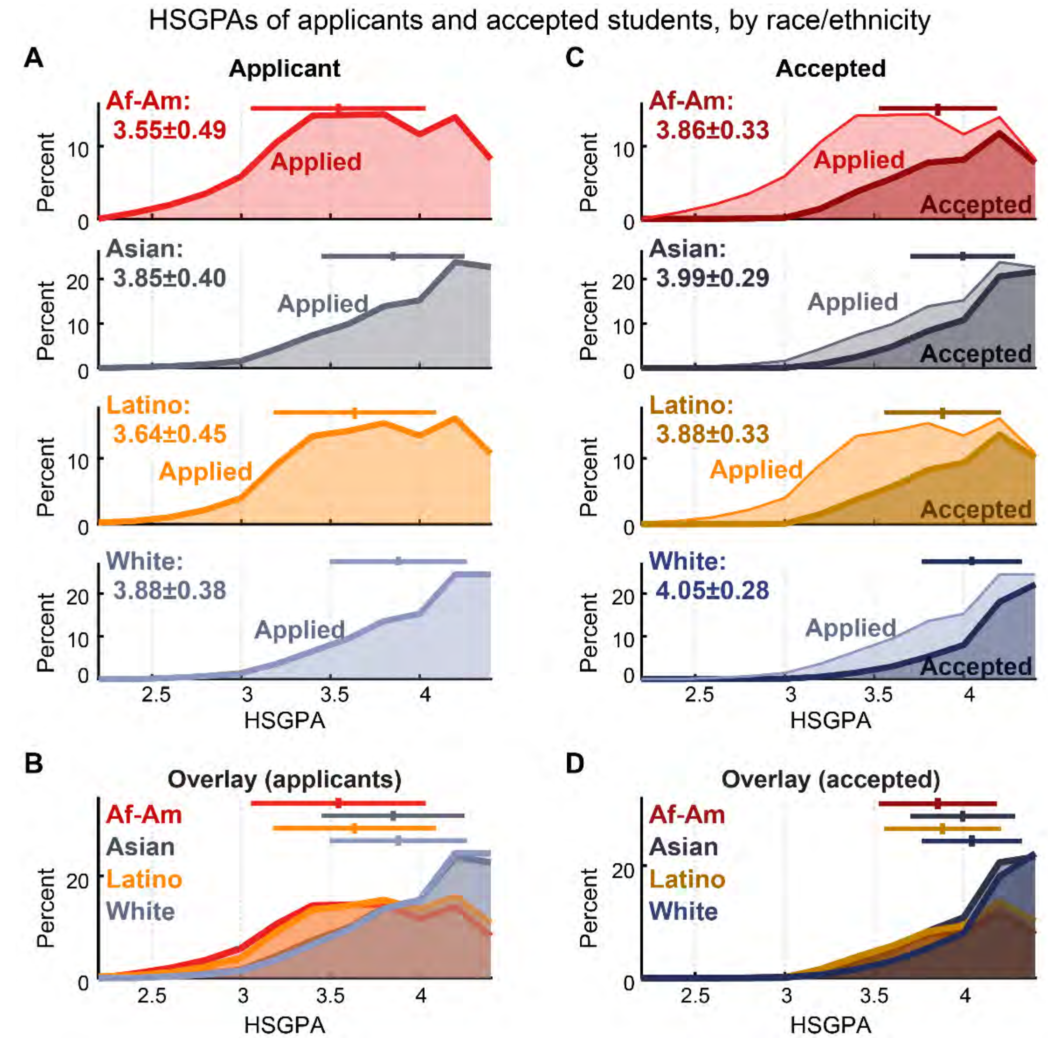


Figure 3C-10. Distribution of High School GPA (HSGPA) among Applicants and Admits to UC in 2018, by Racial/Ethnic Group. Means and standard deviations for each group are represented above each distribution and displayed numerically on each panel. Breakdown of this result for individual campuses is in Appendix V, Figures 1-10 A, which show generally similar trends at the highly selective campuses, and indeed at almost all campuses, with the exception of Riverside and Santa Cruz.

This implies that UC takes low grades quite seriously, and as a result of lower grades, smaller shares of applicants from these two groups are admitted than for the other groups. Importantly, this pattern is *quite* *different from the pattern shown in* **Figure 3C-4**, which showed that differences in SAT score distributions among admitted students were similar to the differences among applicants (implying that UC's comprehensive review process effectively renormed test scores between groups), and is consistent with the results in Section II-B showing that UC's admissions process more heavily weights high school grades than test scores.

The same pattern holds when students are grouped based on family income or family educational history. **Figure 3C-11** shows the distribution of HSGPAs among applicants and among accepted students, for students who are the first in their families to go to college (cyan) or who had a college-going parent (orange), and for students who are from low-income families (blue) or who are not low-income (red). Just as we saw in the ethnicity data, the income and education data also reveal that the differences in HSGPAs between groups of applicants are substantial, but that the differences in HSGPAs among groups of accepted students are smaller. This is quite different from the pattern of SAT score differences between groups, which are equally large among applicants as among accepted students (previously shown in **Figure 3C-6**). This indicates that UC heavily considers HSGPA in admissions, and does not appear to renormalize HSGPA between groups in the same way as SAT.

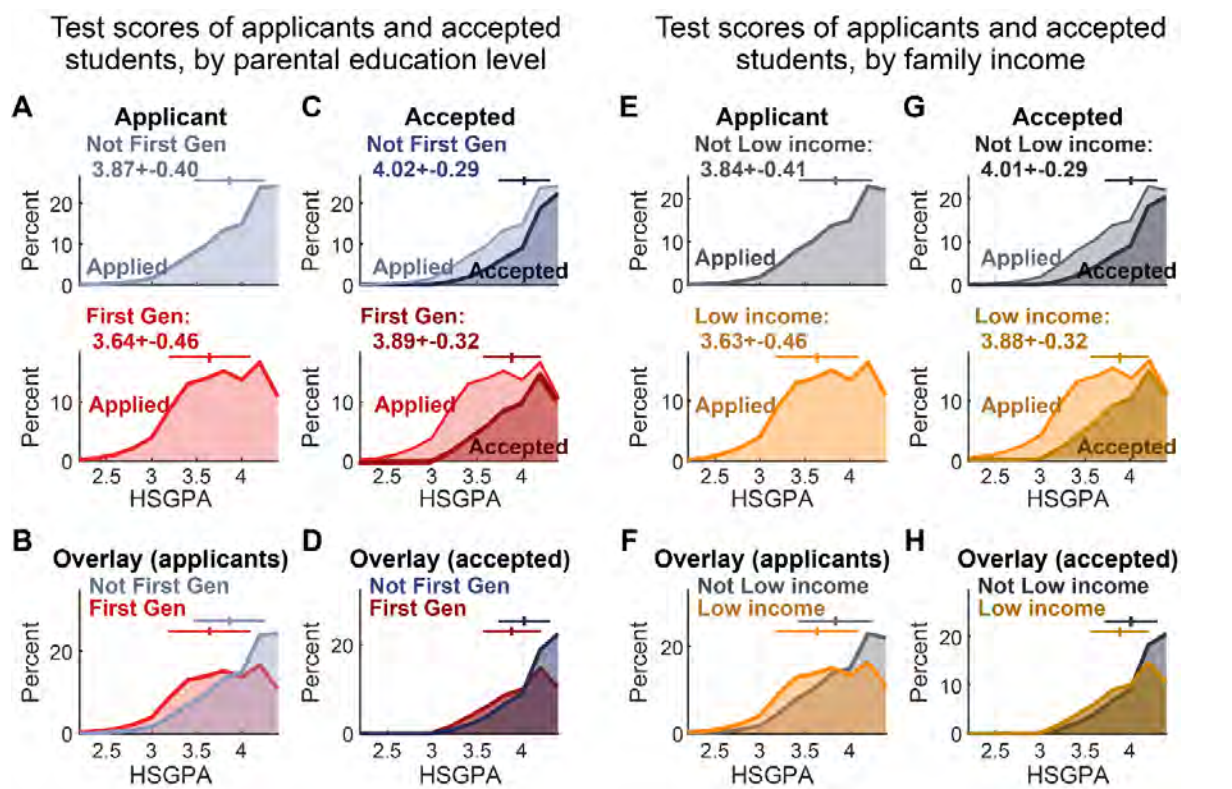


Figure 3C-11. Distributions of SAT scores among applicants and among admits by parental education (left) and family income level (right). Data represent admission decisions for students applying for admission in fall 2018. Distributions of test scores for applicants are shown in paler colors, alone (left columns) and underlaid beneath the distributions of test scores for admitted students (darker colors, right columns). Lines above distributions show means and standard deviations. Means and standard deviations are also printed on each panel. Breakdown of this result for individual campuses is in Appendix V, Figures 11-30 A, which show generally similar trends at the highly selective campuses, and indeed at almost all campuses, with the exception of Riverside and Santa Cruz.

The upshot: lower grades of African-American and Hispanic students appear to play a role in the disparity between the racial mix of high school seniors and UC students. The SAT does not appear to be the only factor, or even the disproportionate factor, in the existing differences in the racial mix of high school graduates completing A-G and enrollees at UC. Differences across racial groups in high school GPA clearly are a substantial factor.

C-7. Summary of results on test scores and representation of minorities at UC. Overall, this section has shown that for URM students overall and the constituent racial/ethnic groups, there exists a disparity between the given group's share in the graduating class in high school and UC admits. The gaps, averaging 22% overall for URM groups combined, are big enough to cause concern, especially in light of the Regental goal that UC reflect the state's diversity.

Do these large gaps result from the way UC admits students? The answer is that UC admissions do play a role in these gaps, but a relatively small role. For URM students overall and also for Latino and African-American students, lower than average UC admission rates contribute about one quarter of the gap. For Native Americans, UC admission rates explain only 4.6% of the under-representation at the University.

So what are the more important factors than UC admissions decisions? For URM students overall as well as for the three URM subgroups, relatively low rates of completion of A-G coursework is the single biggest factor explaining the under-representation. The next most important factor varies by group. For Latino students, relatively low application rates and admission rates are tied as the next most important factors. For African-American students, the second most important factor is UC admission decisions, followed closely by lower than average gaps that originate in UC admission decisions to inter-group differences in the SAT. A comparison of **Figures 3C-4** and **3C-14** above provides clear evidence that SAT scores do not play a predominant role in the different UC admissions rates, but that differences in GPA do play a substantial role.

UC considers 14 factors in admissions. Although we could not examine all 14 factors, our study found that URM applicants to UC on average have lower GPA, have completed fewer A-G courses, and have completed fewer honors courses as well. This means we cannot assume that eliminating the test requirement would raise the URM share of the admissions pool much closer to their share in the pool of high school graduates completing the A-G requirement.

C-8. Summary of results on test scores and demographic indicators overall. These analyses indicate that mean differences in standardized test scores between different demographic groups are often very large. Many of the ways these tests could be used in admissions would certainly produce strong disparate impacts between groups. Yet UC uses the test scores in ways that minimize the differential impacts on admission rates by demographic group. UC weights test scores less strongly than GPA, and through the use of comprehensive review in the admissions process appears to largely renormalize test scores by group. Although it is natural to infer that differences in test scores among applicants must *cause* differences by group in admission rates, this does not seem to be the case for the way that UC uses this information.

D. Do standardized test scores counterbalance the variability among grading patterns at different high schools?

A direct consequence of dropping admissions tests would be that HSGPA would become the only quantitative measure of academic preparation. This is problematic because HSGPA is not an objective measure of achievement. Its relation to objective measures of achievement can vary from teacher to teacher, and from school to school. We can refer to this as *variations in grading standards*. HSGPA is also a less than perfect measure of changes in academic preparation over time, due to evidence of *grade inflation*.

These two problems are related, with variations in standards making it difficult to compare the preparation of two students from different schools in a given year, while grade inflation raises the problem that HSGPA's relation to objective measures of achievement may change over time, and in different ways for different high schools.

Research provides ample evidence of both phenomena.

D-1. Grading standards differ across schools, and wealthy schools typically award higher grades. Betts and Grogger (2003) used a nationally representative sample of 798 U.S. public high schools to examine the relation between individual students' math grades and performance on a standardized test of math achievement.[53] Their model allowed for the possibility that some schools might have more rigorous grading standards, in the sense that the math test scores would on average be higher for a given letter grade. The authors found that the probability that all high schools grade students used the same standard was virtually zero.

Here is one way to get a sense of how big the variations in grading standards were in the data. The standard deviation provides a measure of how much variation there is in a variable. The standard deviation in students' grade 10 math scores was 9.7 points. The standard deviation across high schools in grading standards, measured in grade 10 math scores, was 3.5, which represents about one third of the overall variation in student test scores. This means that high schools with very different levels of math achievement gave similar grades.

More recent data comes from the state of North Carolina, where Gershenson (2018) takes advantage of mandatory end-of-course (EOC) exams and ACT test-taking in that state.[54] He finds that EOC exams predict ACT scores later in high school quite well, but that grades in high school classes are much less predictive. More affluent high schools tended to have higher grades than less affluent schools.

Why might public schools vary so much in what a grade of A or B or C means? The most obvious explanation is that, implicitly or not, teachers curve their grade distributions in part to reflect the academic preparation of their students. At a school with below-average student achievement, teachers are not going to restrict all grades to D and F, just as in a school with above-average achievement teachers are not going to give a grade of A to all students. By ensuring that a certain percentage of students must receive A's regardless of their objective performance, one school with low achievement can institute lower grading standards than another school with higher achievement. It is also likely that schools serving similar groups of students also differ in their grading standards for idiosyncratic reasons.

The consequence is that HSGPA is not a uniform measure of academic performance. This problem is a major concern because if a university used HSGPA alone to measure academic performance many students could be admitted or excluded based on variations in the meaning of grades across high schools.

D-2. Grade inflation is increasing and is statistically biased towards wealthy schools. Grade inflation, in the sense we are using here, refers to changes over time that lead to higher HSGPA relative to objective measures of achievement. National studies have shown increases in HSGPA have been occurring for several decades. For example, Hurwitz and Lee (2017) estimate mean HSGPA holding constant SAT scores for a national sample of students. They find that between 1998 and 2016 HSGPA steadily increased, holding constant SAT scores, by 0.15 grade point. Looking at their raw data, they find

[53] Betts, Julian R. and Jeff Grogger. 2003. "The Impact of Grading Standards on Student Achievement, Educational Attainment, and Entry-Level Earnings," *Economics of Education Review*, (22:4), August, pp. 343-352.
[54] Gershenson, Seth. 2018. "Grade Inflation in High Schools (2005-2016), The Thomas V. Fordham Institute.

that over this period mean HSGPA rose from 3.27 to 3.38. This alone does not mean grade inflation has taken place, but SAT scores for the same period and the same students actually *dropped*, from 1026 to 1002.[55]

Another concern is that changes in grading standards nationwide seem to be working in favor of more advantaged populations. In the preface to this book chapter, the book editors Buckley, Letukas, and Wildavsky succinctly state the major policy concerns raised in the Hurwitz and Lee (2017) analysis:

> *As Hurwitz and Lee explain, as average grades inflate toward a natural "ceiling" (traditionally 4.0, although some high schools have raised this) there is, by mathematical necessity, less variability in grades and therefore less predictive power in the admissions context. But even more important, the authors show conclusively that grade inflation is not occurring evenly in society. On average, they observe greater grade inflation among white, Asian, wealthy, and private school students than among less advantaged students and those in the public schools. In other words,* ***the use of high school grades in admissions is fraught with equity issues and is not a panacea for concerns with standardized tests****.* (Emphasis added by the STTF.)

This is a major finding and should give pause to those who believe that admissions tests should be abandoned in favor of using HSGPA as the main measure of academic preparation. Other studies point in the same direction. For example, the aforementioned study of North Carolina finds that grades have risen in North Carolina over time, but more so in affluent schools. We lack a similar study in California. We do know that the HSGPA of applicants to UC has risen steadily over time, which could indicate grade inflation. But on the other hand, SAT scores among UC applicants have also risen over time. **Figure 3D-1** below shows that by 2018 almost 80% of UC enrollees had a weighted HSGPA of 3.8 or higher. Test scores have risen somewhat but as shown in the figure, the trends are not nearly as clear. This raises concerns about grade inflation. Regardless of whether higher grades in California over time mean higher actual achievement, the data in **Figure 3D-1** show that "grade compression" is severely limiting the information conveyed to admissions officers by the HSGPA measure.

There is ample evidence of grade compression occurring in California high schools. **Figure 3D-1** shows this in one way. **Figure 3D-2** shows the severe grade compression that is making HSGPA less informative in another way. It depicts UC applicants' SAT scores and GPA from two unnamed California high schools. Average achievement as measured by SAT is markedly different at the two schools.

Both show very high HSGPAs with a similar distribution having most students clustered together near the top of the allowable GPA range. This example shows evidence both of variations in grading standards and the reduction in the information provided by grades due to the clustering of students near the top possible GPA values.

[55] Hurwitz, Michael and Jason Lee. 2017. "Grade Inflation and the Role of Standardized Testing," in Buckley, J., Letukas, L., & Wildavsky, B. (Eds.). *Measuring success : Testing, Grades, and the Future of College Admissions.* Johns Hopkins University Press.

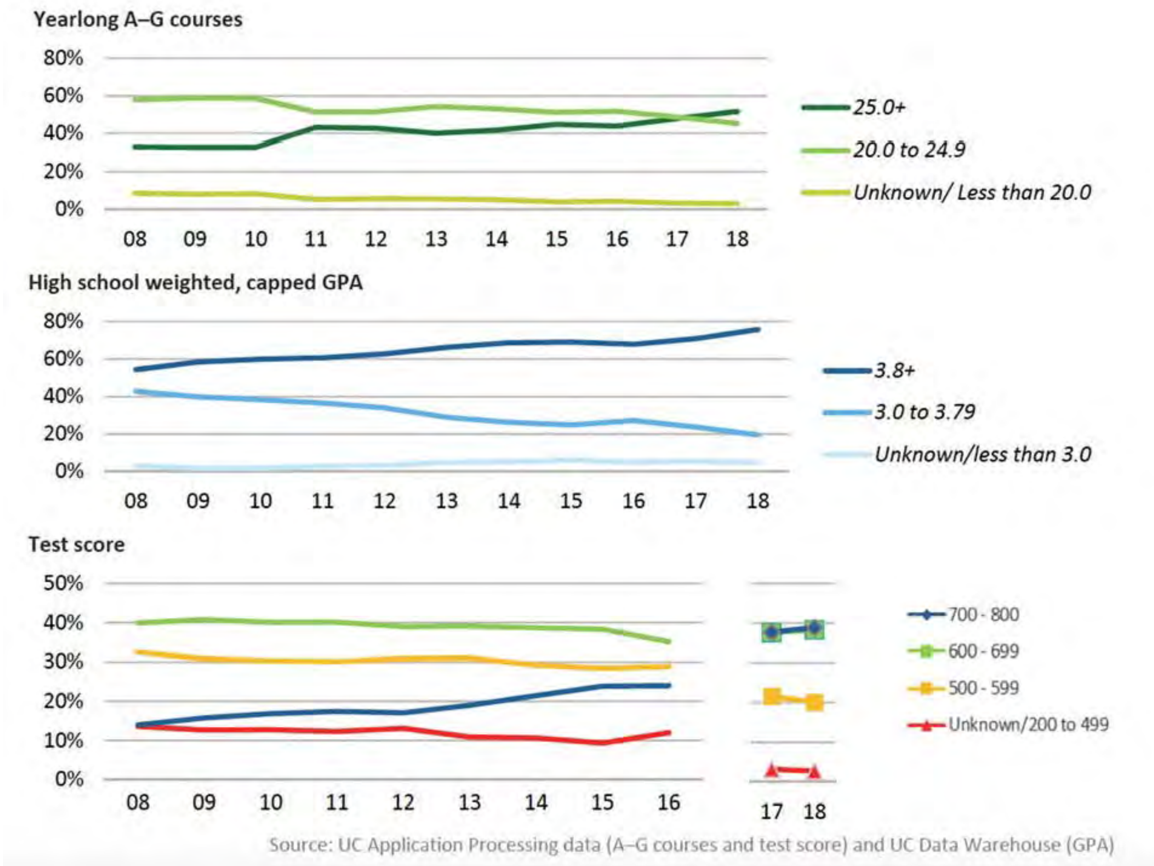


Figure 3D-1 Trends in High School Grades, A-G Courses Completed and Test Scores of Entering UC Freshmen. Source: Figure 1.3.1, UC Office of the President, 2019, "University of California Annual Accountability Report 2019." Note: Test scores show the average of the SAT math and verbal, not the sum.

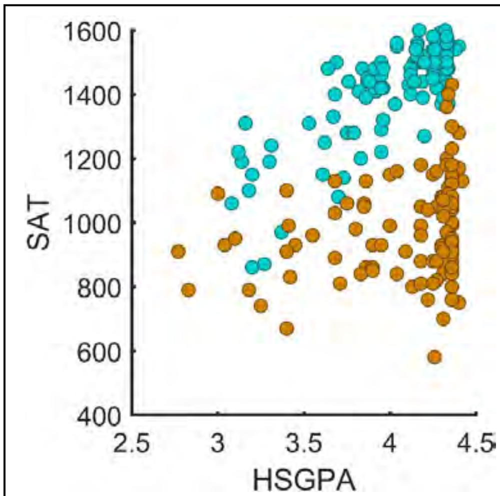


Figure 3D-2 HSGPA and SAT scores for applicants to UC from two California high schools (2018). Students from the two schools are indicated in different colors.

One counter-argument that varying grading standards across schools limits the value of GPA in admissions decisions is that the job of admission officers is in part to get to know all the high schools and their standards, and to make adjustments. This makes sense, but only to a degree. It assumes, for example, that grading standards do not change over time.

We already have national evidence that grading standards have fallen. Even more important, the relative amount of grade inflation at one school relative to another can change. Hurwitz and Lee (2017), in their nationwide study, document that high schools differed dramatically in how much grading standards relative to SAT scores changed between 1998 and 2016. They found something quite different from a gradual predictable increase in GPA inflation across all high schools. The bottom 10% of high schools, ranked by increases in grade inflation, showed mean GPA dropped considerably, from 3.4 to 3.1. The highest decile of schools, meanwhile, showed an increase in mean GPA rising from 3.1 to 3.6 in 18 years. It is not clear that admissions officers, given all of their other duties would be able to detect and remember swings in grading standards until quite a few years had passed.

A related concern is that the predictive power of HSGPA for first-year GPA at UC has declined markedly over the last few years. Recall from section A.4 earlier in this section that between 2007 and 2015 the ability of the SAT to explain first-year UC GPA remained roughly constant, while the predictive of ability of HSGPA fell roughly by a third. This too is consistent with rising GPA among high school applicants not leading to comparable increases in GPA at UC, that is, rising grade inflation and grade compression over time. Ironically, in spite of the current debate about admission tests at UC, this data suggests that we need additional measures of achievement beyond GPA more than at any other time in the last few decades.

E. Additional rationales implicit in UC's current use of tests.

E-1. Test scores are essential to increasing access through enrollment management. UC campuses issue admissions offers each spring, but it is the students and their parents who decide whether to accept them. In order to decide how many freshman admission offers to issue, admissions officers must estimate what fraction of students will accept the offer of admission ("yield rate"). Test scores strongly improve quantitative modeling of yield predictions[56] [57] [58]). That is, the additional information from test scores improves predictions of how many students will enroll. This lets the campuses offer admission to as many students as possible without overcommitting. Without SAT or ACT test scores, UC would lose much of its capability to model yields. As a consequence, UC campuses would have to choose: keep admissions levels the same, likely resulting in more frequent and more severe over-enrollment incidents in which the university is unable to meet its housing and instructional commitments[59]; or choose to reduce the number of offers made, thus reducing access to UC.

E-2. Admissions tests measure core academic skills essential for UC success. As previously discussed, the *sine qua non* for UC admission is completion of 15 A-G courses with a mean grade of B or higher. This allows students to demonstrate proficiency in a broad range of academic disciplines and skills, but as

---

[56] Weiler, W "Factors influencing the matriculation choices of high ability students." Economics of Education Review, Vol. 15, Issue 1, Feb 1996, Pgs 23-36.
[57] Nurnberg, P., Schapiro, M., & Zimmerman, D. "Students choosing colleges: Understanding the matriculation decision at a highly selective private institution." Economics of Education Review, Vol. 31, Issue 1, Feby 2012, Pgs 1-8.
[58] Hill, C. and Winston, G. "Low-income students and highly selective private colleges: Geography, searching, and recruiting." Economics of Education Review. Vol 29, Issue 4, Aug 2010, Pgs 495-503.
[59] https://www.nytimes.com/2017/07/29/us/uc-irvine-acceptance-rejected.html
https://ucsa.org/wp-content/uploads/2018/02/UCSA-Over-Enrollment-Resolution-1.pdf

we have seen, school grading standards vary so widely that a "B" in a mathematics class, for example, does not always indicate the level of preparation needed to succeed at college level academic work, and not all students have access to the same breadth and quality of A-G coursework. The core SAT and ACT tests complement A-G by measuring mastery of English language arts and mathematics, the most fundamental skills needed for University-level academic work. Students in all UC majors must be able to read and answer questions about what they read, and to approach problems quantitatively. These are core skills students must have in order to learn in college level courses, and particularly to benefit from the rigorous academic opportunities a research university such as UC provides.

E-3. Admissions tests are key to administering the "Compare Favorably" requirement for non-resident students. Regental policy, dating back the original Master Plan, requires UC to ensure that undergraduates who are admitted from other states and other countries must be more qualified, on average, than resident undergraduates. This policy is known as "Compare Favorably." UC is audited regularly to ensure compliance, as is appropriate for an institution that is fundamentally a public resource held in trust for the people of California.

Although UC admits students based on 14 comprehensive review factors, public auditors generally evaluate Compare Favorably only using the two factors that are most easily quantified: HSGPA and standardized test scores. Yet as we have seen in Section III-D, course titles, course content, and grading standards vary notably across states and even more so across countries. Thus, UC relies on standardized tests to allow for direct, objective comparisons of non-resident students' qualifications to those of California applicants, using comprehensive review. While nonresidents' GPAs are only 0.15 points higher, on average, than the GPAs of California residents who are admitted (about 0.45 of a "standard deviation," the mean student-to-student difference in HSGPA among the applicant population), nonresidents' SAT scores are about 200 points higher (roughly 1.5 times the standard deviation in SAT).[60]

E-4. Most aspects of opportunity and preparation are not equally distributed between groups; eliminating the use of standardized tests could worsen inequality. UC rightly incorporates aspects of merit that are not purely academic in the admissions process. Leadership, meaningful volunteer work, and significant extracurricular interests indicate that the student is likely to contribute to the campus community in ways beyond simply showing up for class. Challenges the applicant has faced and overcome are evidence that the applicant shows determination or grit, qualities intrinsically worth rewarding and that are also predictive of long-term success.[61] An application that compellingly presents the student's non-academic, difficult to quantify "softer" factors places the student at a significant advantage, compared to an application that minimizes or ignores those factors.

Yet constructing such an application is facilitated greatly by financial resources and a sophisticated understanding of how the admissions system works. This introduces statistical biases that are hard to overcome for low-income students or those who may not be part of a college-going family (i.e. first generation students). Many of the ways an applicant can demonstrate extracurricular interests are either

---

[60] https://senate.universityofcalifornia.edu/_files/committees/boars/Compare-Favorably-Report-to-President-July-2017.pdf

[61] E.G. Gifford, Denise D.; Briceno-Perriott, Juanita; Mianzo, Frank "Locus of Control: Academic Achievement and Retention in a Sample of University First-Year Students." Journal of College Admission, v191 p18-25 Spr 2006. Muenks, K., Wigfield, A., Yang, J. S., & O'Neal, C. R. (2017). How true is grit? Assessing its relations to high school and college students' personality characteristics, self-regulation, engagement, and achievement. Journal of Educational Psychology, 109(5), 599–620 Reviewed in Credé, M., Tynan, M. C., & Harms, P. D. (2017). Much ado about grit: A meta-analytic synthesis of the grit literature. Journal of Personality and Social Psychology, 113(3), 492–511.

costly outright (e.g. sports teams, mission trips)[62] or carry the "opportunity cost" of competing with time spent working for wages, as many low-income students must do.[63] [64] Some require large time commitments on the part of the parents as well, which can be prohibitively difficult for parents who work multiple jobs or jobs with unpredictable schedules. Many advantaged applicants have adults in their lives – well-connected parents, admissions counselors in their schools, or even hired admissions advocates – who can advise them to become involved in leadership or extracurricular activities early, advise them on the balance between breadth and depth of commitment to a single interest, and coach them on how to answer personal insight questions in ways that demonstrate depth of commitment or grit. Even if a disadvantaged student has compelling soft factors, they are unlikely to have the same access to adult assistance in crafting an application that highlights these factors.[65] To quote UC Professor emerita Rebecca Zwick, "…the less specific the demands of university gatekeepers, the more important the role of the candidates' cultural capital – in this case, knowledge about the university culture. Thus, the fuzzier the admissions criteria, the greater the disadvantage suffered by those not already steeped in the academic ethos."[66]

This concern extends to academic features of the application as well. Should a student take an additional honors class (and risk a low grade or lowering their other grades for the term), or take a non-honors class with weaker grading standards? Extra math or an extra year of language, or a course on a new topic to demonstrate academic breadth? Disadvantaged children are less likely to attend schools with ready access to college counselors who can advise them on how to make these curricular choices. Indeed, URM, first-generation, and low-income children are less likely to take these extra courses,[67] less likely to attend schools that even offer these extra courses, and more likely to attend schools in which honors enrollment is capped. Even if the courses are available, in order that the classwork and extracurricular work will be ready to display in the college application, these decisions must be made early in the high school career – earlier than many parents and children, particularly those from low-SES backgrounds or from groups that do not have high levels of college attainment, start planning for college admissions.[68] As discussed in Section II-D, not only are GPAs systematically and significantly higher at wealthy schools, but grade inflation is increasing more rapidly at wealthy schools. All these factors further advantage children of

---

[62] David B. Grusky, Timothy M. Smeeding, C. Matthew Snipp Kaisa Snellman, Jennifer M. Silva, Carl B. Frederick, Robert D. Putnam "The Engagement Gap: Social Mobility and Extracurricular Participation among American Youth" The Annals of the American Academy of Political and Social Science Volume: 657 issue: 1, page(s): 194- 207 2015

[63] https://repository.library.georgetown.edu/handle/10822/1052638

[64] https://pdfs.semanticscholar.org/eac9/ac094f886900ea2aa69d4c0b400c69837b54.pdf

[65] LW Perna, MA Titus "The relationship between parental involvement as social capital and college enrollment: An examination of racial/ethnic group differences" The Journal of Higher Education, 2005

[66] Zwick, R. (2017). "Who Gets In?" Cambridge, MA: Harvard University Press."

[67] This is evident in the patterns of course-taking across different demographic groups. UC "weights" honors and AP courses more heavily than non-honors courses in admissions; honors courses are worth one extra point. In 2018 only 8.5% of Black students and 10.6% of Hispanic/Latino applicants to UC had weighted GPAs above 4.0, compared to 22.7% of Asian-American applicants and 24.6% of Caucasian applicants. The median Black applicant to UC had taken 10 honors courses, the median Hispanic applicant had taken 12, while the median Asian and White applicants had taken 15 and 14, respectively.

[68] Hearn, J. C. (1984). The relative roles of academic, ascribed, and socioeconomic characteristics in college destinations. Sociology of Education, 57(1), 22–30. Cabrera, A.F. & La Nasa, S.M. On the Path to College: Three Critical Tasks Facing America's Disadvantaged. Research in Higher Education (2001) 42: 119

high-SES and college-going families. Indeed, in the past, UC has been sued by the ACLU on behalf of low-income and minority students for considering these factors in admissions.[69] [70]

Considered in this light, the potential for admissions tests to increase opportunity – but only when used appropriately – seems clear.[71] [72] They test only a limited set of skills that are most important for college success, which all high schools teach and which students have other opportunities to learn. Scoring rubrics and university use of the scores are both relatively transparent. A student does not need to begin polishing their CV in eighth grade to achieve a strong admissions test score. Preparation material is freely available through libraries and online services such as Khan Academy. This is not ideal: there is no question that more affluent students have greater access to material that teaches the core academic skills tested by the SAT and ACT, as well as to test-specific preparation resources, just as they have greater access to every other resource that prepares a student to construct a strong college application and to succeed in college once admitted. But the inequities in access to materials needed to score well on a test of basic writing and language skills are more easily remedied than are inequities in access to A-G coursework, school grading standards,[73] and extracurricular opportunity.

[69] Daniel v. State of California (1999) https://www.aclu.org/press-releases/class-action-lawsuit-aclu-says-ca-students-are-denied-equal-access-advanced-placement "There is no reason why such students should be denied the ability to compete equally for admission to California's elite universities," she added, "or to succeed in college degree programs, simply because their schools did not provide an adequate AP program."

[70] Castaneda v. Regents of the University of California (1999)

[71] Wooldridge A, "A True Test" In Defense of the SAT." The New Republic, 1998. "What makes this whole exercise so ironic, of course, is the fact that, until very recently, standardized testing was an instrument of upward mobility—and not a barrier to opportunity. Napoleon used tests to introduce the aristocrat-destroying principle of a "career open to the talents." British Whigs such as Macaulay and Trevelyan used tests to open Oxbridge and the civil service, hitherto playgrounds for the idle rich, to lower-class meritocrats. In the United States, testing played a vital role in forcing the WASP-dominated Ivy Leagues to admit Jews and white ethnics—and then blacks, Asians, and Hispanics. When, in 1933, Harvard's President James Bryant Conant launched a revolutionary scholarship program that awarded money solely on the basis of academic promise, the person he employed to detect that promise was Henry Chauncey, who later went on to found the Educational Testing Service. During a 1955 Supreme Court hearing to debate ways of implementing Brown v. Board of Education, Thurgood Marshall, then a lawyer for the NAACP, relied on testing to explode the segregationist argument that the average black child was so far behind the average white that it made no sense to educate them in the same schools. Why not test all children, black as well as white, he argued, and educate them according to their test results rather than by the color of their skin?"

[72] "Thurgood Marshall: His Speeches, Writings, Arguments, Opinions, and Reminiscences", ed. Mark Tushnet; Lawrence Hill Books, Chicago, IL, 2001.
"A Meeting in Atlanta", The New Yorker, Bernard Taper, March 9 1956.

[73] This was part of the original BOARS rationale for adopting standardized tests: "In addition, their use would tend to reduce the inequities resulting from differences among school marking systems, and thus would be in the interest of the applicant." (UC Academic Senate, 1967 https://oac.cdlib.org/view?docId=hb6p3006xs&brand=oac4&doc.view=entire_text)

IV. Concerns Over Use of Standardized Tests in UC Admissions

As noted in the prior section, there are large disparities in admissions along lines of race, ethnicity, and class. The STTF acknowledges interlocking systems of inequalities, racism and other forms of oppression, and the role they play in limiting educational opportunity, achievement and university admissions. In emphasizing the lack of A-G completion as a significant factor in preparedness and admission to UC, the STTF does not attempt to discount the very real environmental factors that create and contribute to race- and class-based opportunity gaps. This section offers a perspective on the data and analysis presented in Section III above, then situates admission to the University in a broader, historical context to show the importance to all stakeholders of striving for fair access.

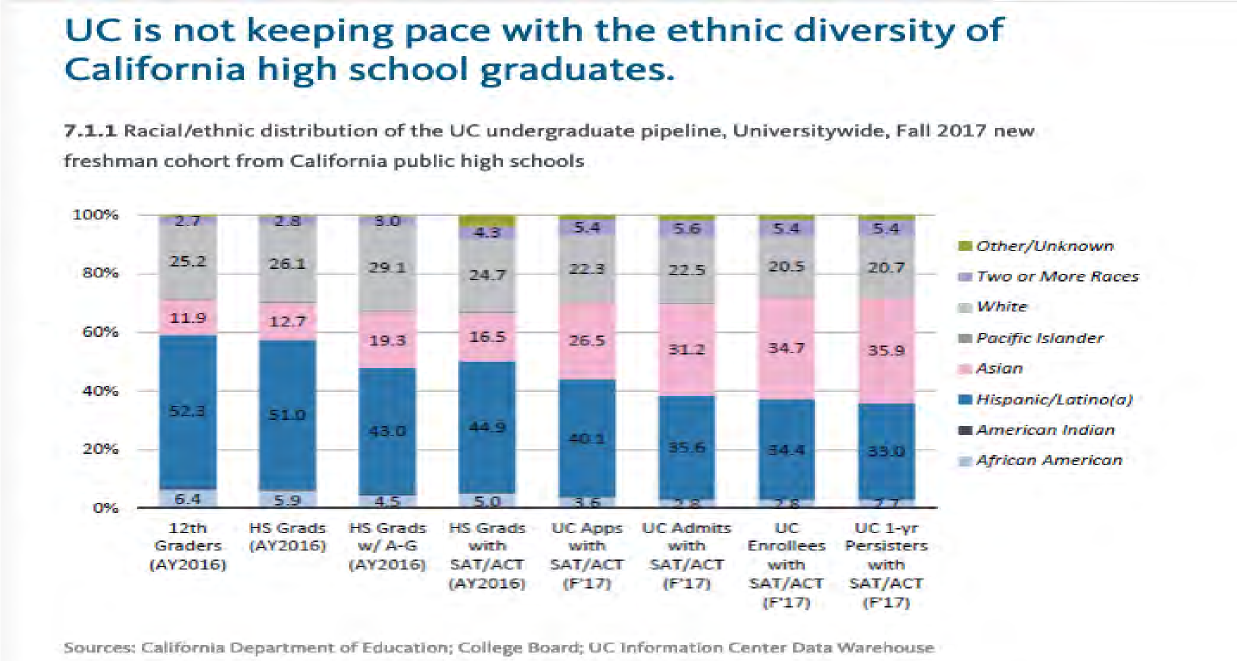
The presentation of data and analyses in the prior section cannot settle questions of policy, nor decide the best balance of practices to meet the obligations and responsibilities of UC as an institution. A sober evaluation and interpretation of the data presented in this report, however, should inform policy decisions by UC regarding admissions and could help all stakeholders, including students and parents, understand better what the implications of various policy choices regarding the continued use of standardized tests for admissions may be. Task Force members debated long and hard how to balance its empirical findings – including that the use of standardized tests for admissions are not the sole, nor even the main, culprit causing disparities in admissions to UC – with the duty of UC, along with other institutional partners, to address the deep structural and institutional causes of disparity in K-12 educational quality and access to higher education across the state. The charge to the Task Force was specific: to consider continued use of standardized tests in UC admissions and make recommendations to the Academic Council of the University. The implications of our empirical findings for what policies might improve fairness of access and our recommendations for how to reach this goal require consideration of the larger historical context when drawing conclusions from our analysis.

Standardized Test Scores and Underrepresentation

While African-American, Latino, and Native American students – who make up the Underrepresented Minority (URM) category as defined by the University of California – account for nearly 60% of high school graduates in the state, they account for about one-third of UC admitted resident freshmen. As shown in the chart below, taken from the annual Accountability Report[74] produced by the UC Office of the President's Institutional Research and Academic Planning (IRAP) team, disparities grow at each stage in the pipeline into the University:

[74] https://accountability.universityofcalifornia.edu/2019/chapters/chapter-7.html

Figure IV-1.



As large as these disparities are, they are based on an overcount of the actual percentage of URM students in the total admitted freshman class. The gap is only between California URM and California non-URM. In fact, the admit pool is about 30% larger than this, including out-of-state and international students. Taken altogether URM admit represented only 26% of the freshman students admitted in 2019. This compares to the nearly 60% of students who were URM in California schools in the 12th grade in the same year. While the data are not available to run analyses for the total pool of admits, it is critical to understand that only about 1 in 4 seats is being offered to California URM students.

Concern has grown in recent months and years that the consideration of standardized test scores in UC's admissions process might be a contributing cause of underrepresentation in particular of African-American and Latino students. The correlation between the test scores and the pattern depicted in Fig. IV-1 above is acknowledged.[75] The lower scores received by African-American and Latino students might be concerning in one way if scores were a major cause of disparate underrepresentation of these students in the pool of admitted students.

However, as made clear in Section III, this correlation should not be confused with causation. The analysis of Section III concludes that in fact, because test scores are considered in the context of comprehensive review, which in effect re-scales the scores to help mitigate between-group differences. No analyses have been conducted for admission of students whose first language is not English, but the data show large discrepancies in SAT/ACT scores for these students as well. (See **Table IV-1**).

---

[75] College Board, 2019 SAT Suite of Assessments Annual Report, https://reports.collegeboard.org/pdf/2019-total-group-sat-suite-assessments-annual-report.pdf. The same pattern is evident in ACT scores as well. See, e.g., https://www.act.org/content/dam/act/unsecured/documents/P_99_999999_N_S_N00_ACT-GCPR_National.pdf.

Table IV-1 Average SAT Scores by Race/Ethnicity and first language
California high school graduates, 2019

Race/Ethnicity	Average Score
Asian	1214
White	1168
Latino	976
Black	948
Native American	950
English not first language	1036

Source: College Board, SAT participation and performance, 2019

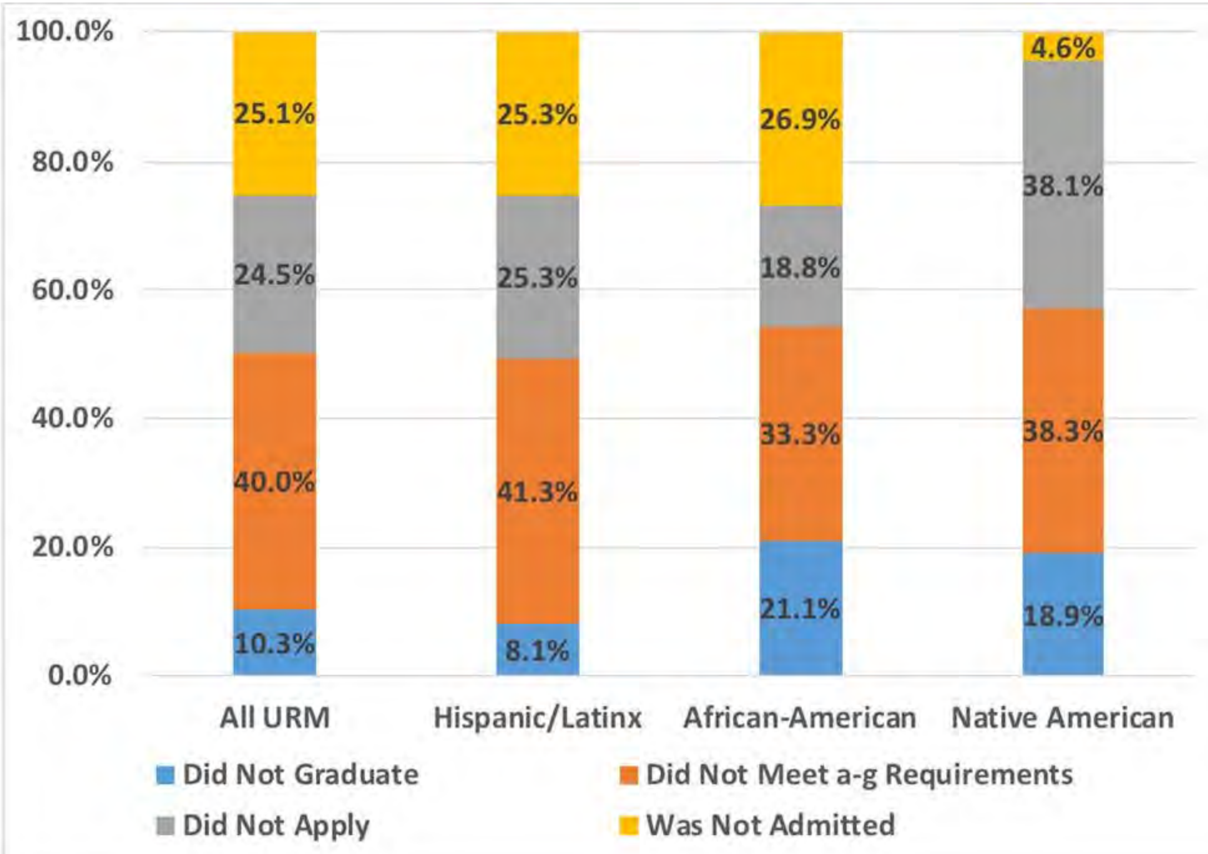
In California, about 80% of these students come from Spanish-speaking homes and so are a subset of the Latino population. Research has shown that standardized tests likely under-estimate these students' capabilities because of lack of sensitivity of the tests to issues of language difference.[76] More than 40% of students begin school in California as English learners.[77] This represents a very large percentage of URMs who are vulnerable to mis-assessment but not accounted for in the existing analyses. To address this issue, one prominent recommendation is to use performance-based rather than traditional standardized tests with students whose first language is not English. One of the STTF's recommendations is to develop a new assessment instrument that incorporates performance-type assessments. If implemented, a new UC-designed test incorporating performance-assessments would not be ready for an estimated nine years. This notional timeline changes the context in which to consider how to respond to the other Task Force recommendations in the meantime as well as establishing the appropriate basis for such policy decisions. The options we see are to maintain the status quo of comprehensive review (including the current role of standardized tests), modify UC's review criteria (as described in several of the Task Force's recommendations), or to drop standardized tests and substantially revise the implementation of the remaining 13 criteria of comprehensive review at each of the nine undergraduate campuses.

The analyses in Section III and the above chart show that standardized test scores are probably not the highest barrier to admission for URMs. See **Figure IV-2** below:

[76] Solano-Flores, G. (2008). Who is given tests in what language by whom, when, and where? The need for probabilistic views of language in the testing of English language learners. *Educational Researcher*, 37(4), 189–199; Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English-language learners. Educational Researcher, 32(2), 3–13; Jamal Abedi & Patricia Gándara (2006).Performance of English Language Learners as a subgroup in large scale assessment: Interaction of research and policy, *Educational Measurement Issues and Practice*, 25, 36-46

[77] California Department of Education, Dataquest, 2019

Figure IV-2 (same as Figure 3C-9) Factors in UC Admissions Process by Race and Ethnicity



The discussion in Section III asserts a few critical points. First, that factors other than standardized test scores disproportionately affect African-American and Latino students in the state; second, that the standardized tests are valuable because they enable more accurate prediction within disadvantaged groups, of likely student outcomes; and third, that UC considers standardized test scores in admissions decisions, in the sense that comprehensive review helps to mitigate mean test score differences among groups. In the paragraphs that follow, we identify sources of concern regarding the interpretation for policy change of each of these assertions. In the main, these concerns result from consideration of the purposes, goals, and mission of the University and the State. These concerns do not rest on the validity of the empirical findings and original analysis completed at great effort (thousands of hours of work analyzing data and interpreting findings by several Task Force members and IRAP staff, including over the holiday seasons, on weekends, while on family vacations, and after hours), but on the meaning attached to these findings and analysis and, in some cases, on the normative assumptions underlying the research questions posed.

Scores on standardized tests do play a role in the University's admissions decisions. The analysis shows that 25 – 27% of the final admission decision is due to whatever happens in the admission process. The Task Force could not ascertain the extent to which SAT/ACT scores play a role in the decision to grant or deny admission, although they do seem to be considered less than HSGPA. The discussion of Section III suggests that the consideration assigned to test scores in admissions decision-making can help to mitigate the effects of between-group differences, but the discussion cannot and does not explain whether that sense of appropriateness extends to questions of the broader context of admissions management for inclusiveness or the challenge of removing systemic inequality in public education. If students from historically excluded groups tend to receive lower scores and those scores are a factor in allocation of precious slots in the University's entering class, how should that factor be used? The data show that UC help to mitigate mean score differences between disadvantaged and advantaged groups of applicants. The data do not show that UC allocates enough slots to students from disadvantaged groups to meets its obligation to reflect the diversity of the state, nor that scores are used appropriately to implement selectivity among students within groups. This is not only an empirical question and the answer will not emerge solely from analysis of data about current and past practices. It must be the result of the weighing of values and sense of fairness.

Further, we also know that implicit bias of admissions officers and potential applicants based on having seen these scores can be a factor.[78] Once the scores are seen they cannot be unseen and, at least in some cases, they are likely to influence the application reader as well students deciding whether to apply. More important, however, is the fact that 25% represents a *large* portion of the total process when considering the very significant underrepresentation of this group. Add to this, the category of "did not apply" almost certainly includes some portion of students who declined to apply after considering their chances with a low test score. The test score may count less in admissions decisions than Task Force members initially believed, but students likely do not know this and it appears that the test (and also likely grades) loom large in their decision whether and where to apply and in their own self-assessment. In assessing these hypotheses about perceptions of what figures in UC eligibility and admissions, more data about what leads students to apply (or not) is needed. Indeed, it seems plausible to hypothesize that since admissions practices are opaque to the public and even to UC faculty, potential applicants conflate eligibility with admission criteria and assume that both are highly correlated with their individual admission probabilities. In sum, combining the 25% in the category of denial of admission to some percent of the category of "did not apply" constitutes an unknown but potentially significant portion of the final outcome that may be associated with the test, even if the weight accorded in UC admissions practices is much smaller than many people think. Policy-makers should consider various possible reforms, from providing better explanation of the actual use of test scores together with incremental adjustment of admissions criteria would be the more effective remedy, all things considered, to problems of implicit bias or self-selection not to apply.

The analysis in Section III shows that URM students are more likely to be admitted with high GPAs and lower test scores than the reverse. In fact, the conclusion is that the test scores do not matter significantly in admission outcomes. One is then compelled to ask, if the potential adverse impact of the scores is compensated for in comprehensive review, why does UC need them? The answer appears to be that they are useful for prediction purposes, to select those students who are most likely to succeed at UC. This is in tension with the statement that the tests do not play a large role in the admission decision – it is difficult to have it both ways. If test scores are weighed in decisions, they adversely affect those with lower scores, who are disproportionately African American and Latino; if they do not, then they are not essential to admissions decisions at all.

The discussion in Section III suggests that through comprehensive review, consideration of test scores *does not* adversely affect URM applicants and *does* facilitate more accurate prediction of future academic performance among students *within* each group disadvantaged prior to application to UC. This report details the issues of the use of test scores for eligibility in various contexts (Statewide, ELC, and ETR) in relation to their use for campus admissions decisions. Because test scores have predictive power among

[78] Hodson, G., Dovidio, J. F., & Gaertner, S. L. (2002). Processes in racial discrimination: Differential weighting of conflicting information. *Personality and Social Psychology Bulletin, 28*, 460–471; Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., & Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, *18*, 36-88.

students within particular disadvantaged groups (e.g. URM, first-generation, low family income, low parental education level), they are used effectively to select among students within each group who are less likely to succeed and to admit students who are more likely to succeed. Thus, the actual use of the test scores is a balancing act at the point of admissions decisions to implement the BOARS principle requiring assessment of preparedness for UC-level coursework, persistence beyond the first year, and eventual successful graduation on the one hand, and on the mandate that UC reflect the diversity of the state on the other hand. The actual use of test scores at subsequent points of evaluating matriculated students who may need additional assistance toward these goals or who may qualify for scholarships or other aid varies at the different campuses. This balancing act does involve comparing students within subgroups of applicants and selecting those with higher probability of success over others from similar educational contexts. Whether the selectivity of UC campuses should be lowered by broadening the range of test scores in the eligibility criteria in order to potentially increase access and equity would be a rebalancing that BOARS, Academic Council and UCOP should consider only in light of the obligation UC would then incur to invest appropriate resources to insure at-risk students not only have a fair chance of admission, but also a fair chance of success once admitted.

The predictive power of the standardized test scores is higher for those student groups who are under-represented at UC, as shown in **Figures 3-3, 3-4**, and **3-5** in Section III. Thus, consideration of test scores allows campuses to select those students from underrepresented groups who are more likely to earn higher grades and to graduate on time. Conversely, consideration of standardized test scores yields weaker predictions for students from higher income families, who are white or of Asian descent, or who are not first-generation. One implication is that consideration of test scores allows greater precision when selecting from URM populations. Whether this is desirable is a normative question.

If, but for consideration of the standardized test score, more students from a URM group would have been admitted in the first place, then the ability to make more precise predictions within that group may not advance the overall cause of inclusion or diversity as an admissions goal. Again, the normative meaning to be attached to the interpretation of the findings matters: is the goal of consideration of test scores to exclude certain *applicants* relative to others who also belong to historically excluded or otherwise disadvantaged groups and who otherwise might access UC? The data in Section III show that UC comprehensive review compensates for disparity in test scores *between* disadvantaged and advantaged groups of *applicants*. However, there remains a 5.5% gap between the pool of URM high school seniors and UC admits that relates to the admissions decisions themselves. As shown in section III-C5, underrepresentation of URM students by about 22% still existed in 2018 between the pool of URM students in grade 12 and the pool of California resident students admitted by UC.[79] About 75% of that gap is attributable to factors that precede admissions. So, about 25% of the remaining gap (or 0.25 x 0.22 = 0.055 or 5.5%) of admissions still reflects a URM underrepresentation due to the admissions decision itself. Section III showed, unexpectedly, that test scores likely make a surprisingly modest contribution to this gap because of the way UC takes into account students' educational context when considering test scores. Differences in GPA, and the key role that GPA plays in UC admissions decisions, appear likely to play a larger role.

Comprehensive review cannot compensate for disparities between applicants and those who do not apply, e.g. because of understandable but false beliefs about how UC campuses actually use test scores for admissions decisions, or because of lack of opportunity or advice or advantage during their K-12 experience, leading students to fail to meet eligibility criteria (e.g. in meeting A-G course requirements, minimum GPA) or to be unaware of opportunities to apply or to estimate their likelihood of admission. A decision to make an admissions policy change at UC should not be driven solely by data and empirical findings. It should be driven by data together with a thoughtful evaluation of what purposes, goals, and

[79] Fall 2018 UC Admit Rates, source (2018_freshman_profile_allcampuses.pdf)

mission are served well by the admissions process using the data it can gather. What the data of Section III do not tell is whether mitigating the quantitative gap in test scores by means of comprehensive review can help us answer normative questions regarding these purposes and goals.

5.5% of 217,036 students admitted to UC campuses for fall 2018 is 11,937 students. A system that may be inequitable for 12,000 students a year may simply be an unjust system. The question, though, is not whether the system is inequitable. We know that it is. Assuming inequity is intolerable; the question is which way of living in an intolerable system while we work to change it should we choose? Dropping the tests altogether would require adjustments to current comprehensive review practices for admissions at all nine undergraduate campuses, with unknown consequences for the extent of the impacts they are likely to have as remedy to the remaining inequity. The simulations discussed in Section III and in the recommendations in Section VI suggest that dropping the tests might make it *harder* in the short (1-3 years) to medium (3-10 years) term rather than easier for UC to make up disparities due to the other factors (such as race, HSGPA, family income, parental education level). These other factors produce 75% of the disparity. Without substantial additional resources to expand the eligibility pool and UC enrollments overall, and disregarding the substantial costs and risks of admissions errors in the short term without adequate preparation of admissions officers and processes, even a well-meaning policy change could fail.

Even in the medium term, if UC were to develop its own test as we recommend below, it would remain unknown until the new test was evaluated on whether UC is doing better or worse by the 12,000 students in the still underrepresented gap between the California high school URM pool and UC admissions pool. Evaluation of a new test would take several admissions cycles beyond test implementation at the very least. So, the question is, should UC risk continuing or worsening the 5.5% underrepresentation gap in the meantime by a qualitative change in admissions policy by dropping the tests or try to address it by incremental fine-tuning as recommended by this report? In blunt terms: yes, the current "system" is unequal and thus intolerable, but a qualitative change without more data, more simulations, and more reflection could make things worse rather than better. No change to UC policy can "fix" the interlocking systems of inequity and oppression that operate in society today. On the other hand, an inequitable system is unjust and the experience of what can be tolerated on *any* time scale depends significantly on who has to do the tolerating and how their own past experiences of oppression and injustice inform their impatience with incremental change. For some, the principle of cancelling any practice that may contribute to injustice, even if its contribution is modest, could be an overriding consideration leading to the conclusion that standardized testing should be dropped from UC admissions considerations, despite the likelihood of considerable disruption and expense in the short term (for example due to inaccurate estimates of yield rates causing significant over- or under-enrollment, reduced graduation rates, etc.). Such a step should not be undertaken, however, without weighing the potential costs to the institution (which the Task Force judges may be considerable) and even the potential harms to UC applicants, *including* those from disadvantaged groups.

Beyond the significance of denial of admission to the thousands of students affected, there is the particular meaning of racial disparities within that group of applicants. Disproportionately affected are students who are members of groups that have been excluded from all manner of opportunity - including educational opportunity - in and by this country for centuries and in and by this state for more than a century. That historical context and the reality of contemporary, ongoing societal discrimination make the normative argument for reforms, up to and including cessation of considering test scores in UC admissions, considerably stronger.

UC already pursues goals other than prediction in admissions. According to the most recent data available, roughly one in two students admitted to UC received an SAT score below 1199, and one in six received a SAT score[80] below 1000.[81] It is unlikely that admission of these students, who are more likely to earn lower grades and to fail to complete, reflects flaws in UC's process. Rather, their admission is evidence that UC pursues goals in addition to admitting the students most apparently likely to earn higher grades and to complete on time. The decision to admit such students reflects a normative judgment born out in the data analysis in Section III; showing that UC admits students with a very wide range of test scores in ways that helps to address diversity and compensate for inequality, in so far as inequalities are manifested in score differentials. As the data also show, disparities between disadvantaged and advantaged groups are not entirely mitigated. UC admissions data do not speak to the context in which high school students decide (or miss the opportunity to decide) whether to apply to UC in the first place.

Prediction of student success raises questions about what constitutes success as well as what the responsibility of a public university is to the people of the state who support it. Success here has been measured in terms of first-year grades, retention and graduation within a certain number of years. Success can be defined differently, such as the longer-term contributions to the society as described by William Bowen and Derek Bok.[82] Such outcomes are certainly less easily measured, but are arguably far more important.

In defending its test recently, ACT's chief executive, Marten Roorda, argued that the test should not be blamed for disparities in scores along lines of race and class. In a statement to the *Los Angeles Times,* Mr. Roorda said: "We don't fire the doctor or throw away the thermometer when an illness has been diagnosed. Differences in test scores expose issues that need to be fixed in our educational system."[83]

UC has addressed these differences in test scores by helping to mitigate score disparities through its comprehensive review policies and two-stage process of determining UC eligibility and campus admission. That is not to say that UC has fixed problems in the K-12 educational system producing the students who apply to UC. The Task Force grappled with these larger issues providing context for our recommendations of changes to UC admissions policy. Members of the Task Force are very concerned that debate about what the tests measure might deflect attention from the much more fundamental discussion we should be having about what is wrong with our educational system. It would be a tragedy to merely fix the instrumental problem of exactly what weight to accord standardized tests in UC admissions and foreclose thoughtful debate about higher education opportunity and disproportionate access along lines of race and class. The evidence-based description of current UC practices should not be mistaken for a normative judgment that because UC is doing well with comprehensive review to address a difficult, bad situation in California, that anyone is off the hook regarding the fundamental problem of unequal access to education in California K-12 schools and institutions of higher learning.

Prediction of likelihood of success at UC in the absence of robust support for students *who are eligible* but somewhat less prepared, raises other questions. Given the enormous inequalities in education in the state of California (discussed further below), what is the responsibility of UC to help ameliorate these

---

[80] SATRM + SATRR, or ACT equivalent

[81] Table 1.3.1, https://accountability.universityofcalifornia.edu/2019/chapters/chapter-1.html#1.3.1. The SAT was redesigned for 2017 and scores reflect the average of the math scores and the evidence-based reading and writing scores. In prior years the percentage was higher but due to change in the tests, 2017-18 data are not directly comparable to prior years.

[82] Bowen, William & Bok. Derek (1998). *The Shape of the River*. Princeton, NJ: Princeton University Press. Former presidents of Princeton and Harvard respectively collected longitudinal data on African-American students, many admitted under affirmative action, from highly selective institutions and found that these students tended to be much more civically involved than their non-black classmates.

[83] Teresa Watanabe, UC Violates Rights of Disadvantaged Students by Requiring SAT for Admission, Lawsuits Say, *Los Angeles Times*, Dec. 10, 2019, at https://www.latimes.com/california/story/2019-12-10/uc-violates-civil-rights-of-disadvantaged-students-by-requiring-sat-for-admission-lawsuit-alleges.

unjust conditions? What is the responsibility of UC to help interrupt the vast inequalities that exist – that is, to prevent the predicted poor outcome? To continue with such major discrepancies in access is to be complicit in the perpetuation of inequality. Some argue that the California Master Plan for Higher Education accounts for this problem by offering admission to the California State University system or the community colleges. Who has access to UC matters greatly in the lives of students and their communities. A college degree provides significant social mobility and a host of other benefits, including better health and marital stability.[84] The average baccalaureate degree holder in 2015 earned 67% more than the individual with only a high school diploma.[85] UC graduates also earn more than their peers who attended less selective colleges once they are in the labor market.[86] But the likelihood of actually completing a degree differs greatly by the selectivity of the college/university one attends. *Controlling for prior grades and test scores*, students who go to more selective colleges and universities complete degrees at much higher rates than those who attend less selective institutions.[87] For example, in 2018, the CSU system posted an all-time high 61% six-year graduation rate across all campuses, but there was large variation among campuses, with several campuses in the 42-48% range.[88] The average graduation rate at the typical CSU is 25-30% lower than at UC (Bleemer, 2018). Moreover, Bleemer studied the effect of ELC enrollment at four UC campuses on graduation rates for URM students in the lowest quartile of test scores versus their peers who attended less selective schools and found

> URM students and students from the bottom SAT quartile of high schools obtain even larger 30 percentage point increases in graduation rates, providing evidence against ability ”mismatch” that could diminish educational attainment (in terms of degree receipt) as a result of attending a more competitive university.[89]

In other words, these students, on average, defied the prediction that they would not do well in the competitive UC environment. There is abundant evidence that attending a more selective institution boosts graduation rates significantly above those at less selective institutions *for similarly prepared students* and this serves to interrupt historic inequities in access to degree attainment for URMs. And there are other advantages to attending a UC campus. UC graduates are more likely to attend graduate school and gain further credentials than students from less selective institutions, and as a result they are more likely to achieve leadership positions, which is of critical value to their communities.

The University's Role in Interrupting Persistent Inequality. The underlying reasons for disparities in test scores should play into UC's decision whether to consider the scores in admissions. If group score differences reflect historic injustices, then continuing to use them contributes to perpetuating the effects of those prior injustices. For example, URMs in particular whose families are more likely to have fewer resources, to K-12 schools that also have fewer resources and that consequently are less able to prepare students for the University, and then uses that lack of preparedness to justify disproportionate exclusion, that is unfair.

---

[84] Ma, Jennifer, Pender, Matea & Welch, Meredith (2016). Education Pays. New York: College Board.
[85] Ibid.
[86] Bleemer, Zachary (2018). The Effect of Selective Public Research University Enrollment: Evidence from California. Center for Studies of Higher Education, UC Berkeley. Research & Occasional Paper Series: CSHE.11.18.
[87]Alon, Signal & Tienda, Marta. (2005) Assessing the "mismatch" hypothesis: Differences in college graduation rates by institutional selectivity. *Sociology of Education*, 78(4), 294–315; William Bowen and Derek Bok (1998) The Shape of the River. Princeton: Princeton University Press.
[88] Gordon, Larry (2019). California State University graduation rates show uneven progress, some backsliding. Edsource**.** https://edsource.org/2019/california-state-university-graduation-rates-show-uneven-progress-some-backsliding/608158
[89] Bleemer, 2018
https://cshe.berkeley.edu/sites/default/files/publications/rops.cshe.11.18.bleemer.effectofselectiveenrollment.9.27.18.pdf

That is the pattern shown in data from 2013, the last year that California ranked its public schools by the now defunct Academic Performance Index ("API"). The table below shows what kind of school, from low-performing (starting on the left with category 1) to high-performing (on the right with category 10), students attend. The rankings were based on test scores and other factors; the table shows the distribution of students by race/ethnicity across API deciles. African-American, Latino and English learner students are clustered at the low end of school performance while white and Asian students are clustered at the upper end, in the schools that are higher performers.

Table IV-4 Percent Race/ethnicity/language by API Decile, California, 2013

	API	1	2	3	4	5	6	7	8	9	10
White		2	4	4	6	18	10	12	15	20	20
Asian		3	4	5	5	5	8	9	12	16	33
African-Amer		14	13	12	11	10	10	9	9	8	4
Latino		12	15	15	13	11	10	9	7	6	3
EL*		14	15	14	13	11	10	8	6	5	4

Source: Gándara[90], 2016, Author's Calculations; * refers to English Learner

The share of white and Asian students in each decile increases as the Academic Performance Index increases, with between 40% and 49% of these students clustered in the top two highest performing deciles. For Black, Latino and English learner students, between 9% and 12% of these students attend the highest performing schools; they are overwhelmingly clustered in the lowest performing schools. In short, opportunities to learn are devastatingly disparate.

The data analyzed by the Task Force in Section III show that UC's comprehensive review practices help to mitigate group score differences in its use of test scores for admissions. We would not endorse a decision to maintain status quo admissions practices on the grounds that inequity in the state's K-12 educational system is the culprit, not that UC has no responsibility to try to remedy inequities created by unfair distribution of primary and secondary educational opportunities *by the state of which UC is a part*, and no responsibility to the students who disproportionately attend under-resourced schools. It would be cold comfort to tell students in groups that attend lower-performing schools and who disproportionately receive lower scores on standardized tests that once K-12 educational equity is achieved, they can expect their representation on UC campuses to increase. That is an argument not to *increase* the role of standardized tests in UC admissions. It is not an argument that, since current UC admissions practice helps to mitigate test score disparities among disadvantaged groups, we have a reason to be complacent, have no duty to reform admissions practices to better reflect statewide diversity and reverse historical injustices, nor have a duty to improve the quality of a UC education and the success of all our students. Rather, we recognize that the University has an immediate obligation to admit and support promising

[90] Gándara, Patricia (2016). The students we share, *Mexican Studies/Estudios Mexicanos,* Oakland: University of California Press, Pp.357-378, p. 370.

students *who are eligible,* but are affected by historic and adverse discrimination, rather than to allow its effects to justify continued exclusion. This is one of the reasons that the University already takes into account the "local context" in admissions, illustrative of UC's commitment to try to enroll a diverse class. However, as this report documents, consistent and persistent disparities (even if the exact magnitudes remain unclear) are evidence that UC should do more to achieve this goal. The goal of admitting students more likely to perform well at UC, as predicted[91] by test scores, should not be singled out as the standard and should not prevent the University from fulfilling its mandate to serve Californians fairly should be judged.

Disparities in test scores should be evaluated in social and historical context.

- Racial inequality runs deep in California. Enormous differences in income and wealth along lines of race mean that parents have very different resources to invest in their children to prepare them for college. Among these material resources are private schools, tutoring, extracurricular classes, sports teams, educational materials (including test prep), and educational experiences like travel. And the differences in these expenditures have been growing over time, widening the expenditure and opportunity gaps between URMs and others.[92] For example, California has the seventh most inequitable economy of the 50 states.[93] The typical white household has $9.72 in wealth for every $1 that a typical African-American family has and $8.26 in wealth for every $1 that a typical Latino family has.[94] African-American and Hispanic households of any given income level typically live in neighborhoods that are substantially poorer than do whites and Asians of the same income level. In fact, lower income white families are more likely than middle-income families of color to live in a middle-class community.[95] This is the result of vast differences in wealth as well as income and the lingering effects of state-sanctioned housing discrimination.[96] This means that African-American and Latino families are locked into housing in neighborhoods that often offer only under-resourced, low performing schools, creating an unending cycle of lack of access to a strong college preparatory curriculum.

- Educational opportunity in the state of California is not equally or equitably distributed across the high school student population. Students who are Latino, who are African-American, whose parents have lesser social and financial capital, disproportionately attend schools that are under-resourced and, in addition and more specifically, are less likely to offer prerequisite coursework (i.e., A-G courses) necessary to gain admission to the University.[97] These students also have extremely limited access to a college counselor who can inform them of their options and enroll them in the available A-G courses. California is 49th in the nation with respect to student-counselor ratios and assigns three

[91] BOARS, Report to the Regents on Admissions Tests and UC Principles for Testing, p. 18. Available at https://senate.universityofcalifornia.edu/_files/reports/hp2mgy_boars-testing_010609.pdf
[92] Sean Reardon (2011) The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations in Gregory Duncan & Richard Murnane (eds) Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances. New York: Russell Sage. Pp. 91-116.
[93] Estelle Sommeiller and Mark Price (2018), *The New Gilded Age: Income Inequality in the U.S. by State, Metropolitan Area, and County*. Washington DC: Economic Policy Institute.
[94] Esi Hutchful (2018). The Racial Wealth Gap: What California Can Do About a Long-Standing Obstacle to Shared Prosperity. Sacramento: California Budget and Policy Center.
[95] Reardon, S.F., Townsend, J., & Fox, L. (2017). A continuous measure of the joint distribution of race and income among neighborhoods. *The Russell Sage Foundation Journal of the Social Sciences, 3, 34-62.*
[96] Richard Rothstein (2017). The Color of Law. New York: Liveright Publishing.
[97] BOARS, Impact of the New Freshman Eligibility Policy at the University of California: A Report to the Regents of the University of California (Nov. 2013), p. 8 (tbl. 2), available at https://senate.universityofcalifornia.edu/_files/committees/boars/Nov52013BOARSReporttoRegents-Final.pdf.

times the recommended number of students to each counselor.[98] Moreover, most counselors have neither the time nor the training to provide college counseling specifically. Affluent families can offset this lack of counseling resources by paying for private counseling. The very unequal assignment to low performing schools is also associated with high teacher and administrator turnover, disproportionate numbers of inexperienced and unqualified teachers, higher absentee rates for educators, and a weaker curriculum.[99] Peers who are educationally ambitious *and headed for college* can be a significant support for each other,[100] but when students are segregated with only other low-income and marginalized peers, this constitutes a weak resource.

- While the current versions of the SAT and ACT tests are purported to assess "academic preparation" rather than "aptitude" (as they once claimed to do), scoring well on these tests is related both to exposure to academic material and to cognitive strategies employed to answer the questions appropriately (which is why "test prep" can increase scores without necessarily increasing academic knowledge). If students have limited exposure in primary and secondary classes to the material tested, as a result of a weak curriculum and/or receiving instruction in low-performing schools, and/or if students have limited experience with the cognitive strategies that the tests are designed to elicit, they are likely to perform less well on the tests. The tests evaluate a narrow slice of human achievement and knowledge that to some extent is class- and culture-bound.[101] Eminent psychologists have demonstrated that tests that are designed to elicit quite different skills from those on the market today can produce novel and accurate ways to solve problems.[102] Students with strong academic preparation and a proclivity for abstract reasoning have an advantage on these tests, but given the same primary and/or secondary school inputs and opportunities or targeted postsecondary support, others may perform equally well in UC classes. Fundamentally, the standardized tests are measuring a student's opportunities to learn, over which the student has little control.

- Stereotype threat is a well-established phenomenon among URMs and women. When black (and other URM) students believe their intellectual skills are being tested, this can activate a concern about validating a negative stereotype about their lesser ability and cause a level of stress that depresses their scores. Claude Steele and Joshua Aronson[103] tested this theory initially with highly talented African-American and white students through a series of studies in which students were told that the test was designed to accurately assess their verbal aptitude. Black students in the verbal aptitude condition responded as predicted, their scores were depressed compared with white students. In the non-threatening condition in which students were told the test was not testing aptitude, black students performed at expected levels. Steele and Aronson, and others, have replicated the studies many times with reasonably consistent results for women in mathematics and for Latinos, among others.[104] These findings strongly suggest that the high stakes of the SAT/ACT by themselves can have the effect of

---

[98] American School Counselor Association (n.d.) State-by-State Student Counselor Ratios. Ten year Trends: 2004-5 to 2014-15. https://www.schoolcounselor.org/asca/media/asca/Publications/ratioreport.pdf

[99] Donald Boyd, Hamilton Lankford, Susanna Loeb and James Wyckoff (2005). Explaining the short careers of high-achieving teachers in schools with low-performing students, *American Economic Review*, 95 (2), 166-171; Charles Clotfelter, Helen Ladd & Jacob Vigdor (2009). Are Teacher Absences Worth Worrying about in the US? Washington DC: Urban Institute; Desiree Carver-Thomas and Linda Darling-Hammond (2017). Teacher turnover: Why it matters and what we can do about it. Palo Alto, CA: Learning Policy Institute

[100] Margaret Gibson, Patricia Gándara and Jill Koyama (2004). School Connections. US Mexican Youth, Peers, and School Achievement. New York: Teachers College Press.

[101] See, for example, Sternberg, Robert (1998). Ability testing, instruction, and assessment of achievement: Breaking the vicious circle, NASSP Bulletin, 82, 4-10.

[102] Ibid.

[103] Claude Steele & Joshua Aronson (1995), Stereotype threat and the intellectual test performance of African-Americans, *Journal of Personality and Social Psychology*, 69(5):797-811.

[104] Aronson, Joshua (2004). The threat of stereotype, *Educational Leadership*, 62 (3), 14-19.

depressing scores even for highly talented URM students who otherwise perform at high levels and are prepared to succeed in college.

- Students from more affluent homes can take advantage of an entire industry dedicated to increasing their test scores. In addition to test preparation materials and classes that can cost on average about $1,000 each, families can purchase coaches or counselors who for several thousand more dollars will work with students with the preparation materials and provide tailor-made assessments and advice. These coaches promise significant gains on the tests and they produce them.[105] For example, Princeton Review guarantees a minimum 200-point gain in scores or your money back.[106] The fact that the evidence is mixed regarding whether test prep can boost performance this much does not mean that test prep is not effective in changing the psychology of the test-taker and may lead some to choose not to take the test in the first place. In extreme cases, some affluent parents have paid "coaches" to falsify test scores in order to gain admission to a selective college.[107] Such incidents not only are corrupt, they undermine students' belief in the fairness of the admissions process.

Students with lower test scores can often rise to the challenge. But many have not had the opportunity to demonstrate what they are capable of and could achieve with the benefit of academic support that UC campuses already offer.[108] The most direct way to interrupt this inequality is through increased and more equitably distributed educational opportunity, including at the university level.

Under the Constitution of the State of California, the University "constitute[s] a public trust, to be administered by the existing corporation known as 'The Regents of the University of California,' with full powers of organization and government."[109] The Regents, in turn, have made the following statement in their Policy 4400 on the mission of the University, worth quoting extensively:

> The diversity of the people of California has been the source of innovative ideas and creative accomplishments throughout the state's history into the present. Diversity – a defining feature of California's past, present, and future – refers to the variety of personal experiences, values, and worldviews that arise from differences of culture and circumstance. Such differences include race, ethnicity, gender, age, religion, language, abilities/disabilities, sexual orientation, gender identity, socioeconomic status, and geographic region, and more.
>
> Because the core mission of the University of California is to serve the interests of the State of California, it must seek to achieve diversity among its student bodies and among its employees. The State of California has a compelling interest in making sure that people from all backgrounds perceive that access to the University is possible for talented students, staff, and faculty from all groups. The knowledge that the University of California is open to qualified students from all groups, and thus serves all parts of the community equitably, helps sustain the social fabric of the State.
>
> Diversity should also be integral to the University's achievement of excellence. Diversity can enhance the ability of the University to accomplish its academic mission. Diversity aims to

---

[105] Park,J.J. & Becks, A.H. (2015). Who benefits from SAT prep? An examination of high school context and race/ethnicity. The Review of Higher Education, 39(1), 1-23.
[106] Astin, Alexander (2016). Are you Smart Enough? Sterling, VA: Stylus Publishing.
[107] See for example, Hussain, Suhauna (2019). College admissions scandal hits UC Berkeley: fake SAT score allegedly submitted. *Los Angeles Times,* March 19. https://www.latimes.com/local/lanow/la-me-uc-berkeley-crew-college-admissions-scandal-20190319-story.html
[108] For example, campuses offer academic support targeted to first generation students. See, e.g., University of California, Irvine, Student Success Initiatives, http://ssi.uci.edu/.
[109] Cal. Const. Art. IX, §9.

> broaden and deepen both the educational experience and the scholarly environment, as students and faculty learn to interact effectively with each other, preparing them to participate in an increasingly complex and pluralistic society. Ideas, and practices based on those ideas, can be made richer by the process of being born and nurtured in a diverse community. The pluralistic university can model a process of proposing and testing ideas through respectful, civil communication. Educational excellence that truly incorporates diversity thus can promote mutual respect and make possible the full, effective use of the talents and abilities of all to foster innovation and train future leadership.
>
> Therefore, the University of California renews its commitment to the full realization of its historic promise to recognize and nurture merit, talent, and achievement by supporting diversity and equal opportunity in its education, services, and administration, as well as research and creative activity. The University particularly acknowledges the acute need to remove barriers to the recruitment, retention, and advancement of talented students, faculty, and staff from historically excluded populations who are currently underrepresented.[110]

The Regents' statement of mission is consistent with arguments by scholars who have written on barriers to access and who have emphasized the affirmative role of higher education not merely to identify but to foster students' abilities.

The Demographics of the State of California Have Changed. The California Master Plan for Higher Education was developed in a state very different from today's California. When the Master Plan was adopted in 1960, California's schools ranked with the best funded in the nation. Economic inequality was at a low point and the state was 92% white,[111] with a small and uncounted Hispanic population bundled in with "whites," estimated to be about 10%.[112] Socioeconomic, racial, and ethnic diversity did not figure in the plan. The Hispanic or Latino population was so small and concentrated in specific areas that the US Census did not consider it important to count for the states. People of Asian descent were similarly uncounted in 1960 but estimated to be about 2% (based on the "non-white" population listed by the US Census Bureau). African-Americans were 6% of the state's population. At the time, the notion of a three-tiered system with access for everyone coming from strong K-12 schools was hailed as an enviable model for the nation, with the top 12.5% of a largely homogenous population admitted to UC where students could prepare for graduate and professional schools. The top 33% were eligible for admission to the CSU, where students could pursue teaching or another applied degree. The community colleges were open access with the right to transfer to the four-year colleges and universities if all prerequisites were satisfactorily completed. The model was a cost effective use of state resources and in its early years the transfer rates from the community colleges reached nearly 50%.[113]

Virtually everything has changed since 1960 *except* the basic features of the Master Plan. In 2018, Latinos comprised nearly 40% of the general population, Asians had grown to 15.3% and African-Americans remained at about 6.5% of the population. The white population stood at 36.8% of the population.[114] Our K-12 public schools are even more diverse, with Latinos at 55%, Asian/Pacific Islanders at 10%, African-

[110] Regents Policy 4400, Regents of the University of California, https://regents.universityofcalifornia.edu/governance/policies/4400.html.

[111] Callan, Patrick (2012). The Perils of Success: Clark Kerr and the Master Plan for Higher Education. In S. Rothblatt (ed) Clark Kerr's World of Higher Education Reaches the 21st Century. NY Springer Table 3.6 p.74

[112] 1970 is the first year in which the US Census reports "Hispanics" in California. In 1970, the Census reported the Hispanic population to be 12%. Allowing for some growth in this population, and the fact that 1960 marked an historic low point in immigration, 10% is a reasonable upper bound estimate.

[113] Douglass, John A. (2004). The Conditions for Admission: Access, Equity and the Social Contract of Public Universities. Palo Alto: Stanford University Press.

[114] US Census Quick Facts, California. https://www.census.gov/quickfacts/CA

Americans at 5.4%, and white students at 23%.[115] The K-12 schools have gone "from first to worst"[116] among the states, economic inequality has skyrocketed, our schools are among the most segregated by race and income in the nation,[117] and the one strategy UC had used at least partially successfully to try to equalize access across racial lines—affirmative action—was outlawed in 1996. In 2017, 59% of California high school graduates were URMs, but across the UC system just 31% of enrollees were URMs. Projections are that without a change in current policy the discrepancies will grow.[118] Latino students disproportionately attend the community colleges where relatively few are successful in transferring to four-year institutions. In fact, UC's entering freshman class is more diverse than its entering transfer class. Getting into UC has become an intense race for limited and coveted slots.

Some would argue that part of the fairness solution lies in expanding opportunity. Rather than strictly rationing the benefits of a UC education, the goal should be expanding access.[119] The difference in the debate this time is that the majority of the state's student population is URM and many of their advocates both from within the University as well as outside have voiced strong objections to the use of the tests.[120]

They are right to voice strong concern. Some members of the Task Force began our investigation with an understanding shared with many members of the public that the large disparities in test scores among various groups were directly reflected in distributions of students admitted because the correlations between test scores and demographic groupings are fairly strong. We learned through close, careful, painstaking study of the actual results of UC's comprehensive review, even as it is implemented in different ways across the nine campuses, helps to mitigate disparities in test scores.

The above examination of the mission of the University in the context of a changing California, and the argument that it is incumbent upon us to do more to interrupt historical inequality, is not meant to take anything away from efforts already undertaken. The University already takes into account additional student characteristics in order to evaluate test scores in context, thereby helping to mitigate some of the potential adverse impact of scores on applicants, and UC does more than many institutions to promote access and socioeconomic mobility.[121] But we can and should do more.

[115] Fingertip Facts on Education in California – CalEdFacts, California Department of Education. https://www.cde.ca.gov/ds/sd/cb/ceffingertipfacts.asp

[116] The Merrow Report: First to Worst, California's schools: America's Future. PBS, Learning Matters, Inc, 2004. This widely distributed video features California educators and policymakers decrying the fact that the state has fallen to the bottom decile of states with respect to funding and points out the totally inadequate facilities that many low-income students attend with broken toilets and windows, no heat or air conditioning, lack of playgrounds. It also contrasts these 'worst" schools with those in affluent neighborhoods where parents contribute to the school's resources, creating huge inequities among "public" schools.

[117] Frankenberg, Erica, Ee, Jongyeon, Ayscue, Jennifer & Orfield, Gary (2019). Brown at 65: Damaging Our Common Future. Los Angeles: Civil Rights Project/Proyecto Derechos Civiles, UCLA.

[118] Carnavale, Anthony, Van Der Werf, Martin, Quinn, Michael, Strohl, Jeff & Rapnikov, Dmitri (2018). Our Separate and Unequal Public Colleges. Washington DC: Georgetown Center on Education and the Workforce.

[119] Ibid.

[120] See, e.g., Watanabe, Teresa, *Drop the SAT and ACT as a requirement for admission, top UC officials say*, Los Angeles Times, Nov. 23, 2019, at https://www.latimes.com/california/story/2019-11-23/uc-officials-recommend-dropping-sat-admission-requirement.

[121] See, e.g., David Leonhardt, The Assault on Colleges and the American Dream, N.Y. Times, May 28, 2017, at SR 6 (and accompanying graphic, at https://www.nytimes.com/interactive/2017/05/25/sunday-review/opinion-pell-table.html, finding that "based on a combination of the number of lower-and middle-income students that a college enrolls and the price it charges these students," the top five institutions nationwide are UC campuses).

V. Questions & Answers

To this point our report has explained how SAT and ACT scores are used at the University of California ("UC" or "the University"), why test scores play an important role in admissions, and how UC's use of tests differs substantially from what many in the public, and indeed at UC, might have believed. The prior section summarized the main concerns about the use of tests.

This section has two goals. First, it will bring the preceding sections together by recapitulating the main concerns about testing expressed in the prior section along with other concerns that Task Force members have heard, and then interrogating the earlier sections of the report to answer these questions. The second goal is to set the stage for the final section of the report, which provides details on the Task Force's main recommendations. Many of these recommendations originated in part from the members' deliberation of the questions set out in this section. By posing these questions here in a direct and sometimes even blunt way, we hope to shed light on how UC admissions uses student information, to illustrate the misconceptions that may have existed in our own and others' beliefs about how UC uses test scores, and to explain the genesis of the Task Force's recommendations.

Q: I have read that the group of students admitted to UC are very different from the students in California high schools, based on race/ethnicity and socioeconomic measures. Is this true? Isn't this evidence that UC admissions procedures somehow unjustifiably reduce admissions rates for some groups?

It is true that the racial mix of students admitted into the freshman class at UC is quite different from the racial mix of high school seniors from the same year. Consider the groups that UC collectively refers to as Underrepresented Minorities (URM): Latino, African-American students and Native American students. In 2017-18, these groups constituted 59.1% of high school seniors in California, but only 37.0% of UC admits among California residents.

The second question is whether this 22.1 percentage point gap arises due to UC admissions decisions. The Task Force considered this in detail. **Figure 3C-7** shows that about one-quarter of the gap is due to the admissions decisions of the UC campuses, but the remaining three-quarters of the gap relates to outcomes that occur before UC admissions officers read files. The single biggest factor is that relatively few of the students in the three underrepresented racial/ethnic groups complete the A-G coursework that both UC and CSU require for students to become eligible to apply. The other key factors include differences in the rates of graduation from high school and in applying to UC.

The fact that admission decisions explain only about one-quarter of the disparity in racial mix of high school seniors and admitted UC freshmen is important, but a reasonable person could wonder whether this contribution, although *relatively* small, might indeed indicate bias of some sort against applicants from some groups relative to others in admissions itself. We will look into this in responses to some of the questions below.

Q: Don't SAT score averages look a lot lower for some groups than others?

Yes, on average, SAT scores are lower for some groups than others, with lower scores for less affluent groups and for underrepresented minorities. As shown in **Table 3C-4**, in 2019, test score gaps were as wide as 234 points between Latino and Asian UC applicants, and 216 points between students who would be the first in their family to graduate from college and those who are not "first-generation".

Q: But don't these large differences in SAT scores across demographic groups cause large variations in UC admissions rates? Isn't that unfair given inequalities in school quality and other inequalities in California society?

This is a reasonable pair of questions. It turns out, though, that differences in average test scores between groups of applicants do not cause highly different admit rates. Campus admissions officers appear to re-norm, i.e., compensate for, SAT scores by comparing a student's test scores not to the average SAT scores across all applicants, but rather to the SAT scores of other students who come from similarly disadvantaged groups. UC does not have a specific cutoff for SAT scores below which admission is not possible. If it did, then all admits would have very high SAT scores, and average score differences between disadvantaged groups and advantaged groups, although large in the applicant pool, would be greatly reduced in the pool of students who were admitted. But this is not what happens. The rightmost column in **Table 3C-4** shows that the test score gaps between racial/ethnic groups, between income groups, and between first- and non-first-generation students among applicants remain virtually as large among admitted students, which means that the test score gaps that are correlated with race, first-generation status, family income, and parental education level are almost fully compensated for in admissions decisions; otherwise the test score gaps in the applicant pool would be different than the gaps in the admitted pool.

Figures 3C-4 through **3C-6** show that for any SAT score, students from disadvantaged groups have a higher probability of being admitted than students from advantaged groups. In other words, how UC uses the test scores appears to be to help identify students within each socioeconomic group who are most likely to succeed.

To re-state this more simply, large inter-group differences in SAT scores do not translate into major differences across student groups in admission rates at UC. This is probably the most important finding the Task Force has made in its data analyses.

The second part of the above question is whether UC is ignoring differences in outcomes that arise in part because of differences in the quality of education available to affluent students and less affluent students. Key here is our finding that UC seems to effectively re-norm the SAT to take into account the context in which each group of students was educated. What this means is that UC takes careful note of inequality in school resources and adjusts its admission decisions to take the local context into account. It also reflects UC's commitment to reflecting the diversity of the state in its admissions decisions by compensating for test score gaps between groups of applicants but using test score differences among individuals within groups of admits to help identify UC students potentially at risk of not succeeding. UC must find a balance between its mandate and value placed on diversity and its obligation to help its at-risk students. The predictive role of the test is used to evaluate underprepared enrolled students and also to select among students within each group (URM, family income, and so forth) those students who already demonstrate academic achievement on a variety of measures, including high school GPA and test scores. It is possible that a small fraction of this narrow role of test scores in UC admissions and evaluation of potential at-risk students is still biased. For that reason, the Task Force makes recommendations to collect more "granular" data about individual questions on the tests to further study potential racial (and other forms of) bias in the tests.

Q: Earlier you said that about one-quarter of the racial gap between UC freshman admits and California high school seniors results from the admissions decisions themselves. You also said that admissions offices effectively re-norm SAT scores, comparing less affluent students to each other and students from high schools that have similar demographic profiles to each other. What explains these gaps, then? Are SAT scores the major factor causing the differences in admission rates between racial/ethnic groups? If not, what other factors matter in the admissions process?

Given the Task Force's findings that the SAT scores are evaluated in a way that effectively renormalizes scores to take disadvantage into account, SAT scores do not appear to play a big role in differences in admission rates **between** disadvantaged and advantaged groups. Other admission factors play a role in reducing the share of disadvantaged groups that are selected in the admissions process (as well as in reducing the share of disadvantaged groups in the admissions pool). A prime example is high school GPA. The report shows that variations in GPA matter more than variations in SAT in explaining admission rates. For example, **Tables 3B1-4** show that, both for UC overall and for individual campuses, changes in high school grade point average (HSGPA) have large effects on admission probability while equivalently-sized changes in SAT have smaller effects on admission probability. **Figure 3C-5** shows that UC admission probabilities rise steeply with HSGPA, and that the relation between admission and GPA is quite similar across racial/ethnic, educational, and socioeconomic status (SES) groups. **Figure 3C-10** shows that disadvantaged groups tend to have significantly lower GPA in high school and that UC admissions does not "renormalize" high school grades nearly as much as it renormalizes SAT scores. Taken together, these imply that GPA likely plays an important role in differences in admission rates across socioeconomic groups.

The comprehensive review process can take into account 14 factors, of which test scores and GPA are only two. The Task Force could not evaluate the role, if any, played by the remaining 12 factors, but it seems quite possible that more advantaged groups of students might have had opportunities that would allow them to be ranked highly on some of the 12 other factors as well. For example, some of these factors relate to how many advanced courses a student takes, and the difficulty of the grade 12 course schedule. Consider also item 11 on the list of 14 factors, which presents a long list of examples of special talents, achievements, and awards, including athletic activities, and signs of leadership that UC considers. Here again it seems plausible that more affluent students have more opportunity to shine in this regard.

Q: Wait a minute, earlier you said that three-quarters of the gap in the racial mix of UC freshman admits and high schools reflects things that happen prior to the admissions decision. Isn't this just an attempt to shift responsibility for any disparity among groups admitted to California's K-12 system?

Section III of our report is simply reporting the data. No blame or shifting of responsibility is being assigned or should be inferred. The report is not saying: "somebody else should fix K-12 education and then UC's current admissions practices will work as we hope they will." That said, we know that despite decades of reform, school resources and school quality vary across California's high schools, as pointed out eloquently in the prior section of this report. It seems likely that these variations have at least some influence on variations across students in their academic progress. What children experience outside of school matters at least as much as what happens within school. The pre-school years are especially important. More affluent parents have access to private pre-schools, and California does not yet have universal pre-school for all needy children. Not all children leave high school with the same academic preparation. We have to know what problem we are trying to solve.

UC does not believe that it can, by its own admission policies, singlehandedly rectify the large differences in academic preparation among high schoolers that relate to family income and race/ethnicity. But given that UC has always faced two mandates – to strive to admit prepared students who at the same time reflect the state's diversity – UC does strive to do its part to amend these inequalities. The way in which UC assesses SAT scores within the context of the student and his or her high school, the Blue and Gold Scholarship program to eliminate tuition for families with limited means, and the many support programs for undergraduates all speak to the diversity aspect of the UC mission. Most of the recommendations the STTF has made derive from the desire to see UC improve its existing efforts.

Q: But let's push further on this. You say that some students arrive at UC ill prepared and often do not graduate. What does UC do to support these students? Couldn't UC solve its own problem of having to rely on test scores and grades to select students who will likely perform well at UC? For instance, can't UC better support admitted students who are predicted to struggle?

This is an excellent point. The statistics we have presented do show ample evidence that UC students with lower test scores and grades have a lower probability of having high grades, persisting past freshman year, and graduating. But these statistics are backward looking, and do not consider the possibility that UC could improve its supports for students, especially in their freshman year.

UC has a wide variety of programs to support students in their freshman and later years. For example, UC campuses provide writing programs for freshmen who have yet to demonstrate adequate mastery of reading and writing. The campuses also invest a lot in orientation programs including some "live-in" programs to help students, especially those who might need additional supports, with a chance to make friends and develop a feel for living on campus well before classes start in the fall.

UC has not, to the best of our knowledge, conducted a rigorous evaluation of these many support programs. The Task Force is thus recommending that UC first catalog these programs, and then perform a rigorous multi-campus evaluation of the impact of these programs, with the goal of expanding programs that work and either fine-tuning or if warranted, shrinking programs that do not prove to be as helpful as hoped. We also propose a robust series of innovative new support programs that in each case should be carefully evaluated. Some of the Task Force's recommendations would likely result in enrollment increases across the system, which would put even more pressure on UC to evaluate the impact of these programs and invest resources to expand the ones that work.

Q: I have heard that SAT scores do not really measure differences in achievement. They are really just measures of family income, right? How could UC ethically use SAT and ACT scores if these tests simply measure income?

There is a lot to unpack here! It is true that SAT and ACT scores are positively correlated with family income, as are many other measures of student achievement. The appendix to this report includes a report from UCOP that documents that the correlation between family income and SAT scores has grown over the last decade, and is now about 0.4, which is a meaningful amount of correlation.

The correlation is problematic because it could be a symptom of a causal relationship that UC might be able to do something about. The positive correlation between scores and family income, however, does not show that family income level causes the scores to be what they are, nor that the test scores cause family income level to be what it is, for any individual student. Also, the pattern of correlation among groups does not show causation for a whole group of students, because both family income and scores could be the result of something else causing both, like race-based oppression that leads some families to have low income and students in those families to have low test scores. For example, students suffering racial oppression might be forced to attend lower quality schools due to economic segregation and that would tend to cause both lower test scores and also reflect lower family income at the same time, but the family income would not be a direct cause of low test scores. What the data show, using other statistical analyses than correlation studies, is that family income may be one contributing cause among several. Then the question becomes: how important a factor is family income compared to the other contributing causes in explaining any disparities in admission rates?

Your second question was, whether it was ethical for UC to use tests that are positively correlated with family income. This is a good question, but what really matters is *how* UC uses these test scores. If UC campuses simply set a single SAT cutoff below which no student could possibly be admitted, without taking into account the role that family income can play, this indeed would be questionable. But what the Task Force's analysis shows differs markedly from this scenario. What we found is that UC admissions practices essentially take into account the entire gap in average SAT scores between the most advantaged and least advantaged groups. This statement is true whether we group students by race/ethnicity, family income, or first-generation status.

So what is the use of SAT scores? Our analysis shows that even within specific socioeconomic groups SAT scores are quite predictive of student performance once at UC. See for example **Figure 3A-3**, which divides students into four equally sized groups based on family income. Even after controlling for family income in this way, and additionally taking into account students' high school GPA, the SAT score is highly predictive of a variety of UC outcomes ranging from first year GPA to the probability of graduating within seven years. Admissions decisions take into account differences in SAT scores *while carefully taking into account the student's socioeconomic context.* This means that any of the other 12 factors in comprehensive review besides test score and high school GPA could override a low test score in individual cases. This seems both ethical and sensible.

Q: But don't wealthier families tend to spend a lot on test prep for their families? I could understand why these families do this, but is this fair to students from less affluent families? Aren't there some things UC could do to level the playing field?

We have three responses to this question. First, it seems likely that more affluent families are more likely to pay for test prep for their high school students. Evidence about the impact of test prep on SAT scores tends to show that these programs boost test scores, but the studies disagree on the degree to which test scores in fact rise. If it were the case that affluent families are wasting money on test prep programs because test scores are not boosted enough to make a difference in UC admissions, access to test prep would not be a reason for UC to drop standardized testing from admissions criteria or even to adjust its current comprehensive review practices using test scores as one factor among 14. On the other hand, perhaps extra confidence is created simply by having had test prep, so students having test prep but not repeating the test do better than they would have done had they not had the test prep. That would be an unfair advantage to the affluent. We cannot answer whether that is the case with data we have evaluated.

On the question about how much test prep boosts scores, we refer the interested reader to Sackett and Huncel (2017).[122] These authors review the literature and find virtually no support for the claims reportedly made by some test prep providers that test prep will boost SAT scores by 100 to 200 points. The exception was one study that found an increase in SAT verbal scores of 121 points, but the study lacked credibility as it used an extremely small sample of only 35 students. Sackett and Huncel report that some studies overstate the impact of test prep by comparing students whose families buy test prep with other students who do not do so, but without accounting for the many other academic activities that the former group of students were also involved in. Sackett and Huncel conclude that the most careful study in the literature suggests that test prep may boost test scores on average by about 14 points on the SAT math and 5 points on the SAT verbal. Such gains are not large enough to make much of a difference to UC admissions decisions. (See for example **Figure 3C-1**, which shows that the probability of admission at UC rises quite slowly as SAT scores rise.)

Second, it is important to remember that admissions readers consider each student's application in the context of his or her own background and the GPA and SAT scores of all other applicants from the same school. This has the important effect of taking into account variations across neighborhoods in practices

---

[122] Sackett, P.R. and N.R. Kuncel. 2017. "Eight Myths about Standardized Admissions Testing," in Buckley, J., Letukas, L., and B. Wildavsky (Eds.), Measuring Success : Testing, Grades, and the Future of College Admissions," Baltimore: Johns Hopkins University Press, pp. 13-39.

like using test prep services. The report shows that SAT scores are effectively re-normed by socioeconomic group. This means that on average, the advantage that is conferred upon more affluent students by test prep is being taken into consideration indirectly by UC. Nonetheless, within socioeconomic groups, test prep is unobserved and could favor one student over another to some degree.

Third, we do think that private test prep services would be far less valuable if the UC admissions test differed from current admissions tests in two important ways. First, suppose that students could have a vast number of practice admissions tests available online, to level the playing field between more affluent families which can buy access to many sample tests and coaching, and less affluent families? Second, suppose that the UC admissions test, when practiced online, generated useful information for students, their parents and their teachers about specific content areas where the student needed to improve? More "diagnostic" tests could benefit students by showing them what specific topics they need to learn more about. Such a test would change the nature of admissions testing in a way that would help all students, regardless of income, to prepare and to learn from practice exams what they need to study. With this in mind, the Task Force recommends the development of a new admissions test to replace the SAT and ACT that would have both of these beneficial features. The final section of this report contains the detailed recommendation.

Q: I have heard that the SAT and tests in general are racially biased. Is this true, and if so what is UC going to do about it?

The Task Force takes this possibility seriously. The definition of bias in a test question that psychometricians use is, roughly speaking, that two groups perform systematically differently on the given question even though the members of the two groups have the same underlying achievement level. For example, suppose that a question on a math test assumed that students had knowledge of how to play a sport, say football. If a math question could be understood only if the student understood the sport that was used for the premise of the question, and the two groups had different knowledge of how that sport worked, this would lead to one group systematically underperforming the other even if on average the two groups had the same level of ability with regards to the given mathematical skill.

Our review of the existing literature suggests that racial bias in the SAT, at least the version of the SAT in place in 1999, is, at most, a minimal problem. A key paper, by Santelices and Wilson (2010) reviews the earlier work in the field and concludes that there is no evidence of SAT racial bias in the math test but that there is mixed evidence for black/white differences on the verbal test.[123] The authors then conduct updated analyses that take into account some criticisms that had emerged of the earlier work. They confirm that the math test is not biased for either of two racial comparisons. Similarly, they conduct analyses of bias on the SAT verbal test and find no evidence of bias in Latino-white comparisons. They find evidence of bias in African-American-white comparisons on the SAT verbal test, but depending on the difficulty of the question, the direction of the bias reverses, with the test biased in favor of whites for easier questions and in favor of African-Americans for more difficult questions.

So, in three of four cases (math for both racial/ethnic comparisons and verbal for Latino/white comparisons) no evidence of racial bias emerges. In the fourth case, black/white comparison on the SAT verbal test, some evidence of bias exists, but the bias is against white students on some questions against black students in other cases. Furthermore, our analysis of the results suggest that for this one, the effects are far too small to explain much of the SAT gap in test scores between black and white students.

[123] Santelices, M.V. and Wilson, M.. 2010. "On the Relationship between Differential Item Functioning and Item Difficulty: An Issue of Methods? Item Response Theory Approach to Differential Item Functioning." Harvard Educational Review Vol. 80 No. 1 Spring, pp. 5-36.

An important qualification of this result is that it uses data from California high school students gathered in 1994 and 1999, a period when the version of the SAT in use was two versions earlier than the current version of the SAT test. The existing evidence is persuasive, yet outdated.

With this concern in mind, the Task Force recommends later in this report that UC request item level student results on both the SAT and the ACT so that UC researchers can update these tests for racial bias. This is an important task that would benefit both students nationwide and the testing services, regardless of which way the findings would go.

This answer considers racial bias at the level of individual test items. It could be that high-stakes standardized tests are racially biased in other ways. One way discussed in Section IV is “stereotype threat.” It could be that if members of one group felt threatened by a test that members of another group did not, it might be the case that this group would answer ALL questions on a test correctly at lower rates. The Task Force did not have time to gather and analyze data bearing on this important question, e.g. to measure the size of the effect of stereotype threat in order to compare it as a possible source of racial bias to other possible sources.

Q: The SAT tests verbal and math abilities. My child is a lot more than the sum of her test scores. Why does UC look so narrowly at the results of a single test? What changes, if any, is UC thinking of to look at my student from a much broader perspective?

This question was on the minds of many in the Task Force. To those who believe that the SAT plays a decisive role both in admission decisions for individual students and in creating disparities in admission rates across socioeconomic groups, we refer them to the above dialogue and the content in Section III of the report. It establishes that admission decisions are much more sensitive to high school grades than to SAT scores. Even more important, it shows compelling evidence that admissions decisions take into account differences across groups in mean SAT scores. Any claim that UC looks narrowly at SAT scores does not turn out to be correct. That does not mean that UC does as well as it could in explaining how UC admissions decisions are made, nor that high school students understand how it works. The report notes that high school guidance counseling services are overburdened and at least anecdotal evidence from UC’s office of admissions suggests that our outreach efforts to high schools, though already extensive, could be improved. One of the STTF’s goals in including the analyses and visualizations in this report was to improve transparency about the UC admissions process.

Furthermore, UC uses 14 criteria in evaluating applications of which the SAT is only one. In fact, UC looks at the “whole person”, and the whole transcript. The 14 items include the number of advanced courses taken, whether the student plans to take a challenging set of courses in grade 12, whether students have shown marked improvement across grades, and, as discussed earlier, UC admissions officers look carefully for evidence of special skills, talents, and leadership activities. Evidence of having overcome a significant life challenge also is among the 14 factors.

Could UC do a better job of looking at the whole person? It seems very possible. For instance, with this in mind, the Task Force recommends later in the report that UC consider adding new factors to the statewide eligibility criteria for guaranteed admission. (The statewide “index” guarantees admission to UC for students who meet or exceed any of a number of possible combinations of GPA and SAT scores.) For example, we recommend that UC consider making these criteria more flexible, perhaps by giving additional consideration to students whose GPA has improved over time.

Q: Why does UC need to use any admissions test at all? Why not just look at high school GPA?

This might be the most important question yet. Given the evidence gathered by the Task Force, we know how the current admissions tests are used and not used in the admissions process. With some confidence, we can predict that the following would happen if UC stopped using admission tests and relied solely on GPA and other aspects of the students' transcript for the academic appraisal of applicants. The following statements are very likely true because our analysis shows that SAT scores predict outcomes even after taking into account GPA and a student's socioeconomic status.

Likely Impacts of Dropping Admission Tests:

- The average student admitted would have a lower first-year GPA, a lower probability of persisting to year 2, a lower probability of graduating within seven years, and a lower GPA upon graduation. The reason for this is that UC would no longer be able to use admissions tests to identify, within socioeconomic groups, the students most likely to succeed.

- UC would have a lessened ability to deliver academic and socio-emotional supports to freshmen because it would no longer be able to identify as accurately the students at risk of faring poorly at UC.

- The impacts on diversity depend largely upon how the remaining 13 admission factors were used in admissions. Our analysis in the recommendations section later in this report suggests that expanding Eligibility in the Local Context, a program that admits the top 9% of graduates in each high school as ranked by GPA, would modestly improve diversity. But countering this is the fact that many underrepresented students who are currently guaranteed admission to UC through the statewide eligibility criteria would probably no longer be guaranteed admission. These students, who are guaranteed admission through the statewide index but not by the Eligibility in the Local Context (ELC) program, by definition have been admitted based on their SAT scores. The numbers of disadvantaged students who would lose guaranteed admission if UC dropped SAT tests is surprisingly large. In 2018, about one quarter of low-income, first-generation and underrepresented minorities who were guaranteed admission to UC earned this guarantee solely by virtue of their SAT scores. African-American and Native American students would be especially hurt by dropping the SAT: among the students guaranteed admission to UC, 40% of African-Americans and 47% of Native Americans won their guarantee because of their SAT scores. **Figure 3B-1** and surrounding text explains these surprising facts.

- The average subsidy that the state of California would provide to California students through subsidized tuition (and complete tuition waivers for lower income students) would need to rise due to a longer time to graduation for the average student. Furthermore, the total subsidy expenditure per Bachelor's graduate would rise, not only due to longer time to graduation, but due to a larger percentage of undergraduates not graduating. These increased subsidies may be socially desirable; we merely note here that there is a financial consequence from dropping admissions tests.

- The UC Regents have long had a policy that out-of-state undergraduates must "compare favorably" to in-state undergraduates. UC takes this requirement seriously, and has used the SAT and ACT test results as objective measures of academic preparation. Comparing resident and non-resident students based solely on GPA would be very difficult and unconvincing, given that grading standards differ in other states, and especially so, in other countries.

The above consequences of using only high school GPA and not admissions tests are quite likely, due to the lessened ability of UC to pick the students within each socioeconomic group most likely to succeed and to identify the students most at-risk, if it did not use the SAT. However, this statement is predicated on UC not making other changes that could somehow compensate for this change in the mix of students admitted. A case in point is that UC could react to a decision not to use admissions tests by evaluating its support programs for undergraduates, with a view to improving the programs and expanding the most successful programs. This reform in theory could prevent graduation rates from falling and UC GPA's from falling were admission tests to be abandoned. The STTF has recommended that UC implement a rigorous evaluation of the impact of its undergraduate support programs on student outcomes, to supplement the other recommendations including the development of a broader admissions test. Clearly, this recommendation could prove at least as valuable in the event that UC dropped admissions testing.

(PARENT A) Q: Doesn't UC, as a publicly supported institution, have a moral obligation to reflect the diversity of the state of California as a whole?

Yes, UC absolutely has a moral obligation to do the best it can to enroll an undergraduate population that reflects the diversity of the state. As mentioned in Part I of this report, UC has been under a stipulation to have its students represent the entire state since as early as 1868!

(PARENT A) Q: But really, how is UC doing? Is it simply admitting affluent students, who tend to be from certain racial/ethnic groups? If so, isn't UC simply perpetuating the income gaps between rich and poor, and between racial and socioeconomic groups, and transmitting these inequalities to the next generation? How is this fair to my child, who has not benefited from being raised in an affluent family?

UC's admissions procedures are designed to ensure fairness to all applicants. As the report has shown, differences in SAT scores that relate to family context are carefully taken into account, so much so that in the admissions pool underrepresented groups have lower average SAT scores than more advantaged groups, and these differences mirror the gaps in test scores in the applicant pool.

In addition, one of the 14 factors that UC considers in making admission decisions is item 13: "Academic accomplishments in light of your life experiences and special circumstances, including but not limited to: disabilities, low family income, first generation to attend college, need to work, disadvantaged social or educational environment, difficult personal and family situations or circumstances, refugee status or veteran status."[124] This item gives students who have faced adversity the opportunity to describe this experience, so that campuses can look at both grades and test scores in the proper context.

There is a third UC practice that illustrates UC's efforts to take disadvantage into account and to be fair to all students. Most students who enroll at UC were guaranteed admission through the statewide eligibility criteria, the ELC criteria, or both. But what happens to students who do not receive a UC admissions guarantee through either path? UC requires that campuses give a full review to all applicants who meet the minimum eligibility criteria, namely having completed the A-G course requirements with grades of C or higher, and who have an overall GPA of at least 3.0. These students are "Entitled to Review" (ETR), which means each campus to which such students have applied, who do not meet either of the guaranteed admissions sets of criteria, must be given a full review. (See Section I of the report for more details.) ETR is not a small program. In 2019, about 40,000 students were ETR, of whom about 17,000 were admitted to at least one UC campus. This amounts to just under 24%of all admitted students.

These three examples show the efforts UC makes to be as fair as possible to students who come from less advantaged situations. But that does not really answer your question of how well UC is doing by way of disadvantaged applicants. It is not the case that UC merely accepts only affluent students. The report provides ample evidence about this.

[124] Guidelines for Implementation of University Policy on Undergraduate AdmisssionAdmission, https://senate.universityofcalifornia.edu/_files/committees/boars/documents/guidelines-implementation-of-ug-admission-rev-7-2019.pdf

Independent reviews of UC's approach could help to answer the question of how UC is doing in a comparative way. One example of this is the *US News and World Report* rankings of national universities on social mobility.[125] The rankings examine students with lower family income who qualify for federal Pell Grants to help them enroll in college. The *US News* Social Mobility rankings compare universities based on the six-year graduation rates of Pell grantees, while giving additional credit to universities where a larger proportion of students hold Pell grants. Another factor is the relative six-year graduation rates of Pell Grant holders and other students.[126] In these rankings of all national universities, UC campuses account for the top three universities, 6 of the top 10 universities, and 8 of the top 30 national universities in the United States. The implication is that relative to other national universities, UC is doing more to enroll lower-income students and to support them successfully through to graduation.

This is great news, but does this mean that UC is doing all it can? Not necessarily. The Task Force has looked at a number of actions that could be taken to increase both the admission rates of disadvantaged students, and their rates of success once enrolled. Among the recommendations that the Task Force is making, with a view to improved diversity and improved success rates, are:

- Expansion of the ELC program, which uses grades only;
- Enhancing the statewide eligibility index to take into account more than SAT scores and GPA, for instance by explicitly giving weight to applicants whose GPA has improved over time;
- Rigorously evaluating UC's many student support programs, and then expanding the most successful programs;
- Performing an updated check for racial bias in existing tests, and also analyzing the current admissions tests both to find out on a more granular level which types of math and verbal questions best predict UC success, and which types of questions account for gaps in test scores across socioeconomic groups;
- Supporting the creation of an entirely new admissions test that would test a broader set of skills than math and verbal achievement, and that would give all students ample opportunities to take practice tests and to learn in a detailed way about specific skills they need to work on. It seems quite possible that by testing a broader set of skills, the new test will improve diversity in admissions as well.

This was a long answer, but we hope it addresses your concern that UC does not care enough about diversity and about disadvantaged students.

(PARENT B) Q: I am sorry, but I do not like this line of questioning about diversity. My student is a white male, from an affluent family, and happens to have excellent grades and test scores. I hear what this other parent has been saying, but it isn't my son's fault that he was born into an affluent white family. Doesn't UC have an obligation to admit the best prepared students? Why all this fuss about the need for diversity?

This is a natural question to ask. We are glad that both of you, Parent A and Parent B, can hear each other's questions and our replies. You are both asking, from different perspectives, about questions of fairness. We hope that you can each put yourselves in each other's shoes, and can see that fairness for one student must be balanced by fairness towards the other student.

The Task Force has worked hard to think through the tension that exists here. On the one hand, the UC Regents have long required UC to admit students who are prepared for the rigorous undergraduate curriculum. But the Regents also want UC to reflect the diversity of the state. The UC response has been

[125] See https://www.usnews.com/best-colleges/rankings/national-universities/social-mobility?_mode=table .
[126] See https://www.usnews.com/education/best-colleges/articles/how-us-news-calculated-the-rankings .

balanced. On the one hand, about two-thirds of admits in 2019 were guaranteed admission under state eligibility guidelines that emphasize GPA and SAT scores. Of these just over half were also ELC, meaning that their within-school GPA ranking also guaranteed them admission. Another twelfth of students were guaranteed admission through ELC only. This group was admitted regardless of their test scores, because their GPA ranking signaled good things about their work ethic and grit. Another quarter of admitted students were not guaranteed admission, but impressed one or more campuses sufficiently to gain admission under the Entitled to Review program. The goal here is to balance measures of academic achievement and the Regental goal of diversity.

Q: I think that the testing culture in our society has gone way too far. Why are we so focused on admissions tests, where just a few points difference could make a big difference in admissions?

We hope that the earlier part of this conversation will have convinced you that at UC, admissions officers do not focus narrowly or solely on SAT scores. Test scores represent one of 14 factors in admission, and the test is used both to predict who will do well, and, among admits, who might most need additional supports starting in their freshman year. We have also pointed out that differences of a few points on the SAT do not make a big difference in admissions decisions in most cases.

Another reason to continue using admissions tests, which we have not talked about much yet, is that in California, HSGPA is becoming less useful as predictor of UC performance (in first-year UC GPA) over time. In fact, at UC it is now the case that SAT predicts performance better than HSGPA. The problem with HSGPA is that grading standards differ across schools and teachers, and are not objective. The other problem is that nationwide HSGPA has been drifting upward for several decades, without concomitant improvements on tests like the SAT and other tests. There is simply less and less information that HSGPA by itself can provide about student prowess. See Part D in Section III for an extended discussion of the issues with relying solely on HSGPA without having a standardized admissions test as a second tool for evaluating student achievement.

Q: Okay, but all this emphasis on tests that produces a couple of test score numbers, without making students as a whole better students. How does this testing effort help students learn, or help their teachers figure out how to provide more support in areas in which students are struggling? Surely there must be a better way.

We agree that undue emphasis on activities such as memorizing techniques to help on admissions tests does not necessarily boost student achievement or enhance learning. This is why the Task Force is recommending the development of an ambitious new form for testing that will test skills more broadly, and that actually could have the effect of improving the quality of schooling. The new test will achieve the latter by giving students and their teachers detailed feedback on specific areas within a subject area where the student needs to improve. Further, the feedback can be frequent, as students can take practice tests online at will, with instant feedback being provided.

Q: Well, suppose you are right that test scores do not matter as much for admissions as the public thinks. Isn't it UC's obligation to explain this better?

This is a valid point, and UC could probably help the public understand more fully the role that admissions tests play. One service that UC already provides in this regard is online information on the range of SAT scores among the middle 50% of admits at each campus. Perhaps UC could make clearer that each campus is willing to accept students with a wide range of SAT scores by also showing the range of SAT scores of the middle 80% of admitted students. The hope of the Task Force is that what we, as individual members, have had the privilege to learn while serving on the Task Force will be widely shared publicly. That is certainly among our goals as we wrote this report.

VI. Reforms the Task Force Considered but Does Not Recommend

The Standardized Testing Task Force (STTF) evaluated two possible reforms to the University's admissions process that have received public attention: 1) adoption of the Smarter Balanced (SBAC) Assessment Consortium's high school assessments in place of the SAT and ACT, and 2) giving applicants to the University of California the option of whether to take the SAT or ACT rather than requiring the tests. We do not at this time recommend either. Below we summarize our concerns about each of these possibilities.

The Task Force does not recommend adopting Smarter Balanced Assessment in lieu of currently used standardized tests.

One alternative to the continued use of the SAT and ACT in University admissions is adoption of the current Smarter Balanced Assessment Consortium's high school assessments as an alternative. While there are a number of positives associated with this alternative, the Task Force does not recommend this option after the analysis described below.

I. Background

Current SBAC assessments were originally developed with grant funding under the Obama administration. At the time, many states, including California, were adopting new college and career readiness content standards in math and English Language Arts (ELA): the Common Core State Standards. Because of the adoption of new standards, new tests aligned to these standards were needed to satisfy federal K-12 accountability requirements. A group of states (California among them), led by Washington as the fiscal agent, jointly developed the SBAC tests. After the federal grant funding ran out in 2014, SBAC was established as a UC-operated program financially supported by the SBAC states' membership fee payments. SBAC was initially housed at UCLA, and since 2017, located at UC Santa Cruz Extension. In California, SBAC tests are managed by the California Department of Education (CDE) with the Educational Testing Service (ETS) currently serving as the main test administration contractor. As required by the federal Every Student Succeeds Act (ESSA) and state education legislation, students in grades 3-8, as well as in the 11th grade, are assessed annually with SBAC tests in California in math and in ELA.

SBAC's management and governance is codified in a set of intergovernmental agreements that the UC signed with all the SBAC states. **Contrary to popular belief, it was not an assessment specifically designed for California, though California plays a major role in the continuing success of SBAC. Governed by the consensus of the states on all major decisions since its inception, it is a system that has coalesced over time to serve the needs of many partners.**

One of the most prominent components of SBAC assessments, which is the subject of the present discussion, is its annual summative test in the 11th grade. In California, SBAC summative tests are administered toward the end of the school year during a weeks-long testing window to capture the cumulative learning achieved by the students over the academic year through instruction. The summative component stands in contrast with SBAC's interim component, which are given several times during the course of instruction to provide more timely feedback to educators. The summative test consists of two parts: an item-by-item Computerized Adaptive Test (CAT), and teacher-led performance tasks. Test scores are generally made available to the parents and students late in the summer or early in the fall of the following academic year.

Purposes of Smarter Balanced Tests

Current SBAC summative tests were originally designed as assessments to measure student achievement and growth with respect to the Common Core to inform program evaluation and school, district, and state accountability. As an accountability test, SBAC's design and test score use are different from typical high-stakes admissions tests.

First, current SBAC test scores carry low stakes for the individual students taking the test. While the California State University (CSU) system accepts SBAC test scores (along with the SAT and ACT scores) to meet conditions of enrollment, they are not used as part of the admissions process. CSU admissions are also not as selective as UC. The current use of SBAC scores at CSU are intended to inform the placement decisions of students in appropriate college-level coursework. The current design and administration procedures of SBAC are not intended for it to withstand high-stakes usage in highly selective UC admissions. More technical discussions on this point will be provided in the sections that follow.

Second, an important distinction between accountability testing vs. admissions testing is that a K-12 accountability test aims to capture differences over larger aggregated groupings (e.g., schools or districts) as opposed to refined differentiation among individuals of varying levels of proficiency or preparedness. It is the aggregated performance on SBAC tests that contributes information into the state's K-12 educational accountability system. **A test designed and optimized for one purpose is not automatically appropriate for another.** The 2014 edition of the Standards for Educational and Psychological Testing stated that:

> *Clear articulation of each intended test score interpretation for a specified use should be set forth, and appropriate validity evidence in support of each intended interpretation should be provided. (p. 23)*[127]

It is incumbent upon the test providers to document validity evidence that supports the purposes for which the scores are being used. Currently available evidence reviewed by the STTF does not support the immediate alteration of the purpose of the SBAC tests from a designated K-12 accountability assessment in California to a UC admissions test.

II. Analysis

The Committee evaluated the possibility of using SBAC test scores in a similar way to how UC currently uses the SAT and ACT. There are many reasons to consider this change. However, the STTF is concerned about a wide range of risks related to the adoption of SBAC to meet UC admissions requirements. Proponents of SBAC use recent data to suggest that adopting SBAC may, a) potentially improve the representativeness of UC matriculants of the broader population (equity); they also make four additional arguments in support of SBAC's adoption: b) the tests assess content accessible from school-based instruction (instructional validity); c) the tests are considered aligned to college and career readiness standards (content alignment); d) the tests are free and administered by teachers in schools (testing experience); and e) it uses a novel CAT format (technical sophistication). Here we address each of these in turn.

[127] American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

a. Equity

In an early administration, when there were no stakes for student performance in relation to UC eligibility or admissions, SBAC high scorers were more racially diverse than SAT/ACT high scorers. An implication is that using SBAC instead of SAT/ACT might allow UC to increase the racial diversity of matriculants. However, the available data showing benefits to diversity is from a low-stakes administration of the test. It is most likely that test-taker and parental behavior would change if the test were to become high-stakes and used in admissions into selective UC campuses. This could not only mean that some of the supposed diversity-related benefits might not be realized, but also that the current, federally mandated purpose for which the SBAC tests were developed would be gravely undermined. The discussions on instructional validity also help explain why adoption of the current SBAC test may in fact be detrimental to fairness and equity due to unequal access to opportunities to learn across K-12 schools.

b. Instructional validity

A main perceived benefit is that the SBAC tests assess content that should already be covered by school-based instruction aligned to the Common Core, which is freely available to all students, unlike paid SAT coaching. As such, teachers' instruction becomes the ultimate test preparation and therefore disparities should become smaller in this criterion-referenced measure. These are good observations, but all tests, independent of their claimed design attributes, almost invariably reflect disparities related to background and prior experience unless there have been effective instructional interventions to mitigate their influence.

Criterion-referenced tests were created with the idea that a specified domain of content and cognitive skills could be codified into consistent standards and that teachers could teach to them. The purpose of such aligned instruction would be to interrupt and to reduce, at least for the lower end of the achievement distribution, the relationship between background and test performance.[128]

For a criterion-referenced test to achieve its stated instructional validity, it should first display sensitivity to instruction, that is, scores on the test should change when learning occurred. This is a property of the measurement instrument that can and should be checked empirically with experimental or well-conducted quasi-experimental studies. A second element of instructional validity relates to the instruction that occurs in the classrooms. From the angle of K-12 assessments, teaching and instruction have been termed "opportunity to learn" in recent times, meaning that opportunities to learn the content to be assessed are made available to students through classroom instruction. Ideally if instruction is aligned to standards in the entire K-12 system, the opportunity to learn should be more evenly available, and a test sensitive to instruction should be able to detect it as a positive change.

Unfortunately, there has been little evidence of instructional validity for most achievement tests. There is a general lack of empirical data regarding the tests' sensitivity to instructional differences. There are inconsistent results on instructional sensitivity from different data collection approaches,[129] including instructional artifacts.[130] Significant portions of standards or content may also be underused in classroom

---

[128] Baker, E. L., Chung, G. K. W. K., Cai, L. (2016). Assessment gaze, refraction, and blur: The course of achievement testing in the past 100 years. *Review of Research in Education, 40*, 94-142.
[129] Ing, M. (2018). What about the ''instruction'' in instructional sensitivity? Raising a validity issue in research on instructional sensitivity. *Educational and Psychological Measurement, 78*, 635-652.
[130] Burstein, L., McDonnell, L. M., Van Winkle, J., Ormseth, T., Mirocha, J., & Guiton, G. (1995). *Validating national curriculum indicators*. Santa Monica, CA: RAND Corporation.

teaching.[131] With present levels of inequalities in K-12 education, the uneven availability of opportunity to learn means that students from less well-resourced schools will be systematically disadvantaged, again raising a more fundamental fairness and equity concern.

An empirical alternative to obtaining detailed evidence about instructional validity could be inferred from increasing levels of performance, a result of a system convergence around the standards and tests. However, K-12 test performance in math and reading has been relatively flat for more than a decade in the US.[132] SBAC has largely replicated that national trend in California. The STTF observed that over five years of administration and implementation of Common Core aligned instruction and assessment, there continued to exist persistent achievement gaps across key student subpopulations along race/ethnicity lines. There have even been year-to-year drops in student performance in the later grades (7 and above).[133]

Finally, the STTF noted that the utility of the test as a predictive measure was much less than previously anticipated. Given that SBAC was designed specifically to be in close alignment to college and career readiness standards and it is directly linked to several years of preparation in the form of high school instruction, it should in theory be a much stronger predictor of typical outcomes than SAT/ACT. From the limited data and analyses that STTF has access to,[134] the STTF's conclusion is that in terms of predicting first-year GPA at UC, SBAC in fact does not do better than the SAT, even when SBAC scores are not currently used in UC eligibility or admissions determination. The last observation is important because in the current analysis, the SAT scores and high school GPA are already range-restricted due to their being used in admissions, leading to attenuated multiple correlation coefficients,[135] but SBAC test scores do not currently suffer from such range restriction. **If SBAC were to become an admissions test for the UC, the multiple correlations reported today would decrease.**

c. Content alignment

Thus far, SBAC tests have been built and used for accountability purposes as a broad achievement measure to capture group differences in achievement levels with respect to college and career readiness content standards (the Common Core). A key claim in favor of SBAC's adoption as new admissions tests for UC is that the SBAC tests are considered aligned to the Common Core. The STTF notes that purposes and conditions of test use are inherently tied to validity. One cannot retrofit an existing test to a new purpose solely because of content alignment, without collecting and evaluating pertinent validity evidence for the new use.

This precept is symmetrical, that is, it applies to using group achievement tests in an individual high-stakes setting (current situation) and to using admissions tests to evaluate school effectiveness. The STTF is not alone in taking this measured position regarding test validation and use of evidence. As a matter of fact, in June 2019, the National Council on Measurement in Education specifically issued a *Position Statement on the Use of College Admissions Test Scores as Academic Indicators in State Accountability*

[131] Opfer, V. D., Kaufman, J. H., Pane, J. D., & Thompson, L. E. (2018). Aligned curricula and implementation of common core state mathematics standards: Findings from the American Teacher Panel. Santa Monica, CA: RAND Corporation.
[132] https://nces.ed.gov/nationsreportcard/data/
[133] https://www.cde.ca.gov/nr/ne/yr19/yr19rel68.asp
[134] Kurlaender, M., & Cohen, K. (2019). Predicting college success: How do different high school assessments measure up? Policy Analysis for California Education.
[135] Meade, A. W. (2010). Restriction of range. In Salkind, N. J. (Ed.) *Encyclopedia of research design* (pp. 1279-1280). Thousand Oaks, CA: Sage.

Systems[136] precisely because under ESSA, states have been given considerable flexibility in using the SAT and ACT as tests for school accountability. These admission tests may or may not be considered valid for the new purpose unless validity evidence is collected and interpreted. If one accepts this as a fair requirement, then the principle of symmetry would imply that SBAC should be subject to the same level of evidence gathering and validation for the intended new use. Content alignment alone is one piece in the overall validation of test score use for a particular purpose and not the only piece.

Finally, it is important to note that content alignment is a matter of degree and not an absolute statement. The STTF notes that claims of alignment to the college readiness content standards have also been made about the ACT, the newly revised SAT, as well as other K-12 measures (e.g., the Partnership for Assessment of Readiness for College and Careers, or PARCC), to varying degrees. Alignment studies that have historically been employed to verify adherence to content standards are based on expert judgment. These studies are not free from possible bias in human judgment and decision-making. Even though judges may be independent and the judging panels may follow clearly defined procedures, there is typically substantial variation/discordance both within and between judges and panels on the degrees to which an item or a test is aligned to a set of standards.

d. Testing Experience

There are many positives associated with the current testing experience of SBAC. It is administered in a familiar environment, the student's own school, and the student does not need to make special arrangements to reach the testing site. It is free to the student and mandatory, unless parents opt out of testing. While there is a clear perception that parental wealth may confer an unfair advantage on test takers' SAT/ACT scores due to intensive paid test preparation, the ability to take the tests multiple times, and having better access to various resources that can make the testing experience more comfortable and manageable, SBAC does not yet currently cast such a negative image. Stereotype threat may also be less of an issue.

There is, however, a scarcity of practice materials relative to what is available freely for the SAT/ACT, even when the SBAC interim assessments, which are primarily intended as tools for teachers to check on student progress,[137] are also counted as practice material. If SBAC were adopted by UC, it is most likely that test preparatory services would develop to serve students whose families have the means to pay for them. Even if one makes the argument that the best test preparation for SBAC is school-based instruction, the lack of access to preparation material and highly skilled instructors still poses fairness questions for students who attend schools that are less well resourced.

Per California state law, the SBAC is administered once in the 11th grade, with a district-determined testing window of at least 25 days after 66% of the instructional days of the year have been completed. As an accountability test, these are entirely sensible choices. But as a high-stakes test that could help determine a student's future career at UC, the testing window leaves students and their families with limited choices when the unexpected happens. Issues such as illness that last for more than a few days, having to move from one part of the state to another, a family emergency that requires out-of-town travel, etc., could cause chaos for the students involved with no clear means of redress that is fair to the individuals involved and to the others at the same school. This issue is especially problematic if the event falls on the last few days of the testing window, or if the time is close enough to June 15, when state-wide availability of SBAC ends. An expanded testing window may address this specific problem, but as discussed below, it will also introduce more pointed concerns related to exposure and validity.

---

[136] https://higherlogicdownload.s3.amazonaws.com/NCME/c53581e4-9882-4137-987b-4475f6cb502a/UploadedImages/Documents/Admission_Statement_06-16-19.pdf
[137] http://www.smarterbalanced.org/assessments/interim-assessments/

In terms of test security, the fact that SBAC is a school-based teacher-proctored test is more advantageous in helping address the problem of cheating by substitution (where someone is paid to take the test for a student). On the other hand, in high-stakes testing, more testing sites also introduces more opportunities for test security protocols to be breached. They are also more difficult to directly monitor. This may become more problematic when the current SBAC tests are delivered over weeks-long testing windows.

Furthermore, because current SBAC is curriculum-based and given at different times of the year for different schools, students may perform better if the test is administered later in the year, systematically advantaging students from later administering sites. But state-wide uniformity in the opening and closing of the testing window would also necessarily take away the control that the local education agencies needed to have for appropriate planning. Students at later administering sites may also gain pre-knowledge of assessment content because of item exposure in Computer Adaptive Testing (CAT). In other words, some students may know some of the items in the CAT before taking the test simply because others have already taken the test and remembered what was on the test. This last point will be further elaborated in the next section.

e. Technical Sophistication

Converting a low-stakes item-by-item computerized adaptive accountability test into a high-stakes admissions test will have unintended consequences that may compromise the validity of the use of the test scores for both its original purpose, and the new purpose. To fully appreciate this concern, it is necessary to understand how adaptive testing works, both theoretically, and in the real world.

In theory, an item-by-item adaptive test chooses the next item from a pool of items (the item bank) using a provisional estimate of the test taker's achievement or proficiency based on answers to items already administered. It also needs the definition of an optimal item to choose next. For example, an optimal item could be defined as an item in the item bank whose level of difficulty matches the provisional proficiency most closely. In this regard, the test is said to be "tailored," with each individual essentially taking a unique form of the test most suitable for the test taker, all put together on-the-fly by an adaptive algorithm.

On paper, adaptive testing is an ideal way to put together a test. Indeed, when the stakes are low, such as in accountability testing, CAT may be helpful. For example, even though each individual CAT is of limited test length, when many CATs are aggregated over students, the result is a set of much more broadly sampled assessment content at the level of schools or districts than traditional fixed form tests can accommodate. The richness of information is helpful for making inferences about school-level or district-level performance.

On the other hand, high-stakes testing at the individual level poses a completely different set of threats to validity. While an item bank may contain hundreds of items, and there is, in principle, a combinatorically large number of possible alternate pathways through the test for test takers of all stripes, a real test must cover a set of required content domains and use a variety of approaches to assess the test taker's knowledge due to design and validity considerations, thereby narrowing the number of possible alternate paths significantly. This leads to a critical problem for item-by-item CAT: item exposure control. An item is called "exposed" whenever a CAT algorithm selects the item for administration for a test taker. The notion is that an exposed item is more amenable to be disclosed to others who may take the test at a later time.

To understand the role exposure plays, it is helpful to revisit the storied history of the computerized adaptive GRE, a high-stakes admissions test for graduate and professional schools. Shortly after ETS started the GRE CAT in the 1990s, the test preparation industry (in both North America and Asia) discovered that by memorizing test items and sharing them after the test, individuals outside of ETS could accumulate enough items from the exposed part of the item bank to compromise the validity of the test. Knowing a portion of the items ahead of time can give the appearance of higher proficiency to the adaptive testing algorithm, thereby pushing the algorithm to give the test taker items that the test taker's true proficiency may never indicate. The size of the item bank only marginally helps limit the impact of item exposure, because one must also simultaneously satisfy content balancing constraints. The ascendance of the internet also enabled the sharing of items at a larger scale than anyone ever anticipated. Ultimately, ETS stopped the GRE CAT in some parts of the world and reverted back to paper-and-pencil testing.

At present, adaptive algorithms used in SBAC rely on more rudimentary methods to control exposure. In general, adaptive tests with rudimentary exposure control methods tend to have highly uneven exposure across items, meaning that a small subset of the available items are seen by a disproportionately larger share of students, increasing the chances that key items may be compromised. If SBAC item exposure is not appropriately controlled and the item bank size remains modest in comparison to earlier high-stakes CAT for admissions purposes, exposure will undoubtedly compromise the validity of the use of the SBAC test scores for its original or new purposes, and possibly leading to more lawsuits. Importantly, this required no technology in the testing sessions, only enough brains capable of remembering some of the items, and the capability to share them. It is substantially easier to share content today than in the 1990s.

f. Lack of Out-of-State Availability or Comparability

Finally, SBAC is not available uniformly in all states, let alone being available at all internationally. The CCSS as adopted by the states nearly always have (small but still noticeable) changes to fit the standards into state-specific contexts. The implementation of SBAC testing itself also varies across states. As a result, it would be difficult to directly compare SBAC scores for out of state students to SBAC scores from in-state students, – thus SBAC scores could not be used to establish that out of state students "compare favorably" to in state students, one of the other major uses of test scores for the accountability of UC to the public.

For these reasons, the STTF does not recommend adopting the currently available SBAC test in UC's admissions process.

The Task Force does not at this time recommend that UC make standardized tests optional for applicants

At this time, the Task Force recommends that UC conduct additional research on the impact of going "test optional" before deciding whether to implement such a policy. We have pragmatic concerns about how campuses would evaluate and compare applicants who submit standardized test scores relative to applicants who do not. We also have concerns about how campuses could best impute, explicitly or implicitly, test scores to applicants. We have ethical concerns about how to treat students in the two groups fairly. Finally, given the state of research, we are uncertain what the effects of going "test-optional" would be on either the preparation or composition of incoming classes.

I. Background

As has been widely reported in the press, several hundred colleges and universities have moved to eliminate the requirement that students include standardized test scores in their applications. Below, we turn to the scholarly assessments of what the effects of these moves have been. At the outset we note, however, that none of these institutions is as large as UC, and we believe that the results of taking such a step here might have effects unlike those described in extant research.

We also note that UC admissions may already use test scores in a manner unlike that of other institutions. For example, currently applicants may be admitted to UC through Eligibility in the Local Context (ELC) without consideration of test scores – although individual campuses consider test scores building their entering classes. UC campus directors of undergraduate admissions informed us that standardized test scores are a useful tool to advance diversity of incoming classes and shared concerns that without use of the scores as a tool, student diversity could actually decline. After all, as more than one admissions director observed, California's high schools produce tens of thousands of graduates with HSGPAs exceeding 4.0, and if HSGPA were to become the dispositive assessment for UC admissions purposes, then the demographics of those students with the highest grades would become the demographics of the UC entering class. That might reduce diversity.

Because of the degree of uncertainty and because we found persuasive the argument made by directors of undergraduate admissions across the campuses that as used at the University, standardized test scores help to identify and admit students who are members of historically excluded groups and/or first generation, we do not recommend making the tests optional.

II. Research

The Task Force reviewed a sampling of the research, guided by a literature review prepared to update a 2007 investigation by BOARS. That literature review is included as one of the appendices to this report. In the following paragraphs we summarize our conclusions based on our review of studies of the effects of making the standardized tests optional for applicants.

Scholarship on the effects of going "test optional" is mixed. The most recent study we reviewed (Saboe and Terrizi, 2019[138]) examined the effects of test-optional admissions on retention and graduation rates at all four-year, public and private, not-for-profit institutions granting BA degrees in the period from 2009 to 2014. This sample included 1,649 colleges that required SAT scores and 127 that had a test-optional policy. The authors report that they "cannot conclude that SAT optional policies have a significant effect on any of our measured outcome variables." They do not find an increase in diversity and find only "limited evidence" that test-optional policies boost applicant volume and improve retention. Similar results are reported by Sweitzer, Blalock, and Sharma, who used a different analytical method attempting to compare liberal arts colleges that switched to test-optional to such institutions that did not. (Sweitzer, Blalock, and Sharma 2018, in Buckley et al., eds.[139])

Syverson, Franks, and Hiss analyzed a sample of 28 colleges and universities including large and small, public and private, nonprofit institutions of varying degrees of selectivity. (Syverson et al. 2018[140]). This study updated prior work on the effects of going test-optional, published in 2014 (Franks and Hiss,

[138] Saboe, M. & Terrizzi, S. (2019). SAT optional policies: Do they influence graduate quality, selectivity or diversity? Economics Letters, 174, 13-17.
[139] Sweitzer, K, Blalock, A, and Charma, D. (2018). [The Effect of Going Test-Optional on Diversity and Admissions.] In Buckley, Jack, Letukas, Lynn, and Wildavsky, Ben, Eds,. Measuring Success: Testing, Grades and the Future of College Admissions (p. 164). Baltimore: Johns Hopkins University Press.
[140] Syverson, S, Franks, V, & Hiss, W. (2018). Defining Access: How Test-Optional Works. Retrieved from NACAC website: https://www.nacacnet.org/globalassets/documents/publications/research/defining-access-report-2018.pdf.

2014[141]). Like Saboe and Terrizi, they find that test-optional admissions increases the number of applications received, but unlike Saboe and Terrizi, Syverson et al. report an increase in the representation of URM students in both the applicant pool and in the entering class. However, three of the four public universities in the sample experienced less growth in URM enrollment than a control group consisting of peer institutions that were not test optional.

Zwick adds variables to the discussion in offering a critique of the Franks and Hiss study, arguing that outcomes differed for students who did and did not submit SAT scores in at least one way not captured by the study: students who did not submit were more likely to pursue majors in non-STEM fields. (Zwick 2018, in Buckley et al., eds.[142]). Zwick criticizes the Franks and Hiss study for failing to offer findings to support the claim that test-optional admissions leads to increased enrollment of URM students. In describing the uncertain effects of going test-optional, Zwick also notes that the family resources devoted to improving performance on standardized tests would presumably be redeployed to improve any other aspect of an applicant's performance necessary to bolster the odds of admission.

The student-level review by Syverson et al. also permitted additional observations about the students who did submit SAT scores ("submitters") and those who did not ("nonsubmitters"). Nonsubmitters were more likely to be first-generation, women, members of historically excluded groups, and/or recipients of federal Pell grants. This study also found that although nonsubmitters were admitted at lower rates, they were more likely to enroll if admitted. Graduation rates at a small, nonprofit institution were higher for underrepresented minority students who were nonsubmitters relative to underrepresented minority students who did submit scores. The study also found that at a large, public institution, undergraduate GPAs of submitters were slightly higher than those of nonsubmitters but that graduation rates were similar.

Belasco, Rosinger, and Hearn analyzed the effects of going test-optional but their study included only relatively small, liberal arts colleges. (Belasco, Rosinger, & Hearn 2014[143]). Their panel covered classes at 180 selective institutions between 1992 and 2010, and the authors found that the colleges that went test-optional did not increase the share of URM or Pell grant-eligible students enrolled, relative to colleges that continued to require the SAT. However, test-optional colleges did report higher SAT scores over time (representing the scores of those applicants who did submit scores). These findings are largely consistent with the conclusions of a series of simulations developed by Espenshade and Chung, who analyzed the possible effects of going test-optional on racial and socioeconomic diversity. (Espenshade and Chung, 2012[144]). They found that a test-optional admissions regime could increase the number of URM and poor and working-class students enrolled. The simulations illustrated that this could be achieved by expanding the share of URM students in the applicant pool and/or by increasing the share of students with lower test scores in the applicant pool. However, the effects are smaller at public universities than at private, nonprofit institutions. This study, while helpful in disaggregating the admissions process used in the simulations, uses the most cautious language in offering predictions: "[G]iven the great variety of postsecondary institutions in the United States – variety in terms of size, endowment, mission, geography, degree of selectivity, and whether private or public, among others – it is difficult to generalize to all

---

[141] Hiss, W, Franks, V. (2014). Defining Promise: Optional Standardized Testing Policies in American College and University Admissions. Retrieved from NACAC website: http://www.nacacnet.org/research/research-data/nacac-research/Documents/DefiningPromise.pdf.
[142] Buckley, Jack, Letukas, Lynn, and Wildavsky, Ben, Eds,. Measuring Success: Testing, Grades and the Future of College Admissions. Baltimore: Johns Hopkins University Press, 2018.
[143] Belasco, A. S., Rosinger, K. O., & Hearn. J. C. (2014). The Test-Optional Movement at America's Selective Liberal Arts Colleges: A Boon for Equity or Something Else? Educational Evaluation and Policy Analysis, pp. 1-18.
[144] Espenshade, T. J., and Chung, C. Y. (2012). [Diversity outcomes of test-optional policies.] In J. A. Soares (Ed.), SAT wars: The case for test-optional admissions (p. 177). New York: Teachers College Press. Retrieved from https://www.princeton.edu/~tje/files/Espenshade_Diversity_outcomes_chapter_SAT%20Wars_Soares.pdf

campuses about the likely consequences of a move to a test-optional admissions policy." We share this conclusion.

III. Summary

Multiple studies found that one effect of going test-optional was to increase the number of applicants. This is not a priority for the University at this time. After all, while there is much concern about UC admissions, it does not arise because critics believe the pool of applicants is too small. We anticipate that although some advocates of reform to admissions will call for making SAT and ACT scores optional, we believe that such a move could have significant, unanticipated, and undesirable effects on the profile of matriculating classes.

The research reviewed did not go into great detail about the methods used by admissions officers. For example, the studies did not explore how admissions offices evaluated applicants who submitted SAT scores relative to applicants who did not, nor about whether or how admissions offices imputed scores to applicants who chose not to provide them. These are significant questions if the admissions process is to be transparent and fair. We believe that our concerns could be possibly addressed by an appropriate process but thoughtful design would take time. At this point, we are unsure how best to design a fair, test-optional admissions process.

For these reasons, we do not recommend making the SAT and ACT optional for applicants for admission to UC at this time. The potential effects of making these tests optional deserve considerably more research than has yet been conducted, certainly at a university the size of UC. We suggest that UC undertake such study.

VII. Recommendations of the Task Force

The Task Force recommends the following reforms.

Potential Reforms that the University of California ("UC" or "the University") Could Implement in the Near Term

[The Task Force] recommends that in the near term BOARS consider reforming the UC admissions process by updating the UC eligibility criteria, with the goal of reducing reliance on test scores, maintaining or improving diversity, while ensuring admitted students are prepared. In this section we recommend that BOARS consider two changes to the UC Eligibility Criteria.

Background

There is a long history at the University of California of reconsidering the respective role of grades and test scores. Most recently, in 2012 UC implemented a number of reforms to its admissions procedures. Overall, these reforms increased the influence of GPA relative to test scores in determining the eligibility pool for UC admissions. The reforms increased emphasis on the Eligibility in the Local Context (ELC) approach in which the top-ranked students in each high school, ranked by GPA, become eligible to attend UC, subject to meeting the UC A-G course requirements and an overall GPA cutoff of 3.0. The reforms decreased emphasis on testing, first by dropping the SAT subject tests, and second, by reducing the proportion of the eligibility pool that would come from the "statewide context," which is based on students meeting various combinations of GPA and SAT or ACT score. As reported in a BOARS (2013) study of these reforms:

> The changes implemented for 2012 involved 1) elimination of the SAT subject tests as a requirement for eligibility; 2) a decrease from 12.5% to 9% of California high school graduates who were identified as eligible in the statewide context; 3) an increase from 4% to 9% of graduates who were identified as eligible in the local (within-school) context (ELC); and 4) the introduction of a new category of "entitled to review" (ETR) applicants who are assured a comprehensive review of their application at all campuses to which they apply, but who are not eligible for referral to another campus for admission. The 9% statewide and 9% local eligibility was expected to result in approximately 10.5% of California high school graduates being identified as eligible for referral to a campus with available space if they are not admitted to a campus to which they apply.[145]

The BOARS analysis indicated that in the first two years of the program, the eligibility pool was slightly larger than predicted. Significantly, academic preparation of eligible students remained strong while diversity improved somewhat.

Given the Task Force's mandate to study the ways in which tests are used in the admissions process at UC, it seems natural to explore the possibility of further reforming the relative number of students who become eligible through the "statewide" versus "local" contexts.

---

[145] Impact of the New Freshman Eligibility Policy at the University of California. A Report to the Regents of the University of California,
https://senate.universityofcalifornia.edu/_files/committees/boars/Nov52013BOARSReporttoRegents-Final.pdf

Recommendation 1. Review and update components of the statewide eligibility index.

The University *admits* students to campuses based on fourteen factors but bases *systemwide eligibility* on only two, HSGPA and test scores. We recommend that BOARS consider whether, among the 14 factors, we might identify additional quantitative features that could be included in the statewide eligibility index. For instance, improvement in GPA over the course of high school is already a factor that campuses use in admissions. Looking for students who improve their achievement over time could help UC find students on an upward trajectory.

Recommendation 2. Expand Eligibility in the Local Context.

The Task Force recommends that BOARS consider reforming the UC admissions process by adjusting the relative importance of the "statewide" versus "local" eligibility criteria. Specifically, BOARS should carefully evaluate expanding the Eligibility in the Local Context (ELC) pool by admitting more than the top 9% of students in each high school, when ranked by GPA. The Task Force recommends leaving the statewide index formula unchanged but expanding ELC for two reasons. First, it would be divisive to exclude some prepared students who would have been admitted under the current statewide eligibility rules, but who would have been excluded as cutbacks were made to the size of the statewide pool to accommodate replacements from the ELC pool. Second, defining and defending growth in the undergraduate population that would come from expanding ELC over a period of several years would send a signal to Californians about the high value that UC places on becoming more representative of the state over time.

In expanding the ELC cutoff beyond the top 9% of students, the Task Force recommends that BOARS take the following factors into consideration:

1. Any change to ELC or statewide eligibility should be implemented judiciously, with the goal of preventing the overall eligibility pool from contracting or expanding in dramatic ways. This is important because sudden large increases in enrollment would lead to shortages of spaces in classrooms and dormitory beds, while on the other hand sudden large decreases in enrollment would needlessly reduce access to postsecondary education at UC.

2. If the size of the eligibility pool is kept constant, then expanding ELC will necessarily displace some students who would have been "statewide" eligible under the current system. This would be unfortunate and politically divisive. A superior approach would be for UC to expand enrollment over time, dedicating new slots in the eligibility pool to the expanded ELC program, while leaving the current statewide eligibility criteria unchanged. We note that such growth would be consistent with an additional goal, beyond our purview but subject of popular discussion, of increasing the population of college-educated Californians.

3. In performing simulations of possible reforms, a factor to consider is the likely acceptance rates from students who would become eligible or ineligible for admissions to UC. This analysis of the yield rate will be crucial to control the overall enrollment levels at UC.

4. Any move to expand ELC while keeping the size of the "statewide" eligibility pool constant should be accompanied by a vigorous request for full funding by the legislature of the new slots to support the overall growth in enrollment that would result.

5. BOARS should monitor the academic progress of the pool of students admitted specifically through any expansion of ELC.

6. One practical issue that UC already faces with its 9% ELC program is that over time the share of high school seniors who satisfy this requirement can change. Several school districts including the two largest districts in California have recently made completion of A-G courses a graduation requirement (although these districts have allowed those with D grades to graduate, while UC requires grades of C.) We believe that BOARS will need to monitor A-G completion rates on an annual basis. If ELC is expanded, this could incentivize more students to complete the A-G courses with the required grades. This would be a desirable outcome, given that failure to complete the A-G coursework currently excludes many students. But such an increase in numbers makes annual monitoring of yields from the program even more crucial.

Likely Impacts on the Student Eligibility Pool: Some Examples

This Task Force had neither the time nor the resources to conduct detailed simulations. But with the assistance of key staff at UCOP, we have learned of the results of some early simulations, which are quite informative. Below we present results from three variants of one simulation that in turn reduce enrollment, keep enrollment constant or expand enrollment in a way designed to maintain the number of students who were admitted to UC by "Index Only," that is, through the statewide eligibility criterion. This is just one of many examples that BOARS could consider.

We are grateful to staff at UCOP for performing the simulations outlined below.

1) A Simulation that Would (Inadvertently) Lower Enrollment

At one extreme, in determining eligibility one could rely solely on ELC and cease to use the statewide eligibility that focuses on test score and GPA combinations. One such simulation examined the impact on the eligibility pool of expanding ELC to admit students in the top 12.5% of their classes and suspending the use of "Index" eligibility based on SAT and GPA data. The simulation, using applications for fall 2017 to fall 2019, shows the following:

1) The diversity of the newly eligible students is higher along several dimensions than that of students who were "Index Only" eligible.
2) However, many disadvantaged students and Underrepresented Minority Students (URMs) who were eligible based on the statewide SAT/GPA criteria would no longer have been eligible for admission under this system.
3) The high school GPA and SAT scores of those newly eligible would be both somewhat lower than for the students who were eligible under the Index.
4) This reform would have reduced the eligibility pool over the three years from 167,534 to 120,368, a drop of 28%, or about 47,000 students over three years. So, on average, about 16,000 students a year who were "Index Only" eligible would not be eligible under the simulated expanded ELC program.

2) A Simulation that Would Keep Enrollment Constant

An immediate conclusion from this simulation is that to keep enrollments at historical levels, the UC might have to extend ELC considerably beyond the top 12.5% of each class when ranked by GPA to meet its enrollment targets.

Alternatively, and more in line with what we propose, it would make sense for UC, if it chose to expand ELC from the top 9% to the top 12.5%, to continue to use the statewide GPA/SAT index to admit students, but to make the GPA/SAT criteria more stringent.

Of the 59,327 “Index Only” students who would not have been eligible under a top 12.5% ELC program over 2017-2019, UC would want to make up the 47,000 student shortfall over three years by making the top 47,000 of the 59,327 “Index Only” students eligible in the state context. This would mean taking roughly the top 80% of the “Index Only” students.

In other words, one feasible approach would be to expand ELC from the top 9% to the top 12.5% of each high school class while reducing the number eligible under statewide eligibility rules by about 20%.

In **Table 7-1** the left column shows the demographic makeup during 2017-19 under the existing 9x9 program, and the right column shows the projected makeup had UC expanded ELC to the top 12.5% and reduced the statewide eligibility index pool by the same number with the goal of keeping enrollment constant. In total, the ELC pool would have grown by 12,161 students and the statewide pool would have fallen by about the same number.

Table 7-1 The Demographic Makeup of the UC Eligibility Pool in 2017-19 under the Current Policy and under an Alternative Policy with the ELC Pool Increased to the Top 12.5% with Corresponding Drops in the Size of the Statewide Index Pool

	Current Policy (9x9)	**ELC 12.5% + Adjusted Statewide Index**
Ethnicity		
African-American	3.8%	4.0%
Native American	0.5%	0.5%
Asian	36.3%	34.6%
Latino	28.3%	31.3%
Pacific Islander	0.3%	0.3%
White	27.5%	26.1%
Unknown (U.S.)	3.3%	3.2%
First Generation College		
First Generation	35.5%	38.8%
Not First Generation	62.6%	59.4%
Unknown	1.8%	1.8%
Low-Income Status		
Low-Income	29.2%	31.9%
Not Low-Income	64.2%	61.8%
Unknown	6.6%	6.3%
Gender		
Female	56.3%	57.2%
Male	41.4%	40.6%
Other/unknown	2.2%	2.2%

The table shows moderate changes in the demographic mix of the eligibility pool, with growth of about 2% in share of the pool that is low-income, and growth of about 3% in the shares of students who are

Latino or first-generation. All of these changes represent a move towards better reflecting the diversity of California. In this particular simulation, some very good applicants who in 2017-19 would have become statewide eligible but were not eligible through the ELC would not have automatically become eligible for admission.

What about the academic preparation of those deemed eligible in 2017-2019 under the current 9X9 system, and that of those who would be eligible under a top 12.5% ELC system with an adjusted statewide eligibility index designed to keep overall eligibility constant? **Table 7-2** shows basic statistics on the GPA and SAT of those in the actual eligibility pools and under the simulation.

Table 7-2 Academic Preparation of the Actual 9 x 9 Eligibility Pool for 2017-2019 and the Simulated Eligibility Pool Using An Adjusted Statewide Index Eligibility and/or ELC 12.5% Eligible

	High School GPA			**SAT**		
Group	Mean	25th Percentile	75th Percentile	Mean	25th Percentile	75th Percentile
9 X 9 Eligible Applicants	4.05	3.90	4.23	1325	1220	1440
Simulated Index Eligible and/or ELC 12.5% Eligible	4.04	3.90	4.23	1310	1190	1440

Note: Simulation raises the rigor of the statewide eligibility index to keep the overall eligibility pool at historical levels.

In part because the 12,000 newly eligible students in the simulation represent about 7.3% of the historical eligibility pool, changes in academic preparation are modest. The table shows that mean high school GPA and SAT scores, and also their 25th and 75th percentile values, are quite similar under the actual and simulated scenarios.

Most students would continue to be eligible under both the adjusted statewide eligibility and the adjusted 12.5% ELC criteria. That said, it is important to recognize that the expanded ELC and the statewide eligibility index are selecting students in different ways. One way to see this is to compare those eligible either through statewide eligibility or by ELC, but not by both. **Table 7-3** below shows the academic preparation of the subsamples of the simulated eligibility pool who would have been admitted **only** through the adjusted statewide eligibility criterion or admitted **only** by the expanded 12.5% ELC criterion.

Table 7-3 Academic Preparation in the Simulation of Those Who Would Have Been Eligible Only Through the Adjusted Statewide Index or Only Through the 12.5% ELC Criterion

	High School GPA			**SAT**		
	Mean	25th Percentile	75th Percentile	Mean	25th Percentile	75th Percentile
Adjusted Statewide Index Eligible Only	3.97	3.85	4.11	1397	1340	1460
ELC 12.5% Eligible Only	3.82	3.68	4.00	1083	1010	1160

The table shows that those who would have been eligible only through the adjusted statewide eligibility index on average had high school GPAs of 3.97 compared to 3.82 for those who would have been eligible only under the 12.5% ELC option. The corresponding numbers for mean SAT score were 1397 and 1083. Similar gaps emerge when we look at the 25th and 75th percentiles.

3) *A Simulation that Would Increase Enrollment, By Expanding the ELC Pool While Maintaining the Size of the Statewide Index Pool*

Some may express concerns that in the above simulation, the only way to keep the eligibility pool constant was to exclude some students who were eligible in 2017-2019 under the existing statewide eligibility index.

Another approach would be to maintain the same rules for statewide eligibility that currently exist, while increasing the ELC pool and thereby overall UC enrollment. This would have left the number who were statewide eligible constant over 2017-19, while the ELC eligibility pool would have increased. Overall, the total eligibility pool over the period would have grown by the projected growth in the ELC pool, from 167,534 to 179,695, or an increase of 7.3%. Thus, this approach implies UC growing its enrollment by about 7.3% to accommodate the ELC growth.

This growth in total enrollment would of course occur over four years, because new cohorts of freshmen would gradually take the place of smaller earlier cohorts. The increase is certainly meaningful, but not unprecedented. Between fall 2009 and fall 2019, for example, UC undergraduate enrollment rose from 177,453 to 226,125, a rise of 27.4% across a decade.[146]

We emphasize that we believe BOARS will need to take the lead in deciding whether expanding the ELC percentage is advisable and if so by how much. The simulation referred to above provides one of many possible ways of expanding ELC, and does not constitute a formal recommendation. Rather, we intend it as a useful illustration of likely consequences for one specific example of such an expansion. *That said, the Task Force is recommending a meaningful increase.*

Analysis of UC admissions steps that may contribute to disparities in access.

[146] Calculations based on total fall undergraduate enrollment obtained from https://www.universityofcalifornia.edu/infocenter/fall-enrollment-glance.

Our preliminary analysis indicates that the disparities along lines of race, ethnicity, and socioeconomic status evident in the undergraduate population of the University are a function of multiple factors, including but not limited to the SAT and ACT. For example, we note that disparities in access to and completion of A-G courses account for a disproportionate lack of UC eligibility for students who are members of underrepresented groups. Differences in A-G completion rates across schools also significantly reduce the yield of students through ELC. This research matters because long-term solutions to promote better representation of historically underrepresented groups will only be achieved by understanding the multiple factors that impede better representation in UC.

Recommendation 3. Further analyze factors contributing to disproportionate representation.

The Task Force strongly recommends further analysis of the admissions process to identify those points in the admissions process, and the factors at work at each such point, that contribute to disproportionate declines in representation of students who belong to populations that have historically been excluded from higher education opportunity at UC.

Recommendation 4. Study and expand student academic support services.

Background

This report has established that both the existing admissions tests and high school GPA have the ability to predict undergraduate success along a number of important dimensions. The data analyses underlying that work necessarily used historical data, and are backward-looking.

Of course, though, history need not be repeated if institutions evolve. What if UC could find ways to improve undergraduate outcomes, especially for those who enter UC with lower SAT scores and HSGPAs? This could lessen the predictive power of admissions tests and grades by virtue of improving outcomes for at-risk students. Some Task Force members wondered whether UC as a whole could do a better job than it already does supporting undergraduates. Supporting students in their freshman and sophomore years may be especially crucial, as this is when many students are still adjusting to the rigors of postsecondary education, and the emotional stresses related to moving away from home, dealing with the financial burdens of postsecondary education, and making major decisions about career paths. Ideally, then, undergraduate support programs should provide both academic and socio-emotional support.

This line of thinking quickly led us to the outlines of a proposal i) to catalog existing support programs and what we know about these programs' effectiveness, ii) to evaluate these support programs rigorously, and iii) to design and evaluate new support programs as needed. The ultimate goal of this work would be to improve and expand the system of support programs at UC, in the process materially improving the undergraduate success of admits whose high school academic profiles suggest they could struggle at UC. The proposal appears below, followed by a discussion of some potential pitfalls and opportunities.

The Task Force recommends a multi-pronged effort to study, fine-tune and expand the system of academic and socio-emotional supports that UC provides to its undergraduate students. The efforts should begin by identifying and learning about existing programs, and then proceed quickly to a second phase that would evaluate existing programs and fine-tune these programs as needed. In the third phase, UC would design and evaluate new support programs as needs are identified. As a whole, this venture would involve the steps listed below.

Phase 1

1) Systematically catalog the support programs, both academic and socio-emotional, that UC campuses currently provide to undergraduates, especially to those who are predicted to be less likely to succeed, based on GPA and SAT scores.
2) Learn what UC staff already know about the impact of these programs on student outcomes. In addition to looking at formal and informal evidence of overall impacts, look for evidence about whether there exist specific groups of students who gain more from a given program.

Phase 2

3) Based on what is learned in Phase 1, conduct formal evaluations of the programs that campuses view as the most important. Goals would include evaluation of i) overall student impacts and ii) identification of student subgroups that benefit relatively more or less from a given program.
4) Using the evidence from step 3), expand the most successful programs, and if warranted by evidence target some programs quite selectively towards student subgroups that benefit the most from specific programs.

Phase 3

Phase 2 is likely to find gaps in what support programs currently provide to students. Where unmet needs are identified, UC should develop new programs. This phase should include the following elements. Notably, the first suggestion below could be carried out as early as Phase 1.

5) Survey the literature on academic and socio-emotional support programs to identify new types of support programs that hold promise.
6) Based both on this literature review and what Phases 1 and 2 reveal about unmet needs at UC, develop new support programs, and evaluate them.
7) Using the evidence from step 6), maintain and possibly expand the new programs that prove successful, while modifying or curtailing the new programs for which evidence fails to find a benefit for student outcomes.

For all the evaluation efforts, UC should use rigorous methods such as randomized controlled trials, or strong quasi-experimental methods if necessary. These quantitative evaluation tools should be combined with qualitative methods that use interviews, focus groups and related methods to learn directly from student participants and program staff which aspects of a program work and which need re-thinking.

Discussion

Funding — Funding for the support programs would need to come from UC. There is a strong possibility that some of the evaluation work, especially if it involved randomized controlled trials, could be funded through grants from the U.S. Department of Education, given that its Institute of Education Sciences (IES) has announced last year and again this year that a major new area for research it is willing to fund involves the development and rigorous evaluation of reforms designed to improve postsecondary education.

Buy-In from Campuses — It would take support from throughout the administration of each campus for this to work, as well as careful and diplomatic engagement with the administrators of existing support programs.

Research Expertise — There exist numerous ladder rank faculty with expertise in education and either quantitative or qualitative social science research. In addition, it seems possible that some Lecturers with Security of Employment, depending on their area of expertise, might find some aspect of the program development and evaluation work to be exciting and relevant to the creative activity parts of their duties. UC may also need to make a number of strategic faculty hires to help design evidence-based support programs evidence wherever possible. For example, there has been a great deal of recent work by psychologists on socio-emotional supports such as mindfulness based interventions and their impact on mental health and sense of well-being, including in school- and college-age populations.

Evaluation of New versus Existing Programs — Evaluating new or existing but limited programs is easier than evaluating full-scale programs because in the former cases, it is easier to develop application lists and randomize who enters the programs from the application lists, thereby creating a valid comparison group as part of a Randomized Controlled Trial (RCT). Large support programs with almost universal reach may not be amenable to evaluation this way. One possibility in this latter case is that if large programs use HSGPA or SAT/ACT scores to determine eligibility cutoffs, a strong quasi-experimental design known as Regression Discontinuity could be used to choose a valid control group, by comparing those just barely eligible to participate with other students who are on the other side of the eligibility cutoff. Both of these approaches, RCT's and Regression Discontinuity, have typically been reviewed positively in past IES grant competitions.

Relevance — It may strike some readers as odd that a Task Force on admissions testing would be concerned about improving supports for undergraduate students. But the link is clear – our review of evidence shows variations in undergraduate performance that can be predicted by numerous measures, including grades and test scores, and these variations can reflect inequalities in school systems and society more broadly. Improved support systems to help students UC admits who are predicted to have a lower chance of succeeding in college overcome challenges would lead to a larger and probably more diverse population of college graduates in California. At the same time, it could lessen somewhat the need for UC to use admissions tests and grades heavily in the admissions decision.

Updated Analysis of Current Standardized Tests

Recommendation 5. Obtain data to perform item-level analysis of current standardized tests.

The Task Force strongly recommends that UC obtains the necessary data and test items to perform an updated analysis of current standardized tests. This recommendation contains two distinct but interconnected pieces. The first piece is an updated differential item functioning (DIF) analysis, conducted with more modern methods, e.g., with explanatory models (De Boeck & Wilson, 2004), using more recent data. The second piece involves a two-phased feature analysis of the test items. The qualitative review of items is combined with a quantitative analysis of student item response data using the extracted features that may be related to cognition, instruction, prior preparedness, and language to better understand the disparities in student performance on current standardized tests.

DIF Analysis

First, to refresh prior studies that were based on data from two decades ago, UC should request the most recent operational student-level data from College Board and ACT in the form of raw item responses, as well as necessary applicant and school background information as they are available. With data, the UC can provide an unbiased and definitive answer to the question of whether the tests and the items making up the tests are indeed biased against subgroups in ways that the assessment community do not fully appreciate or understand. This is a legitimate concern that can be immediately addressed with currently available statistical and psychometric techniques, given the extensive revisions made recently to the SAT, as well as the age and relevance of currently available research

Feature Analysis

Feature analysis is an integrated process involving the qualitative rating of assessment items/tasks against a set of attributes (along the lines of content, language, and cognition) in a particular context (UC-level college readiness), followed by a subsequent quantitative analysis to determine how these attributes contribute to or subtract from successful task performance.

As described by Baker et al. (2015)[147], one of the earliest references to the idea of feature analysis can be attributed to Gordon (1970) in the Report of the Commission on Tests: II.[148] Gordon mentioned qualitative analysis of assessments to emphasize "description and prescription," that is, the qualitative description of cognitive functions leading to the prescription of the learning experiences required to more adequately ensure academic success. Gordon suggested that existing instruments can be examined with a view towards categorization and interpretation to determine whether data can be reported in qualitative ways, in addition to traditional quantitative ways. For instance, Gordon mentioned that response patterns can be reported differently for information recall or vocabulary. He also referred to features such as problem solving, expression, and information management, among many others.

To accomplish this, UC should also request the items/tasks themselves from College Board or ACT. The items should be rendered in a human-reviewable form containing complete information as they were presented to the students in operational administration, i.e., prompts, texts, graphs, response options, etc. The results of this study will provide immediate answers to where the subgroup performance differences on the standardized tests may come from. The results of the analysis may also inform UC in the development efforts related to Recommendation 6 immediately below.

Longer Term Reform: Development of a New Assessment System

The STTF believes that it is possible to create fundamentally different kinds of assessments that satisfy all BOARS principles on testing, while retaining predictive power, providing improved feedback to students and teachers, and broadening the demographic range of students UC is able to admit. Over the longer term, we believe that the University should study the development of a completely new suite of high school student assessments by "designing in" the principles developed by BOARS that govern UC admissions, instead of attempting to "retrofit" any existing assessment programs under those principles. UC should commit to evaluating the fairness and predictive power of any newly developed assessments that satisfy UC testing principles even if they do not resemble current admissions tests in content, format, administration, and the number or types of scores available for use by UC to gauge an applicant's achievement or potential.

---

[147] Baker, E. L., Madni, A., Michiuye, J. K., Choi, K., & Cai, L. (2015). Mathematical Reasoning Project quantitative analyses resultsQuantitative Analyses Results: Grades 4, 8, and 11. Los Angeles: University of California, Los
Angeles, National Center for Research on Evaluation, Standards, and Student Testing. Available at http://www.smarterapp.org/documents/FeatureAnalysis_2015-09-25.pdf

[148] Report of the Commission on Tests: II. Briefs. College Entrance Examination Board, New York, NY. (1970). Available at https://files.eric.ed.gov/fulltext/ED045704.pdf

Recommendation 6. Develop a new assessment that will be continuously accessible to students and that will assess a broader array of student learning and capabilities than any of the currently available tests.

This new suite of assessments should provide more feedback to schools and to students, and may enable UC to admit classes of students more representative of the diversity of the state.

1. Background

Current standardized assessments measure a relatively narrow slice of academic achievement or potential, late in an aspiring applicant's high school career, with minimal feedback and time for corrective action even if desired. A new system should allow the longitudinal measurement and formative development of a broader range of skills and attributes, related to an applicant's success at UC, reflecting the principle that the University "…should select those students who demonstrate a strong likelihood that they will persist to graduation." [149]

In particular, based on the research reviewed by the Task Force, we believe that UC will be able to articulate a set of skills and attributes that 1) can be measured in a uniform and quantitative way, 2) are likely to be predictive of success at UC, and 3) could potentially show smaller disparities than current measures along the lines of race, ethnicity, and socioeconomic status. **Of course, the tests should continue to evaluate math and reading achievement due to their foundational importance in college and career success.** The formats of the tasks on the new assessments can and should also take full advantage of technological advances to let the aspiring applicants exercise a wider range of knowledge, skills, and reasoning than multiple-choice questions do.

2. Can We Build These Measures?

So far a number of large-scale assessments have explored technology-based performance assessments in low-stakes settings. The most prominent examples include NAEP Problem Solving in Technology-Rich Environments[150] and NAEP Science Interactive Computer Tasks[151] under the National Assessment of Educational Progress (NAEP) program, and PISA 2012 Creative Problem Solving Tasks[152] and PISA 2015 Collaborative Problem Solving Tasks[153] under OECD's Program for International Student

[149] Guidelines for Implementation of University Policy on Undergraduate Admission, https://senate.universityofcalifornia.edu/_files/committees/boars/documents/guidelines-implementation-of-ug-admission-rev-7-2019.pdf

[150] Bennett, R. E., Persky, H., Weiss, A. R., & Jenkins, F. (2007). *Problem solving in technology-rich environments: A report from the NAEP Technology-Based Assessment Project* (NCES 2007–466). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

[151] National Center for Education Statistics (NCES). (2012). *The nation's report card: Science in action: Hands-on and interactive computer tasks from the 2009 science assessment* (Report No. NCES 2012–468). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

[152] Organisation for Economic Co-operation and Development (OECD). (2014). *PISA 2012 Results: Creative problem solving: Students' skills in tackling real-life problems (Volume V)*. Paris: OECD Publishing. Retrieved from: http://dx.doi.org/10.1787/9789264208070-en

[153] Organisation for Economic Co-operation and Development (OECD). (2017). *PISA 2015 Results: Collaborative problem solving (Volume V)*. Paris: OECD Publishing. Retrieved from: https://dx.doi.org/10.1787/9789264285521-en

Assessment (PISA). Another related project is SimScientists,[154] [155] which represents the collaboration of six states to study simulation-based science assessments. The results of these projects suggest some promising directions that the proposed new development can take.

NAEP Problem Solving in Technology-Rich Environments and NAEP Science Interactive Computer Tasks feature distinct attempts at building fundamentally different kinds of assessment experience. It is not a coincidence that NAEP chose to examine interactive problem solving in their entry into the development of a newer, fully-digitized generation of assessment products that will contribute to our Nation's Report Card. One out of 10 adults, averaged across 24 countries, reported encountering at least one non-routine, complex problem per day requiring 30 minutes or more to solve.[156]

Note that these examples illustrate the promise of *large-scale* technology-based performance assessments. Numerous *small-scale* studies around technology-based performance assessments have been studied beginning in the 1990s – often by UC researchers. For example, Baker led a research program on the use of computer-based concept mapping to measure students' conceptual understanding. Students created concept maps (networks) to represent their understanding of how concepts were related, and also wrote essays on the same topic explaining how the concepts were related. Concept mapping was tested in numerous studies with different content and with students spanning elementary to adults,[157], [158] with statistically significant relations found between concept map scores and essay scores (r = .4 to .7). Stevens[159] developed interactive multi-media exercises (IMMEX) and its use as a Web-based simulation with middle and high school students. While intended for instructional use, IMMEX also served assessment purposes. Cognitive process validation showed that students' cognitive processes relating to their simulation performance, with productive reasoning processes positively related to successful simulation performance (r = .7) and unproductive processes relating to unsuccessful simulation performance (r = .5).[160] These are two concrete examples of research conducted by UC faculty and intended to illustrate the maturity and predictive power of carefully designed technology-based performance assessments. For a review of technology-based performance assessments and issues in general, see Clauser, Margolis, and Clauser (2016)[161] and Katz, LaMar, Spain, Zapata-Rivera, Baird, and Greiff (2017).[162]

---

[154] Quellmalz, E. S., Timms, M. J., Silberglitt, M. D., & Buckley, B. C. (2012). Science assessments for all: Integrating science simulations into balanced state science assessment systems. *Journal of Research in Science Teaching*, *49*, 363-393.
[155] Quellmalz, E. S., Davenport, J. L., Timms, M. J., DeBoer, G. E., Jordan, K. A., Huang, C.-W., & Buckley, B. C. (2013). Next-generation environments for assessing and promoting complex science learning. *Journal of Educational Psychology, 105*, 1100–1114. doi:10.1037/a0032220
[156] Organisation for Economic Co-operation and Development (OECD). (2013). *OECD skills outlook: 2013: First results from the survey of adult skills*. Paris: OECD Publishing. Retrieved from: http://dx.doi.org/10.1787/9789264204256-en
[157] Chung, G. K. W. K., Harmon, T. C., & Baker, E. L. (2001). The impact of a simulation-based learning design project on student learning. *IEEE Transactions on Education, 44*, 390–398.
[158] O'Neil, H. F., & Chung, G. K. W. K. (2011, April). *Use of knowledge mapping in computer-based assessment*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
[159] Beal, C. R., & Stevens, R. H. (2010). Improving students' problem solving in a web-based chemistry simulation through embedded metacognitive messages. *Technology, Instruction, Cognition & Learning, 8*, 255–271.
[160] Chung, G. K. W. K., de Vries, L. F., Cheak, A. M., Stevens, R. H., & Bewley, W. L. (2002). Cognitive process validation of an online problem solving assessment. *Computers in Human Behavior, 18*, 669–684.
[161] Clauser, B. E., Margolis, M. J., & Clauser, J. C. (2016). Issues in simulation-based assessment. In F. Drasgow (Ed.), *Technology and testing: Improving educational and psychological measurement* (pp. 49–78). New York, NY: Routledge.
[162] Katz, I. R., LaMar, M. M., Spain, R., Zapata-Rivera, J. D., Baird, J. A., & Greiff, S. (2017). Validity issues and concerns for technology-based performance assessment. In *Design Recommendations for Intelligent Tutoring System-Volume 5: Assessment Methods* (pp. 209–224). U.S. Army Research Laboroatry.

We propose that the UC look into the development of simulations and realistic performance tasks that can assess creative problem solving, inductive and deductive reasoning, and analytical capability because these attributes are hallmark features of what has come to be expected out of higher education. Macro trends such as automation and the shift to knowledge-based economy have generally favored careers that involve nonroutine tasks, and the attributes listed above are among those that are required in this new economic reality.[163]

To build the new assessments, it is important to follow assessment development principles under the Evidence-Centered Design (ECD) framework[164]. ECD has become the de facto assessment design process, with its core principles developed and refined over 20 years of research in numerous assessment contexts,[165], [166], [167], [168], [169], [170], [171] including large-scale standard testing, English proficiency testing, performance assessments, simulations, and games. The essential design element in ECD is to first establish a good representation of the construct being measured. The construct in this case is the set of academically relevant and perhaps even disciplinary-specific knowledge, skills, and attributes (KSAs) that the UC deems important for an applicant's success as a UC undergraduate. For the new assessment instruments to be useful in UC admissions, the development should begin with the creation and vetting of these KSAs with strong input from its faculty and students.

Then the assessment developers, led by UC faculty, must decompose the construct into increasingly more refined and measurable elements. These elements may be interconnected and one may depend on another. This results in a graphical representation of the domain – complete with nodes and links that illustrate the theory of what underlies the construct. Then the task of the assessment designer becomes that of finding ways that will elicit from students the evidence that can be used to tap into the various elements and ultimate to form an overall measure. This design approach results in an explicit, external representation of the target content and skills that can be reviewed by various UC stakeholders. In addition, this external representation, if available to the public, can serve as roadmap of what knowledge and skills are expected of UC applicants. We believe an assessment constructed this way would predict outcomes for UC matriculants better than current tests because the assessment will have specific UC needs built in.

Another benefit of designing the new assessment in this manner is that rather than giving feedback to students in the form of a few numbers aimed at refined differentiation of individuals, detailed feedback can be made available to students on their performance in different areas, for example the processes

---

[163] Autor, D. (2010). The polarization of job opportunities in the U.S. labor market: Implications for employment and earnings. Washington, DC: Center for American Progress and The Hamilton Project.
[164] Mislevy, R. J. (2013). Evidence-centered design for simulation-based assessment. *Military Medicine, 178* (supplement 10), 107–114.
[165] Mislevy, R. J. (1994). Evidence and inference in educational assessment. *Psychometrika*, *59*, 439–483.
[166] Mislevy, R. J. (1996). Test theory reconceived. *Journal of Educational Measurement, 33,* 379–416.
[167] Mislevy, R. J. (2018). Sociocognitive Foundations of Educational Measurement. New York, NY: Routledge.
[168] Mislevy, R. J., Almond, R. G., Yan, D., & Steinberg, L. S. (1999, July). Bayes nets in educational assessment: Where the numbers come from. In *Proceedings of the fifteenth conference on uncertainty in artificial intelligence* (pp. 437–446). Burlington, MA: Morgan Kaufmann.
[169] Mislevy, R. J., Corrigan, S., Oranje, A., DiCerbo, K., Bauer, M. I., von Davier, A., & John, M. (2016). Psychometrics and game-based assessment. In F. Drasgow (Ed.), *Technology and testing: Improving educational and psychological measurement* (pp. 23-48). New York, NY: Routledge.
[170] Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice, 25,* 6–20.
[171] Haertel, G., Wentland, E., Yarnall, L., & Mislevy, R. J. (2012). Evidence-centered design in assessment development. In *Handbook on Measurement, Assessment, and Evaluation in Higher Education* (pp. 277–296). New York, NY: Routledge.

within problem solving, or within reading and math. An example is PISA 2012, wherein the core problem solving processes are defined as: (a) Exploring and understanding; (b) Representing and formulating; (c) Planning and executing; and (d) Monitoring and reflecting. Significant individual and group differences were found across these problem solving processes. More nuanced feedback not only would give students diagnostic evidence on specific areas in which the student needs to improve over time, but once admitted to the UC this information would allow campuses to provide tailor-made supports to students rather than generic remedial classes that are inefficient and stigmatizing.

3. How Do Scores Relate to Student's Background Characteristics?

The PISA 2012 Problem Solving assessment offers an example of why we think alternatives to traditional standardized test may yield fewer disparities along the lines of race, ethnicity, or socio-economic status. In the PISA 2012 Problem Solving assessment, items were designed to focus on measuring distinct problem solving processes. Generally speaking, the data suggest that there is substantial variation in problem solving skills both within and between schools. On average across participating Organisation for Economic Co-operation and Development (OECD) countries, the variance in student performance that is observed within schools amounts to 61% of the OECD average total variance in student performance. The remainder is attributable to differences in student performance between schools. In the US, the decomposition of test score variance (for math or reading) into between and within school components generally results in around 75% of the variance attributable to the within school part. The ratio of variance unique to problem solving versus variance shared with core academic subjects assessed by PISA (reading, math, and science) is roughly 1:2 for both the within school and between school components, suggesting that problem solving is distinct from the core academic subjects. In other words, it is unlikely that one can recombine the academic subject tests to measure problem solving well. Perhaps unsurprisingly, the largest correlation in student performance is between math and problem solving (0.81); the smallest is between reading and problem solving (0.75). These correlations may appear large when standalone, but they are appreciably smaller than the correlation observed among student performance in the core academic subjects of reading, math and science (often above .85). Finally, the relation of socio-economic status with problem-solving performance is weaker than its relation with general academic proficiency in math, reading or science.[172]

In a more recent study, PISA administered Collaborative Problem Solving Tasks in 2015.[173] Unlike the assessments of the core PISA subjects, the PISA 2015 collaborative problem solving assessment is not a direct measure of individual differences, but the design of these performance tasks and the results provide useful guidance for the present proposal. Again, the relationship between socio-economic status and collaborative problem solving performance is weaker than its relation with performance on other academic subjects. On average across OECD countries that participated in the collaborative problem solving assessment, a one-point increase in students' socio-economic status (as measured by a standardized index used by PISA in all countries) is associated with a 13-point improvement in collaborative problem solving performance, compared to between 17 and 19 points in the three core PISA subjects. A one-point increase in schools' socio-economic profile is associated with a 59-point improvement in collaborative problem-solving performance compared to between 66 and 73 points in the three core PISA subjects. A number of countries in PISA 2015 had sizeable immigrant populations. There appears to be no significant relationship between immigrant status and the uniquely collaborative aspects of the problem solving assessment, after the relationship between diversity and socio-economic profile has been accounted for.

---

[172] Organisation for Economic Co-operation and Development (OECD), 2014, p 94.
[173] Organisation for Economic Co-operation and Development (OECD), 2017, p. 93.

SimScientists[174] [175] focuses on middle school science (Ecosystems, and Force & Motion). The project studied the psychometric quality, feasibility, and utility of simulation-based science assessments designed to serve teacher's formative assessment[176] needs during a unit and to provide summative evidence of end-of-unit proficiencies.[177] The goal is to test whether interactive assessments may be more effective than traditional, static assessments at discriminating student proficiency across 3 types of science practices: (a) identifying principles (e.g., recognizing principles), (b) using principles (e.g., applying knowledge to make predictions and generate explanations), and (c) conducting inquiry (e.g., designing experiments). The results suggest that static assessments are not as effective as interactive assessments for differentiating between factual knowledge and the ability to apply that knowledge in meaningful contexts. The interactive assessments were more effective at measuring students' ability to engage in inquiry practices.[178] Critics have contended that English Learners and students with disabilities may be overwhelmed by complex, dynamic simulations. In this study, evidence suggests that these populations of students were better able to show their skills and potential on the simulation-based assessments than on the conventional multiple-choice test items.[179]

Based on evidence so far, we expect that the new system being envisioned can be designed in such a manner that inequalities in socioeconomic status and availability of information or resources in K-12 education do not correlate as strongly with the results. This is because the new assessments aim to quantify a student's capacity to learn new material, rather than manifesting what the student has not had the opportunity to learn. For this reason, we hope and expect that UC will look into the development of this new set of assessments in place of currently used standardized tests.

4. Outstanding Issues

The following issues are technical problems that have to be solved before implementing these new assessments at a full scale, but we expect that they can be solved to UC's satisfaction over a development period of several years, followed by pilot testing and refinement.

a. *Age Range and Stakes*

The oldest students targeted in NAEP, PISA, SimScientist, and now TIMSS are 15 years old. These are also low-stakes assessments so the same potential criticism raised against the conversion of Smarter Balanced into a high-stakes assessment is applicable here. However, the new assessment system will be designed from the ground up for UC's admissions purposes. It can be designed to accommodate high stakes use and more expanded age range. On the other hand, retrofitting a statewide accountability test for the new purpose will be more difficult.

b. *Response Format*

Even for the state-of-the-art computer-based assessments, the tasks are "still quite structured and rigid, with complex forms of learning either being neglected or measured in ways that may not correspond to

---

[174] Quellmalz, E. S., Timms, M. J., Silberglitt, M. D., & Buckley, B. C. (2012).
[175] Quellmalz, E. S., Davenport, J. L., Timms, M. J., DeBoer, G. E., Jordan, K. A., Huang, C.W., & Buckley, B.C. (2013).
[176] In simple terms, formative assessments are intended to gauge a student's learning during ongoing instruction, so that effective feedback and intervention may be provided. Summative assessments, as the name suggests, provide a comprehensive summary of what student has achieved after a unit, a semester, or perhaps a year.
[177] Quellmalz, E. S., Timms, M. J., Silberglitt, M. D., & Buckley, B. C. (2012), p. 1
[178] Quellmalz, E. S., Davenport, J. L., Timms, M. J., DeBoer, G. E., Jordan, K. A., Huang, C.-W., & Buckley, B. C. (2013), p, 1112
[179] Quellmalz, E. S., Timms, M. J., Silberglitt, M. D., & Buckley, B. C. (2012), p. 387

real-life situations."[180] (p. 167). Many task formats, in the end, use multiple-choice items to collect data. Students can manipulate variables in a simulation and see how the system changes as a result, but then a prompt is presented and students are asked to choose an answer. This is generally true for the NAEP science tasks and for SimScientist. The range of formats so far appears to be based on what is available in the standard UI/UX toolkit[181] [182]. Standard User Interface elements such as sliders, drop down, drag and drop, text input boxes, and so on are used. More custom response inputs are rare except in game-based measures or in concept mapping exercises where the drawing of a graph with nodes/edges are produced at the end of the task. Efforts should be made to expand the possible response formats, e.g., with touchscreens. This would require a close collaboration between item designers and computer programmers that many large-scale assessment programs cannot easily accommodate.

c. *Automated Scoring*

This is related to the problem of Response Format. How to make sense of process data (e.g., clicks or interactions with touchscreens) from interactive tasks to produce or supplement outcome scores remains a major topic of research. Scoring of open-ended responses appears to be a major bottleneck in efforts to scale up the assessments. It must be tackled so that tasks more similar to real-life situations can be created and used. Essay and speech scoring seem to be the only areas that received serious and sustained attention. Automated scoring needs to be explainable – statistical prediction alone is not sufficient. Black-box (prediction of human judged scores only) methods for high stakes purposes are problematic because they are not transparent. In a high-stakes setting, transparency helps with fairness. UC is home to foundational research on cognitive AI, and much of the recent development has resided within computer science. Assessment is one area in which UC's capacity in computer science may help bring needed change.

d. *Automated Generation of Assessment Content*

Automated task/item generation[183] [184] refers to the generation of tasks or items that all share underlying design features but different surface features, e.g., generating items with different coefficients with the same functional form in an equation solving task. Automated task/item generation is a necessity here because the new assessment system must be able to provide open access to practice material at any time by any student. The same automated engine will be used to generate the proctored assessment content, different from the openly available practice content. With always-on availability of these authentic practice content, students can and should be encouraged to try out the assessments for as many times as they desire. By doing this, students can inoculate themselves against any negative testing experience, obtain feedback about their preparation, or maintain skills and knowledge. This is only fiscally possible if enough of the content generation is automated. It also takes automated generation of assessment content to efficiently scale up the assessment system to meet the needs of the entire State's applicant pool (as well as allowing assessment of private and out of state students).

---

[180] O'Leary, M., Scully, D., Karakolidis, A., & Pitsia, V. (2018). The state-of-the-art in digital technology-based assessment. *European Journal of Education*, *53*, 160-175.
[181] Parshall, C., Harmes, J. C., Davey, T., & Pashley, P. (2010). Innovative Items for Computerized Testing. In W. J. van der Linden & C. A. W. Glas (Eds.), Elements of Adaptive Testing (pp. 215-230): Springer New York.
[182] Scalise, K., & Gifford, B. (2006). Computer-based assessment in e-learning: A framework for constructing" intermediate constraint" questions and tasks for technology platforms. The Journal of Technology, Learning and Assessment, 4(6).
[183] Embretson, S. E., & Kingston, N. M. (2018). Automatic item generation: A more efficient process for developing mathematics achievement items? *Journal of Educational Measurement*, *55*, 112–131.
[184] Gierl, M. J., & Lai, H. (2012). The role of item models in automatic item generation. *International Journal of Testing*, *12*, 273–298.

e. *Security and Authenticity*

Security for the versions of the assessment that are proctored and which are considered in the UC admissions process should be maintained. Student access to proctored assessment sessions should be increased, to allow a chance to take the test in more than one grade at the very least. Because we expect that the assessment system will yield information useful and predictive at other institutions of higher education, the scores will become as portable as current standardized tests.

5. A Notional Timeline

We do not make a recommendation now about who should govern, finance, develop, evaluate, maintain, and administer these assessments. An endeavor of this scale will not be inexpensive and will likely require the collaboration of a number of stakeholders. Experience also suggests that the startup and long-term success of a new assessment system such as the one proposed here will likely involve the public, private, and philanthropic sectors.

However, to ensure more focused consideration of the feasibility of the new assessment and its usefulness for UC admissions, we provide the following notional development timeline as a current best estimate. With the stakes involved, there will be considerable anticipation of the timely availability of this new assessment system. The timeline, though objectively and justifiably reasonable, given the scope and importance, may appear to be impracticably long to key observers and decision-makers. Some of us are concerned that momentum may be lost during the long development cycle without assurance from the University to initiate and to support its development and evaluation, not to mention a clear commitment to adoption of the new assessment system, once accepted, to replace current standardized tests. Members of the Task Force were divided on whether to recommend cessation of consideration of standardized test scores sooner than this timeline would contemplate – likely *before* the completion of development of this new suite of assessments. The disagreement within the Task Force mirrors that outside the University and reflects different views of the relative risks and benefits of possible, additional steps that UC could take to ensure that the entering classes more closely reflect the diversity of the State.

<table>
<tr><th>Months</th><th>Task</th><th>Brief Description</th></tr>
<tr><td>0-24 months</td><td>1. Solicit Input and Build Consensus</td><td>Before any large-scale assessment development work can begin, there must be a period of fact-finding and consensus building among the key stakeholders. We propose that the UC should spend a period of 24 months to detail, in the form of a public-facing framework document, what the core knowledge, skills, and attributes this new assessment system should reflect. Needless to say, UC faculty and students, UC alumni, and the broader California public should have their voices represented in the process. Part of this work will have to be qualitative and inductive, but it is necessary so that we may understand what are the missing components from our current thinking. Part of this work will involve synthesis of knowledge from research, broadly construed. The ultimate goal would be the concrete definitions of a set of cognitive, content, and linguistic features that a new assessment should reflect.</td></tr>
<tr><td>0-24 months</td><td>2. Analyze Current Technology</td><td>UC should examine the state of the art in assessment technology, assessment methods, as well as the state of readiness in the assessment field outside of research and academic settings. In a series of RFIs, information about automated item generation, item authoring, item rendering, test security, test delivery, and automated item scoring should be gathered and analyzed by UC researchers with experience in large-scale assessment development, psychometrics, statistics, and computer science and computer engineering.</td></tr>
<tr><td colspan="3">Check Point 1
At the end of 24 months, the UC should have enough information to make a decision on whether the continuing buildout of this new system of assessments will likely meet UC's goals. If UC chooses to move ahead, a sunset date of the current standardized tests can be decided at this time.</td></tr>
<tr><td>25-36 months</td><td>3. Development of a Prototype</td><td>Prototypes for a couple of select content areas or select kinds of new response formats will be developed. Using iterative development, small low-stakes tryouts should already have occurred by the end of this task.</td></tr>
<tr><td colspan="3">Check Point 2
At this point, UC will have enough evidence to determine whether the full system build out is viable from a time and financial standpoint. The option to stop remains open at this point, but if the decision is to move ahead with full scale development, a sunset date of the current standardized tests should be decided at this time for ethnical and practical reasons.</td></tr>
<tr><td>37-72 months</td><td>4. Full Scale Development</td><td>The system will be built and a large field test will be carried out. If adequate resources are available, this task is the part of the project where more dollars can translate into less time until completion.</td></tr>
<tr><td>73-96 months</td><td>5. Operational administrations</td><td>The assessment enters into a provisional stage and is ready for use by campus admissions offices.</td></tr>
<tr><td colspan="3">Check Point 3
Pending two years of successful data gathering and positive evaluation outcomes, the new assessment enters into an operational stage, and is maintained actively for future years.</td></tr>
</table>

UNIVERSITY OF CALIFORNIA Academic Senate

STTF Roster and Acknowledgements

STTF Co-Chairs
Henry Sánchez, Pathology (UCSF)
Eddie Comeaux, Education (UCR)

Members
Eva Baker, Education (UCLA)
Julian Betts, Economics (UCSD)
Li Cai, Education (UCLA)
Darlene Francis, Public Health (UCB)
Patricia Gandara, Education (UCLA)
Jonathan Glater, Law (UCI)
James Griesemer, Philosophy (UCD)
Andrea Hasenstaub, Otolaryngology – Head and Neck (UCSF)
George Johnson, Mechanical Engineering, (UCB)
Mona Lynch, Criminology, Law & Society (UCI)
Andrew Maul, Education (UCSB)
Yolanda Moses, Anthropology (UCR)
Kip Tellez, Education (UCSC)
Haim Weizman, Chemistry & Biochemistry (UCSD)
Anne Zanzucchi, Writing Studies (UCM)
Wendy Rummerfield, Graduate Student Representative (UCI)

STTF Writing Subcommittee
Jonathan Glater, Chair
Julian Betts
Li Cai
Patricia Gandara
Andrea Hasenstaub
George Johnson

Staff
Michael LaBriola, Academic Senate Assistant Director
Ken Feer, Academic Senate Analyst

Special Thanks
The STTF thanks its Writing Subcommittee, whose extraordinary work, sustained dedication, and original analyses of new data made this report possible. We also thank UC Director of Institutional Research and Academic Planning (IRAP) Tongshan Chang, and Director of Undergraduate Admissions Han Mi Yoon-Wu, who along with IRAP Analysts Matt Reed and James Hershey, provided invaluable analytic support to the STTF. In addition, STTF member James Griesemer's expert and nuanced editing of the report helped bridge differences of opinion across the STTF. Lisa Garcia and Liz Terry in the Office of Student Affairs also provided special editing assistance. Senate Assistant Director Michael LaBriola ensured the Task Force had robust and informed discussions; the Report benefited from his editing assistance at a crucial juncture. Finally, we thank the testing experts and education leaders who acted as consultants to the STTF during the last year:

- The College Board and ACT
- Saul Geiser, Research Associate, UC Berkeley Center for Studies in Higher Education,
- UC Campus Directors of Undergraduate Admissions
- California State Board of Education President Linda Darling-Hammond
- Professor of Education Michal Kurlaender, UC Davis
- Michelle Center and Stephanie Gregson, California Department of Education
- Smarter Balanced CEO Tony Alpert
- Professor Anthony Carnevale, Georgetown University Center on Education and the Workforce
- Professor of Education William E. Sedlacek, University of Maryland
- Professor of Psychology Sheldon Zedeck, UC Berkeley
- Testing Consultant Jay Rosner
- Professor of Education Emeritus Rebecca Zwick, UC Santa Barbara
- The UC Board of Admissions and Relations with Schools

VIII. List of Appendices

Appendix I: Relationship of the SAT/ACT to College Performance at the University of California (IRAP Report)

Appendix II: Tables of Regression Coefficients

Appendix III: Percentage of Applicants Reporting SAT Scores Who Were Admitted by Campus, High School GPA (HSGPA) Range and SAT Range. Fall 2018 California Resident Applicants.

Appendix IV: Interpreting R^2 as Percent Explained Variance

Appendix V:

- **Figures 1-10:** Applicant and admittee HSGPAs, test scores, and admissions rates for all campuses together (Figure 1) and each campus individually (Figures 2-10), broken down by racial/ethnic subgroup.
- **Figures 11-20:** Applicant and admittee HSGPAs, test scores, and admissions rates for all campuses together (Figure 1) and each campus individually (Figures 2-10), broken down by family educational history.
- **Figures 21-30:** Applicant and admittee HSGPAs, test scores, and admissions rates for all campuses together (Figure 1) and each campus individually (Figures 2-10), broken down by family income category.
- **Tables 1-3:** Applicant and admittee HSGPAs and test scores, by demographic group and campus.
- **Tables 4-6:** Applicant and admittee HSGPAs and test scores with respect to highest-scoring groups, by demographic group and campus.

Relationship of the SAT/ACT to College Performance at the University of California

University of California Office of the President
January 2020

Executive Summary

For many years, high school grade-point average (HSGPA) and standardized test scores (i.e. SAT/ACT) have served as the two most important and relied upon college admissions criteria. However, there is constant debate about whether HSGPA and SAT/ACT are fair measures of college preparation and valid predictors of college success, and whether SAT/ACT should remain as a requirement for college admissions.

Supporters of standardized tests argue that SAT/ACT evaluates college-preparedness on an equal standing despite inconsistent grading systems throughout the nation's high schools. While HSGPA compares a student to the rest of his/her school, SAT/ACT scores compare the student to the rest of the country. Opponents of standardized tests argue that SAT/ACT is biased towards white, upper middle class students from high performing schools. Focusing on HSGPA in college admissions may help foster racial and socioeconomic diversity in the college student population.

Previous research generally treats standardized test results as measures of college preparation. The literature surrounding the predictive validity of standardized test scores on student success is mixed. Many studies found a statistically significant association between SAT/ACT scores and college outcomes including first-year college GPA, first-year retention, four-year graduation, and graduation GPA. On the other hand, research also shows that demographic predictors weakened the ability of SAT/ACT scores to be effective predictors of college success.

In light of previous research, this report presents the results of separate studies on the relationship among HSGPA, standardized tests, and demographics, and explanatory power of test scores for college success based on UC's application and enrollment data. Findings from this study are generally consistent with those from prior research.

Analysis of the relationship among standardized tests and high school GPA, UC's Analytical Writing Placement Exam (AWPE), and demographics shows:

- **Demographics are stronger predictors of SAT/ACT scores than of HSGPA.** The share of variance in SAT/ACT scores accounted for by demographic factors (parental income, parental education, and ethnicity) increased from a low of 26 percent in the late 1990's to 43 percent in 2016. The share of variance in HSGPA accounted for by the same demographic factors increased from 5 percent in the late 1990's to 11 percent in 2016. About one-third of the increase for SAT/ACT scores can be explained by disparities between CA high schools, with the remainder explained by changes in the composition of applicants to UC. Campus-specific estimates do not support increasing racial bias on the tests as an explanation for increase.
- **Standardized test scores are positively correlated with high school GPA but at moderate level.** When controlling for socioeconomic factors (family income and parental education), the correlations between SAT/ACT scores and high school GPA fall between .30 to .51 across high schools and fall between .31 to .56 within high schools. The lowest correlations happen between SAT/ACT writing scores and high school GPA in 2018.

From 2005 to 2018, there has been an increase in the correlations between SAT Reading/Math as well as ACT Composite with high school GPA. Overall, the SAT and ACT writing tests are less correlated with HSGPA than SAT Reading/Math or ACT Composite.

- **New SAT EBRW (**Evidence-Based Reading and Writing**) and ACT** ELA (English Language Arts) **scores have a strong association with passing UC's Analytical Writing Placement Exam (AWPE), but SAT Essay scores have only a moderate association with passing AWPE.** The analysis of the relationship between the new SAT and new ACT and UC's Analytical Writing Placement Exam (AWPE) shows that high scores on the SAT EBRW and ACT ELA are associated with passing the AWPE. However, on the SAT Essay, even at the top score of 24, only 69 percent passed the AWPE. Based on the results of data analysis, UCOPE has accepted the new ACT ELA and SAT EBRW scores but not the new SAT Essay scores as methods of meeting UC's Entry Level Writing Requirement (ELWR).

Analyses of the relationship between standardized tests and college success show that standardized tests add value to the prediction of college outcomes beyond HSGPA alone.

- **SAT/ACT scores and HSGPA are both moderate predictors of student college GPAs, and weak to moderate predictors of student retention and graduation.** Between 2001 and 2015, SAT Reading/Math scores account for 13 to 21 percent of the variance in freshman GPA, and 15 to 17 percent of the variance in graduation GPA. ACT Composite scores generally account for 14 to 22 percent of the variance in freshman GPA, and 17 to 19 percent of the variance in graduation GPA. In comparison, HSGPA accounts for 13 to 21 percent of the variance in freshman GPA, and 15 to 18 percent in graduation GPA.
- Without controlling for student demographics, SAT/ACT scores are a stronger predictor of freshman GPA when compared to HSGPA, but have almost the same explanatory power of graduation GPA, first year retention and graduation. After controlling for student demographics, HSGPA and test scores have the same explanatory power of the freshman GPA for 2015, the latest year included in this study, but HSGPA is a stronger predictor of the first year retention, graduation GPA and four-year graduation.
- **Supplementing HSGPA with SAT/ACT scores increased the explanatory power of pre-admission measures on college success metrics.** Models that combined both SAT/ACT and HSGPA account for an additional 5 to 11 percent of the total variance of first-year GPA when compared to models that only use HSGPA scores. Similarly, combined HSGPA and SAT/ACT models account for an additional 3 to 11 percent of variance associated with UC graduation GPA when compared to models that only use HSGPA.
- **Adding SAT/ACT writing to SAT/ACT scores does not increase the explanatory power of pre-admission measures on college success.** SAT and ACT writing scores account for an additional 1 to 2 percent of the variance associated with most student success metrics (i.e., freshman GPA and freshman retention), with the exception of

graduation GPA, where SAT and ACT writing scores account for just as much, if not more variance associated with graduation GPA, when compared to SAT Reading/Math and ACT composite.

- **HSGPA and SAT scores are associated with course performanc**e. They are independently important explanatory factors for first-year students' course performance relative to their peers in similar courses, with the latter growing in relative importance as SAT predictive validity improves and HSGPA variation declines.

In summary, this report presents what literature and UC data show about the validity of standardized test scores in terms of academic preparation and college success. Should UC decide to continue to use standardized test scores as part of its admissions evaluation criteria? To answer this question, a series of simulations will have to be done to further examine what admissions outcomes would look like based on new criteria to evaluate students' academic achievements. Institutional Research and Academic Planning (IRAP) at UCOP will continue to work on analyses and simulations.

I. Introduction

In 1968, the University of California (UC) introduced a new policy that required all freshman applicants to submit scores from the SAT I or the ACT and three College Board Achievement Tests (later called SAT II: Subject Tests and now called SAT Subject Tests). The original purposes of this requirement were to identify applicants' strengths and weaknesses, to improve UC's ability to identify students that are likely to succeed at the university, and to adjust for differences in high school grading practices.[1] Since then, standardized test scores have played an important role in evaluating applicants' academic preparation for college and predicting students' success at UC. One important use of test scores at UC is to create a sliding scale eligibility index along with high school Grade Point Average (HSGPA). Beginning with the Class of 1979, such an index has been used to select the top 12.5 percent of the California public high school graduating class in order to meet the guidelines of the California Master Plan for Higher Education.[2] In 2000, UC adjusted the eligibility index to include SAT II scores in the index and to weigh them more heavily relative to the SAT I/ACT scores effective for the fall 2003 entering freshman class because the SAT I was a relatively weaker predictor of academic performance in the freshman year based on UC's 1996 study.[3] UC, however, has never stopped examining the validity of standardized tests in evaluating applicants' academic preparation and predicting college success. Based on a series of studies and evaluations, UC has revised the requirement for the SAT and the ACT several times to meet its admissions goals over the years. One of the most important revisions since 2000 was eliminating the requirement for the SAT Subject Tests (formerly called SAT II: Subject Tests) effective for the 2012 entering freshman cohort.

The College Board and ACT have also changed their test structures and content several times during the time period from 2001 to 2018 covered in this study, which may have had an impact on how universities decided to use these test scores in admissions. In 2005, the College Board announced a major change to the SAT effective in 2006 to reflect the importance of clear and succinct writing, which is not only considered a skill to be used in college courses, but also one necessary for success in a wide range of careers. To meet this goal, the College Board changed the SAT I to the SAT Reasoning, which included three tests, Critical Reading, Math, and Writing (formerly a subject test called SAT II Writing), and changed the SAT II: Subject Tests to SAT Subject Tests. These changes were mainly attributed to a

Versions of the SAT

UC Academic Years 2001-2005

- SAT I: Verbal and Math
- SAT II : Writing

UC Academic Years 2006-2016

SAT Reasoning

- Critical Reading
- Math
- Writing

UC Academic Years 2017-2018

- Evidence-Based Reading and Writing
- Math
- Essay

[1] Dorothy A. Perry, Michael T. Brown, & Barbara A. Sawrey. (2004). Rethinking the Use of Undergraduate Admissions Tests: The Case of the University of California. In Rebecca Zwick (Ed.), *Rethinking the SAT: The future of standardized testing in university admissions*. (pp. 103-124). New York and London: RoutledgeFalmer.

[2] University of California Office of the President, Institutional Research and Academic Planning. California Master Plan for Higher Education. Retrieved on August 31, 2018 from https://www.ucop.edu/institutional-research-academic-planning//content-analysis/academic-planning/california-master-plan.html.

[3] University of California Eligibility and Admissions Study Group. Final Report to the President. (2004, April). https://senate.universityofcalifornia.edu/_files/committees/boars/documents/studygroup_final0404.pdf.

series of debates over aptitude versus achievement tests in college admissions sparked by UC President Richard C. Atkinson's advocacy to eliminate the SAT I in UC admissions.[4] After about a decade, the College Board made changes in the content, format, and scoring of the SAT Reasoning again in 2016. The new SAT includes three parts: 1) Evidence-Based Reading and Writing (EBRW), 2) Math, and 3) Essay with three subscores: Reading, Analysis and Writing. The new test was first administered in the spring of 2016, and currently costs $64.50 with the essay and $47.50 without it, though some school districts subsidize these fees.[5] A validity study by the College Board based on a sample of participants from 15 four-year institutions show that the redesigned SAT is as predictive of college success as the prior SAT, that redesigned SAT scores improve the ability to predict college performance beyond high school GPA alone, and that there is a strong, positive relationship between redesigned SAT scores and grades in matching college course domains.[6] However, the 2016 changes were not done without controversy. For example, the new exam's wordy math questions may unfairly penalize students because of their language burden.[7] When commenting on this, Jamal Abedi, a UC Davis professor who specializes in educational assessments stated, "The problem is going to mostly affect English-language learners."

ACT added an option of a 30-minute direct writing test to their examination beginning in February 2005. Based on an early study focusing on the value of using the ACT Writing test in placing students in composition and related courses, the writing test score added value to the accuracy of course placement decisions over and above the ACT English test.[8] Performance on the writing test in combination with performance on the English test was incorporated into the Combined English/Writing (or English with Writing) score. Ten years later, in September 2015, ACT introduced changes to the design of this writing test, with modifications to the writing task, scoring rubric, and score reports. ACT states that many elements of the writing task remain similar to those of the previous task, both emphasizing argumentative writing skills that are essential for college and career success, but the new writing task is more consistent with the writing skills emphasized in

Versions of the ACT

UC Academic Years 2001-2005
- ACT Composite

UC Academic Years 2006-2015
- ACT Composite
- ACT English with Writing

UC Academic Years 2016- 2018
- ACT Composite
- ACT English Language Arts

[4] Zwick, R. (2004). College Admissions Testing in California: How Did the California SAT Debate Arise. In Rebecca Zwick (Ed.), *Rethinking the SAT: The future of standardized testing in university admissions*. (pp. 101-102). New York and London: RoutledgeFalmer.
[5] College Board. (2018.) SAT Suite of Assessments: Test Fees. Retrieved on October 26, 2018 from https://collegereadiness.collegeboard.org/sat/register/fees.
[6] Shaw, E., Marini, J., Beard, J., Shmueli, D., Young, L., and Ng, H. (2016). The Redesigned SAT Pilot Predictive Validity Study: A First Look. Retrieved on August 31, 2018 from https://collegereadiness.collegeboard.org/pdf/redesigned-sat-pilot-predictive-validity-study-first-look.pdf.
[7] Dudley, R. (2016). Despite warnings, College Board redesigned SAT in way that may hurt neediest students. *Reuters*. Retrieved on August 31, 2018 from https://www.reuters.com/investigates/special-report/college-sat-redesign/.
[8] ACT. ACT Writing Test Technical Report. (2009). Retrieved on August 31, 2018 from https://www.act.org/content/dam/act/unsecured/documents/TechReport.pdf.

the Common Core State Standards (CCSS) and other standards developed by ACT.[9] The new writing test score is combined with scores on the English and Reading tests and reported as the English Language Arts (ELA) score. The current test costs $67.00 with the writing section and $50.50 without it, but some school districts subsidize these fees.[10]

Despite the efforts the College Board and ACT made over the years to improve the validity of their standardized tests in predicting college success, the number of colleges using Test Optional Policies (TOPs) in higher education admissions has dramatically increased in recent years. According to a study by the National Association for College Admission Counseling (NACAC), more than 1,000 institutions had officially deemphasized standardized tests in admissions by spring 2018 when the report was released.[11] The report further emphasizes that the momentum indicates that undergraduate admissions is moving away from heavy reliance on measures increasingly "deemed to provide a narrow assessment of human potential." The study also indicates that almost all institutions included in the study increased representation of underrepresented groups (URGs) among applicants and enrollees and there are no signs of academic slide in terms of college GPAs and graduation rates.

In June 2018, the University of Chicago announced that it would not require domestic freshman applicants to submit standardized test scores any more. There were several reasons for this decision, which, according to an article published in the *Chicago Tribune*, "marks a dramatic shift for the South Side university..."[12] The University's leaders have long wanted to increase diversity and hoped this decision would prevent students from assuming that anything less than an outstanding test score automatically takes them out of the running. The Undergraduate Admissions Dean at the University also said that there was a big industry of test preparation, which served higher-income students very well.

Although few universities (e.g., University of New England) have followed the University of Chicago's decision to stop requiring ACT and SAT scores for prospective undergraduates, many universities eliminated the requirement for the SAT and the ACT Writing since spring 2019. Harvard University announced it was dropping the requirement in March, followed by Dartmouth in April, Yale and the University of San Diego in June, then Princeton, Stanford, Brown, Duke, and the University of Michigan in July.[13] By November 2018, only 12 universities

[9] The ACT College and Career Readiness Standards available online at http://www.act.org/content/dam/act/unsecured/documents/ACT_RR2015-4.pdf and the 2011 NAEP Writing Framework available online at https://www.nagb.gov/content/nagb/assets/documents/publications/frameworks/writing/2011-writing-framework.pdf.

[10] ACT. (2018). The ACT Test: Current ACT Fees and Services. Retrieved on October 26, 2018 from http://www.act.org/content/act/en/products-and-services/the-act/registration/fees.html.

[11] Syverson, S., Franks, V., Hiss, W. (2018). Defining Access: How Test-Optional Works. Retrieved on August 31, 2018 from https://www.nacacnet.org/globalassets/documents/publications/research/defining-access-report-2018.pdf.

[12] Rhodes, D. (2018.) University of Chicago to stop requiring ACT and SAT scores for prospective undergraduates. Retrieved on August 31, 2018 from http://www.chicagotribune.com/news/local/breaking/ct-university-chicago-sat-act-20180614-story.html.

[13] The Princeton Review. (2018). Who Requires SAT and ACT Essays (and why they shouldn't). *The Score (blog).* Retrieved on August 28, 2018 from https://princetonreview.blog/2018/03/18/it-is-time-to-eliminate-the-sat-and-act-optional-essays/.

still required the SAT Essay score for admissions, among which nine were UC's campuses, according to a list published by CompassPrep.[14] The other three schools requiring the SAT Essay were the United States Military Academy, Martin Luther College, and Soka University of America.

What decision should UC make? An article published by *Inside Higher Ed* in July, 2018 used an eye-catching title to emphasize that "For fate of SAT Writing Test, watch California." Is this true? Given the fact that all other research universities dropped the SAT/ACT writing requirement for admission and only three non-UC institutions continue to require the SAT or ACT Writing, there is no doubt that if UC drops this requirement, the list of institutions requiring the writing tests will be quite short. However, as Henry Sanchez, the former chair of the Board of Admissions and Relations with Schools (BOARS), noted in an interview with *Inside Higher Ed*, UC has a different situation from some of the colleges dropping the requirement recently. Nearly 200,000 high school seniors apply to UC now with UCLA receiving more than 113,000 applications for fall 2018 freshman admission, the largest number among all colleges and universities throughout the country. To make a wise and evidence-based decision about whether to sustain or drop this requirement for admissions, UC would benefit from a better understanding of what existing research tells us and what UC data show about SAT and ACT scores in terms of measuring college preparation and predicting student success nationwide and at UC specifically and whether or not SAT and ACT scores can help UC to admit students to meet its educational philosophy.

In this report, Institutional Research and Academic Planning (IRAP) at the University of California Office of the President (UCOP) reviews current research on the validity of the SAT and ACT and the ability of the tests to predict college success, and uses application and enrollment data from academic years 2001 to 2018 to address two sets of questions as follows:

1. How do SAT and ACT scores relate to other measures of applicants' academic preparation such as high school grades? Do socioeconomic factors unrelated to a student's academic potential affect SAT and ACT scores?

2. How well do SAT and ACT scores predict college success as measured by freshman GPA, first-year retention, four-year graduation, and graduation GPA at UC? Do the SAT and ACT writing tests add any explanatory power to predict college success at UC, above and beyond what is predicted by HSGPA and SAT Reading and Math or ACT composite? Does the ability of SAT and ACT scores to predict college success vary by student characteristics?

This report summarizes the findings of current research on relationships among standardized tests (the SAT and ACT), high school grades, and demographics, and explanatory power of the tests on college students' success, and also analyzes UC data to examine what they show about applicants' college preparation and student success in terms of the test scores. The report is organized into five sections. Section I describes the research methodology. Section II summarizes findings of the existing research on the validity of the SAT and ACT. Section III

[14] CompassPrep. (2018). ACT Writing and SAT Essay Requirements. Retrieved on May 13, 2019 from https://www.compassprep.com/act-writing-and-sat-essay-requirements/.

examines relationships between the SAT/ACT scores and high school academic preparation measured by high school weighted, capped GPA and UC's Analytical Writing Placement Exam (AWPE) scores. Section IV discusses the relative validity of the SAT and ACT in predicting student success measured by the freshman GPA, one-year retention, four-year graduation, and graduation GPA at UC. The last section summarizes the major findings of this study.

II. Methodology

This study uses UC's application and enrollment data to examine the relationship between the SAT/ACT and other college preparation measures of applicants to UC and college performance of students enrolled at UC. The analysis focuses on California resident applicants and enrollees. The SAT and ACT scores are derived from official and self-reported scores. If official scores are missing, self-reported scores are used.

The current study also examines differential impacts of both SAT and ACT by campus, intended discipline (both broad area and STEM/Non-STEM), family income, parental education level, Pell status, and high school ranking based on the Academic Performance Index (API). These analyses are conducted as previous research has indicated that the predictive validity of SAT/ACT varies by student demographics.

As noted earlier, UC accepts both SAT and ACT scores. To understand how related each of these two tests is to college preparation and performance, this study analyzes the SAT and ACT separately. In other words, two separate sets of statistical models are built, one set with the SAT predictors and the other with the ACT predictors. Also, to examine what additional variance of student success at UC the SAT and ACT Writing can explain beyond the SAT Math and Reading scores, or the ACT Composite scores, separate models in both sections are estimated with a predictor of the sum of the SAT Math and Reading or the ACT composite, the SAT/ACT Writing, or two predictors of the sum of the SAT Math and Reading and the SAT Writing, or ACT Composite and ACT Writing. The College Board and ACT have made significant changes to their test structure over time. Therefore, a longitudinal analysis was conducted to better understand how the changes of test structures, content, and scoring may have affected their validity.

SAT Models
1. Sum of Reading (EBRW since 2017) and Math Scores
2. Writing (Essay since 2017)
3. Sum of Reading (EBRW since 2017) and Math Scores and Writing Score (Essay since 2017)

ACT Models
1. Composite
2. English with Writing (Writing subscore since 2017)
3. Sum of Composite and English with Writing

To examine the relationships between standardized test scores and other measures of college preparation, this study first provides a profile of the SAT and ACT takers among UC applicants from 2001 to 2018 (see Appendix A for details), then examines the percent of variance associated with SAT/ACT scores and high school GPA that can be attributed to factors students have no control over (i.e., family income, parental education and race/ethnicity). Finally, a series of partial correlations were calculated to examine the relationship between SAT/ACT scores and high school grades, controlling for selected demographic factors. Analyses throughout this report use high school weighted and capped GPA[15] for high school grades.

The methodology used to analyze relationships between standardized test scores and college performance at UC is similar to that used in a paper by Saul Geiser and Roger Studley published in 2002, which examined predictive validity and differential impact of the SAT I and SAT II at UC.[16] The findings in Geiser's research will be described later in this report. However, since their research included enrollment data from 1996 to 1999, this study does not replicate their research; instead, it conducts similar analyses using enrollment data from 2001 to 2015 in order to compare results with their findings. As many other studies (see Section III for a summary of previous research findings) claim that freshman GPA (or first-year GPA) is by far the most frequently used outcome variable measuring college success, this study employs the freshman GPA as a criterion to measure college success, and also examines how well the SAT/ACT predicts first-year retention, four-year graduation, and graduation GPA.

This report presents some results for all years from 2001 to 2018, and other results, especially those about relationships of the SAT and ACT to college preparation and performance for selected years including 2001, 2005, 2007, 2012, 2015, and 2018. These years were chosen because 2001 was the first year UC adopted the Eligibility in the Local Context (ELC) program, 2005 was the last year students took the old SAT and ACT, 2007 was the second year students submitted new SAT and ACT test scores, 2012 was the first year UC revised the admissions eligibility structure with top nine percent of high school graduates from a school qualified for a guaranteed admission, 2015 was the last year before the current version of the ACT, and 2018 is the third year students submitted new ACT scores and the second year students submitted new SAT scores.

III. What Previous Research Findings Show about the SAT and ACT

Although there is no single definition or measure of college preparation or college readiness, researchers, national associations of education, and test developers usually suggest that standardized test scores in conjunction with other measures available during high school can act as proxies for performance in college courses and careers. Among the possible measures are SAT/ACT test scores, high school degree completion, high school GPA, taking challenging high school courses, and performance in high school courses.[17]

[15] High school Grade Point Average (HSGPA) used in this analysis is an honors-weighted, capped GPA where extra points up to eight semesters, no more than four in the 10th grade are added to the GPA. The UC's admission's website provides more detailed information about how a weighted, capped GPA is calculated, http://admission.universityofcalifornia.edu/freshman/requirements/gpa-requirement/index.html.

[16] Geiser, S., & Studley, R. (2002). *UC and the SAT: Predictive validity and differential impact of the SAT I and SAT II at the University of California*. Educational Assessment, 8(1), 1-26.

[17] Maruyama, Geoffrey. (2012). Assessing College Readiness: Should We Be Satisfied With ACT or Other Threshold Scores? *Educational* Researcher, 41:7, 252 – 261. Retrieved on August 28, 2018, from http://journals.sagepub.com/doi/abs/10.3102/0013189X12455095.

Green and Winters developed a measure of public high school college readiness to reflect the minimum standards of the least selective four-year college.[18] The standard includes earning a regular high school diploma, completing a minimum set of course requirements, and being able to read at a basic level (scoring at or above the basic level on the National Assessment of Educational Progress [NAEP] reading assessment).

The National Center for Educational Statistics (NCES) measured college readiness based on a student's high school GPA, senior class rank, National Education Longitudinal Study (NELS) 1992 test scores, and SAT/ACT college entrance scores.[19]

The ACT is designed to measure academic skills and knowledge taught in high school and required for first year college courses.[20] Therefore, it is not surprising that previous studies (by ACT and others) have found that factors such as high school coursework and high school GPA affect ACT scores. ACT scores are influenced by achievement in core subject areas of high school courses.[21] High school GPA accounts for 31% of the variance in ACT scores, more than high school coursework, high school characteristics, non-cognitive factors, and demographic characteristics. Non-cognitive factors such as parental involvement or perceptions of education affect ACT scores via their impact on high school GPA.[22] Completing more AP courses is associated with higher ACT scores.[23] Taking an AP math course and taking more AP courses are associated with meeting benchmark scores on the ACT.[24] Taking and passing AP exams is associated with higher ACT scores, but simply enrolling in AP courses is not.[25] The SAT is also moderately correlated with HSGPA.[26]

[18] Green, J.P., & Winters, M.A. (2005). Public high school graduation and college-readiness rates: 1991-2002. Manhattan Institute. Retrieved on August 10, 2018, from https://www.manhattan-institute.org/html/public-high-school-graduation-and-college-readiness-rates-1991-2002-5911.html.

[19] Berkner, L., & Chavez, L. (1997). Access to postsecondary education for the 1992 high school graduates. (NCES 98-105). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

[20] Ferguson, Richard. (2004). Achievement versus Aptitude in College Admissions. In Zwick, Rebecca. *Rethinking the SAT*. New York and London: RoutledgeFalmer.

[21] Allen, J. (2015). *Influence of Achievement in Core High School Courses on ACT Scores*. Retrieved on August 28, 2018 from https://www.act.org/content/dam/act/unsecured/documents/2015-Tech-Brief-Influence-of-Achievement.pdf.

[22] McNeish, D., Radunzel, J., Sanchez, E. (2015). A Multidimensional Perspective of College Readiness: Relating Student and School Characteristics to Performance on the ACT®. Retrieved on August 28, 2018 from http://www.act.org/content/dam/act/unsecured/documents/ACT_RR2015-6.pdf.

[23] Anderson, K. (2016). The Effectiveness of Advanced Placement Courses in Improving ACT® Scores for High School Students. Retrieved on August 28, 2018 from http://www.kylestevenanderson.com/uploads/7/0/5/8/70582975/kyle_anderson_edd_dissertation_pdf.pdf.

[24] Mo, L., Yang, F., Hu, X., Calaway, F., & Nickey, J. (2011). ACT test performance by Advanced Placement students in Memphis City schools, *The Journal of Educational Research*, 104, 354–359. Retrieved on August 28, 2018 from https://www.tandfonline.com/doi/abs/10.1080/00220671.2010.486810.

[25] Warne, R., Larsen, R., Anderson, B., Odasso, A. (2015). The Impact of Participation in the Advanced Placement Program on Students' College Admissions Test Scores, *The Journal of Educational Research*, 108, 400-416. https://www.tandfonline.com/doi/full/10.1080/00220671.2014.917253.

[26] Shaw, E., Marini, J., Beard, J., Shmueli, D., Young, L., and Ng, H. (2016). The Redesigned SAT Pilot Predictive Validity Study: A First Look. Retrieved on August 31, 2018 from https://collegereadiness.collegeboard.org/pdf/redesigned-sat-pilot-predictive-validity-study-first-look.pdf.

Recently, a paper was published by researcher Saul Geiser, where he examined the correlation between demographic predictors (i.e., parental education, family income and race/ethnicity) and SAT/ACT scores in UC California resident freshman admissions from 1994-2011.[27] Geiser found that demographic predictors accounted for more than a third of the variance associated with SAT/ACT scores among UC applicants, and that race/ethnicity had become the strongest of those three predictors. This report includes results of the analysis replicated based on slightly different methods than Geiser used for his analysis.

A large volume of research has also examined relationships between standardized tests and student success in college. However, findings are not always consistent across research. This could be because different research might use different measures to evaluate college success, utilize different methods to analyze data, examine different samples, and control different demographic variables in modeling. Most of the research on standardized tests and academic success focuses either on the SAT alone or combines SAT and ACT scores. Very few studies look at the tests separately. Combining scores on the two tests is often done as SAT and ACT scores are highly correlated. Based on UC's application data, the Pearson correlation coefficient between the sum of SAT Reading and Math scores and the ACT composite scores was 0.91 in 2001, 0.89 in 2006, and 0.93 in 2016.

Many studies that use first-year GPA (FYGPA) in college as a measure of college success have found a strong correlation between test scores and FYGPA. In 1960, Fishman and Pasanella reviewed 147 studies that included the SAT as a predictor of FYGPA, finding that the correlation between SAT scores and high school record with FYGPA ranged from moderate to strong (0.34 to 0.82).[28] In 1989, Morgan of the College Board analyzed the predictive validity of the SAT on first-year GPA and found that the correlation between SAT scores and FYGPA declined over the years, but there was less change for private institutions, small institutions, and more selective institutions.[29] More recently, Hezlett and colleagues performed a meta-analysis of approximately 3,000 validity studies, with more than one million students. They found that the SAT is a valid predictor of FYGPA, with correlations ranging from moderate to strong (0.44 to 0.62).[30]

A report published in 2011 by the National Bureau of Economic Research found that the English and math sections of the ACT were more strongly correlated with student success than were the reading and science sections.[31] Specifically, the ACT math and English sections only were found to be more strongly related to first-year GPA than the science and reading sections only. While

[27] Geiser, S. (2015). The Growing Correlation Between Race and the SAT Scores. *UC Berkeley Center for Studies in Higher Education.* Retrieved from https://cshe.berkeley.edu/publications/growing-correlation-between-race-and-sat-scores-new-findings-california-saul-geiser on August 3, 2018.
[28] Fishman, J.A., & Pasanella, A.K. (1960). *College admission selection studies*. Review of Educational Research, 30(4), 298–310.
[29] Morgan, R. (1989). Analysis of the Predictive Validity of the SAT and High School Grades from 1976 to 1985. (College Board Research Report No. 1989-7). New York: The College Board.
[30] Hezlett, S.A., Kuncel, N., Vey, M.A., Ahart, A.M., Ones, D.S., Campbell, J.P., & Camara, W.J. (2001, April). The effectiveness of the SAT in predicting success early and late in college: A comprehensive meta-analysis. Paper presented at the annual meeting of the National Council on Measurement in Education, Seattle, WA.
[31] Bettinger, E.P., Evans, B.P., and Pope, D.G. (2011). Improving College Performance and Retention. The Easy Way: Unpacking the ACT Exam. *National Bureau of Economic Research.* Retrieved from http://www.nber.org/papers/w17119.pdf on August 7, 2018.

not directly challenging these findings, ACT does contest the interpretation of these findings. They argue that most students score similarly on all four tests, and those that differ may provide admissions officers with important additional information.[32]

As major fields have their own unique characteristics with different grading standards, there are likely to be differences in the predictive validity of the SAT for cumulative college GPA by academic program. Shaw and colleagues of the College Board showed that correlations between SAT and cumulative GPA were of moderate strength for most majors (0.50-0.60). The strongest correlations tended to be found in STEM (Science, Technology, Engineering, and Mathematics) fields. The weakest correlations were found for undeclared students ($r = 0.42$).[33] Bridgeman, Pollack, and Burton of the College Board further pointed out high school GPA is a slightly poorer predictor of cumulative college GPA for African American and Hispanic students. SAT scores are moderately correlated with cumulative college GPA (.50) over four or more years. For all minority group students, both male and female, SAT scores predict college success about as well as they do for white students.[34]

Some research findings indicate that a strong relationship exists between the SAT score and college retention, over the past 30 years. By analyzing a national dataset including individual level data on nearly 150,000 students from 106 colleges and universities, Mattern and Patterson of the College Board found a strong correlation between SAT performance and retention to second year.[35] They found that 64 percent of students in the lowest SAT score band returned for their second year, compared to 96 percent of students in the highest SAT score band.[36] The relationship between SAT scores and retention to the third and fourth years was also examined, and similar results were found.[37, 38]

Research has also shown that SAT scores predict graduation as well as HSGPA. Burton and Ramist[39] of the College Board found that high school record had an uncorrected correlation of .29 with college graduation, while each of the individual SAT sections had an uncorrected correlation of .27, and the best combination of the two sections had an uncorrected correlation of

[32] Mathews, J. (2011). Report finds 2 of 4 tests in ACT Poor Predictors of College Success. *The Washington Post.* Retrieved from https://www.washingtonpost.com/blogs/class-struggle/post/report-finds-2-of-4-tests-in-act-poor-predictors-of-college-success/2011/07/19/gIQAOiPFOI_blog.html?utm_term=.f1505f48e84a on August 13, 2018.
[33] Shaw, E., Kobrin, J., Patterson, B., and Mattern K. (2012). The Validity of the SAT for Predicting Cumulative Grade Point Average by College Major (College Board Research Report No. 2012-6). New York: The College Board.
[34] Bridgeman, B., McCamley-Jenkins, L., & Ervin, N. (2000). Predictions of freshman grade-point average from the revised and recentered SAT I: Reasoning Test. (College Board Research Report No. 2000-1). New York: The College Board.
[35] Mattern, K. and Patterson, B. (2009). *Is Performance on the SAT Related to College Retention?* (College Board Research Report No. 2009-7). New York: The College Board.
[36] Mattern, K. and Patterson, B. (2009). *Is Performance on the SAT Related to College Retention?* (College Board Research Report No. 2009-7). New York: The College Board.
[37] Mattern, K. D., & Patterson, B. F. (2011a). *The relationship between SAT scores and retention to the third year: 2006 cohort* (College Board Statistical Report No. 2011-2). New York: The College Board.
[38] Mattern, K. D., & Patterson, B. F. (2011b). *The relationship between SAT scores and retention to the fourth year: 2006 cohort* (College Board Statistical Report No. 2011-6). New York: The College Board.
[39] Burton, N., & Ramist, L. (2001). *Predicting success in college: SAT studies of classes graduating since 1980* (College Board Research Report No. 2001-2). New York: The College Board.

.33. Mattern, Patterson and Wyatt[40] of the College Board found that students with higher SAT scores are more likely to graduate, and graduate in a timely manner (i.e., four years), even after controlling for HSGPA, institutional characteristics, and institutional selectivity.

Geiser and Studley[41] examined the relationship between SAT scores and freshman GPA based on the records of 77,893 students who entered UC between Fall 1996 and Fall 1999. The study found that SAT II subject tests were better predictors of student success at UC when compared to the SAT I aptitude tests. They suggested that using the achievement tests (SAT II subject tests) for college admissions could be valuable to the prediction of college success with clarity in admissions standards and close linkage to the high-school curriculum.

However, other research has questioned the predictive validity of SAT scores as it relates to student success within the first year, and beyond. Using a dataset containing over 400,000 individual records from 176 colleges with various levels of selectivity, researchers sought to better understand if the SAT was predictive of first-year GPA.[42] Evidence suggests that the SAT, while overall predictive of first-year GPA, was less predictive across subgroups. For example, when comparing African American and White students, the SAT Critical Reading either over- or under-predicted first-year GPA at 20 percent of colleges and universities.[43]

Furthermore, using UC data, Geiser[44] found that demographic predictors weakened the ability of SAT/ACT scores to be effective predictors of five-year graduation rates, and in contrast, the explanatory power of HSGPA was not found to be affected by demographic controls. Geiser and Santelices found that HSGPA was consistently a stronger predictor of student success when compared to SAT scores, and that its power increased as students progressed through their UC career.[45]

Similar to what was found by Geiser and Santelices, recently, an article posted in *Forbes* concluded that HSGPA matters more than SAT/ACT scores when it comes to student success.[46] For example, even students with high SAT/ACT scores (above 1,100, with ACT concorded to the SAT scale) have a low expected six-year graduation rate (51 percent), if their HSGPA is

[40] Mattern, K., Patterson, B. and Wyatt, J. (2013). *How Useful Are Traditional Admission Measures in Predicting Graduation Within Four Years?* (College Board Research Report No. 2013-1). New York: The College Board.
[41] Geiser, S., & Studley, R. (2002). *UC and the SAT: Predictive validity and differential impact of the SAT I and SAT II at the University of California*. Educational Assessment, 8(1), 1-26.
[42] Aguinis, H., Culpepper, S.A. & Pierce, C.A. (2016). Differential Prediction Generalization in College Admissions Testing. *Journal of Educational Psychology (7).* 1045-1059.
[43] Jaschik, S. (2016). Faulty Predictions? *Inside Higher Education.* Retrieved from https://www.insidehighered.com/news/2016/01/26/new-research-suggests-sat-under-or-overpredicts-first-year-grades-hundreds-thousands on October 12, 2018.
[44] Geiser, S. (2016). A Proposal To Eliminate the SAT in Berkeley Admissions. *UC Berkeley Center for Studies in Higher Education.* Retrieved from https://cshe.berkeley.edu/publications/proposal-eliminate-sat-berkeley-admissions on August 31, 2018.
[45] Geiser, S. & Santelices, M.V. (2007). Validity of High School Grades in Predicting Student Success Beyond the Freshman Year: High-School Record v. Standardized Tests as Indicators of Four-Year College Outcomes. *UC Berkeley Center for Studies in Higher Education.* Retrieved from https://cshe.berkeley.edu/sites/default/files/publications/rops.geiser._sat_6.13.07.pdf on August 31, 2018.
[46] Cooper, P. (2018). What Predicts College Completion? High School GPA Beats SAT Score. *Forbes.* Retrieved from https://www.forbes.com/sites/prestoncooper2/2018/06/11/what-predicts-college-completion-high-school-gpa-beats-sat-score/#3c6910c04b09 on August 17, 2018.

relatively low (3.00-3.32). In contrast, students with average SAT/ACT scores (900-990) and a high HSGPA (3.67-4.00) have a much higher expected graduation rate (62 percent). However, the study used a sample of students who attended a group of less selective four-year public colleges and universities. This report uses the same method to analyze timely graduation rates based on UC's longitudinal data.

In summary, studies that examined relationships between standardized tests and other college preparation measures such as HSGPA suggested positive relationships between them. The literature surrounding the predictive power of standardized test scores on student success is mixed. The College Board and supporters of standardized tests argue that they offer a check against grade inflation, assess skills needed not just for higher education, but the workplace as well, and that in combination with HSGPA provide a more "complete" assessment of students' academic potential. To support their argument, some research does suggest that both the SAT and the ACT are related to student success within the first year, and that the first year is important for student success, overall.

IV. What UC's Data Show about the SAT and ACT

Test Takers and Average Test Scores of UC California Resident Applicants

Appendix A provides a profile of UC California resident applicants who submitted SAT and/or ACT scores in their applications. From 2001 to 2013, more than 90 percent of UC CA applicants submitted SAT scores. However, there has been a declining trend in freshman applicants submitting SAT scores sine earlier 2010s. In 2018 one year after the College Board launched their new SAT, only about 86 percent of UC CA freshman applicants submitted SAT scores, including SAT Math, EBRW, and Essay scores. On the other hand, the share of CA applicants submitting ACT Composite scores increased steadily from 28 percent in 2001 to 47 percent in 2011 then fluctuated between 44 percent and 53 percent in 2012 through 2018. It is also important to note that the share submitting ACT ELA scores was substantially lower than the share submitting ACT Composite scores in 2017 and 2018 (at 46 percent and 38 percent respectively). It is hard to understand why about 15 percent of those who submitted ACT Composite scores did not submit ACT ELA scores. Were there any reasons for or obstacles to this? More analyses are needed to answer this question. However, to determine applicant's eligibility for guaranteed admissions, UC has to use SAT scores for these applicants. If they did not provide a full score of SAT tests either, they would be ineligible for guaranteed admissions under the current ETR (Entitled to Review) policy, which requires a full score of either SAT or ACT including the writing test.

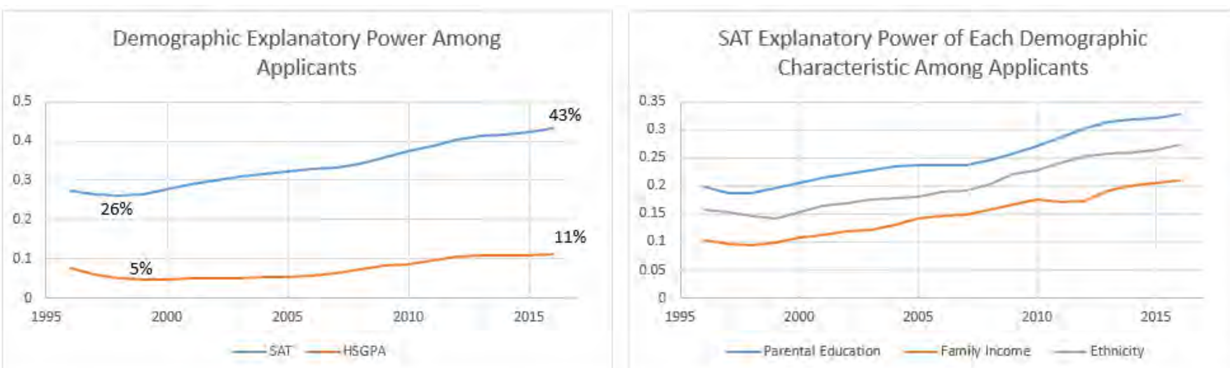
The average SAT I Verbal/Math scores ranged from 1163 to 1177 and average SAT II Writing scores ranged from 573 to 581 in 2001 to 2005 (Appendix B). The average SAT Critical Reading/Math scores dropped from 1161 in 2006 to 1117 in 2016 and the SAT Writing scores dropped as well from 572 in 2006 to 550 in 2016. Students seem to have higher new SAT Math/EBRW scores in 2017 and 2018 relative to recent SAT Critical Reading/Math averages (1203 and 1167 respectively).

Despite the changes over time to the ACT, the average ACT composite scores have stayed largely the same, at 24 or 25 in all years between 2001 and 2018 except 2017, when the average was 26. Excluding 2017 and 2018, when very few students submitted scores, English with Writing was 23 or 24 in all years (2006 through 2016). The average for the new ELA scores was 25 in 2017 and 2018 and the average writing subscore was 8 (on a 2-12 scale).

The Relationship between Demographic Characteristics and SAT/ACT Scores

Figure 1 below shows an apparently disturbing trend in the proportion of variation in UC applicants' SAT scores that are accounted for by fixed student characteristics. The left panel of the figure shows that annual linear regressions of UC applicants' SAT scores on three such characteristics—parental income, parental education, and ethnicity—show that the amount of demographically-explained variation in applicants' SAT scores (as measured by R^2) has increased from 26 percent in the late 1990s to 43 percent in 2016, suggesting that nearly half of the variation captured by the SAT score across UC applicants reflects student background characteristics, not all of which are directly observed by UC's admissions offices. Meanwhile, the proportion of explained variation of applicants' high school GPAs (HSGPAs) has remained

Figure 1: Demographics' Explanatory Power for SAT and HSGPA



Note: R^2 from annual OLS regressions of applicants' SAT score or high school GPA on comprehensive parental education indicators, ethnicity indicators, and family income (and an indicator whether family income is reported), combined (left) and one at a time (right). Sample restricted to California-resident freshmen.

at a far lower level, rising from 5 percent in the 1990s to 11 percent in 2016. Figure 1's right panel breaks out each permanent characteristic separately, showing that all three demographics each explain an additional 10 percentage points of SAT variation since the 1990s.[47]

[47] A number of modeling assumptions are necessary in the production of this chart, and our choices differ from those made by the previous version of this brief (September, 2018) as well as from Geiser (2015). First, consider the three demographic characteristics analyzed in this study:

1. Parental Income: Both our earlier version and Geiser include only log CPI-adjusted parental income as their measure of income. This technique implicitly drops two important groups of applicants from the sample: (a) applicants who report 0 parental income, since the log of 0 is non-finite (about 4% of the sample), and (b) applicants who do not report parental income on their applications, usually because they do not intend to receive financial aid, indicating high-income households (about 12% of the sample).

There are at least two well-known interpretations of this important finding. The first is that the SAT is increasingly racist, classist, or otherwise biased against students from disadvantaged backgrounds over time, such that despite their unchanged average aptitude over the past 25 years, disadvantaged applicants' SAT performance has deteriorated (unlike their HSGPA performance, which may not face the same systematic issues of bias). The second is that increased residential segregation by ethnicity and class, combined with increased educational disparities across California secondary schools, have led to actual average declines in college preparation among disadvantaged groups, as accurately reflected by their declining SAT scores (but perhaps not by HSGPAs, which are normed within increasingly-disparate high schools).

A third possible explanation is that the trend is driven by changes in the composition of youths in the estimation sample, most likely as a result of changing UC admissions policies. Expanding admissions policies that favor disadvantaged applicants, for example, could increase the explanatory power of demographic characteristics among admits by enlarging the pool of low-SAT high-disadvantage students in the sample, which would mechanically increase demographics' correlation independent of the exam's bias or changes in applicants' average latent aptitude.

Arbitrating between these alternative explanations for the **"Demographic Testing Trend", or DTT**, is crucial to understanding how the SAT's informativeness has evolved over the past 25 years. Advocates in favor of the first explanation (the "**Bias Explanation**") include Saul Geiser, who argues that the DTT invalidates use of the SAT in the absense of race-based affirmative

Omitting these samples may mechanically decrease the correlation between income and SAT, since they represent the two extremes of income where the covariance with test scores may be highest. In order to maintain these samples, this analysis includes three measures of parental income in each regression model: log CPI-adjusted parental income (replaced as 0 when missing or infinite), an indicator for missing income, and an indicator for zero income. This change likely explains the higher proportion of SAT variation explained by the presented estimates.

2. Parental education: Our earlier version included only an indicator for whether one parent has a college degree, while Geiser may have included an ordered integer measuring the more-educated parent's highest level of education. Both of these measures simplify a high-dimensional student feature—the educational level of their parent—into a highly-parametric summary. This analysis includes indicator variables for each combination of educational background held by the applicant's parents, using the full available information set. This change may partly account for the aggregate increase in explanatory power of demographics for applicants' SAT score.
3. Ethnicity: Our earlier version and Geiser include only an indicator for whether the applicant is from an underrepresented group, including Black, Chicano/Latinx, or Native American. This analysis includes indicators for every observed ethnicity, or 15 in all. This may also contribute to the general increase in demographics' explanatory power for SAT scores.

The added value of including these multi-dimensional measures of students' background characteristics is that they more fully specify each student's background, leading to more explanatory power and avoiding possibly-important model restrictions that could challenge interpretation (especially in the case of parental income). The disadvantage of using multi-dimensional measures is that there is no longer a single standardized regression coefficient associated with each measure, making it impossible to directly compare the degree to which each contributes to their mutual absorption of SAT variation. As a result, rather than presenting regression coefficients, we show the degree to which each individual characteristic (as measured multi-dimensionally) alone can explain variation in applicants' SAT scores.

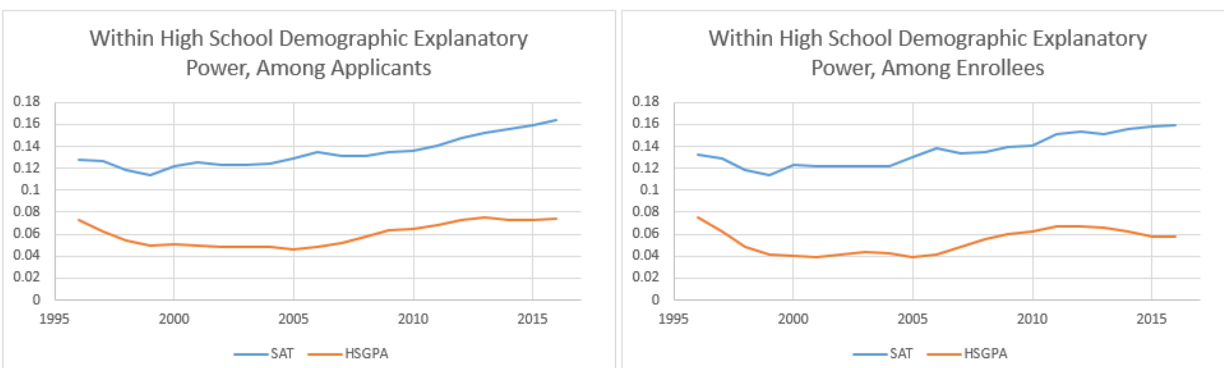
Finally, all three analyses use the same definition of applicants' SAT score: the sum of the mathematics and reading components of the SAT exam.

action because it increasingly favors White and Asian applicants (Geiser, 2015). Advocates for the second explanation (the "**Real-Disparities Explanation**") include the College Board, which argued in its recent meeting with the Task Force that "performance on the SAT differs across subgroups, which largely reflects educational differences in high schools". The third explanation (the "**Compositional Explanation**") has drawn less vocal support.

This section presents evidence that approximately one-third of the DTT is explained the Real-Disparities Explanation, with two-thirds explained by the Compositional Explanation. Figure 1 shows that the Real-Disparities Explanation explains 31 percent of the DTT. Figures 2 and 3 motivate the Compositional Explanation, showing the magnitude of UC compositional changes over the past 25 years, and Figure 4 shows that the Compositional Explanation fully explains the remaining net DTT trend. Figure 4 directly challenges the Bias Explanation as even playing a secondary role in explaining the DTT.

First, we rerun the regressions that estimate the cross-student DTT including high school fixed effects, which eliminate cross-high-school variation that likely reflects increasing educational disparities in California secondary schools. We show analysis for both the population of UC applicants and the population of UC enrollees, restricting the sample to California-resident freshman applicants. If the Real-Disparities Explanation were to be accurate in the strong sense (that is, if the real disparities were *exclusively* across high school, without increasing disparities within high school as well), then the results would show no trend in demographics' explanatory power for applicants' or enrollees' SAT scores.

Figure 2: Within High School DTT Trend



Note: R^2 (excluding fixed effects) from annual OLS regressions of applicants' or enrollees' SAT score or high school GPA on comprehensive parental education indicators, ethnicity indicators, and family income (and an indicator whether family income is reported), including fixed effects by origin high school. Sample restricted to California-resident freshmen.

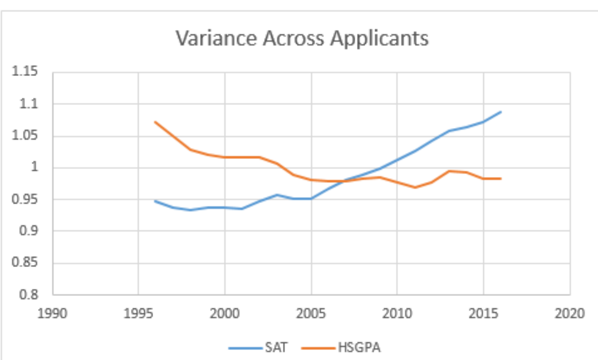
Figure 2 above shows that this is not the case. While far less variation in SAT scores can be explaned by demographics when only comparing students to others from their same high schools, there is still a notable upward trend in demographics' explanatory power, from 11.3 percent at its trough in 1999 to 16.4 percent in 2016 among UC applicants (representing a 31 percent decline in the proportional increase in explanatory power since 1996). Demographics' explanatory power for

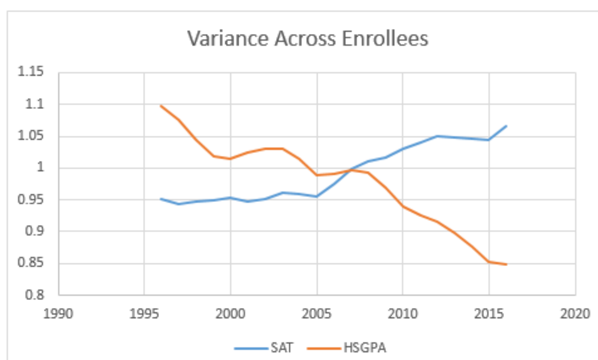
HSGPA increased from 4.9 to 7.4 percent in the same period, suggesting even faster proportional growth (a 51 percent increase, relative to 45 percent). A similar trend holds among UC enrollees.[48]

One important feature that these figures share with the original DTT chart is the decline in demographics' explanatory power between 1995 and 1999. These declines likely reflect that period's phasing out of UC's affirmative action program, which decreased underrepresented minorities' UC enrollment by at least 700 students per year across all campuses (Bleemer, 2019) and led to parallel declines in UC URM applications. As a result, the Compositional Explanation clarifies that demographics' explanatory power in the period declined as the affirmative-acted students—who tended to have lower SAT scores that could be explained by their demographic characteristics—ceased enrolling at UC. This observation provides the first circumstantial evidence supporting the role of student composition in regulating the correlational relationship between demographics and SAT performance.

In order to further motivate the Compositional Explanation, we next plot the annual variance in SAT scores and HSGPAs across all UC campuses. Figure 3 shows that the amount of variance in SAT scores among UC applicants has been swiftly rising since about 2005, while the amount of variance in HSGPA fell in the late 1990s and early 2000s and has persisted at the lower level. The trends among UC enrollees are even more pronounced; variation in SAT scores among UC applicants has increased by more than 10 percent since 1996, while variation in HSGPAs has fallen by more than 20 percent. These trends likely reflect two important admissions policies—Eligibility in the Local Context and Holistic Review—that have substantially replaced affirmative action since the 1990s in enrolling disadvantaged applicants. As various UC campuses increase their numbers of low-SAT high-HSGPA students (the latter of which is measured relative to the lower-preparedness high schools from which the disadvantaged applicants are pulled), SAT scores are increasingly varying across the campuses' student bodies.

Figure 3: Change in Annual SAT and HSGPA Standard Deviations Since 1995





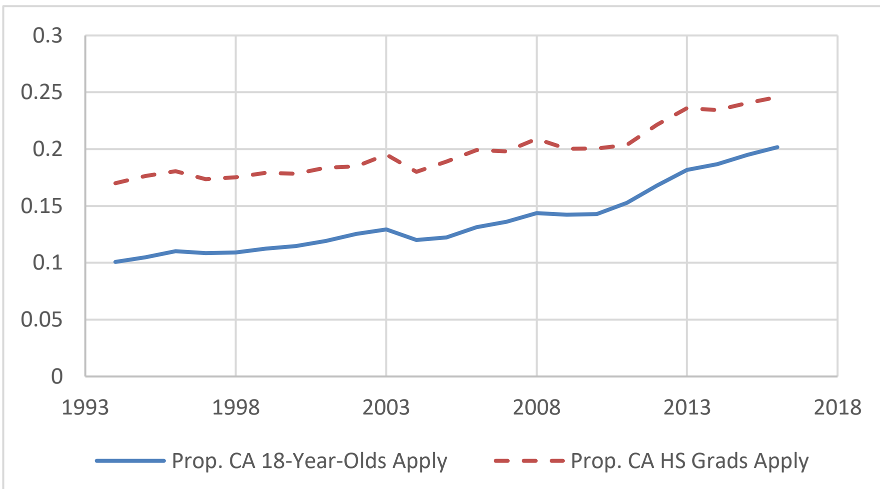
Note: Annual standard deviation in SAT and HSGPA of UC applicants and enrollees. SAT and HSGPA are normed to have standard deviation 1 on average across all years. Plot shows two-year moving averages.

[48] Figures 2 and 5 present 'projected R^2" measures from the relevant annual linear regression estimates of SAT on demographic characteristics. To be more specific, these models include high school fixed effects, which themselves (importantly) absorb some cross-school variation in SAT performance, and there's no reason to include that explanatory power in the reported R^2. Instead, I merely report the R^2 of *projected* SAT scores, after differencing out the high-school-specific averages, regressed on demographic characteristics. This is a standard technique implemented using the fixed effect linear regression package *felm* in *R*.

In other words, the increased variance is by design, an artifact of admissions policies which intentionally target lower-SAT applicants. We will return below to another explanation for the increase in SAT variation over time: the expansion of the Riverside and Merced campuses.

The next plot visualizes the proportion of California high school students who apply to at least one University of California campus. One of the chief successes—and political challenges—of UC's Comprehensive and Holistic Review admissions programs has been their encouragement of applications from high school graduates who would previously have not applied to UC because of their poor perceived likelihood of admission. Figure 4, pulled from Douglass and Bleemer (2018), shows that the proportion of 18-year-olds in California who apply to at least one UC campus has doubled since 1995, from about 10 percent to about 20 percent. Some of this

Figure 4: Proportion of California Residents Who Apply to UC

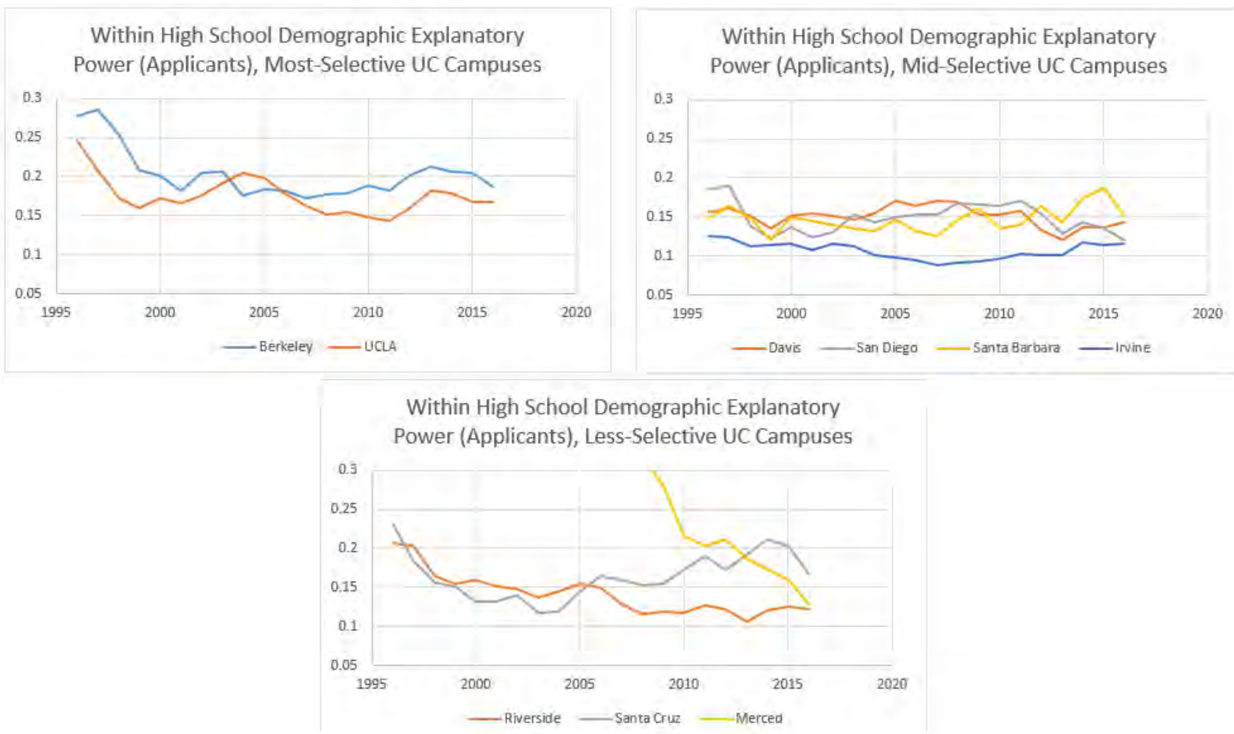


Note: The proportion of California 18-year-olds and California high school graduates who apply to at least one UC campus in each year since 1994. The annual number of California 18-year-olds is as estimated by the California Department of Finance, which also reports the annual number of high school graduates in the state.

increase comes from increasing high school graduation rates, but even among graduates the proportion of applicants has increased by about 8 percentage points, to almost 25 percent. This change in application behavior has surely dramatically altered the composition of UC applicants, and is also reflected in UC's students as a result of changing admissions policies. These charts showing increasing SAT variation and increasing broad application behavior by California youths strongly suggest that compositional changes in UC applicants and enrollees are central factors in explaining the DTT: after all, UC has spent the past 20 years bolstering admissions policies that favor the lower-SAT disadvantaged applicants who would mechanically increase the SAT-demographics correlation.

Finally, and perhaps most consequentially, we replicate Figure 2 by campus (for applicants). Under either the Bias Explanation or the Real-Disparities Explanation, we would expect that the predictive power of the SAT has increased consistently at every UC campus, either because of consistent bias or consistently-varying disparities across applicants. In fact, Figure 5 shows a very different pattern. Only a single campus, Santa Barbara, manifests any increase in demographics-SAT correlation, and even there the increase in demographics' explanatory power for the SAT is slight (from 14.8 to 15.2 percent). Most other campuses have faced almost no change in demographics' explanatory power since the end of affirmative action in the late 1990s (like Irvine, Davis, and UCLA) or have actually experienced declines in demographics' explanatory power (San Diego, Riverside, and most notably Merced). What would cause these patterns?

Figure 5: Within High School DTT Trend by Campus



comprehensive parental education indicators, ethnicity indicators, and family income (and an indicator whether family income is reported), including fixed effects by origin high school. Estimated separately for each UC campus. Sample restricted to California-resident freshmen.

The answer is the Compositional Explanation. Campuses with high explanatory power—especially Merced, which in the late-2000s had demographic explanatory power in the 30-40 percent range—have grown, in enrollees but especially in applicants (who wouldn't otherwise have applied to UC campuses). The end of affirmative action pushed demographics' explanatory power down, especially at the Berkeley and UCLA campuses where that program was most effective, and the end of the old ELC program in 2011 also appears to have compressed demographics' explanatory power at the campuses where that program was most effective (San Diego, Davis, and Irvine). Meanwhile, all of the campuses were growing more selective on average, compressing their HSGPA distributions, but also instituting disadvantaged-focused admissions programs that purposely admitted students whose low test scores were offset by measures of disadvantage that mechanically strengthened the correlation between demographics and the SAT, both among enrollees and the applicants the the programs encouraged. These results are very difficult to reconcile with the Bias Explanation, and strongly suggest that the net DTT after the Real-Disparities Explanation can be fully explained by the Compositional Explanation.

As a result, we conclude that about one third of the DTT can be explained by the Real-Disparities Explanation, since only comparing students who enrolled at the same high school explains 31 percent of demographics' explanatory power for SAT scores. The remainder appears to be explained by the Compositional Explanation, which is motivated by measured compositional changes within and across UC campuses in their applicant and student bodies and confirmed by the absense of a DTT trend within 7 out of 8 UC campuses, which nevertheless show patterns strongly consistent with the implementation of disadvantage-favoring admissions programs. Indeed, the recently-increasing DTT-demographics correlation at many campuses is best understood as a manifestation of their admissions policies' successes in attracting the disadvantaged lower-SAT students that they are designed to attract.

Standardized Tests and Academic Preparation

This section examines the correlations between SAT/ACT scores and high school GPA as well as the relationship between SAT/ACT scores and UC's Analytical Writing Placement Exam (AWPE) score. SAT Writing refers to the SAT II writing for 2001 through 2005, SAT Writing for 2006 through 2016, and SAT Essay for 2017 and 2018. SAT Reading/Math refers to SAT I Verbal and Math for 2001 to 2005, SAT Critical Reading and Math for 2006 to 2016, and SAT Evidence-Based Reading and Writing (EBRW) and Math for 2017 and 2018. ACT Writing refers to the ACT Combined English/Writing (CEW) scores for 2006 through 2016 and the writing (essay) subscore for 2017 and 2018. ACT Composite combines the English, Reading, Math, and Science multiple choice sections. CEW combined the English multiple choice section with the writing (essay) section. On the new ACT, the writing subscore is combined with the English and Reading multiple choice sections to produce an English Language Arts (ELA) score.

Table 1 shows the across and within school correlations between SAT/ACT and high school GPA controlling for socioeconomic status (family income and parental education). Generally, the within school correlations between SAT scores and high school GPA are slightly higher than the across school correlations. This finding is consistent with what Zwick and Green (2007) found in their analysis.[49] They concluded that one of the primary reasons is that grading stringency varies across both high schools and courses. This can potentially result in small between-school variability in high school grades. Their study also indicates that between-school variance in mean test scores is likely to be substantial than within-school variance. For both across and within correlations between SAT and HSGPA, we see an increasing trend from 2005 to 2018, except that those between SAT writing and high school GPA dropped in 2018, which might be due to the new SAT essay (starting in 2017) with a different scoring scale. Across the years, the correlations between SAT Writing/Essay and HSGPA are consistently lower than those between SAT Reading/Math and HSGPA. Overall, the correlations between SAT and HSGPA are moderate at around .50.

[49] Zwick, R. and Green, J.G. (2007). New Perspectives on the Correlation of SAT Scores, High School Grades, and Socioeconomic Factors. *Journal of Educational Measurement.* Retrieved from https://onlinelibrary.wiley.com/doi/full/10.1111/j.1745-3984.2007.00025.x, May 3,2019

Similarly, the within school correlations between ACT and high school GPA are higher than the across school correlations. Generally, there is an upward trend of correlations between ACT and HSGPA across years, except for 2018 when there is a drop from 2015. ACT Writing is also less correlated with HSGPA than ACT Composite.

Table 1. Partial Correlation of Standard Test Scores and High School GPA

		2001	2005	2007	2015	2018
SAT Reading/Math	Across schools	0.47	0.45	0.48	0.50	0.51
	Within schools	0.53	0.51	0.53	0.55	0.56
SAT Writing (SAT Essay for 2018)	Across schools	0.44	0.44	0.46	0.48	0.39
	Within schools	0.49	0.49	0.50	0.52	0.41
Sum of SAT Reading/Math and SAT Writing	Across schools	0.48	0.47	0.50	0.52	n/a
	Within schools	0.54	0.54	0.56	0.57	n/a
ACT Composite	Across schools	0.47	0.46	0.46	0.50	0.47
	Within schools	0.51	0.51	0.51	0.55	0.53
ACT Writing (ACT ELA for 2018)	Across schools	n/a	n/a	0.43	0.46	0.30
	Within schools	n/a	n/a	0.47	0.49	0.31
Sum of ACT Composite and ACT Writing	Across schools	n/a	n/a	0.47	0.50	n/a
	Within schools	n/a	n/a	0.52	0.55	n/a

Note: All correlations control for family income and parental education.

A previous analysis also looked at the relationship between the new SAT and new ACT and UC's Analytical Writing Placement Exam (AWPE) for the Fall 2017 incoming class.[50] The analysis found that 76 percent of those at or above a threshold of 690 for SAT Evidence Based Reading and Writing (EBRW) passed the AWPE (Table 2). On the old SAT, more than 75 percent of those meeting the old threshold of 680 on SAT Writing would have passed the AWPE. On the SAT Essay, even at the top score of 24, only 69 percent passed the AWPE, and only two percent of test-takers achieved that score.[51] For ACT, 75 percent or more of students scoring 30 or higher on ACT English Language Arts (ELA) would have passed the AWPE. This is similar to the pattern for the old threshold using the old ACT Combined English/Writing scores. The report recommended setting a threshold for passing the Entry Level Writing Requirement (ELWR) of 690 on SAT EBRW, leaving the threshold of 30 on the ACT ELA in place, and not setting a new threshold using SAT Essay.

[50] University of California Office of the President, Institutional Research and Academic Planning. (2017). New SAT and ACT tests and the Entry Level Writing Requirement.

[51] A later analysis looked at the SAT Essay sub-scores, showing that in Fall 2017, 65% of those with a top Writing score of eight passed the AWPE and only 5% of test-takers achieved this score; 73% of those with a top Analysis score of eight passed the AWPE and only 2% of test-takers achieved this score; 65% of those with a top Reading score of eight passed the AWPE and only 5% of test-takers achieved this score.

Table 2. Number of test-takers by AWPE score status and SAT EBRW, SAT Essay, or ACT ELA score, Fall 2017

SAT EBRW score	Failed AWPE		Passed AWPE		Total	
>= 690	185	24%	572	76%	757	100%
< 690	6,754	59%	4,778	41%	11,532	100%
Total	6,939	56%	5,350	44%	12,289	100%
SAT Essay score	Failed AWPE		Passed AWPE		Total	
= 24	15	31%	34	69%	49	100%
< 24	6,924	57%	5,316	43%	12,240	100%
Total	6,939	56%	5,350	44%	12,289	100%
ACT ELA score	Failed AWPE		Passed AWPE		Total	
>= 30	n/a	n/a	n/a	n/a	n/a	n/a
< 30	6,939	56%	5,350	44%	12,289	100%
Total	6,939	56%	5,350	44%	12,289	100%

Note: This table includes all incoming freshmen who submitted a statement of intent to register (SIR) and had both an AWPE score and an SAT EBRW, SAT Essay, or ACT ELA score (as applicable). Students with a 30 or higher on the ACT ELA did not have to take the AWPE.

On the basis of the data analysis described above, the University Committee on Preparatory Education (UCOPE), a committee of UC Academic Senate, approved a new threshold of 680 on SAT EBRW for passing the Entry Level Writing Requirement (ELWR), left the threshold of 30 on ACT ELA in place, and did not set a new threshold using SAT Essay. Note that SAT EBRW is based on multiple choice sections of the test while ACT ELA is based on multiple choice sections and the writing/essay section. SAT Essay is a stand-alone essay section.

In summary, the partial correlations between test scores and HSGPA after controlling for SES are moderate. This means that standardized test scores and HSGPA can both measure part of students' knowledge and skills but there are different aspects of students' academic preparations that can only be measured by test scores and HSGPA separately. In addition, the writing tests are less correlated with HSGPA than Reading/Math or composite tests, which means SAT/ACT writing tests may provide additional information about students' qualifications that HSGPA cannot provide. High scores on the SAT EBRW and ACT ELA are associated with passing the UC's AWPE writing exam, but this is not necessarily true for the SAT Essay.

Standardized Tests and Student Success

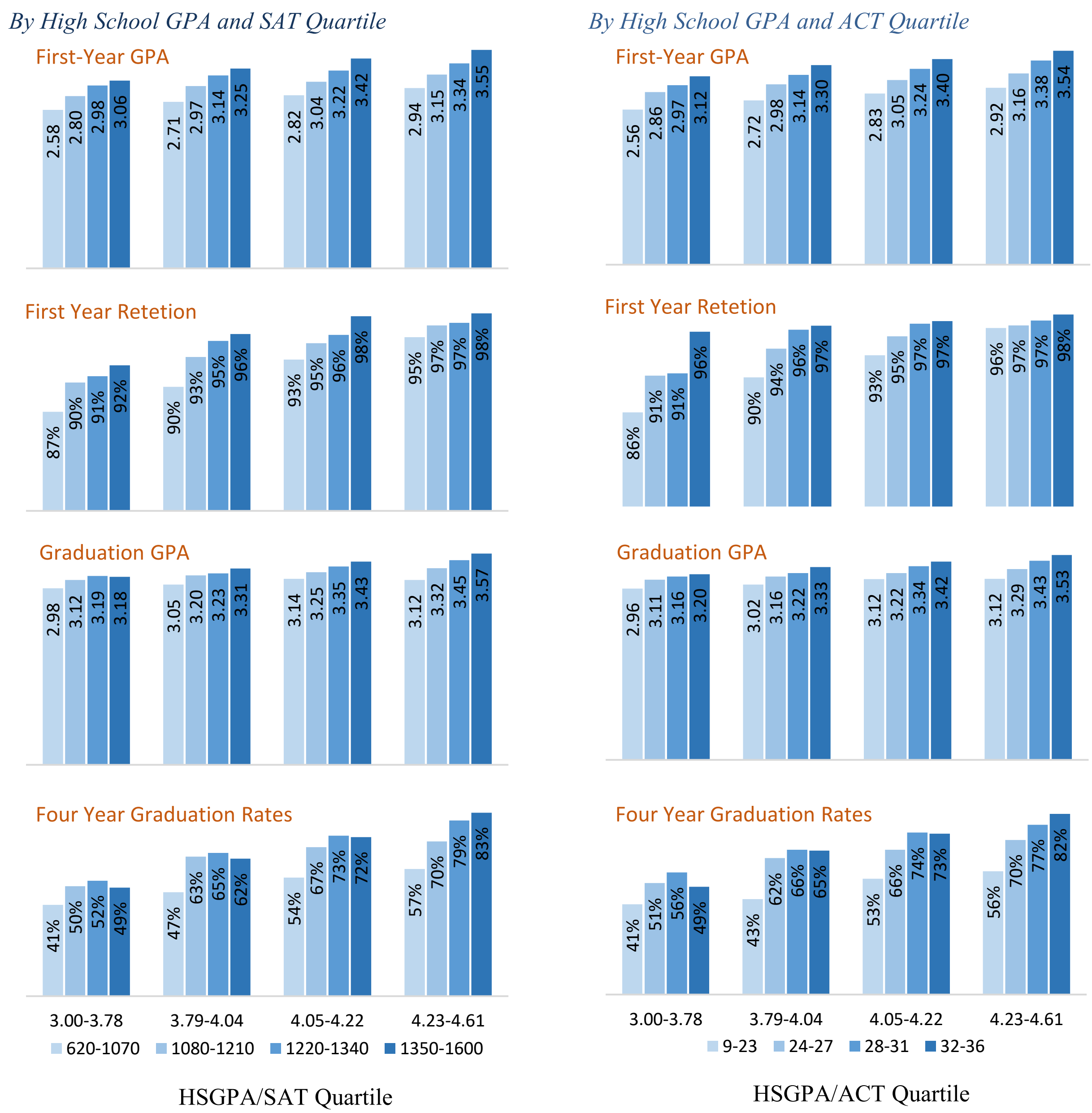
This section examines the relationship between the SAT/ACT and college success at UC through four outcome measures: (1) student first-year college GPA; (2) first-year retention; (3) four-year graduation; and (4) college graduation GPA. The section first presents results of descriptive analysis and then inferential statistics to show how well standardized test scores predict college success.

Figure 6 shows mean first-year GPA, first year retention rates, four-year graduation rates and mean graduation GPA of four-year graduates by HSGPA and SAT Reading/Math or ACT Composite quartile. Not surprisingly, students with a higher HSGPA and a higher SAT Reading/Math score or ACT Composite score (i.e., both in the highest quartiles) tend to have better performance on all four measures than students in the lowest test score and HSGPA quartiles. On average, their first-year GPA is higher than the GPA of those in the lowest HSGPA and test score quartiles by one point, or about one third; their first year retention rate is almost 100 percent, compared to about 90 percent for those in the lowest quartiles. The difference in graduation GPA between those in the highest and the lowest quartiles is smaller, but it is still about a half point. On average, students in the highest quartiles are twice as likely to graduate within four years.

In addition, Figure 6 shows that students in the same HSGPA quartile, but in a higher test score quartile are more likely to perform better, especially in terms of freshman GPA, first year retention rates, and four-year graduation rates than those in a lower test score quartile. For example, 83 percent of students in the highest HSGPA (4.23-4.61) and the highest SAT (1350-1600) quartiles graduate within four years, compared to 57 percent of those in the same HSGPA quartile, but with a lower SAT score (620-1070). This indicates that test scores do provide additional value beyond HSGPA to explain differences in students' performance at UC.

Results further show that students with a higher test score (1350-1600) and a lower HSGPA (3.00-3.78) seem to have a better freshman GPA, but a low first year retention rate when compared to students with a lower SAT/ACT score (620-1070) regardless of their HSGPA quartile. However, the first year retention rate graphs show the opposite, indicating that students in the highest HSGPA quartile and the lowest test score quartile are retained at a rate of 95 percent compared to 92 percent for those in the lowest HSGPA quartile and the highest SAT/ACT quartile. Similar to what was found for first-year GPA and retention rates, on average, students with higher test scores, but lower HSGPA, tend to have a higher graduation GPA and a lower four-year graduation rate, compared with those with a higher GPA, but a lower score on SAT Reading/Math or ACT Composite. This finding suggests that HSGPA may be slightly more predictive of first-year retention and four-year graduation, while standardized test scores may be slightly more predictive of both first-year and graduation GPA. Regression models will be developed in the following sections to further examine if this conclusion based on descriptive statistics holds and how HSGPA and test scores are related to college success.

Figure 6. First-Year GPA, First-Year Retention Rates, Four-Year Graduation GPA, and Four Year-Graduation Rate by HSGPA and SAT Reading/Math or ACT Composite Quartile



Standardized Tests and Academic Preparation: Aggregate First-Year GPA

Table 3 shows the percent of variance in freshman GPA accounted for by HSGPA and test scores based on a series of regression models. Results indicate that HSGPA, SAT Reading/Math, SAT Writing, ACT Composite, and ACT Writing are all moderate predictors of freshman GPA at UC. HSGPA accounted for 17 to 20 percent of variance in the freshman GPA prior to 2007, and then 15 percent in 2012 and 13 percent in 2015, the lowest ever while the variance accounted for by test scores has increased over the time from 13 percent in 2001 to 20 percent in 2015.

Table 3: Percent of Variance in Freshman GPA Accounted for by HS GPA and the SAT/ACT

Model	2001	2005	2007	2012	2015
(1) HSGPA	17%	18%	20%	15%	13%
(2) SATRM	13%	17%	19%	21%	20%
(3) SAT Writing	15%	18%	19%	20%	19%
(4) SAT TOTAL (SATRM + Writing)	17%	19%	21%	22%	22%
(5) HSGPA + SATRM	22%	26%	27%	26%	25%
Standardized Coefficients: HSGPA	*0.31*	*0.32*	*0.29*	*0.24*	*0.23*
SATRM	*0.25*	*0.29*	*0.31*	*0.36*	*0.36*
(6) HSGPA + SATTOTAL (SATRM + Writing)	24%	27%	28%	27%	26%
Standardized Coefficients: HSGPA	*0.30*	*0.30*	*0.30*	*0.22*	*0.22*
SATTOTAL	*0.28*	*0.32*	*0.33*	*0.38*	*0.38*
(7) ACT Composite	16%	21%	20%	23%	22%
(8) ACT Writing	n/a	n/a	19%	18%	19%
(9) ACT Total (Composite + Writing)	n/a	n/a	22%	23%	22%
(10) HSGPA + ACT Composite	22%	27%	27%	27%	26%
Standardized Coefficients: HSGPA	*0.28*	*0.31*	*0.29*	*0.22*	*0.23*
ACT CMP	*0.29*	*0.30*	*0.33*	*0.38*	*0.39*
(11) HSGPA + ACT Total (ACT CMP and ACT Writing)	n/a	n/a	29%	27%	26%
Standardized Coefficients: HSGPA	n/a	n/a	*0.28*	*0.22*	*0.22*
ACT Total	n/a	n/a	*0.35*	*0.38*	*0.38*
(12) HSGPA + SAT Total + Demographics	30%	32%	34%	32%	32%
Standardized Coefficients: HSGPA	*0.39*	*0.39*	*0.38*	*0.34*	*0.34*
Standardized Coefficients: SATTOTAL	*0.25*	*0.26*	*0.28*	*0.34*	*0.34*

There are multiple reasons for the declining trend in explanatory power of HSGPA. Since 2001, more and more campuses have adopted a comprehensive review process in admissions, including several that have adopted holistic review.[52] Using this process, campuses look beyond grades and test scores to evaluate students' qualifications for admission. However, several aspects of UC admissions policy still emphasize high school grades and test scores. The minimum requirement for UC admissions eligibility is to complete a minimum of 15 college-preparatory courses with a letter grade of C or better; the Eligibility in the Local Context (ELC) program purely replies on HSGPA in UC-approved coursework completed in the 10th and 11th grades; the statewide eligibility indices were created based on test scores and HSGPA; UC Regents increased the minimum HSGPA required for UC freshman eligibility from 2.80 to 3.00, effective for the fall 2007 entering class. Therefore, HSGPA still plays a significant role in eligibility in admissions at

[52] All campuses have used comprehensive review since 2002 and six campuses have adopted holistic review as their method of implementing comprehensive review: Berkeley (starting 2002), Los Angeles (2007), Irvine (2011), San Diego (2011), Davis (2012), Santa Cruz (2012).

UC. In addition, rapid growth of the qualified applicant pool pushed up the overall selectivity at all UC campuses. Analyses showed that recent UC admits have higher HSGPA than admits in prior years. All these together may have led to less variability in HSGPA of enrolled students. This change along with more variability in test scores caused by educational disparity among California K-12 schools, and constant variability in the freshman GPA at UC over the years may be some of many reasons that may have contributed to the change in the explanatory power of HSGPA and test scores in the freshman GPA at UC. (See the previous section "The Relationship between Demographic Characteristics and SAT/ACT Scores.")

Two sets of models with the sum of SAT Reading/Math and SAT Writing (Model 4) or the sum of ACT Composite and ACT Writing (Model 9) were developed to examine how much additional variance writing scores accounted for beyond SAT Reading/Math or ACT Composite. The reason to use the sum of scores on two tests is because SAT Reading/Math (ACT Composite) and SAT Writing (ACT Writing) are highly correlated (e.g., r=.85 for the freshman entering cohort in 2015). Therefore, there would be a collinearity issue if two measures were entered in the same regression model. Results indicate that in 2001, adding SAT Writing to SATRM increased the variance explained from 13 percent to 17 percent, or by four percentage points, but since 2005, it has only increased the explained variance by about two percentage points (difference in the explained variance between Model 2 and Model 4). Similarly, adding ACT Writing scores to ACT Composite scores does not increase the explanatory power at all (e.g., the difference in variance between Models 7 and 9). It is concluded that in the most recent year, adding writing scores to reading/math or composite scores does increase the explanatory power in explaining variation of freshman GPA, but the increase is not substantial.

In addition, four models were developed to examine how much additional variance in the freshman GPA standardized test scores account for beyond HSGPA. As showed by Models 5, 6, 10, and 11, adding SAT Reading/Math or ACT Composite to the model in recent years (e.g., in 2015) doubled the variance accounted for by HSGPA alone. However, adding SAT Total or ACT Total to the HSGPA models hardly changed the variance accounted for by the HSGPA and SATRM or ACT Composite models (difference in variance between SAT Models 5 and 6, and ACT Models 10 and 11). Similar to what has been found previously, it is concluded that writing scores do not add any additional value in predicting student's freshman GPA beyond HSGPA and SAT Reading and Math tests or the ACT composite test. Also, the standardized coefficients in these multivariate regression models indicate that test scores are stronger predictors for freshman GPA than HSGPA, especially for the 2012 and 2015 entering cohorts.

The analysis in previous sections of this report indicates that student characteristics (parental education, family income and race/ethnicity) account for 26 percent variation in applicants' SAT scores in the late 1990s and more than 40 percent in recent years. The explanatory power of these three factors in HSGPA has also increased from five percent in 2000 to 11 percent in recent years. Thus, it is helpful to examine the relationship between HSGPA and/or SAT Total and freshman GPA after controlling for student demographics. We ran regression models adding in student demographics such as campus affiliation, major discipline, first-generation status, family income, and high school API quintile (Model 12 in Table 3). Results show that controlling for demographics increased explained variation of freshman GPA by six percentage points (difference between Model 6 and Model 12). Results further show that after controlling for demographics, the explanatory power of HSGPA increased from 0.22 to 0.34 in terms of standardized coefficients in the 2015 model. The standardized coefficients of the SAT total decreased from 0.38 to 0.34 in the 2015 model, which indicates a decrease of the explanatory power of the SAT total in freshman GPA. This implies that SAT scores are more associated with the demographics than HSGPA, consistent with earlier findings.

The influence of varying student characteristics on the relationship between HSGPA and SAT/ACT scores on first-year GPA is further examined based on 2015 entering students. Detailed model coefficients are presented in Appendix C. Results are generally consistent with those from the overall models (Table 3), indicating that standardized test scores were generally the stronger predictor of first-year GPA than high school GPA by campus, discipline, race/ethnicity, first generation, family income, and CA API ranking. For some groups, HSGPA predicted first year GPA better. For example, HSGPA appeared to be a stronger predictor than test scores for African American students, White students, students with family income over 161k, and etc.

We also examined the explanatory power of the SAT Essay score and subscores on each of three components (Reading, Analysis, and Writing), using the most recent 2017 freshman entering cohort data. Results (Table 4) show that the Essay score only accounts for nine percent of variation in freshman GPA, which is much less explanatory power than SATRM (21 percent) and the SAT Writing score (20 percent) in earlier years. Each of the three components accounts for six to eight percent of variation in freshman GPA with the Essay Analysis being the strongest predictor. Adding the Essay Total to HSGPA only increased the explanatory power by one percentage point. With limited explanatory power added by the Essay scores, it is worth considering the social costs of additional writing tests. It is not ideal that students spend lots of time preparing for various tests while they could be focusing their energy on more important academic and social activities that could benefit them in the future.

Table 4: Percent of Variance in Freshman GPA Accounted for by HSGPA and the SAT

Model	2017
(1) HSGPA	16%
(2) SATRM	21%
(3) Essay Total	9%
(4) Essay Reading	6%
(5) Essay Analysis	8%
(6) Essay Writing	7%
(7) HSGPA + Essay Total	19%

Standardized Tests and Academic Preparation: First-Year Course-Level Performance

The previous analyses of the relationship between application components and freshman GPA are limited by the substantial variation in average grades across campuses, departments, and even individual professors. For example, students enrolled in first-year courses in Engineering departments tend to have higher-than-average SAT scores but earn lower-than-average grades, not as a result of poor performance but rather due to the higher standards and lower grade curve enforced by Engineering departments nationwide. As a result, Engineering departments decrease the positive correlation between SAT scores and freshman grades, though the decrease is uninformative about actual student performance.

In order to "control for" students' freshman course selections, we disaggregate our analysis from the level of freshman GPAs to the individual-course level. The resulting data contains the numerical grades (from 0 to 4) received by each student in each course. Courses are weighted by their number of units, and each student is given equal (total) weight. Unlike the analysis above, these "within-course" estimates condition on the specific courses taken by each student (that is, each linear regression is estimated using course-semester fixed effects). The resulting estimates can be interpreted as the amount of variation *within each course a student completes in their first year*—that is, only comparing the student's performance to that of the other students in their same class—that can be explained by each application component. As a result of data availability, we estimate course-level results from 2001 to 2016 for students at only three UC campuses: Berkeley, Davis, and Riverside. These regressions do not control for any fixed student characteristics, though controlling for characteristics like race and gender hardly changes the findings.

Table 5: Percent of Within-Course Freshman GPA Explained by HS GPA and SAT (3 Campuses)

Model	2001	2005	2009	2013	2016	Full Sample
(1) HSGPA	7.9%	7.0%	7.6%	5.4%	5.1%	6.8%
(2) SATRM	8.6%	8.3%	9.8%	10.6%	11.9%	9.2%
(3) SATW (Writing)	7.0%	6.6%	7.5%	8.1%	8.8%	7.3%
(4) SAT TOTAL (SATRM + SATW)	9.0%	8.7%	10.3%	11.1%	12.2%	9.7%
(5) HSGPA + SATRM	14.3%	14.1%	15.3%	14.5%	15.8%	14.5%
Standardized Coefficients: HSGPA	*0.23*	*0.24*	*0.25*	*0.24*	*0.24*	*0.19*
SATRM	*0.24*	*0.25*	*0.26*	*0.29*	*0.30*	*0.17*
(6) HSGPA + SAT TOTAL	14.4%	14.2%	15.5%	15.0%	16.2%	14.7%
Standardized Coefficients: HSGPA	*0.22*	*0.23*	*0.24*	*0.23*	*0.24*	*0.18*
SAT TOTAL	*0.25*	*0.25*	*0.27*	*0.3*	*0.31*	*0.20*

Note: Reported R^2 from the projected performance outcome net course-specific fixed effects, which also capture campus effects. Courses are weighted by units earned and then normalized to give each student equal weight.

Table 5 shows the percent of within-course freshman grade variation that can be explained by high school GPA (HSGPA) and SAT scores. In 2001, HSGPA and SAT each explained about 8-9 percent of within-course performance variation, but reflecting a similar trend in the aggregate GPA results, the two slowly diverge over the following years; by 2016, HSGPA only explains about 5 percent of variation (likely reflecting the decline in cross-student HSGPA variation as UC becomes more selective), while SAT explains 12 percent of variation. Moreover, HSGPA and SAT strongly complement each other; the inclusion of both nearly sums to two R^2 values, implying that they explain independent components of students' first-year course performance (see Model 6 compared to sum of Model 1 and Model 4). The SAT Writing exam has lower explanatory power that has only slightly increased in the past 15 years. Estimates of standardized coefficients show that both HSGPA and SAT explain separate components of first-year course performance, with a one-standard-deviation increase in SAT scores correlating to a 0.3 point increase (almost a full grade step, e.g., A- to A) in each course taken by first-year students.

We conclude from this analysis that both HSGPA and SAT scores are independently-important explanatory factors for first-year students' course performance relative to their peers in similar courses, with the latter growing in relative importance as SAT predictive validity improves and HSGPA variation declines.

Disaggregating freshman GPAs to the course level also enables more precise analysis of the explanatory power for student performance in different academic disciplines, which might provide further insight into the specific values of each measure of students' academic preparedness. We identify first-year courses taught in four areas—Humanities, Social Sciences, Natural Sciences, and Engineering—and estimate the proportion of variation in course performance across areas that can be explained by SAT and HSGPA.

Table 6: Percent of Within-Course 2016 GPA Explained by HSGPA and SAT by General Area

Model	Humanities	Social Sciences	Natural Sciences	Engineering
(1) HSGPA	4.8%	6.5%	4.7%	2.3%
(2) SATR (Reading)	7.4%	12.6%	6.9%	3.5%
(3) SATM (Math)	5.0%	10.2%	13.5%	7.4%
(4) SATW (Writing)	7.5%	11.6%	7.9%	5.3%
(5) HSGPA + SATR + SATM	12.2%	19.1%	18.9%	11.4%
Standardized Coefficients: HSGPA	*0.19*	*0.24*	*0.30*	*0.26*
SATR	*0.17*	*0.21*	*0.10*	*0.07*
SATM	*0.05*	*0.12*	*0.35*	*0.27*
(6) HSGPA + SATR + SATM + SATW	13.1%	19.7%	19.1%	12.0%
Standardized Coefficients: HSGPA	*0.19*	*0.24*	*0.30*	*0.26*
SATR	*0.11*	*0.16*	*0.06*	*0.02*
SATM	*0.03*	*0.10*	*0.33*	*0.24*
SATII	*0.10*	*0.10*	*0.07*	*0.10*

Note: Reported R^2 from the projected performance outcome net course-specific fixed effects, which also capture campus effects. Courses are weighted by units earned and then normalized to give each student equal weight.

Table 6 shows expected explanatory patterns. Disaggregating the SAT into its Reading and Mathematics components, we find that the Reading component explains more variation in Humanities and Social Science course performance (7-13%) while the Mathematics component explains more variation in Natural Science and Engineering course performance (7-14%). SAT scores explain substantially more variation than HSGPA alone; in Engineering, for example, high school GPA explains 2.3 percent of performance, but its combination with the SAT components explains more than 12 percent of performance. The SAT Writing out-performs HSGPA in all four areas, explaining between 5 and 12 percent of variation, but its addition to HSGPA and the two primary SAT components only explains substantial variation (almost 1%) in the Humanities.

The scaled coefficients reported in Table 6 show that a one standard deviation increase in HSGPA is associated with increases in students' performance in each of their first-year courses by 0.19 grade points in the Humanities and 0.30 grade points in the Natural Sciences. A one standard deviation increase across students' SAT scores yields even stronger correlations when holding HSGPA fixed, by 0.22 points in the Humanities and 0.45 points in the Natural Sciences, with non-negligible contributions by each SAT component in the areas of their respective strengths. We conclude that both HSGPA and SAT are independently-important explanatory factors for first-year students' course performance across academic disciplines, with SAT particularly important in cases like Engineering courses where HSGPA alone has lower explanatory power.

Finally, we identify two course types of particular interest: freshman writing courses, which are required for most enrollees (unless they satisfy their Entry Level Writing Requirement via examination), and Organic Chemistry, the key gatekeeper course for the popular pre-medical track at each campus. Data for these courses is available at all campuses for 2016, though noise in the nine-campus course data results in lower R^2 values across the board.

Table 7: Percent of Within-Course 2016 GPA Explained by HSGPA and SAT

Model	Freshman Writing	Organic Chemistry
(1) HSGPA	1.7%	3.7%
(2) SATR (Reading)	2.0%	3.0%
(3) SATM (Math)	0.9%	5.8%
(4) SATW (Writing)	2.1%	3.4%
(5) HSGPA + SAT TOTAL	3.8%	8.3%
Standardized Coefficients: HSGPA	0.14	0.22
SAT TOTAL	0.16	0.27

Note: Reported R^2 from the projected performance outcome net course-specific fixed effects, which also capture campus effects. Courses are weighted by units earned and then normalized to give each student equal weight.

Table 7 shows that, unsurprisingly, the SAT Reading and Writing components each explain substantial variation in freshman writing course performance, while the SAT Math component explains the largest share of variation in Organic Chemistry performance. High school GPA also explains a significant share of performance variation in each course type, and when both admissions components are included simultaneously, once again high school GPA and the SAT scores explain similar amounts of variation in course performance. As in the full course analysis, these results suggest that each of the four admissions components—HSGPA, SAT Math, SAT Reading, and SAT Writing—provide valuable and differentiated information about expected UC student performance in their chosen first-year coursework.

Standardized Tests and Academic Preparation: Longer-Run Outcomes

Table 8 presents the percent of variance in graduation GPA accounted for by HSGPA and test scores. Results show that HSGPA and test scores are weak predictors of graduation GPA. They have almost equal explanatory power, each accounting for less than 20 percent of variance in graduation GPA. Adding SAT or ACT writing scores to SAT Math/Reading scores does not increase predictive value of Graduation GPA beyond SAT Reading/Math or ACT Composite scores. For example, for 2010 entering cohort, SAT Reading/Math accounts for 16 percent of variance in graduation GPA, while the sum of SAT Reading/Math and Writing scores (SAT Total) accounts for 18 percent, up by only two percentage points.

When HSGPA and SAT Reading/Math or ACT Composite scores were entered into the same models, the percent of variance accounted for by the two predictors together increased about five percentage points or one third of explanatory power by the HSGPA model alone. However, replacing SAT Reading/Math or ACT Composite scores with the sum of SAT Reading/Math and SAT Writing scores (SAT Total) or the sum of ACT Composite and ACT Writing scores (ACT Total) in these models does not change explanatory power. For example, for the 2010 entering cohort, the HSGPA and SATRM model accounts for 23 percent of variance in graduation GPA, while the HSGPA and SAT Total model accounts for 24 percent of variance. The two models are basically the same. This means that again writing scores do not contribute to predictive value of graduation GPA beyond HSGPA and SAT Math/Reading or ACT Composite. This is largely due to a high correlation between SAT Reading/Math and SAT Writing scores or ACT Composite and ACT Writing scores.

After controlling for campus and student characteristics including campus affiliation, major discipline, first-generation status, family income, and high school API quintile, HSGPA has more explanatory power of UC graduation GPA than the SAT Total in terms of standardized coefficients (Model 12 in Table 8). This means that SAT scores are more associated with the demographics than HSGPA, consistent with previous findings.

The influence of varying student characteristics on the relationship between HSGPA and SAT/ACT scores on graduation GPA is also examined based on 2010 entering students. Detailed model coefficients are also presented in Appendix D. Similar results were found that standardized test scores were the stronger predictor of graduation GPA than high school GPA by campus, discipline, race/ethnicity, first generation, family income, and CA API ranking. However, the difference in the explanatory power between HSGPA and SAT/ACT scores became smaller for graduation GPA than for first year GPA. For some groups, HSGPA predicted graduation GPA better. For example, HSGPA appeared to be a stronger predictor of graduation GPA than test scores for students at Davis, students majored in Arts, White students, not first generation students, students with family income over 107k, and etc.

Table 8. Percent of Variance in UC Graduation GPA Accounted for by HSGPA and the SAT/ACT

Model	2001	2005	2007	2010
(1) HSGPA	16%	17%	18%	16%
(2) SATRM	15%	15%	17%	16%
(3) SAT Writing	18%	19%	19%	18%
(4) SAT Total (SATRM + Writing)	18%	18%	19%	18%
(5) HSGPA + SATRM	22%	24%	25%	23%
Standardized Coefficients: HSGPA	*0.29*	*0.32*	*0.31*	*0.27*
SATRM	*0.27*	*0.28*	*0.28*	*0.29*
(6) HSGPA + SATTOTAL (SATRM + Writing)	27%	26%	26%	24%
Standardized Coefficients: HSGPA	*0.27*	*0.29*	*0.29*	*0.26*
SATTOTAL	*0.31*	*0.32*	*0.31*	*0.32*
(7) ACT Composite	18%	19%	18%	17%
(8) ACT Writing	*n/a*	*n/a*	20%	17%
(9) ACT Total (Composite + Writing)	*n/a*	*n/a*	19%	23%
(10) HSGPA + ACT Composite	23%	26%	25%	22%
Standardized Coefficients: HSGPA	*0.24*	*0.27*	*0.28*	*0.25*
ACT CMP	*0.33*	*0.34*	*0.31*	*0.31*
(11) HSGPA + ACT Total (ACT CMP and ACT Writing)	n/a	n/a	26%	23%
Standardized Coefficients: HSGPA	*n/a*	*n/a*	*0.26*	*0.24*
ACT Total	*n/a*	*n/a*	*0.34*	*0.33*
(12) HSGPA + SAT Total + Demographics	32%	33%	34%	31%
Standardized Coefficients: HSGPA	*0.38*	*0.40*	*0.40*	*0.37*
Standardized Coefficients: SATTOTAL	*0.31*	*0.29*	*0.29*	*0.32*

As shown in Table 9, both HSGPA and test scores are very weak predictors of first year retention and four year graduation. The models with HSGPA and test scores only account for about 2.5 percent of variance in first year retention and eight percent of variance in four year graduation. In addition, the standardized coefficients show that there is no significant difference in explanatory power of predicting first year retention and four year graduation between HSGPA and test scores. The percent of variance accounted for by these predictors has remained about the same since 2001, so the results from the logistic regression models based on 2015 entering cohort for first year retention and 2010 for four year graduation are presented here in the table.

Again after controlling for student demographics, HSGPA is still a stronger predictor for first year retention and four-year graduation, with more explanatory power than SAT scores in terms of standardized coefficients (Model 5 in Table 9). This finding further indicates that SAT scores are more associated with the demographics than HSGPA, consistent with previous findings.

Table 9. Statistics of Logistic Regression Models Predicting First Year Retention and Four Year Graduation by HSGPA and SAT/ACT Scores

Model	First Year Retention (2015 Entering Cohort)	Four Year Graduation (2010 Entering Cohort)
(1) HSGPA + SATRM (R-squared)	2.3%	7.5%
Standardized Coefficients: HSGPA	*0.22*	*0.23*
SATRM	*0.22*	*0.17*
(2) HSGPA + SATTOTAL (SATRM + Writing, R-squared)	2.4%	7.9%
Standardized Coefficients: HSGPA	*0.20*	*0.21*
SATTOTAL	*0.22*	*0.19*
(3) HSGPA + ACT Composite (R-squared)	2.5%	7.7%
Standardized Coefficients: HSGPA	*0.20*	*0.19*
ACT CMP	*0.23*	*0.20*
(4) HSGPA + ACT Total (ACT CMP and ACT Writing, R-squared)	2.4%	8.1%
Standardized Coefficients: HSGPA	0.20	0.18
ACT Total	0.22	0.22
(5) HSGPA + SAT Total + Demographics	3.0%	12.0%
Standardized Coefficients: HSGPA	*0.19*	*0.25*
Standardized Coefficients: SATTOTAL	*0.14*	*0.16*

Conclusion

This study reviewed literature about the validity of the SAT and ACT and examined the relationship between these tests and college preparation of UC California resident applicants and college performance of enrolled students at UC.

The following conclusions emerged from the literature review:

- High school GPA and test scores are moderately correlated. Previous research also indicates that socioeconomic status is associated with SAT and ACT scores.

- SAT/ACT test scores are positively associated with college success in terms of freshman GPA, graduation GPA, first-year retention, and graduation. In addition, research demonstrates the increment in the predictive validity afforded by the SAT/ACT over HSGPA. On the other hand, evidence also suggests that standardized test scores are less predictive across certain subgroups.

Analysis based on UC data shows that:

- The majority of UC applicants submit SAT scores, while more and more students submit ACT scores. The average test scores changed slightly over years. SAT/ACT scores are more associated with demographic factors than HSGPA. The explanatory power of parental education, the URG status, and parents' income has been growing with parental education continuing to be the strongest predictor of SAT/ACT scores. This possibly could be a problem as the UC seeks to find race neutral ways to achieve racial diversity at selective campuses.[53] SAT/ACT scores may have a disproportionate effect on admit rates for URG students.

- Similar to what previous research findings show, test scores are moderately correlated to high school GPA. Apparently, there is some overlap between what the tests measure and what HSGPA measures, but not that much. The College Board and ACT hold that SAT and ACT scores are designed to measure knowledge and skills in high school subject areas and readiness for first year college courses. High school GPA likely represents both knowledge and skills and also behavior (such as completing homework).

- The analysis of the relationship between the new SAT and new ACT and UC's Analytical Writing Placement Exam (AWPE) shows that high scores on the SAT EBRW and ACT ELA are associated with passing the AWPE. However, on the SAT Essay, even at the top score of 24, only 69 percent passed the AWPE. Based on the results of data analysis, UCOPE has accepted the new ACT ELA and SAT EBRW scores but not the new SAT Essay scores as methods of meeting UC's Entry Level Writing Requirement (ELWR).

- Regarding college success, findings from UC's data analysis are generally consistent with those from previous research. SAT scores and HSGPA are both moderate predictors of student first-year GPA and graduation GPA. They are weak predictors of student retention and graduation. HSGPA predicted first-year GPA slightly better than SAT scores from 2001 to 2005 before controlling for student characteristics. Over the years, the explanatory power of HSGPA has been decreasing. For graduation GPA, HSGPA and test scores have the same explanatory power. HSGPA and test scores are not strong predictors of first-year retention and four-year graduation. However, after controlling for student characteristics, HSGPA and test scores have the same explanatory power of freshman GPA, but HSGPA is a stronger predictor of the first year retention, graduation GPA and four-year graduation.

- Models that combine HSGPA along with standardized test scores predict student success better than ones that only use one or the other. In contrast, the addition of the SAT/ACT writing tests did little to improve the prediction of student success above and beyond HSGPA and SAT Reading/Math and ACT composite scores.

[53] UC Office of the President. (2015). UC Files Amicus Brief in Affirmative Action Case. Retrieved from https://www.universityofcalifornia.edu/press-room/uc-files-amicus-brief-affirmative-action-case on October 16, 2018.

In summary, this report presents what literature and UC data show about the relationship between standardized test scores and high school GPA, and predictive value of HSGPA and test scores in student success as measured by freshman GPA, first year retention, graduation GPA and four-year graduation. IRAP will do a series of simulations to further examine what eligibility pool and admissions outcomes would look like if UC used different measures to determine eligibility and admit students.

Appendix A: Test Takers and Average Test Scores of UC California Applicants from 2001 to 2018

Table A1: Test Takers and Average Test Scores, UC California Applicants, 2001 to 2018

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
CA Freshman Applicants	59,747	62,903	66,774	63,852	65,851	71,011	74,509	80,029	81,113	82,341	85,187	93,460	99,447	100,077	103,259	105,547	111,869	120,026
% Submitted Scores																		
SAT I Verbal Math	98.4%	98.4%	98.4%	98.3%	98.3%	88.4%	0.8%											
SAT II Writing	96.9%	96.8%	96.1%	96.0%	94.8%	48.5%	0.2%											
SAT I & SATII Writing	96.0%	95.8%	95.0%	94.9%	93.8%	48.1%	0.1%											
SAT Verbal Math						98.3%	97.9%	97.5%	97.0%	94.6%	94.0%	93.2%	91.5%	89.3%	86.7%	84.6%	19.5%	0.8%
SAT Writing						98.5%	97.9%	97.5%	97.0%	94.6%	94.0%	93.2%	91.5%	89.3%	86.7%	84.6%	20.4%	0.8%
SAT Reading Math Writing						98.1%	97.9%	97.5%	97.0%	94.6%	94.0%	93.2%	91.5%	89.3%	86.7%	84.6%	20.4%	0.8%
SAT Math and EBRW (Evidence-Based Reading and Writing)																	73.2%	86.4%
SAT Essay																	73.2%	86.4%
ACT Composite	27.6%	27.7%	32.7%	32.9%	33.2%	34.9%	35.1%	39.4%	44.0%	46.2%	47.3%	45.5%	46.4%	49.1%	52.2%	53.4%	52.5%	44.3%
ACT English with Writing						29.5%	32.8%	37.8%	43.1%	45.4%	46.9%	45.0%	45.9%	48.7%	51.7%	50.2%	0.2%	0.0%
ACT ELA (English Language Arts)																	45.5%	37.6%
ACT Writing subscore																	27.9%	37.4%
Average Scores																		
SAT I Verbal Math	1171	1163	1164	1172	1177	1162	1144											
SAT II Writing	575	578	573	581	579	649	559											
SAT I & SATII Writing	1751	1746	1744	1760	1766	1816	1727											
SAT Verbal Math						1161	1159	1153	1154	1157	1153	1137	1135	1136	1128	1117	1241	1327
SAT Writing						572	570	569	571	573	572	566	560	562	556	550	610	643
SAT Reading Math Writing						1734	1729	1722	1725	1729	1725	1703	1695	1698	1684	1667	1845	1952
SAT Math and EBRW (Evidence-Based Reading and Writing)																	1167	1203
SAT Essay																	16	16
ACT Composite	24	24	24	24	24	24	24	25	25	25	25	25	25	25	25	25	26	25
ACT English with Writing						24	24	24	24	24	24	23	23	24	24	24	29	27
ACT ELA (English Language Arts)																	25	25
ACT Writing subscore																	8	8

Appendix B: Average Test Scores of UC California Applicants by Demographic Characteristics, Selected Years

Table B1: SAT Reading/Math and Writing by Parents' Highest Level of Education

Year	No High School		Some High School		High School Graduate		Some College		2 Year College		4 Year College		Post Graduate Study		Unknown	
	SAT RM	SAT W	SAT RM	SAT W	SAT RM	SAT W	SAT RM	SAT W	SAT RM	SAT W	SAT RM	SAT W	SAT RM	SAT W	SAT RM	SAT W
2001	978	479	1008	492	1079	526	1105	542	1125	550	1182	577	1249	617	1221	602
2005	981	483	1008	491	1082	528	1107	543	1135	555	1191	582	1264	625	1229	604
2007	967	472	998	485	1069	522	1087	533	1107	542	1177	578	1253	620	1210	594
2012	949	471	977	485	1049	519	1068	528	1104	544	1183	588	1271	637	1118	549
2015	952	468	973	479	1040	510	1061	522	1091	536	1181	583	1279	635	1114	544
2018	1052	15	1069	15	1126	15	1145	15	1179	16	1260	17	1342	17	1192	16

Table B2: SAT Reading/Math and Writing by Parents' Income (2001 dollars)

Year	$0-$56k		$56k up to $112k		$112k up to $169k		$169k+		Unknown	
	SATRM	SATW	SATRM	SATW	SATRM	SATW	SATRM	SATW	SATRM	SATW
2001	1140	559	1180	578	1222	601	1251	617	n/a	n/a
2005	1084	531	1184	581	1236	608	1255	619	1248	616
2007	1068	521	1167	573	1213	599	1243	617	1244	616
2012	1049	520	1185	588	1243	620	1280	644	1208	607
2015	1042	512	1184	583	1249	619	1297	647	1205	601
2018	1127	15	1251	17	1311	17	1358	18	1256	16

Table B3: SAT Reading/Math and Writing by Race/Ethnicity

Year	African American		American Indian		Hispanic/Latinx		Asian/Pacific Islander		White	
	SATRM	SATW	SATRM	SATW	SATRM	SATW	SATRM	SATW	SATRM	SATW
2001	1035	520	1169	570	1047	521	1182	566	1215	603
2005	1042	526	1184	587	1049	524	1198	575	1230	610
2007	1027	516	1160	576	1032	514	1187	570	1213	603
2012	1015	509	1164	581	1013	507	1206	594	1215	607
2015	1025	512	1149	569	1008	499	1214	595	1217	602
2018	1120	15	1225	16	1099	15	1298	17	1288	17

Note: SAT Writing refers to SATII Writing for 2001 and 2005 and SAT Writing for 2007, 2012 and 2015, and SAT Essay for 2018. International students are included in each racial/ethnic category based on their self-reported race/ethnicity.

Table B4: ACT Composite and Writing by Parents' Highest Level of Education

Year	No High School		Some High School		High School Graduate		Some College		2 Year College		4 Year College		Post Graduate Study		Unknown	
	ACT CMP	ACT W	ACT CMP	ACT W	ACT CMP	ACT W	ACT CMP	ACT W	ACT CMP	ACT W	ACT CMP	ACT W	ACT CMP	ACT W	ACT CMP	ACT W
2001	19	n/a	20	n/a	22	n/a	23	n/a	23	n/a	25	n/a	26	n/a	25	n/a
2005	20	n/a	20	n/a	22	n/a	23	n/a	24	n/a	25	n/a	26	n/a	26	n/a
2007	20	19	20	20	22	22	23	23	23	23	25	25	27	26	26	25
2012	20	19	21	19	22	21	23	22	24	23	26	25	28	26	25	23
2015	20	19	21	19	23	21	23	22	24	23	27	25	29	27	25	23
2018	20	19	21	20	23	21	24	22	25	23	28	26	30	28	26	24

Table B5: ACT Composite and Writing by Parents' Income (2001 dollars)

Year	$0-$56k		$56k up to $112k		$112k up to $169k		$169k+		Unknown	
	ACT CMP	ACT W	ACT CMP	ACT W	ACT CMP	ACT W	ACT CMP	ACT W	ACT CMP	ACT W
2001	23	n/a	24	n/a	25	n/a	26	n/a	n/a	n/a
2005	22	n/a	25	n/a	26	n/a	26	n/a	26	n/a
2007	22	21	25	24	26	25	27	26	27	26
2012	22	21	26	25	28	26	29	27	27	25
2015	22	21	27	25	28	27	29	28	27	26
2018	23	22	27	26	29	27	30	29	29	27

Table B6: ACT Composite and Writing by Race/Ethnicity

Year	African American		American Indian		Hispanic/Latinx		Asian/Pacific Islander		White	
	ACT CMP	ACT W	ACT CMP	ACT W	ACT CMP	ACT W	ACT CMP	ACT W	ACT CMP	ACT W
2001	21	n/a	25	n/a	21	n/a	24	n/a	26	n/a
2005	21	n/a	25	n/a	21	n/a	24	n/a	26	n/a
2007	21	21	25	25	21	21	25	24	26	26
2012	22	21	26	25	22	20	26	24	27	26
2015	22	22	26	25	22	21	27	25	28	27
2018	23	22	27	26	22	21	29	27	29	27

Notes: ACT Writing refers to ACT English with Writing for 2007, 2012, and 2015 and ACT English Language Arts for 2018. International students are included in each racial/ethnic category based on their self-reported race/ethnicity.

Appendix C: Standardized Regression Coefficients for HSGPA and SAT/ACT Scores in UC Freshman GPA Models, 2015 Freshman Entering Cohort

	SAT Models			ACT Models		
	HSGPA	SAT_RM	% Variance Explained	HSGPA	ACTCMP	% Variance Explained
UC System	0.23	0.36	25%	0.22	0.38	26%
By Campus						
Berkeley	0.20	0.36	22%	0.22	0.32	19%
Davis	0.18	0.44	27%	0.17	0.44	27%
Irvine	0.31	0.36	28%	0.32	0.38	30%
Los Angeles	0.16	0.39	23%	0.17	0.41	25%
Merced	0.29	0.29	18%	0.29	0.28	18%
Riverside	0.24	0.25	11%	0.22	0.28	12%
San Diego	0.24	0.38	21%	0.22	0.39	21%
Santa Barbara	0.29	0.32	21%	0.28	0.36	24%
Santa Cruz	0.22	0.31	16%	0.22	0.34	18%
By Enrolled Majors						
Arts	0.21	0.28	15%	0.15	0.40	22%
Engineering/Computer Sciences	0.16	0.42	26%	0.17	0.40	25%
Humanities	0.30	0.35	31%	0.32	0.35	33%
Life Sciences	0.25	0.42	31%	0.23	0.44	31%
Physical Sciences/Math	0.28	0.36	27%	0.26	0.37	28%
Professional Fields	0.24	0.38	26%	0.22	0.39	25%
Social Sciences/Psychology	0.29	0.34	26%	0.26	0.39	29%
Undeclared/All Others/Unknown	0.25	0.31	20%	0.25	0.33	21%
By Race/Ethnicity						
African American	0.27	0.25	18%	0.28	0.26	20%
Asian	0.22	0.35	22%	0.19	0.38	22%
Latino(a)	0.23	0.24	14%	0.23	0.25	15%
White	0.28	0.23	18%	0.28	0.22	28%
By First Generation Status						
Not First Generation	0.26	0.29	21%	0.24	0.29	20%
First Generation	0.22	0.29	22%	0.22	0.30	17%
By Family Income						
$0-$53,999	0.22	0.30	18%	0.22	0.32	19%
$54,000-$106,999	0.22	0.31	19%	0.21	0.32	20%
$107,000- $160,999	0.26	0.28	20%	0.24	0.28	19%
$161,000 or higher	0.27	0.26	20%	0.27	0.25	19%
Missing	0.24	0.27	19%	0.07	0.01	16%
By High School API Quintile (2013)						
1 through 3	0.22	0.20	12%	0.25	0.20	13%
4 through 7	0.25	0.26	19%	0.25	0.27	19%
8 and higher	0.30	0.26	22%	0.26	0.28	21%
Private	0.27	0.26	21%	0.30	0.23	20%
Public Missing	0.21	0.34	22%	0.23	0.36	24%
Missing Other	0.17	0.25	11%	0.18	0.15	8%

Appendix D: Standardized Regression Coefficients for HSGPA and SAT/ACT Scores in UC Graduation GPA Models, 2010 Freshman Entering Cohort

	SAT Models			ACT Models		
	HSGPA	SAT_RM	% Variance Explained	HSGPA	ACTCMP	% Variance Explained
UC System	0.27	0.29	23%	0.25	0.31	22%
By Campus						
Berkeley	0.20	0.30	18%	0.20	0.33	20%
Davis	0.34	0.30	23%	0.33	0.32	24%
Irvine	0.24	0.27	14%	0.22	0.26	12%
Los Angeles	0.20	0.32	18%	0.17	0.31	16%
Merced	0.26	0.33	17%	0.34	0.32	23%
Riverside	0.25	0.25	13%	0.21	0.27	13%
San Diego	0.29	0.32	20%	0.25	0.34	20%
Santa Barbara	0.27	0.28	17%	0.24	0.32	18%
Santa Cruz	0.31	0.30	21%	0.29	0.34	23%
By Enrolled Majors						
Arts	0.35	0.19	20%	0.29	0.22	17%
Engineering/Computer Sciences	0.30	0.28	24%	0.27	0.31	24%
Humanities	0.29	0.30	26%	0.28	0.34	28%
Life Sciences	0.30	0.35	30%	0.27	0.36	27%
Physical Sciences/Math	0.30	0.33	26%	0.28	0.36	27%
Professional Fields	0.29	0.27	23%	0.30	0.31	27%
Social Sciences/Psychology	0.30	0.29	25%	0.26	0.32	24%
Undeclared/All Others/Unknown	0.24	0.32	22%	0.24	0.31	21%
By Race/Ethnicity						
African American	0.22	0.25	16%	0.26	0.22	16%
Asian	0.29	0.26	21%	0.26	0.26	18%
Latino(a)	0.24	0.27	18%	0.24	0.28	18%
White	0.30	0.19	17%	0.27	0.20	15%
By First Generation Status						
Not First Generation	0.31	0.20	19%	0.29	0.21	18%
First Generation	0.24	0.27	18%	0.23	0.28	17%
By Family Income						
$0-$53,999	0.25	0.28	19%	0.23	0.29	18%
$54,000-$106,999	0.27	0.25	18%	0.25	0.25	18%
$107,000- $160,999	0.30	0.19	17%	0.29	0.21	18%
$161,000 or higher	0.36	0.16	20%	0.31	0.17	17%
Missing	0.32	0.21	21%	0.31	0.20	19%
By High School API Quintile (2013)						
1 through 3	0.24	0.23	15%	0.24	0.21	14%
4 through 7	0.26	0.26	19%	0.23	0.28	19%
8 and higher	0.34	0.18	20%	0.32	0.19	20%
Private	0.31	0.25	24%	0.29	0.29	24%
Public Missing	0.21	0.39	26%	0.12	0.40	22%
Missing Other	0.31	0.20	19%	0.34	0.26	25%

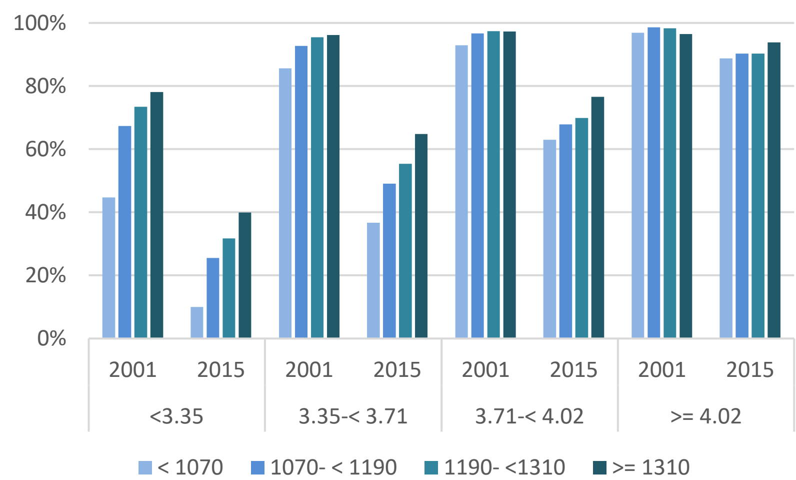
Appendix E: Explanation for Change of Variance in Freshman GPA Accounted for by High School GPA and Test Scores

Over the past 10 years, the amount of freshman GPA variation explained by high school GPA (HSGPA) has fallen significantly. Table 3 shows that, in terms of aggregate freshman GPA, HSGPA's explanatory power fell from 17% in 2001 and 20% in 2007 to 13% in 2015. There are a large number of possible explanations for this decline. This section provides the results of some preliminary analysis conducted to examine the change.

Change of the Admit Pool

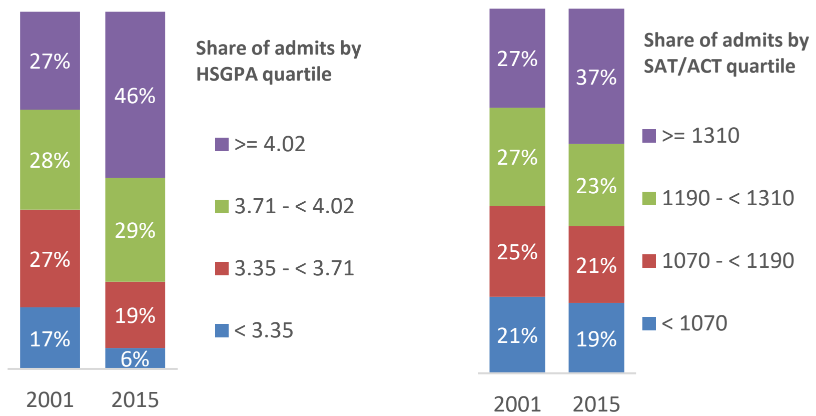
The change of variation in HSGPA and test scores may be related to the admit pool. While always skewed toward the top quartiles of SAT/ACT and HSGPA, the distribution of admits has changed over time (Figures E1 and E2). In particular, admits became more concentrated in the top quintiles of HSGPA between 2001 and 2015. This change has definitely resulted in less variation in HSGPA and may be related to the diminishing power of HSGPA to predict UC freshman GPA during this time period. Almost all applicants with high HSGPA are admitted, but the admit rates of those with high SAT/ACT depend on HSGPA. This also suggests that HSGPA is given more weight in the admission process than SAT/ACT.

Figure C1. Admit Rate by HSGPA and SAT/ACT Quartiles



Note: Thresholds for HSGPA and SAT/ACT quartiles were constructed from 2001 data and used for all years for consistency.

Figure C2. Share of Admits by HSGPA and SAT/ACT Quartiles



Note: Thresholds for HSGPA and SAT/ACT quartiles were constructed from 2001 data and used for all years for consistency.

High School Grade Inflation, Enrollment in Low-Grade and Larger Freshman Courses

There are a large number of possible explanations for the decline in the explanatory power of HSGPA over time, many of which can be adjudicated between using course-level grade performance.

Table C1 shows that at the freshman course level, HSGPA explanatory power has fallen by 18.8 percent since 2012, from 4.3 percent to 3.5 percent (note that the SAT and HSGPA explain less

Table C1: Changes in Explanatory Power of HSGPA for Freshman Grades

	2012-2013		2016-2017		Δ	
	FE R^2	GPA R^2	FE R^2	GPA R^2	FE R^2	GPA R^2
Overall		**4.3%**		**3.5%**		**-18.8%**
High School	5.6%	**5.3%**	5.8%	**4.0%**	3.5%	**-25.0%**
SAT	5.2%	**2.6%**	4.5%	**2.2%**	-12.8%	**-15.3%**
Course Department	9.4%	**5.8%**	11.5%	**5.1%**	23.2%	**-12.6%**
Course Size	4.5%	**4.9%**	4.4%	**4.0%**	-2.1%	**-18.1%**

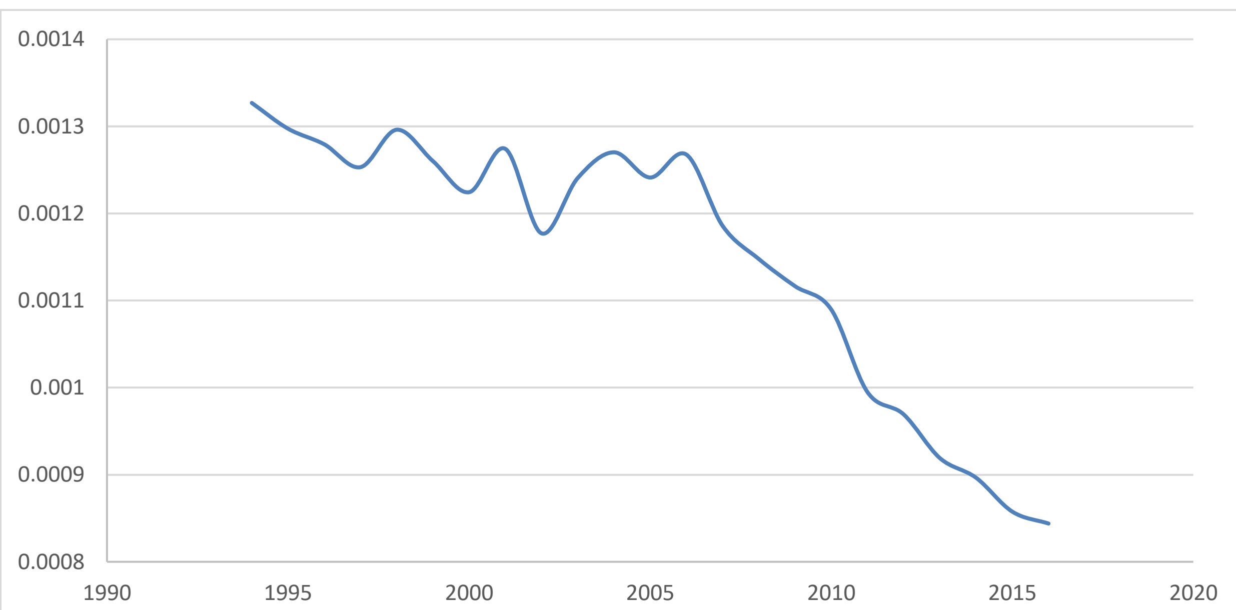
Note: "FE R2" measures the percent of variation explained by each additional group of fixed effects (FE), like FE's for every high school or every SAT score. "GPA R2" measures the percent of remaining variation explained by HS GPA. "Δ" shows the percent change from 2012 to 2016. Valid explanations for the decline in HS GPA explanatory power for freshman grades should decrease the GPA R2, since some of the decline would be explained away by the FEs.

course-specific variation than they do freshman aggregate GPA variation, since the latter are averaged over many courses and thus present a smoother and more-easily-modeled distribution of student performance). This baseline decline is conditional on gender, ethnicity, and campus. Each row of Table C1 adds an additional set of controls, removing one dimension of grade variation and testing how that information's removal impacts the degree to which HSGPA's explanatory power has declined since 2012. Controls that lead to a decline in the change in HSGPA's R^2 are interpreted as playing a role in HSGPA's explanatory decline.

The second row of Table C1 tests the hypothesis that HSGPA's explanatory power is declining because UC is enrolling students from a wider range of high schools with varying GPA standards, muddying the informational content of HSGPA. Indeed, Figure C3 shows that the concentration of high schools from which UC has enrolled students (as measured by a Normalized Herfindahl Index) has indeed been declining since about 2007 after a period of relative stability, implying that UC has been taking students from increasingly-dispersed high schools, likely as a result of its Eligibility in the Local Context program and the expansion of its Merced campus. However, when the relationship between HSGPA and freshman GPA is measured conditional on high school fixed effects—that is, only comparing students who attended the same high school—the decline in the explanatory power of HSGPA grows to 25.0 percent. We conclude that the wider distribution of UC students' high schools plays no role in the decline in HSGPA's explanatory power, since the within-high-school decline is even larger than the across-high-school decline.

The third row tests the hypothesis that HSGPA has less variance among students with the same SAT score, which would likely arise if the distribution of HSGPAs was increasingly compressed as a result of high schools' grade inflation. Controlling for fixed effects for every SAT score, the

Figure C3: High School Concentration of First-Year UC Students



Note: Annual Normalized Herfindahl Index of high schools attended by freshman California-resident UC students, measuring the degree to which students come from a concentrated set of schools.

decline in HSGPA explanatory power falls slightly to 15.3 percent. We conclude that a small part of HSGPA's decline in explanatory power results from its diminished relationship with SAT score, likely resulting from high schools' grade inflation.

The fourth row of Table C1 tests whether the decline in HSGPA explanatory power is the result of which departments freshmen choose to enroll in. If high-HSGPA students are increasingly enrolling in low-grade courses like Engineering and Economics, then HSGPA's explanatory power would decline (since it would appear that high-HSGPA students were performing poorly, when in fact they're merely taking courses that award lower grades). Indeed, this appears to be the largest identified factor in the decline in HSGPA's explanatory power, explaining about 1/3 of the decline. Moreover, a full 10 percent of freshman grade variation can be explained by the department in which each course was taken, and the explanatory power of departments has increased by 25 percent since 2012. Cross-department variation is a key explanation for changes in freshman GPA evaluation in the past five years.

Finally, if students have tended towards taking larger freshman courses with more formal test-based examination (like multiple choice exams), then HSGPA would lose explanatory power because student performance would increasingly resemble SAT examination as opposed to the multi-modal examinations provided in high school. The UC Accountability Report shows that the availability of small classes at UC has fallen dramatically over the past 5 years, limiting freshman course options. Indeed, controlling for course size shows that size explains a small part of the decline in HSGPA explanatory power.

Overall, we conclude that almost half of the decline in HSGPA's ability to explain freshman course performance can be attributed to changes in freshman course enrollment across departments (the largest observed factor), HSGPA compression as a result of grade inflation and increased UC selectivity, and increases in the average size of freshman courses. While about half of the explanatory decline remains unexplained, we conclude that increased dispersion in the high schools sending students to UC does not play a role in HSGPA's explanatory decline, since within-high-school explanatory power has fallen even faster than overall explanatory power (leaving no room for a decline across high schools).

In summary, the change of admit pool and thus enrollment may result in less variability of HSGPA and a slightly greater variability of test scores. This change may be related to increase of eligible applicants, admission process, and student enrollment behavior. Further analyses need to be conducted to test this hypothesis. Regardless of how it happened, it definitely explains part of HSGPA's decline and test scores' increase in explanatory power. In addition, high school grade inflation and enrollment in low-grade course and large freshman courses also explain part of the decline in HSGPA explanatory power in course performance.

HSGPA versus SAT as a predictor of UC freshman grades across demographic groups, 2010-2012

Category	N	GPA-only models		SAT-only models	
		Slope	R^2	Slope	R^2
All	90504	0.39	0.16	0.46	0.21
<$30K	20178	0.34	0.12	0.41	0.17
$30-$60K	19232	0.36	0.13	0.42	0.18
$60-$120K	19307	0.36	0.13	0.39	0.15
$120K+	21339	0.41	0.17	0.35	0.12
Asian	37899	0.38	0.15	0.42	0.18
Black	3652	0.29	0.09	0.37	0.14
Hispanic	22522	0.33	0.11	0.39	0.15
White	23046	0.38	0.14	0.34	0.11
Not 1st Gen	47626	0.39	0.15	0.37	0.14
First Gen	40405	0.34	0.12	0.4	0.16

Table 1. High school GPA (HSGPA) compared to SAT as a predictor of freshman grades at UC for different demographic groups, 2010-2012. SAT is calculated as SATRR + SATRM. Both SAT and HSGPA are normalized (z-scored) prior to ordinary least-squares regression.

HSGPA <u>plus</u> SAT to predict UC freshman grades across demographic groups, 2010-2012

Category	N	Combined models		
		GPA slope	SAT slope	R^2
All	90504	0.25	0.36	0.26
<$30K	20178	0.23	0.33	0.22
$30-$60K	19232	0.23	0.34	0.23
$60-$120K	19307	0.25	0.29	0.20
$120K+	21339	0.32	0.21	0.20
Asian	37899	0.26	0.32	0.24
Black	3652	0.19	0.30	0.17
Hispanic	22522	0.22	0.31	0.20
White	23046	0.29	0.22	0.18
Not 1st Gen	47626	0.29	0.26	0.21
First Gen	40405	0.23	0.32	0.21

Table 2. High school GPA (HSGPA) combined with SAT to predict freshman grades at UC for different demographic groups, 2010-2012. SAT is calculated as SATRR + SATRM. Both SAT and HSGPA are normalized (z-scored) prior to ordinary least-squares regression.

HSGPA <u>versus</u> SAT as predictors of graduation GPAs at UC across demographic groups, 2010

Category	**N**	**GPA-only models**		**SAT-only models**	
		Slope	**R^2**	**Slope**	**R^2**
All	25523	0.39	0.15	0.4	0.16
<$30K	5154	0.34	0.12	0.36	0.13
$30-$60K	5040	0.36	0.13	0.36	0.13
$60-$120K	5457	0.37	0.14	0.34	0.11
$120K+	5519	0.41	0.16	0.32	0.1
Asian	11077	0.39	0.15	0.38	0.14
Black	869	0.32	0.1	0.34	0.12
Hispanic	5239	0.33	0.11	0.36	0.13
White	6917	0.38	0.14	0.31	0.09
Not 1st Gen	13876	0.4	0.16	0.33	0.11
First Gen	10273	0.34	0.11	0.35	0.12

Table 3. High school GPA (HSGPA) compared to SAT as a predictor of GPA at graduation for different demographic groups, 2010. SAT is calculated as SATRR + SATRM. Both SAT and HSGPA are normalized (z-scored) prior to ordinary least-squares regression.

HSGPA <u>plus</u> SAT to predict graduation GPAs at UC across demographic groups, 2010

Category	N	Combined models GPA slope	SAT slope	R^2
All	25523	0.28	0.29	0.22
<$30K	5154	0.25	0.28	0.19
$30-$60K	5040	0.26	0.27	0.19
$60-$120K	5457	0.28	0.23	0.18
$120K+	5519	0.33	0.18	0.19
Asian	11077	0.29	0.26	0.21
Black	869	0.23	0.26	0.16
Hispanic	5239	0.23	0.27	0.17
White	6917	0.3	0.19	0.17
Not 1st Gen	13876	0.31	0.2	0.19
First Gen	10273	0.24	0.27	0.17

Table 4. High school GPA (HSGPA) combined with SAT to predict GPAs at graduation for different demographic groups, 2010. SAT is calculated as SATRR + SATRM. Both SAT and HSGPA are normalized (z-scored) prior to ordinary least-squares regression.

Percentage of Applicants Reporting SAT Scores Who Were Admitted by Campus, High School GPA (HSGPA) Range and SAT Range. Fall 2018 California Resident Applicants.

- The SAT used is the maximum sum of the reading and math sections. No concordance from ACT scores was used.
- The means and standard deviations used were computed for each campus or the system as a whole.
- Cells corresponding to fewer than 30 students are blank.
- In the tables below:
  - "High" = more than one standard deviation above the mean
  - "Mid-High" = between the mean and one standard deviation above the mean
  - "Mid-Low" = between the mean and one standard deviation below the mean
  - "Low" = more than one standard deviation below the mean

System

	Percent of Applicants Admitted				
	High SAT	Mid-High SAT	Mid-Low SAT	Low SAT	Net
High HSGPA	95%	95%	94%	84%	95%
Mid-High HSGPA	85%	75%	73%	64%	75%
Mid-Low HSGPA	65%	51%	43%	28%	43%
Low HSGPA	41%	25%	17%	9%	16%
Net	86%	68%	52%	27%	58%

Berkeley

	Percent of Applicants Admitted				
	High SAT	Mid-High SAT	Mid-Low SAT	Low SAT	Net
High HSGPA	54%	42%	43%	30%	47%
Mid-High HSGPA	36%	20%	18%	14%	23%
Mid-Low HSGPA		2%	2%	2%	2%
Low HSGPA					1%
Net	36%	17%	11%	4%	16%

Davis

	Percent of Applicants Admitted				
	High SAT	Mid-High SAT	Mid-Low SAT	Low SAT	Net
High HSGPA	88%	91%	94%	92%	90%
Mid-High HSGPA	48%	42%	45%	50%	45%
Mid-Low HSGPA	3%	3%	5%	6%	4%
Low HSGPA					1%
Net	58%	38%	27%	15%	34%

Irvine

	Percent of Applicants Admitted				
	High SAT	Mid-High SAT	Mid-Low SAT	Low SAT	Net
High HSGPA	80%	68%	65%	42%	73%
Mid-High HSGPA	40%	33%	26%	13%	30%
Mid-Low HSGPA	9%	4%	2%	1%	3%
Low HSGPA					1%
Net	53%	30%	16%	4%	25%

STTF Report Appendix III

Los Angeles

Percent of Applicants Admitted					
	High SAT	Mid-High SAT	Mid-Low SAT	Low SAT	Net
High HSGPA	59%	34%	30%		45%
Mid-High HSGPA	19%	8%	9%	5%	11%
Mid-Low HSGPA		1%	2%	1%	2%
Low HSGPA					1%
Net	31%	10%	7%	2%	12%

Merced

Percent of Applicants Admitted					
	High SAT	Mid-High SAT	Mid-Low SAT	Low SAT	Net
High HSGPA	99%	99%	98%	91%	99%
Mid-High HSGPA	98%	97%	90%	57%	90%
Mid-Low HSGPA	83%	74%	62%	19%	59%
Low HSGPA	16%	15%	10%		9%
Net	91%	82%	66%	25%	68%

Riverside

Percent of Applicants Admitted					
	High SAT	Mid-High SAT	Mid-Low SAT	Low SAT	Net
High HSGPA	97%	96%	91%	61%	94%
Mid-High HSGPA	94%	84%	57%	13%	70%
Mid-Low HSGPA	77%	48%	11%	1%	25%
Low HSGPA	27%	7%			3%
Net	90%	69%	33%	5%	49%

San Diego

Percent of Applicants Admitted					
	High SAT	Mid-High SAT	Mid-Low SAT	Low SAT	Net
High HSGPA	75%	75%	73%	57%	74%
Mid-High HSGPA	38%	33%	31%	20%	32%
Mid-Low HSGPA	13%	8%	6%	2%	6%
Low HSGPA					1%
Net	50%	32%	19%	6%	27%

STTF Report Appendix III

Santa Barbara

Percent of Applicants Admitted					
	High SAT	Mid-High SAT	Mid-Low SAT	Low SAT	Net
High HSGPA	90%	74%	69%	42%	80%
Mid-High HSGPA	55%	32%	31%	17%	35%
Mid-Low HSGPA	11%	7%	8%	2%	6%
Low HSGPA					1%
Net	63%	31%	19%	5%	29%

Santa Cruz

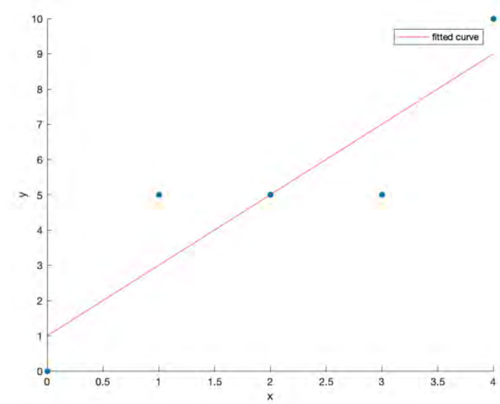
Percent of Applicants Admitted					
	High SAT	Mid-High SAT	Mid-Low SAT	Low SAT	Net
High HSGPA	97%	93%	61%	30%	89%
Mid-High HSGPA	93%	57%	35%	15%	54%
Mid-Low HSGPA	36%	18%	13%	1%	14%
Low HSGPA					1%
Net	84%	48%	22%	4%	39%

Interpreting R^2 as Percent Explained Variance

In regression, R^2 is a measure of how well a model fits a set of data. Sometimes, it is referred to as the "proportion of explained variation." If the data is perfectly fit by the model, then $R^2 = 1$; if the model doesn't' fit the data at all then $R^2 = 0$; This brief describes where the language of "explained variation" comes from and hopefully will provide a better understanding of how to interpret R^2.

Consider a very simple set of data y = {0, 5, 5, 5, 10}. If we know nothing about where this data came from, it would be natural to at least calculate the mean, which is just $\bar{y} = 5$. The "residual" is the difference between each data point and the mean, which in this case is r = {-5, 0, 0, 0, 5}. The "variance" of the data from the mean is called the "null variance" (no model, just calculated from the mean) and is the sum of the square of the residuals. In this case, the null variance var_0= 50.

Now suppose that we can associate another variable x with y in a way that we can express y as a function of x: $y = f(x)$. For our data, suppose that x = {0, 1, 2, 3, 4}. We can examine the relationship between the two variables by plotting one against the other, and let's now fit this data with a straight line that minimizes the variance between the predicted y's at each x and the actual y's.



Doing so gives $f(x) = 1 + 2x$, which gives predicted y's of y' ={1, 3, 5, 7, 9}. The residuals for this model are then r' ={-1, 2, 0, -2, 1}. The variance is again the sum of the squares of the residuals, which in this case is model variance var_m = 10. This is obviously smaller than the null variance of 50, and as a percentage, the model reduces the variance by 80%, or we might say that the model "explains" 80% of the originally observed variation.

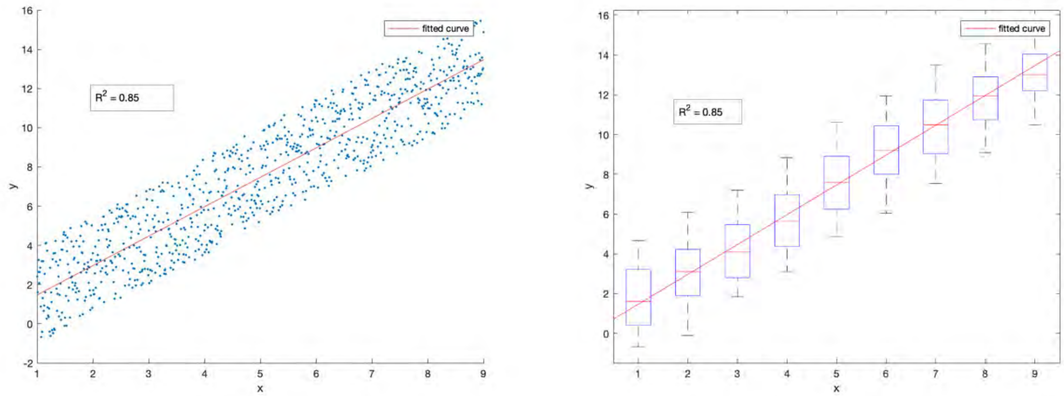
In general for this sort of linear regression, R^2 can be computed as the normalized difference between the null and the model variances,

$$R^2 = \frac{var_0 - var_m}{var_0} = 1 - \frac{var_m}{var_0}.$$

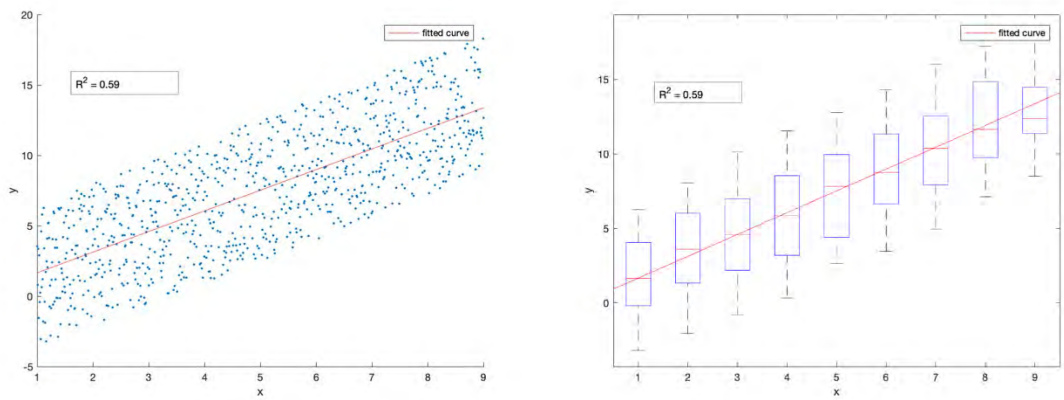
For our example, R^2 = 0.80.

STTF Report Appendix IV

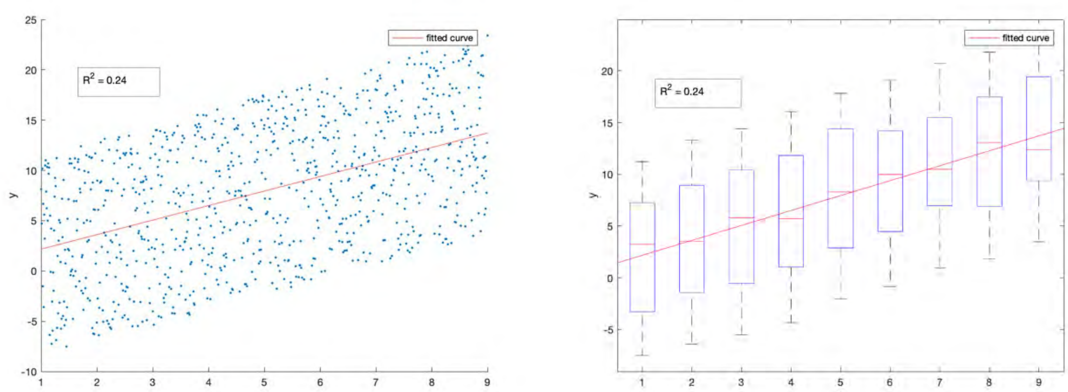
Now let's consider some larger sets of data in order to get a visual sense of how much scatter is associated with varying levels of R^2. Below are three simulations created by adding random "noise" of increasing magnitude to the same straight line. In all three cases, a scatter plot and a box-and-wire plot are shown, along with the fitted line. The vertical scales are different due to the different amounts of scatter in the y data.



Noise introduced into y at the level of ± 2.5.

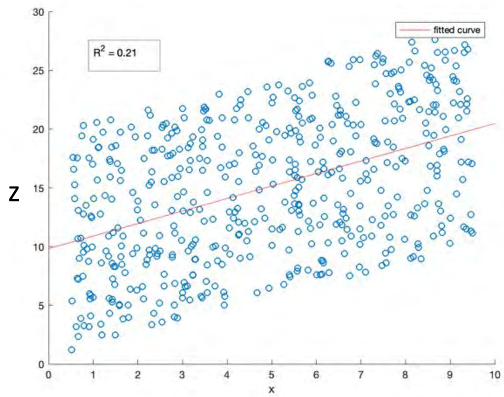


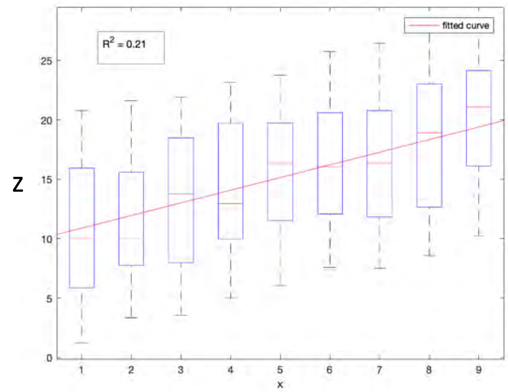
Noise introduced into y at the level of ± 5.



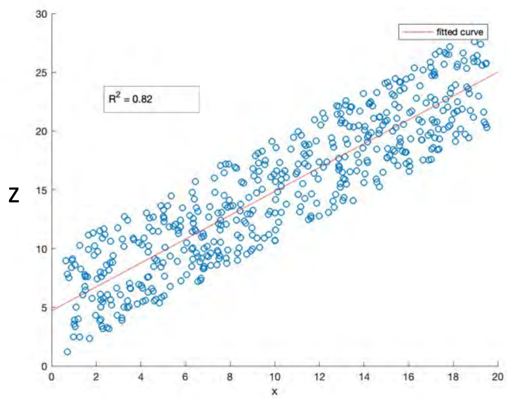
Noise introduced into y at the level of ± 10.

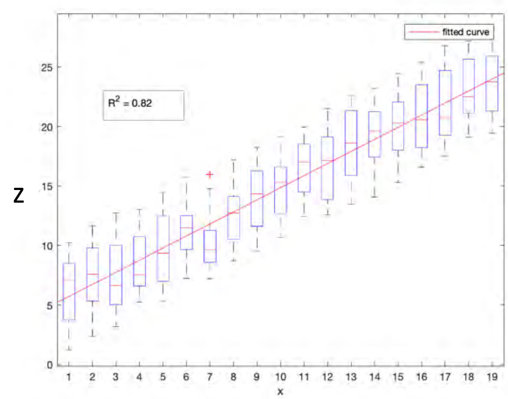
As a final example, we consider data that may be a function of two independent variables: $z = f(x, y)$. Noise is again introduced and various functional fits are examined against the data. First consider z with only x as an independent variable:





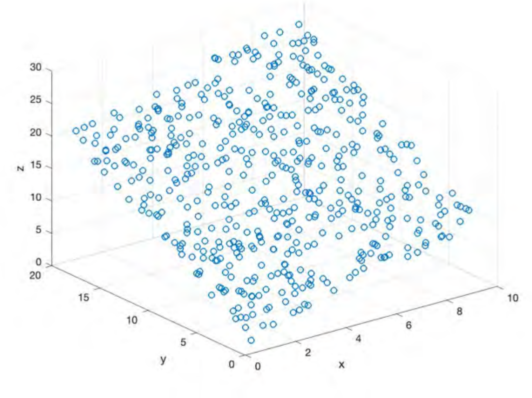
Now consider z with only y as the independent variable:

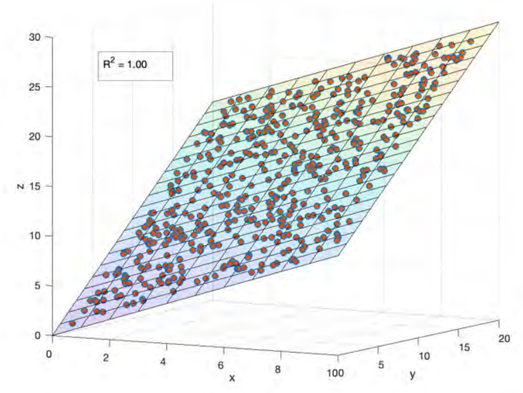




If we can only use one independent variable to predict z, it is clear that the fit using y (R^2 = 0.82; maximum residual ≅ 5) does a better job than the fit using x ($R^2 = 0.21$; maximum residual ≅ 10).

If we fit z using a function that is linear in both x and y, we find that the residual is zero and $R^2 = 1$. This is because the noise was introduced into x and y not into z, but we couldn't tell that by examining either of the 1D models. The second figure below shows that all of the points lie on a single plane.





Appendix V. Campus by campus study of relationships between HSGPA, test scores, and demographics in admissions.

- **Figures 1-10: Applicant and admittee HSGPAs, test scores, and admissions rates for all campuses together (Figure 1) and each campus individually (Figures 2-10), broken down by racial/ethnic subgroup.**
- **Figures 11-20: Applicant and admittee HSGPAs, test scores, and admissions rates for all campuses together (Figure 1) and each campus individually (Figures 2-10), broken down by family educational history.**
- **Figures 21-30: Applicant and admittee HSGPAs, test scores, and admissions rates for all campuses together (Figure 1) and each campus individually (Figures 2-10), broken down by family income category.**
- **Tables 1-3: Applicant and admittee HSGPAs and test scores, by demographic group and campus.**
- **Tables 4-6: Applicant and admittee HSGPAs and test scores with respect to highest-scoring groups, by demographic group and campus.**

Legend for Figures 1-30

Aa: Rows show HSGPAs for applicants from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group.

Ab: HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group.

Ad: HSGPAs for accepted students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As A-B but for SAT scores instead of HSGPA.

To preserve student privacy, combinations with fewer than 30 accepted students are blanked out. Groups that UC considers disadvantaged or underrepresented are indicated in red or orange.

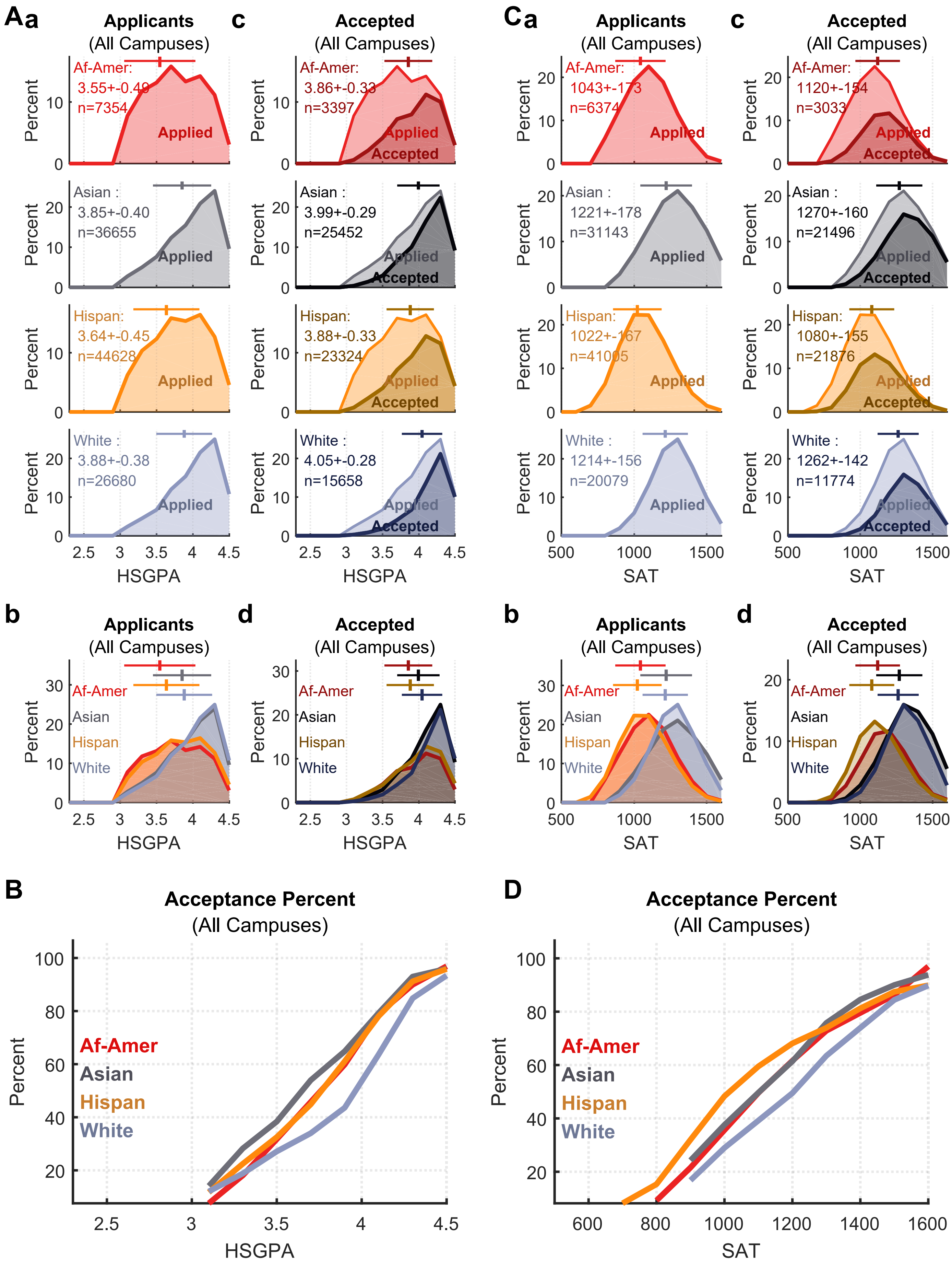
Tables 1-3: Applicant and admittee HSGPAs and test scores, by demographic group and campus.

Legend for Tables 1-3: Left: Each row shows, for one campus and one demographic group, the mean HSGPAs for applicants and for admitted students, the change in mean grades between the pool of applicants and the pool of accepted students, and the size of the change normalized to units of standard deviations (calculated across the entire applicant pool) – a measure of how strongly the campus selects based on that qualification. Normalized changes are color coded (red = strongest selection, white = no selection). **Right**: As Left, but for SAT. Note that 1) the HSGPAs of admitted students are typically substantially higher than the HSGPAs of applicants, 2) the SATs of admitted students are also typically higher than the SATs of applicants, 3) measured selectivity based on test scores is generally much smaller than the selectivity based on grades, across different demographic groups and campuses, and 4) for the groups most underrepresented at UC, the contribution of selection on HSGPA exceeds the contribution of selection on test scores.

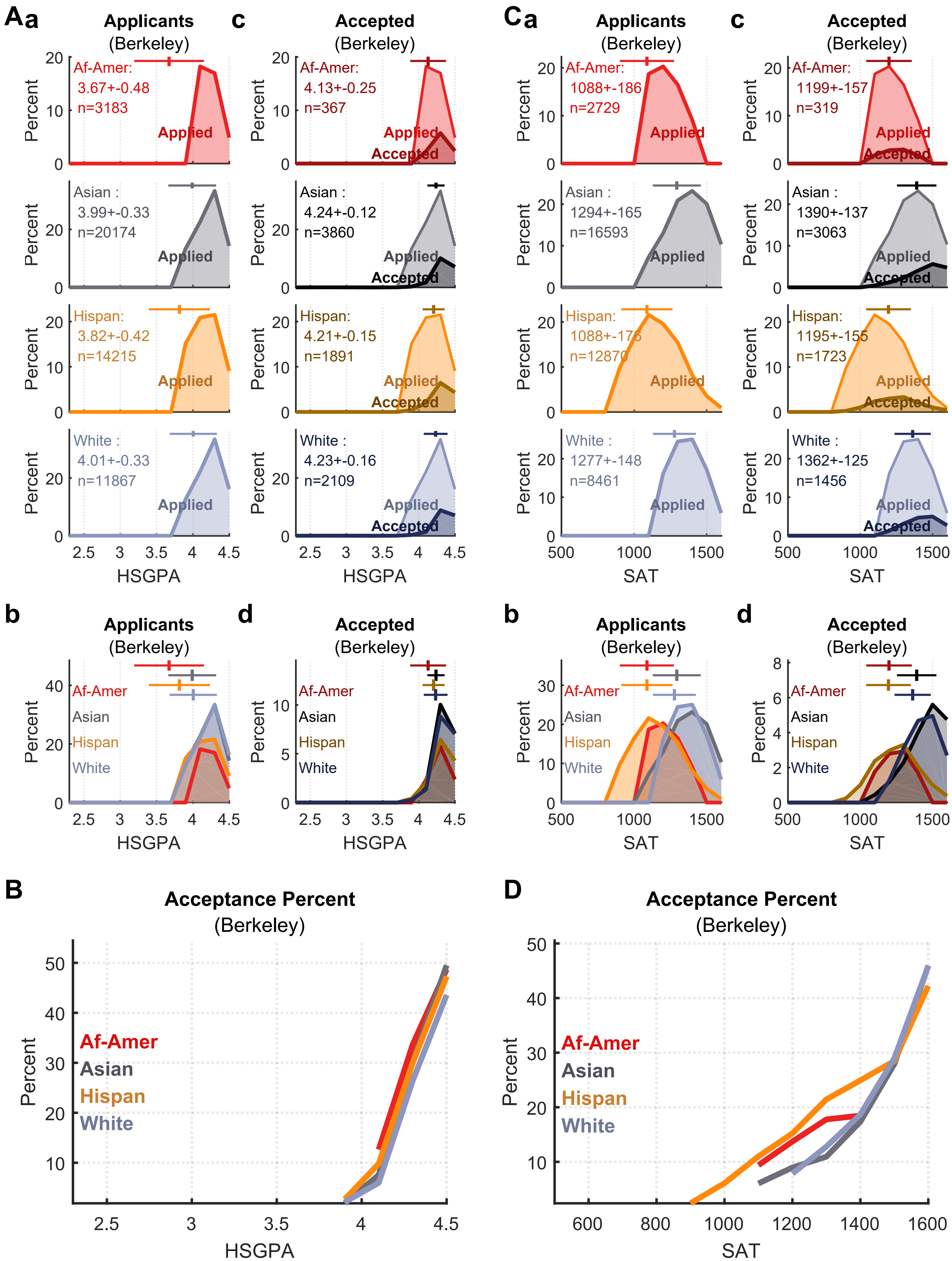
Tables 4-6: Applicant and admittee HSGPAs and test scores with respect to highest-scoring groups, by demographic group and campus.

Legend for Tables 4-6: Left: Each row shows, for one campus and one demographic group, mean HSGPAs for applicants, the difference in mean HSGPA (green) between applicants in that group and the group with the highest test scores, mean HSGPAs for admitted students, the difference in mean HSGPA between admitted students in that group and the group with the highest test scores (green), and the ratio of the Accepted difference vs the Applicant difference. If the difference in scores between two groups of accepted students is exactly the same as the difference in scores between applicants in those groups, (i.e., uncompensated difference of 0), that is consistent with the hypothesis that the UC admissions process effectively compensates for the mean differences between groups. **Right**: As Left, but for SAT (differences in blue). Note that 1) the remaining, uncompensated differences in HSGPAs are typically much larger than uncompensated differences in test scores (i.e., the left green columns are darker than the right green columns, but the left blue columns are similar to the right blue columns), 2) the uncompensated differences in HSGPA are largest for underrepresented groups.

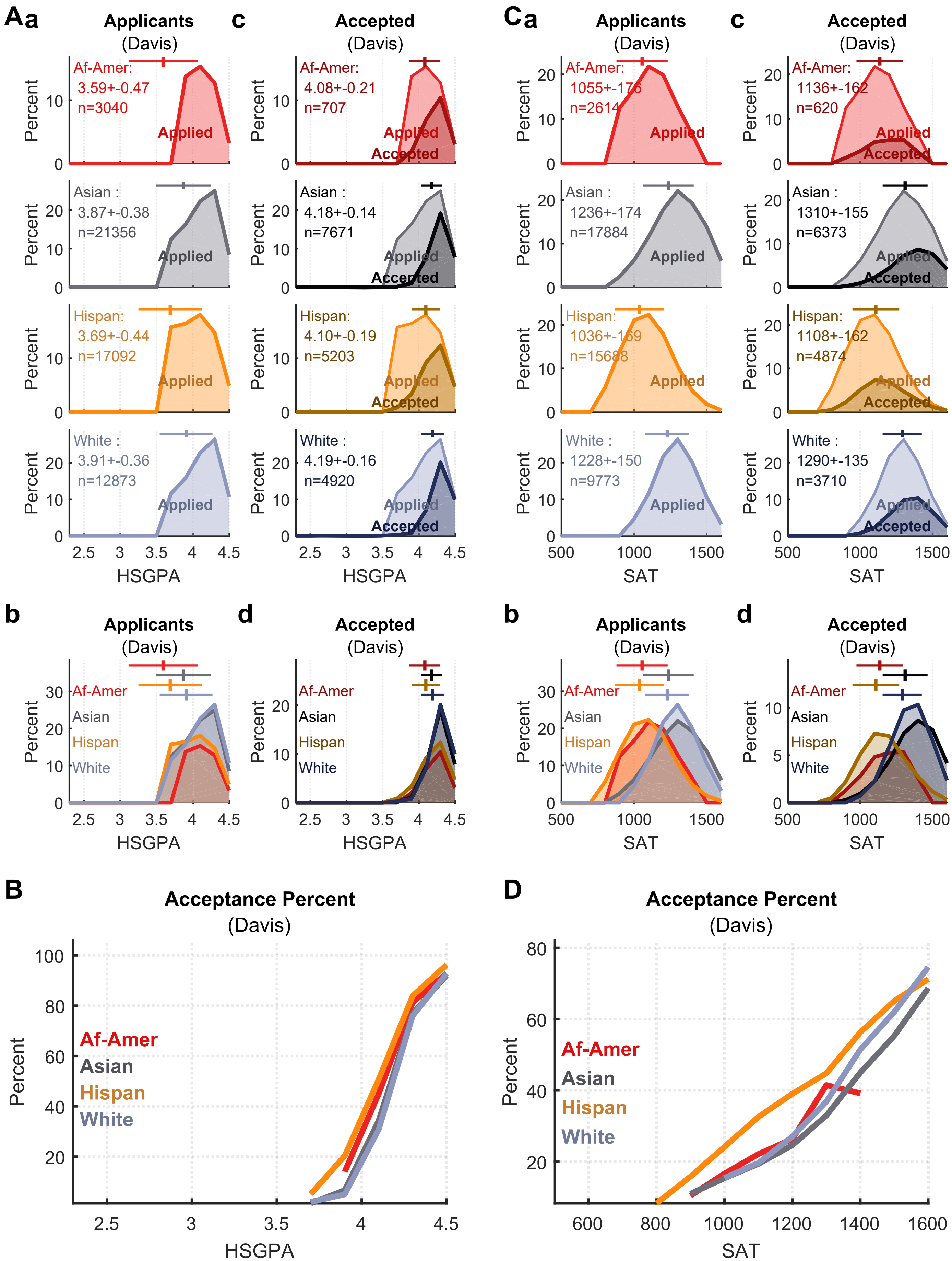
Appendix V. Figure 1. Race/ethnicity, HSGPAs, and test scores in the applicant and admit pools for all UC campuses combined.



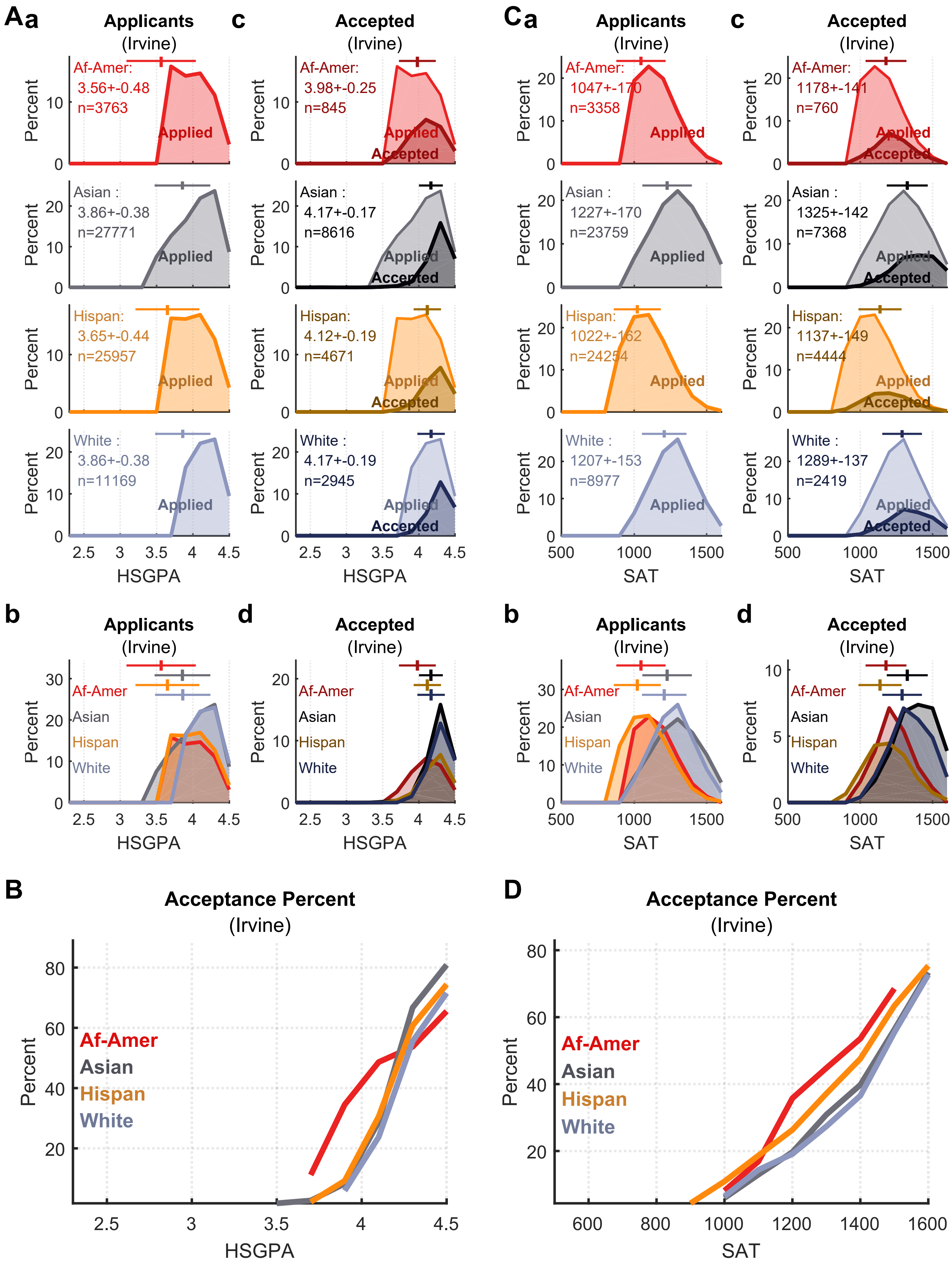
Appendix V. Figure 2. Race/ethnicity, HSGPAs, and test scores in the applicant and admit pools for UC Berkeley.



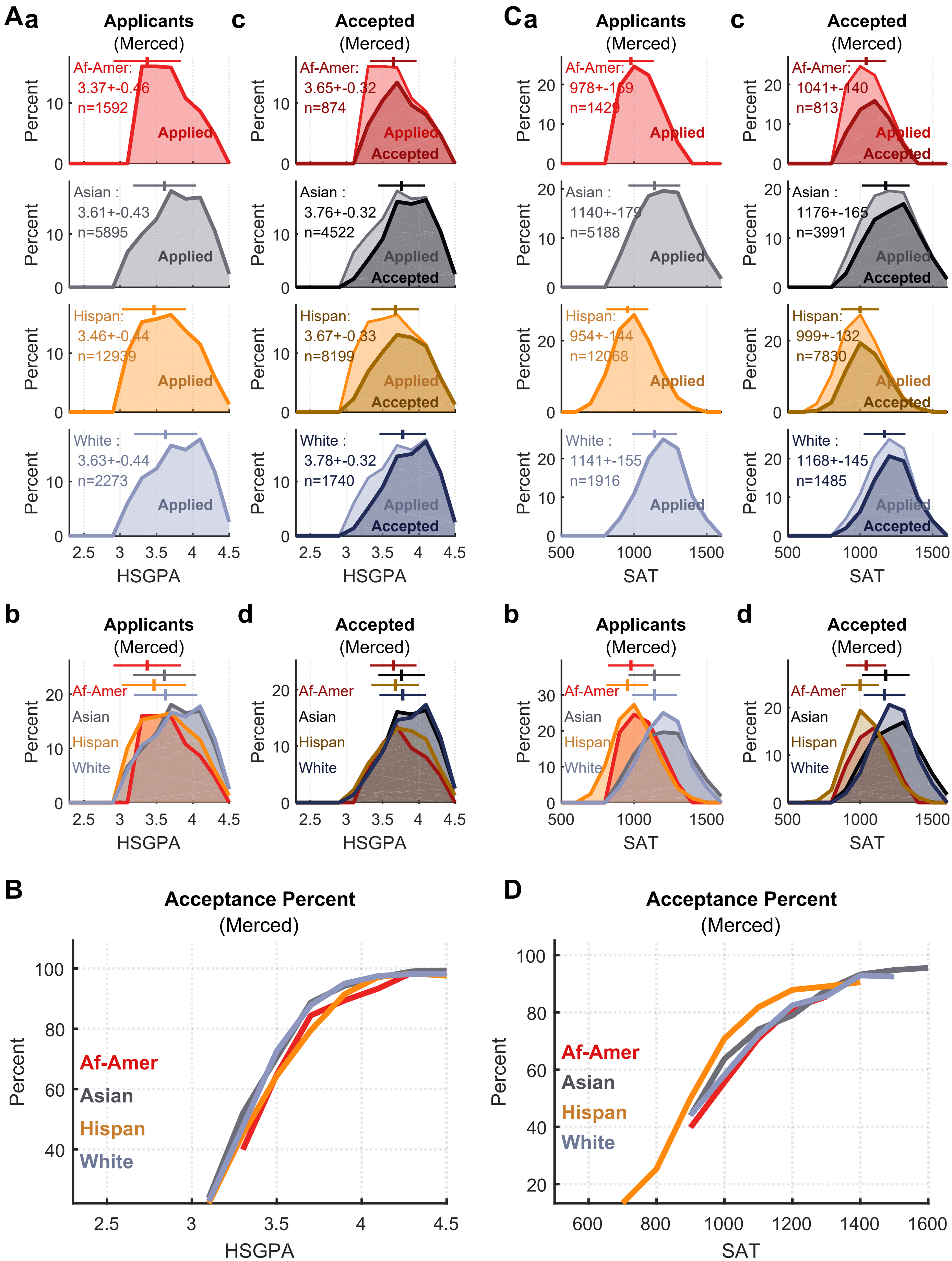
Appendix V. Figure 3. Race/ethnicity, HSGPAs, and test scores in the applicant and admit pools for UC Davis.



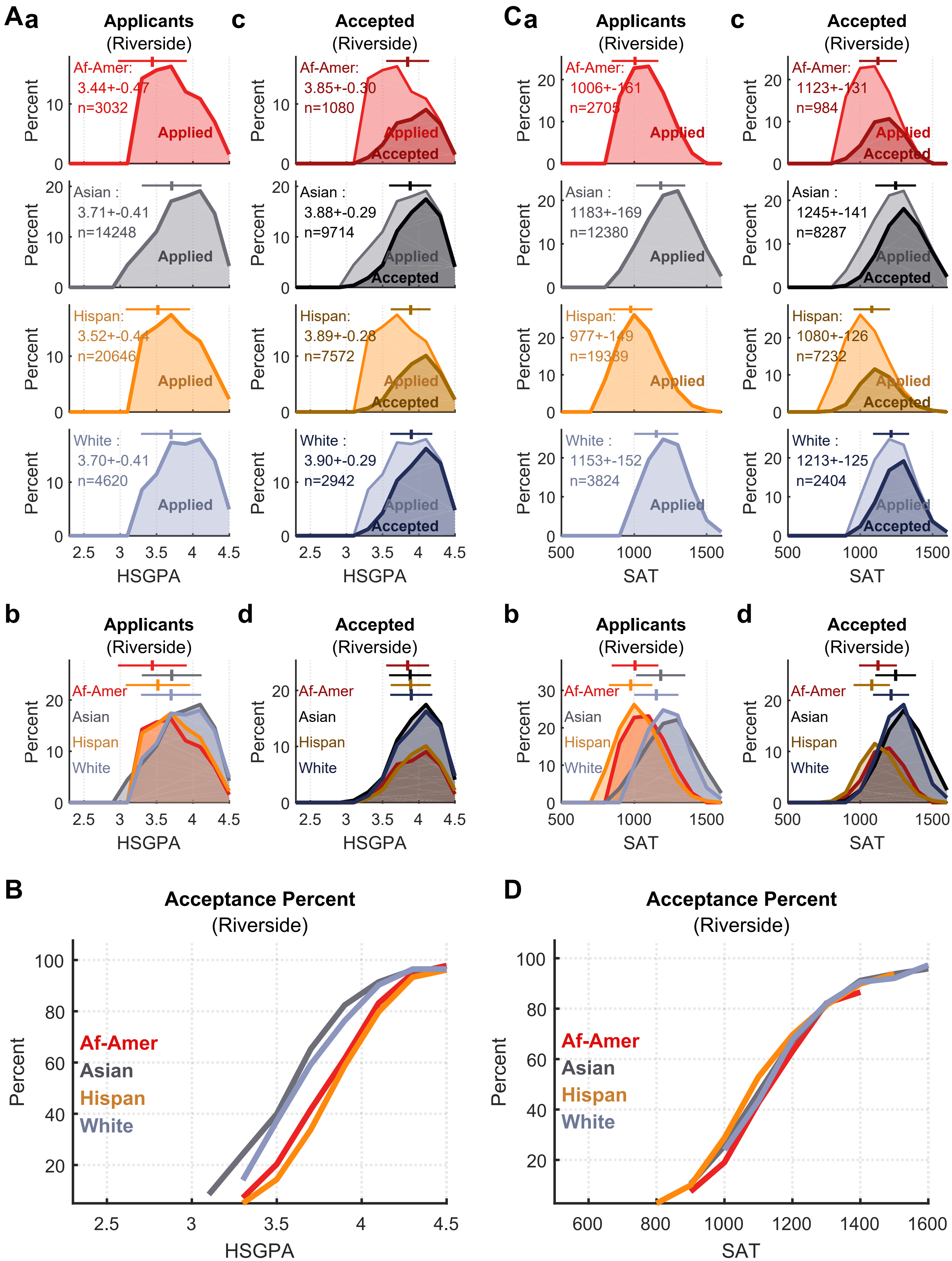
Appendix V. Figure 4. Race/ethnicity, HSGPAs, and test scores in the applicant and admit pools for UC Irvine.



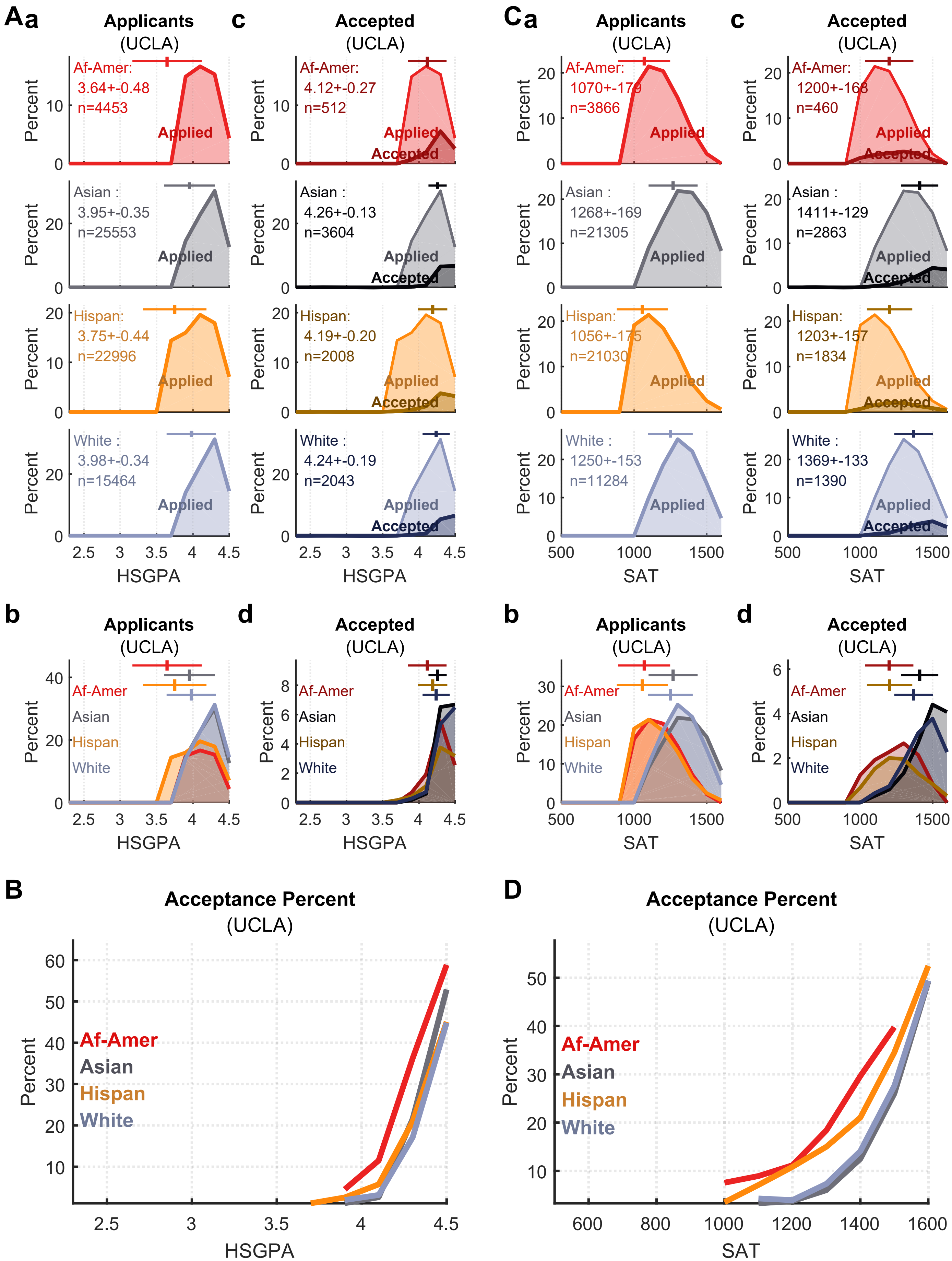
Appendix V. Figure 5. Race/ethnicity, HSGPAs, and test scores in the applicant and admit pools for UC Merced.



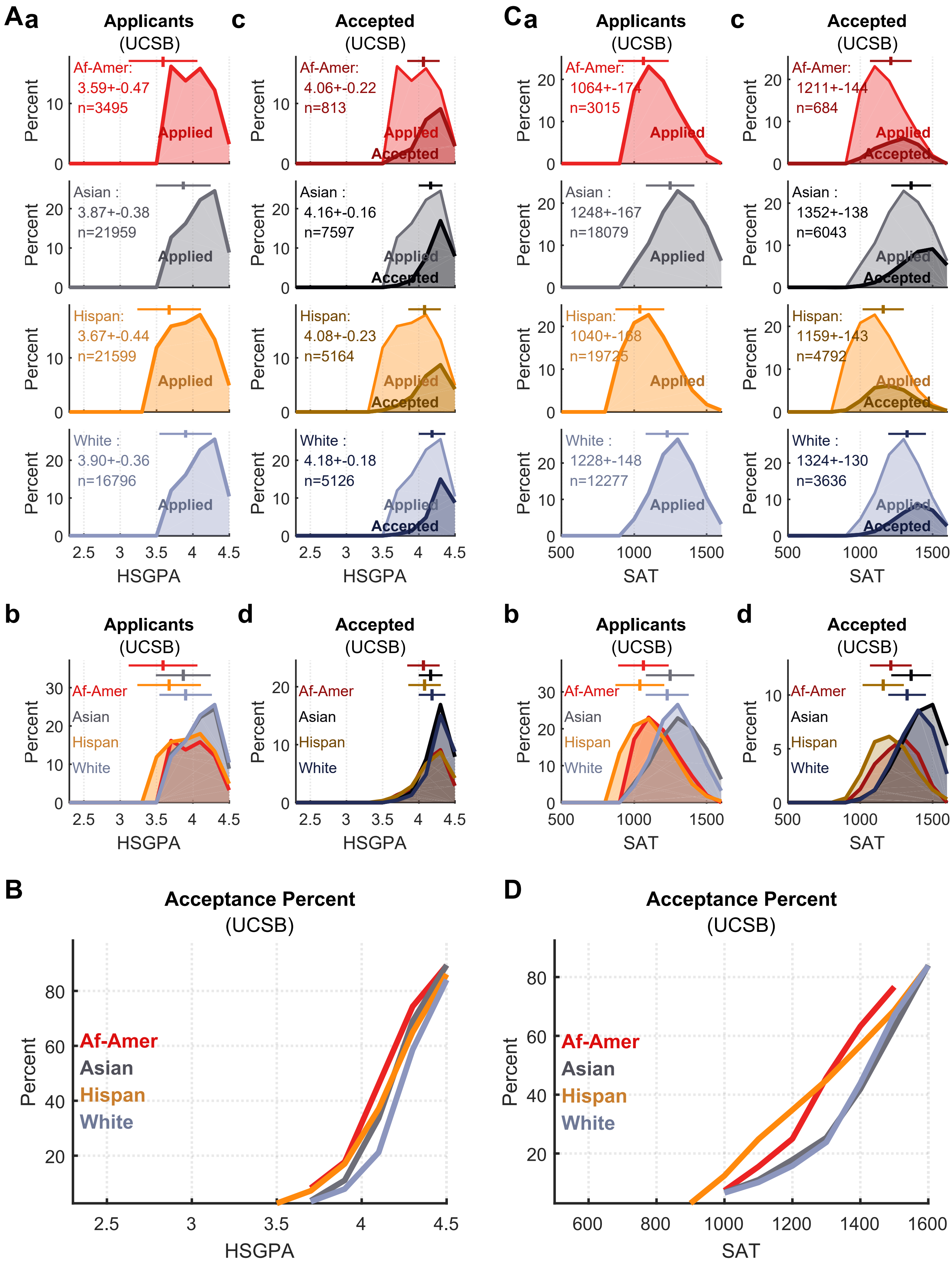
Appendix V. Figure 6. Race/ethnicity, HSGPAs, and test scores in the applicant and admit pools for UC Riverside.



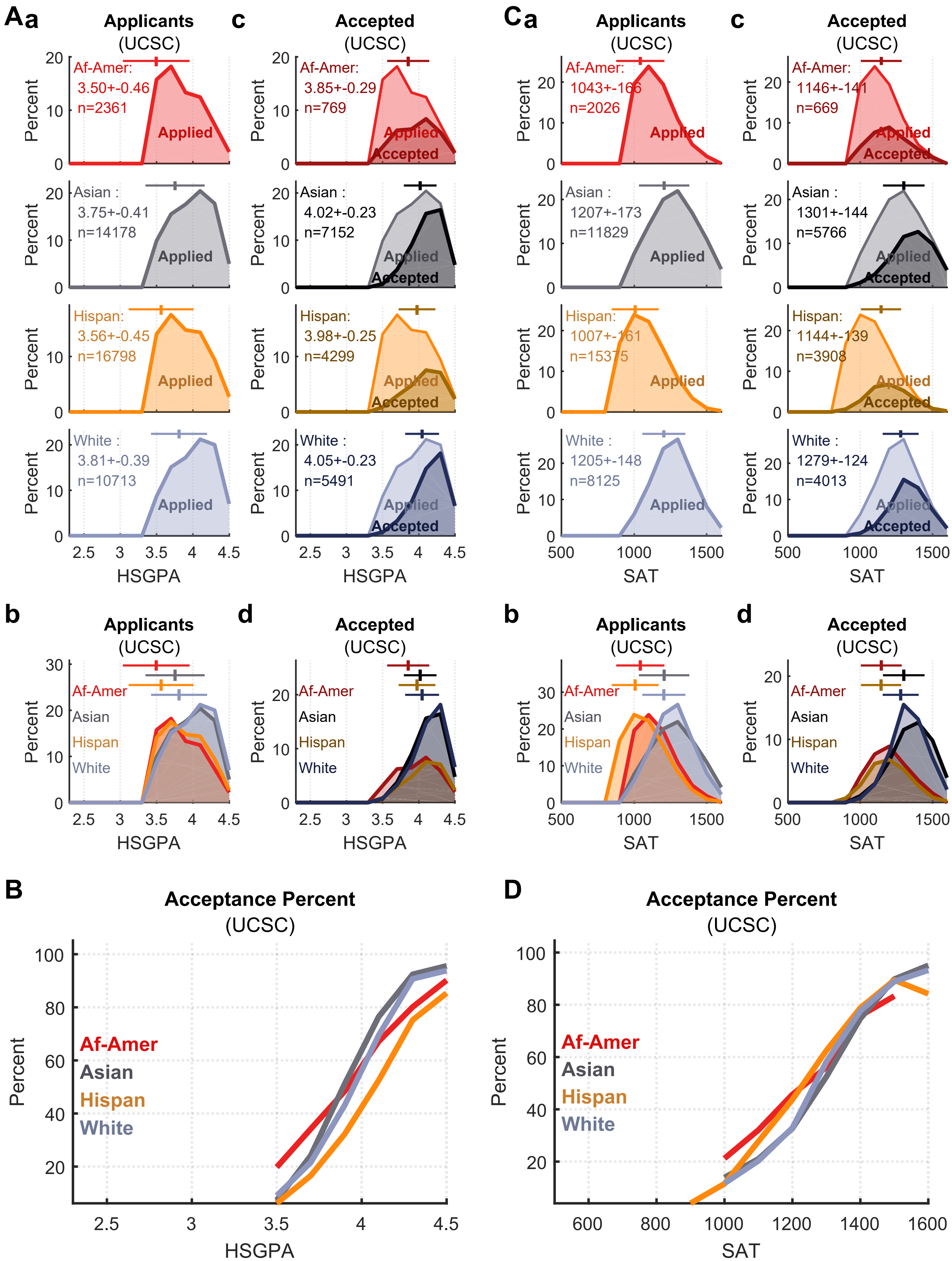
Appendix V. Figure 7. Race/ethnicity, HSGPAs, and test scores in the applicant and admit pools for UCLA.



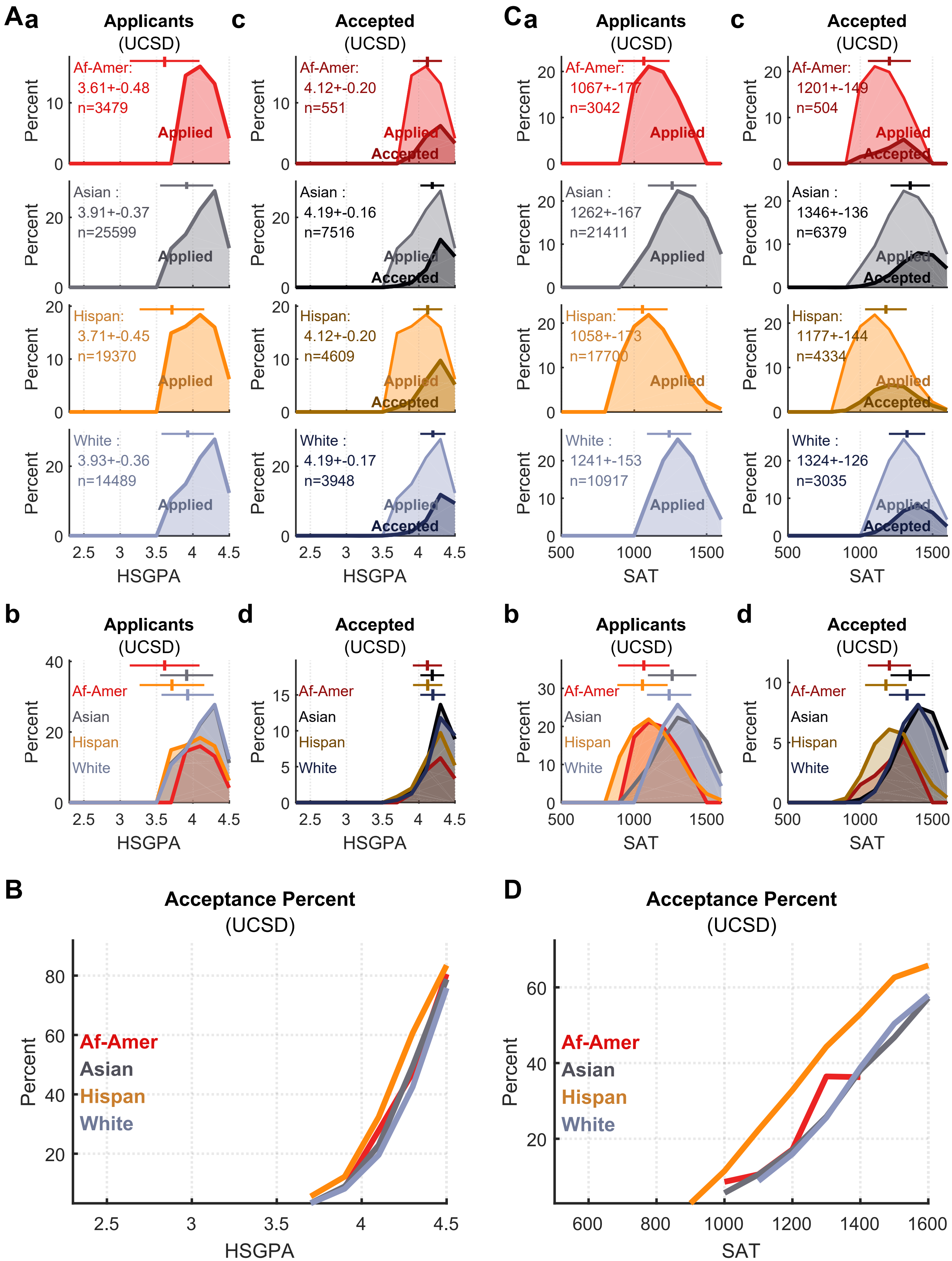
Appendix V. Figure 8. Race/ethnicity, HSGPAs, and test scores in the applicant and admit pools for UC Santa Barbara.



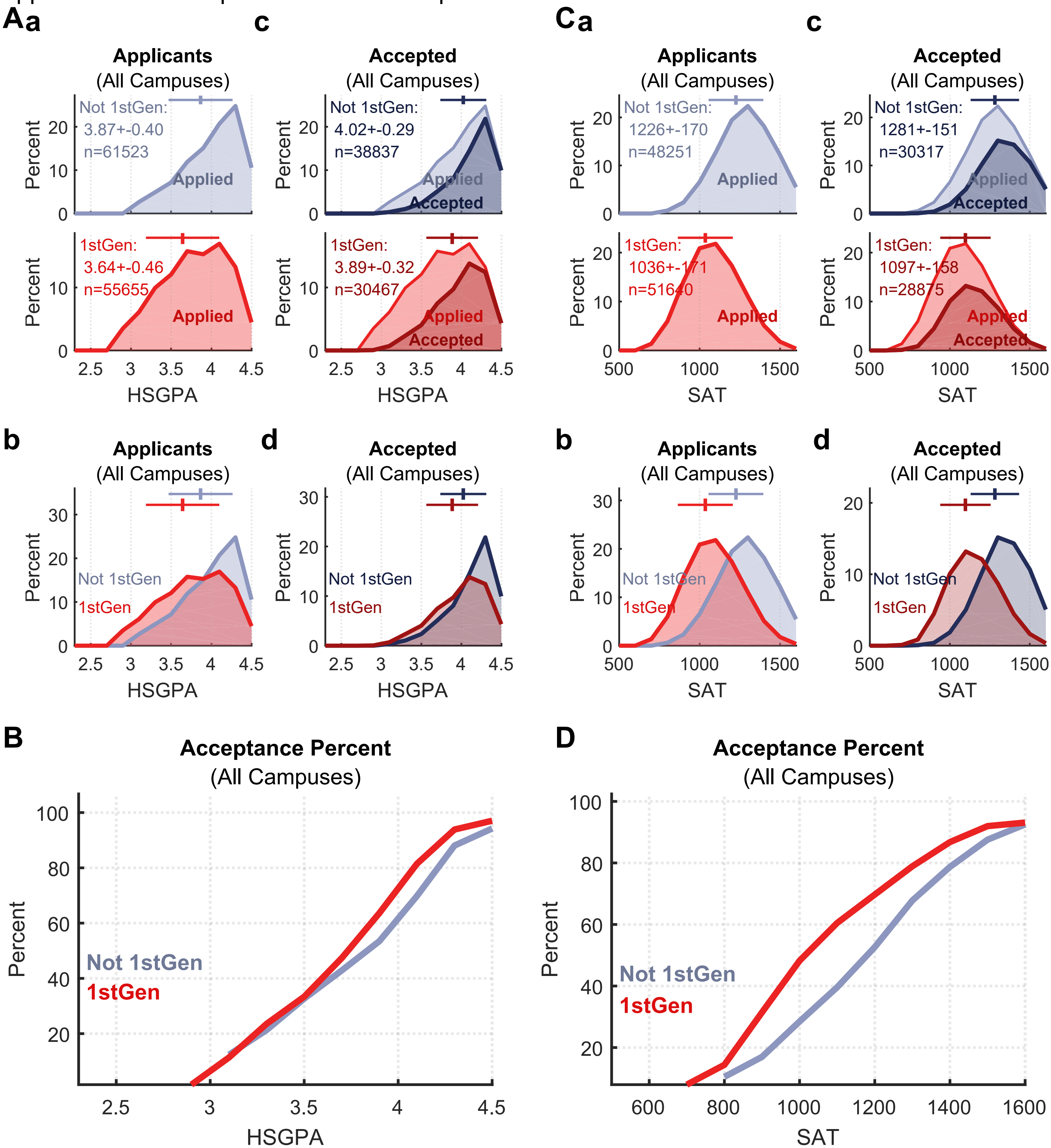
Appendix V. Figure 9. Race/ethnicity, HSGPAs, and test scores in the applicant and admit pools for UC Santa Cruz



Appendix V. Figure 10. Race/ethnicity, HSGPAs, and test scores in the applicant and admit pools for UC San Diego.



Appendix V. Figure 11. Family income, HSGPAs, and test scores in the applicant and admit pools for all UC campuses combined.



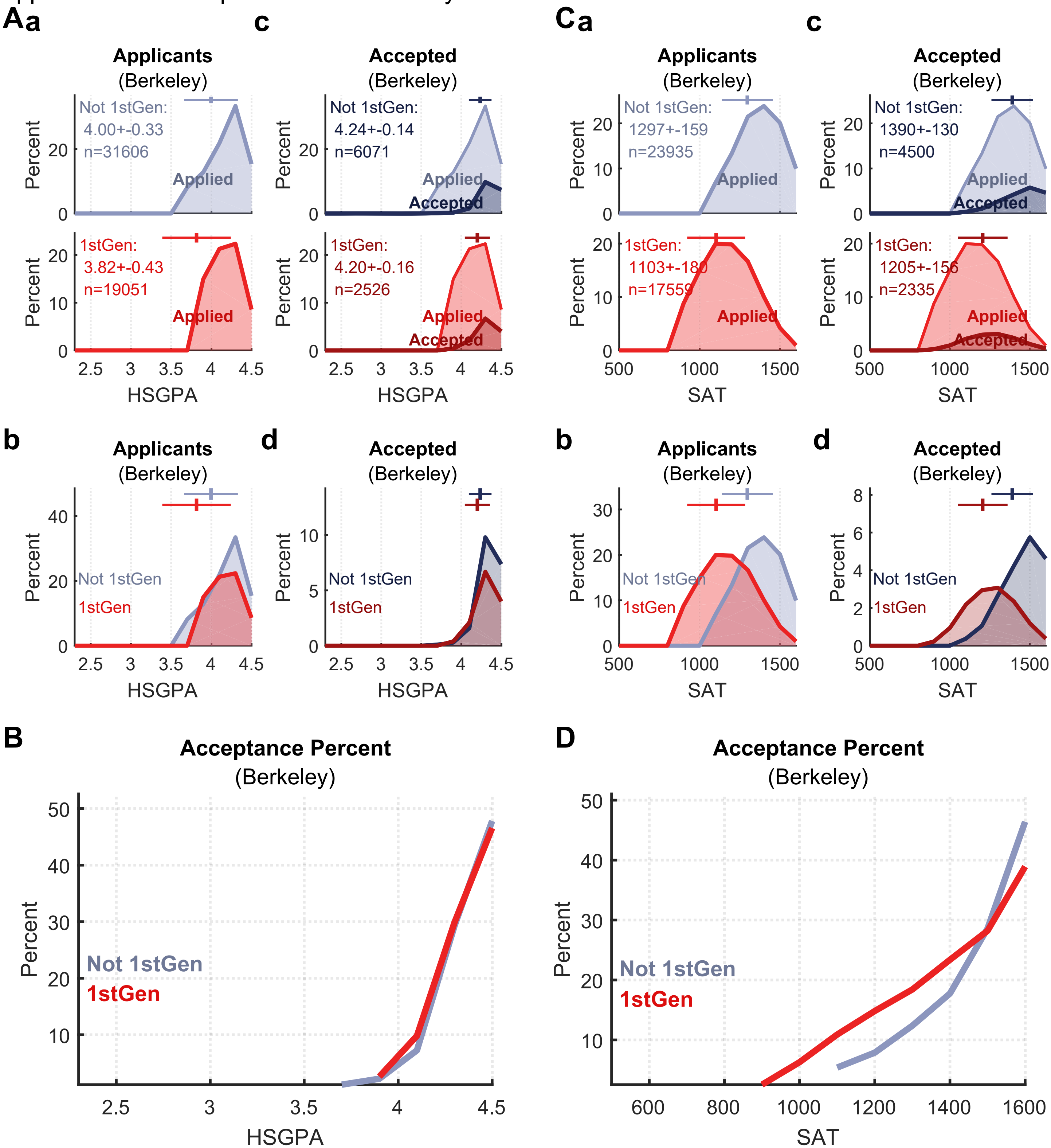
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 12. Family income, HSGPAs, and test scores in the applicant and admit pools for UC Berkeley.



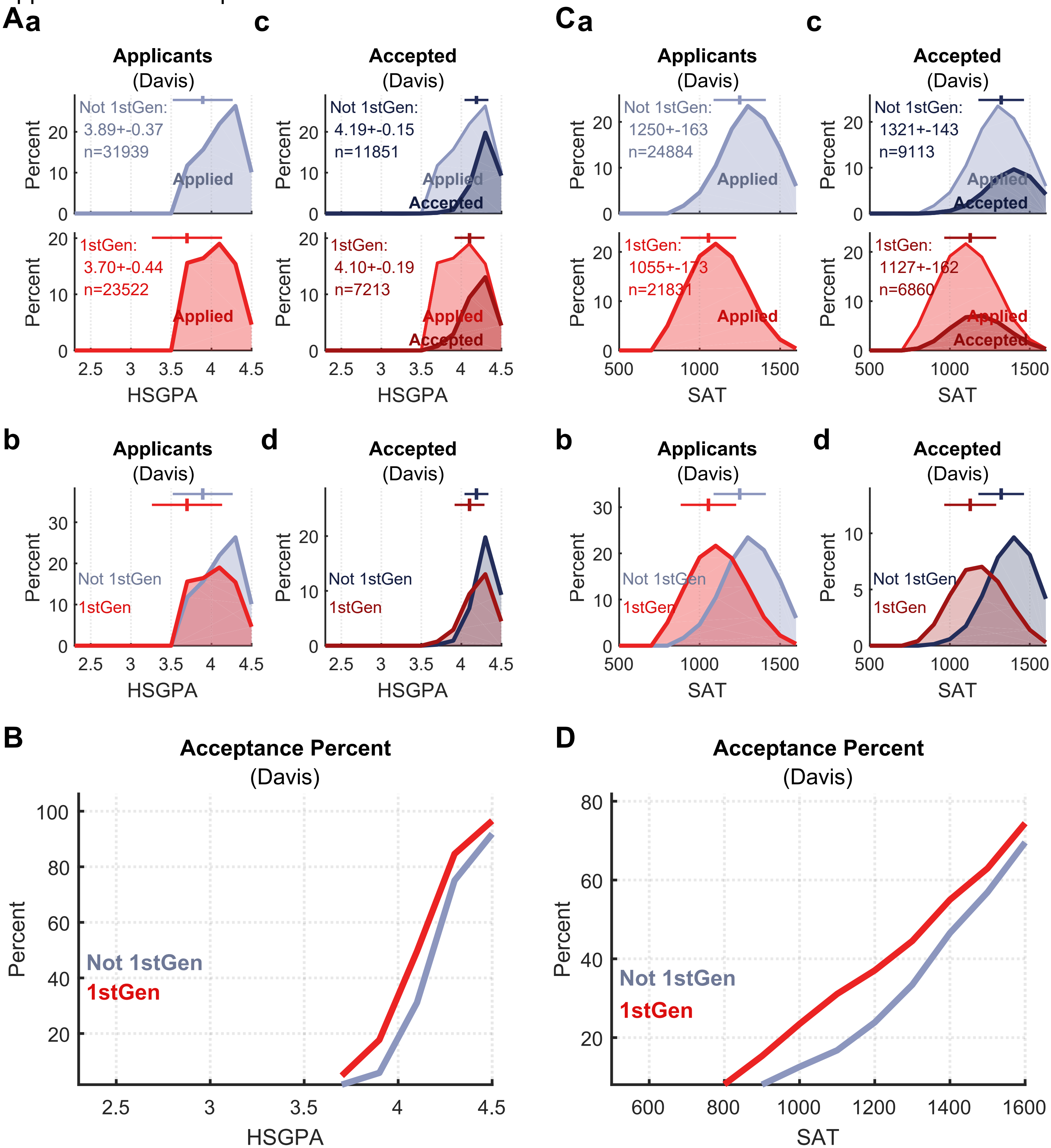
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 13. Family income, HSGPAs, and test scores in the applicant and admit pools for UC Davis.



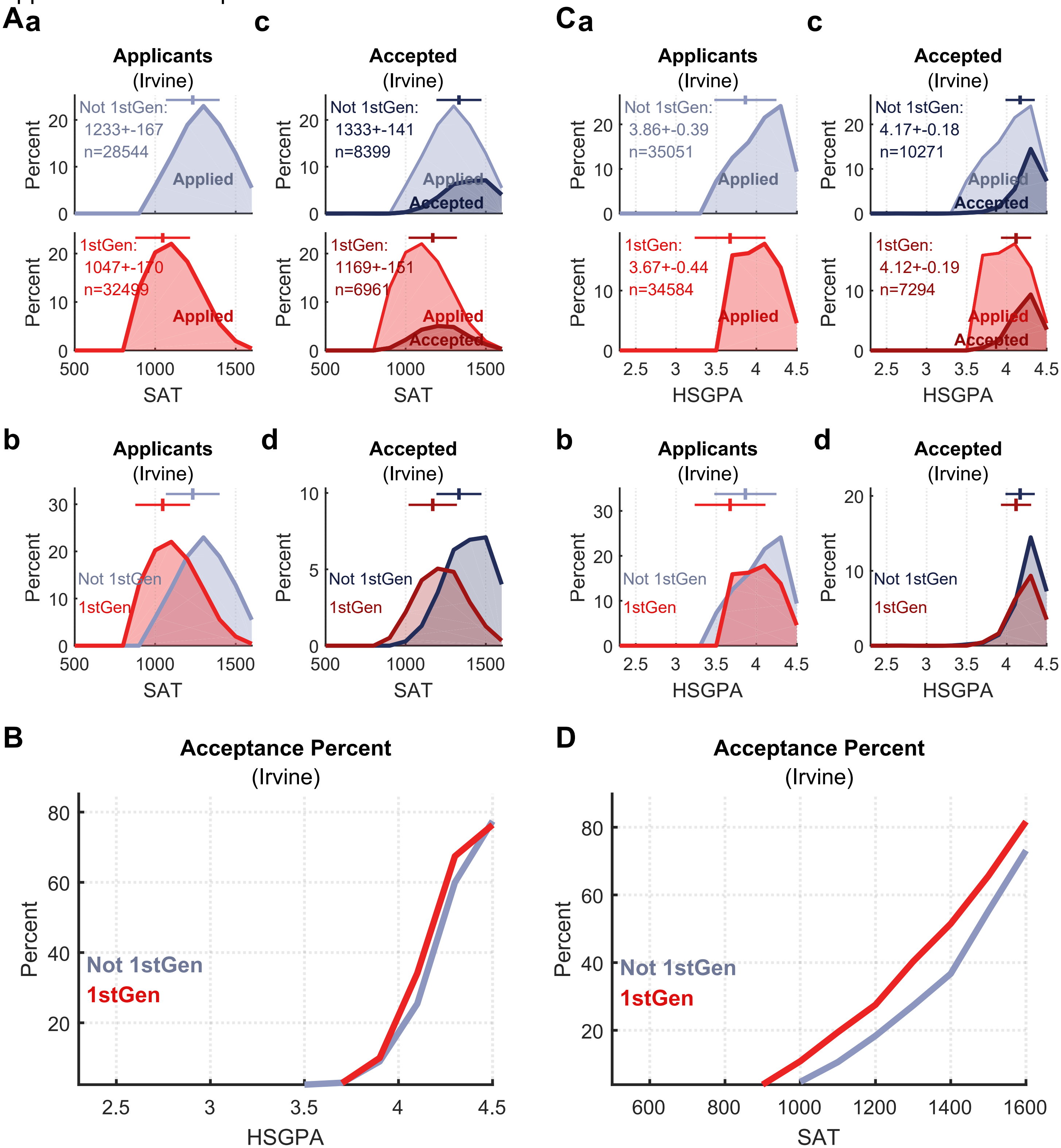
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 14. Family income, HSGPAs, and test scores in the applicant and admit pools for UC Irvine.



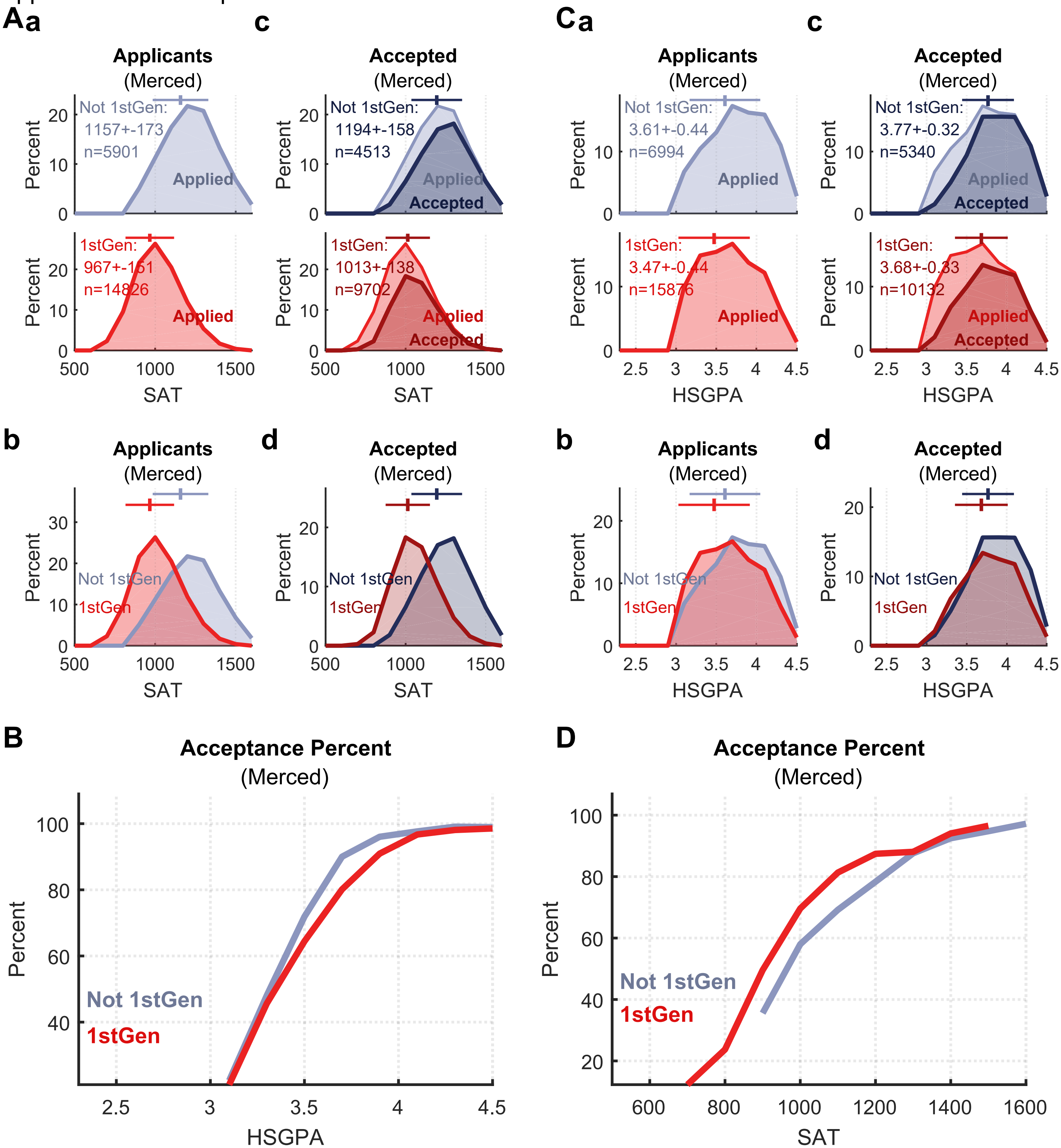
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 15. Family income, HSGPAs, and test scores in the applicant and admit pools for UC Merced.



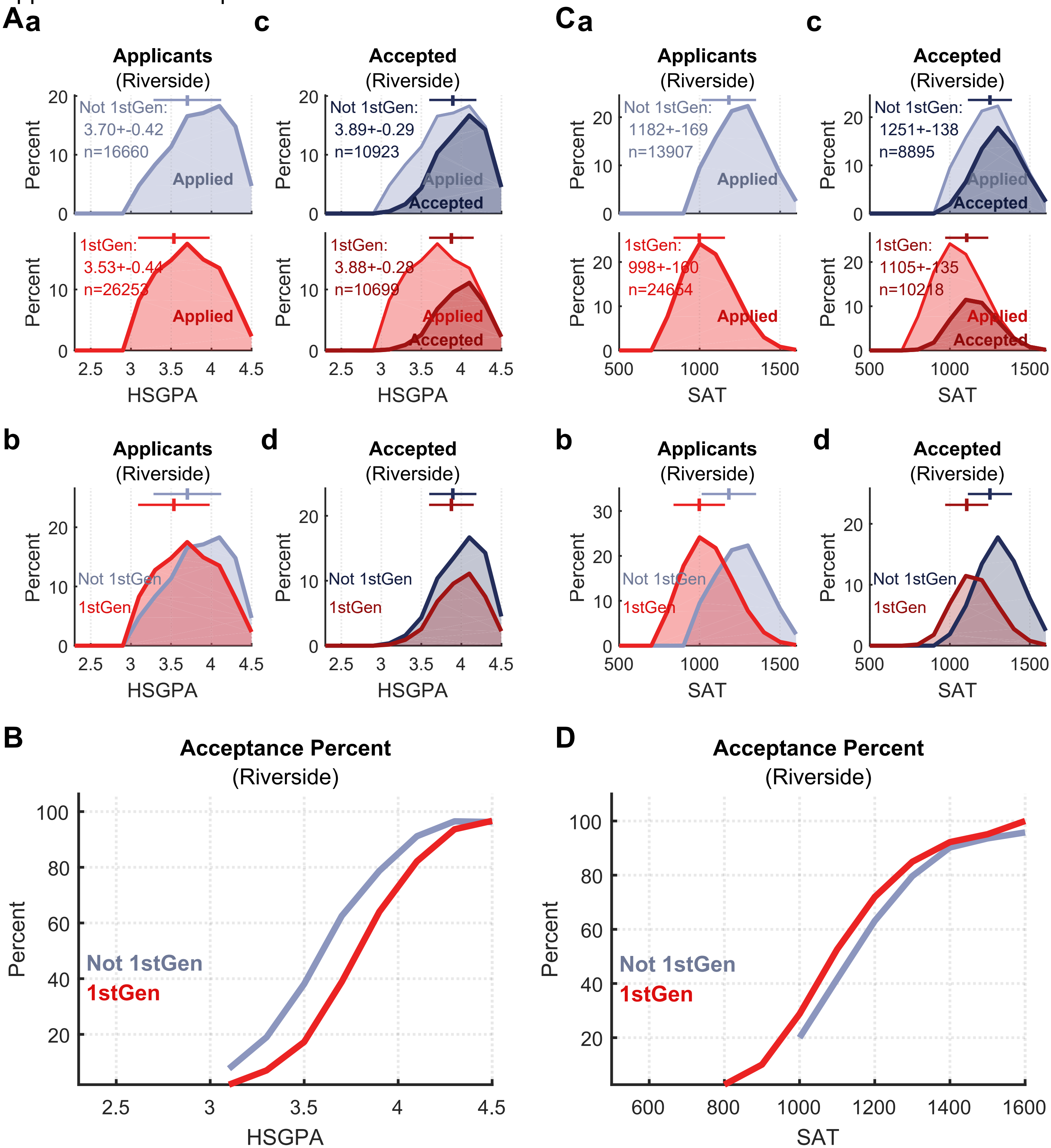
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 16. Family income, HSGPAs, and test scores in the applicant and admit pools for UC Riverside.



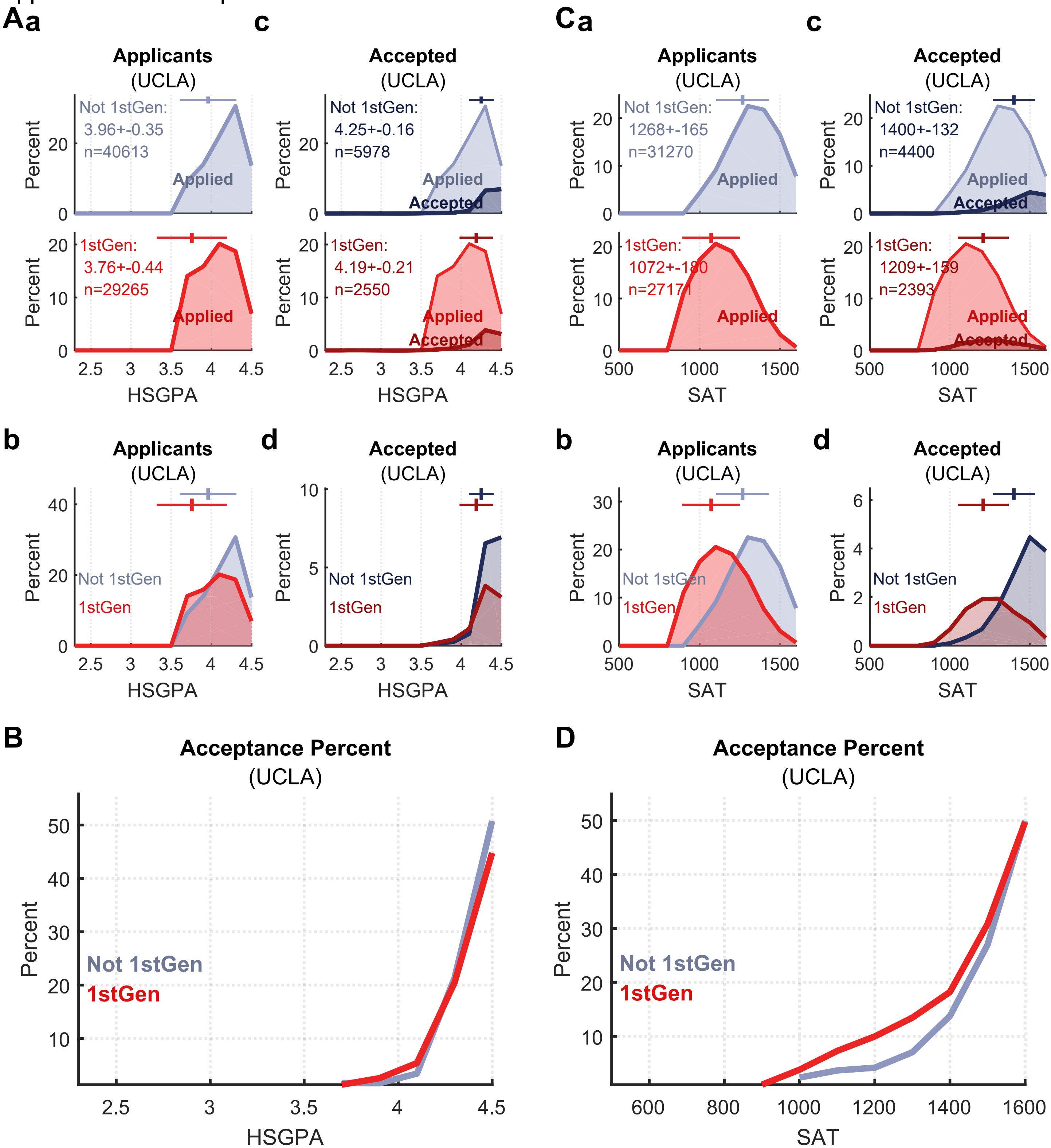
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 17. Family income, HSGPAs, and test scores in the applicant and admit pools for UCLA.



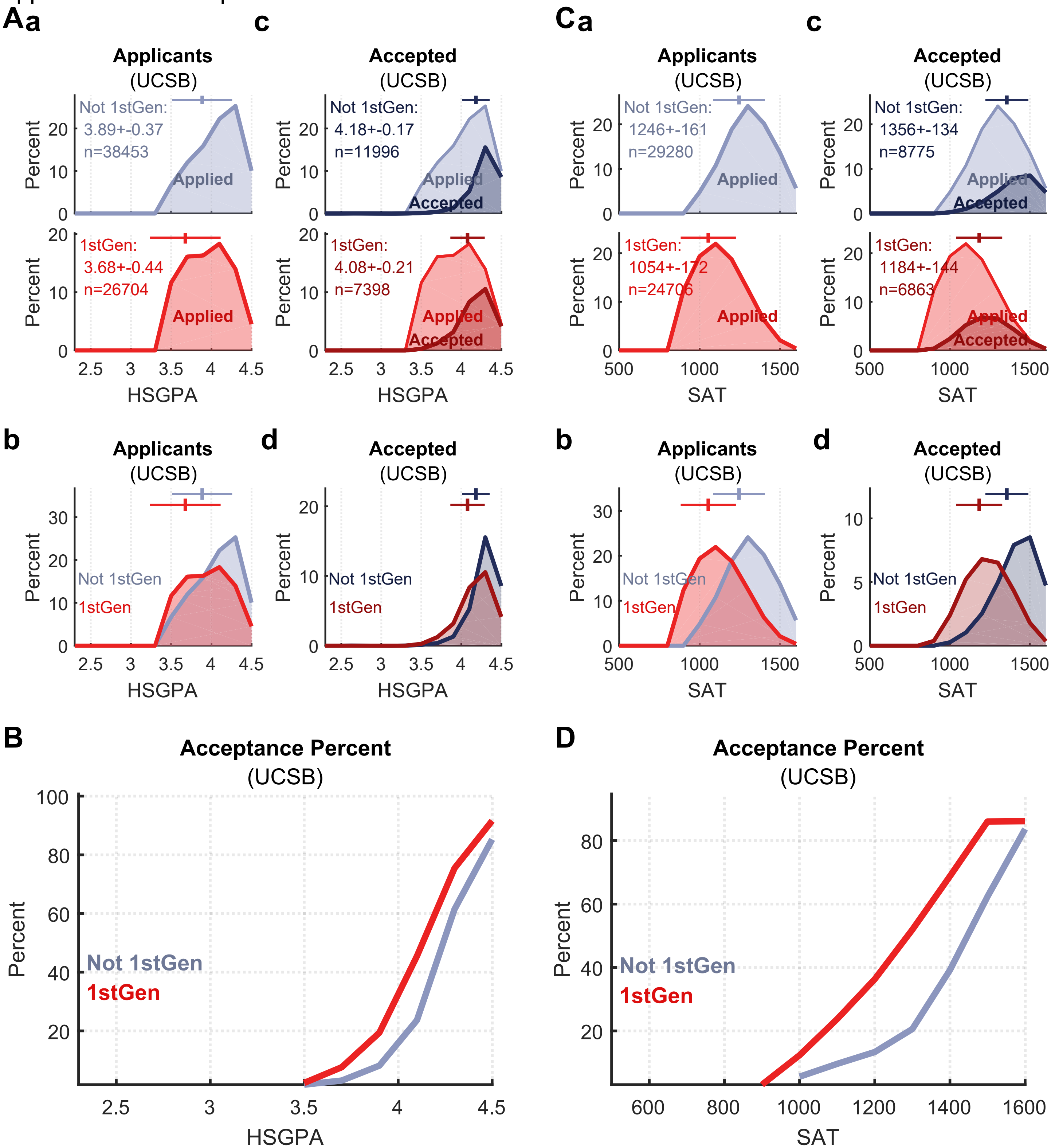
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 18. Family income, HSGPAs, and test scores in the applicant and admit pools for UC Santa Barbara.



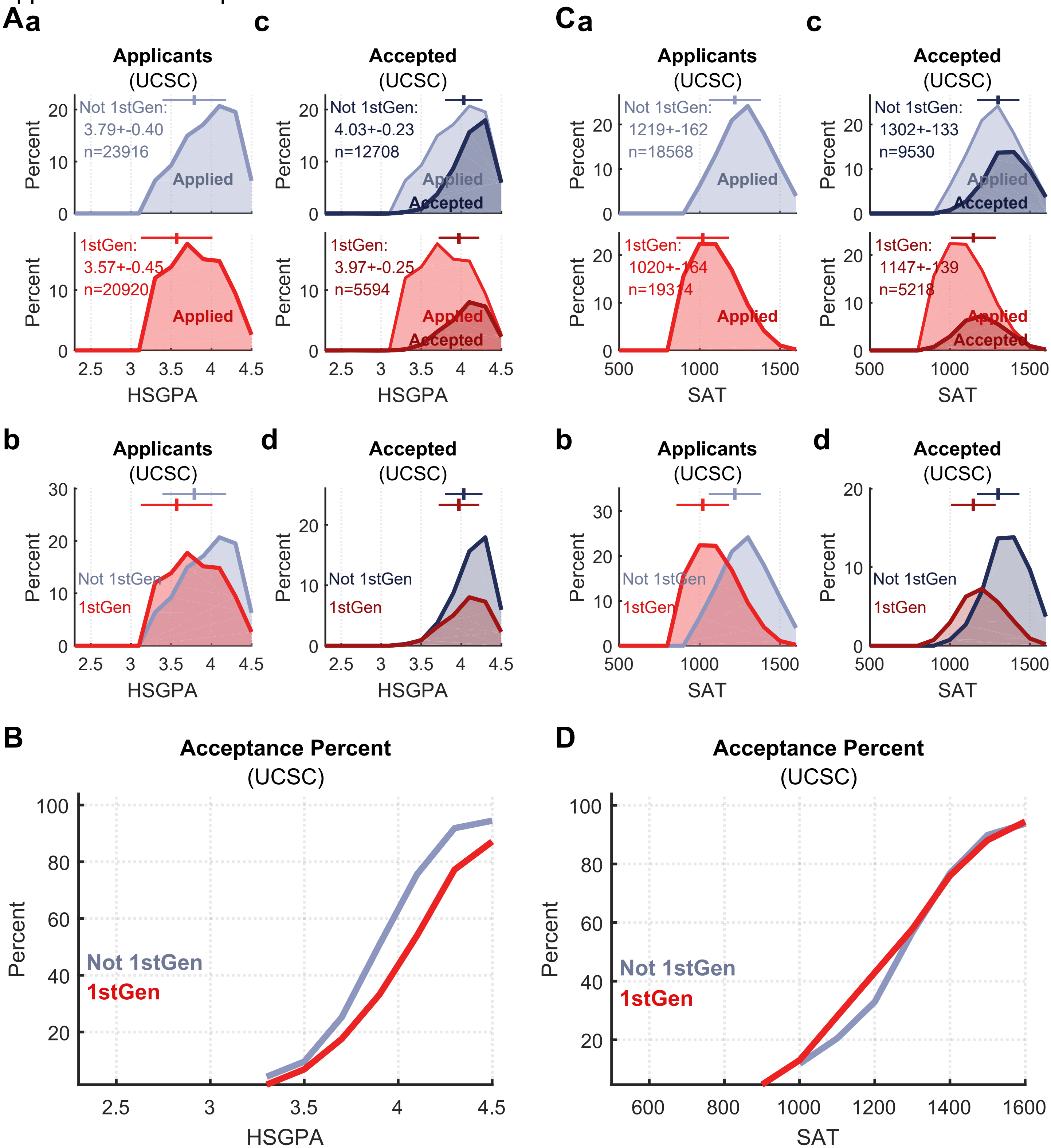
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 19. Family income, HSGPAs, and test scores in the applicant and admit pools for UC Santa Cruz



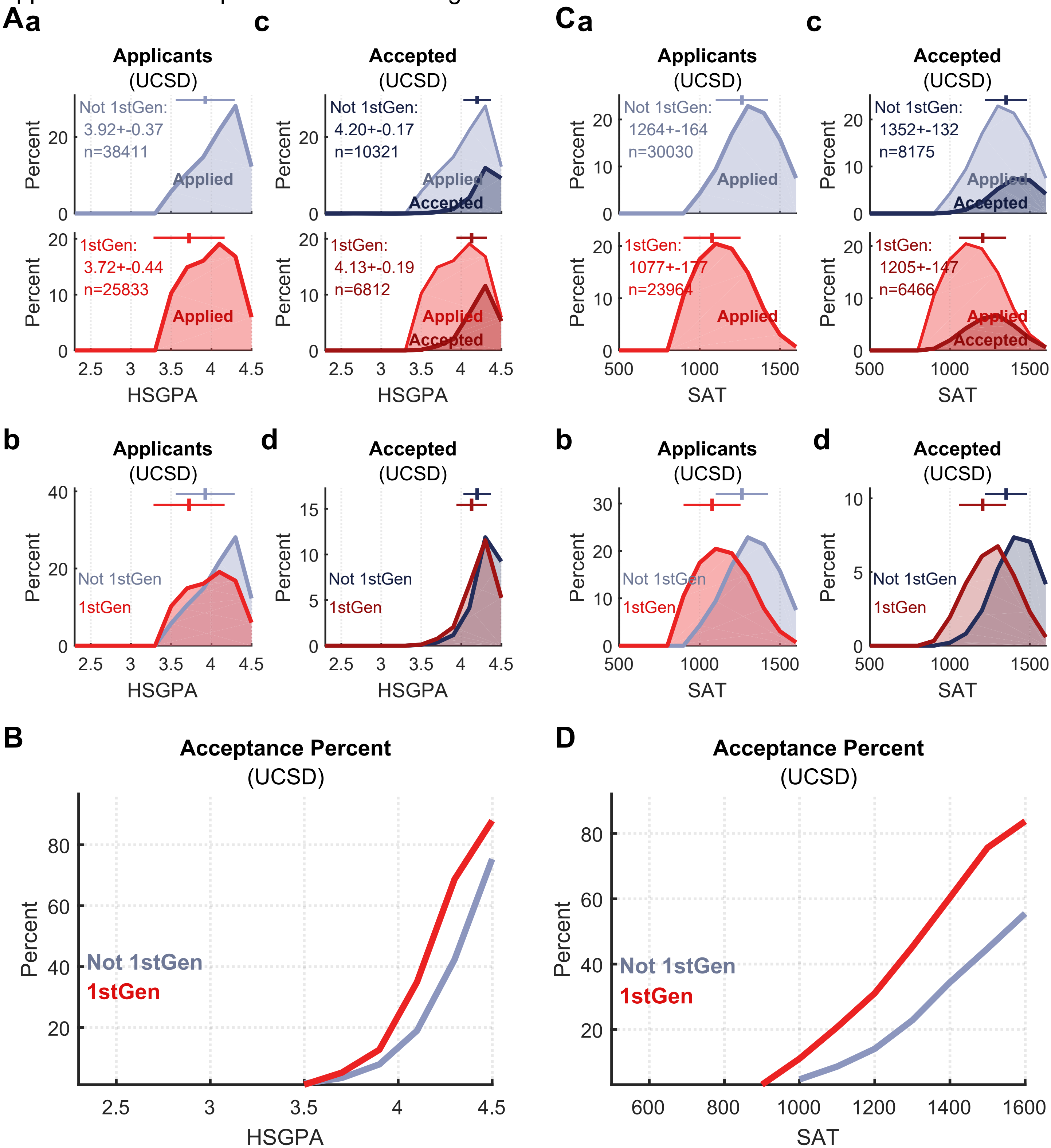
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 20. Family income, HSGPAs, and test scores in the applicant and admit pools for UC San Diego.



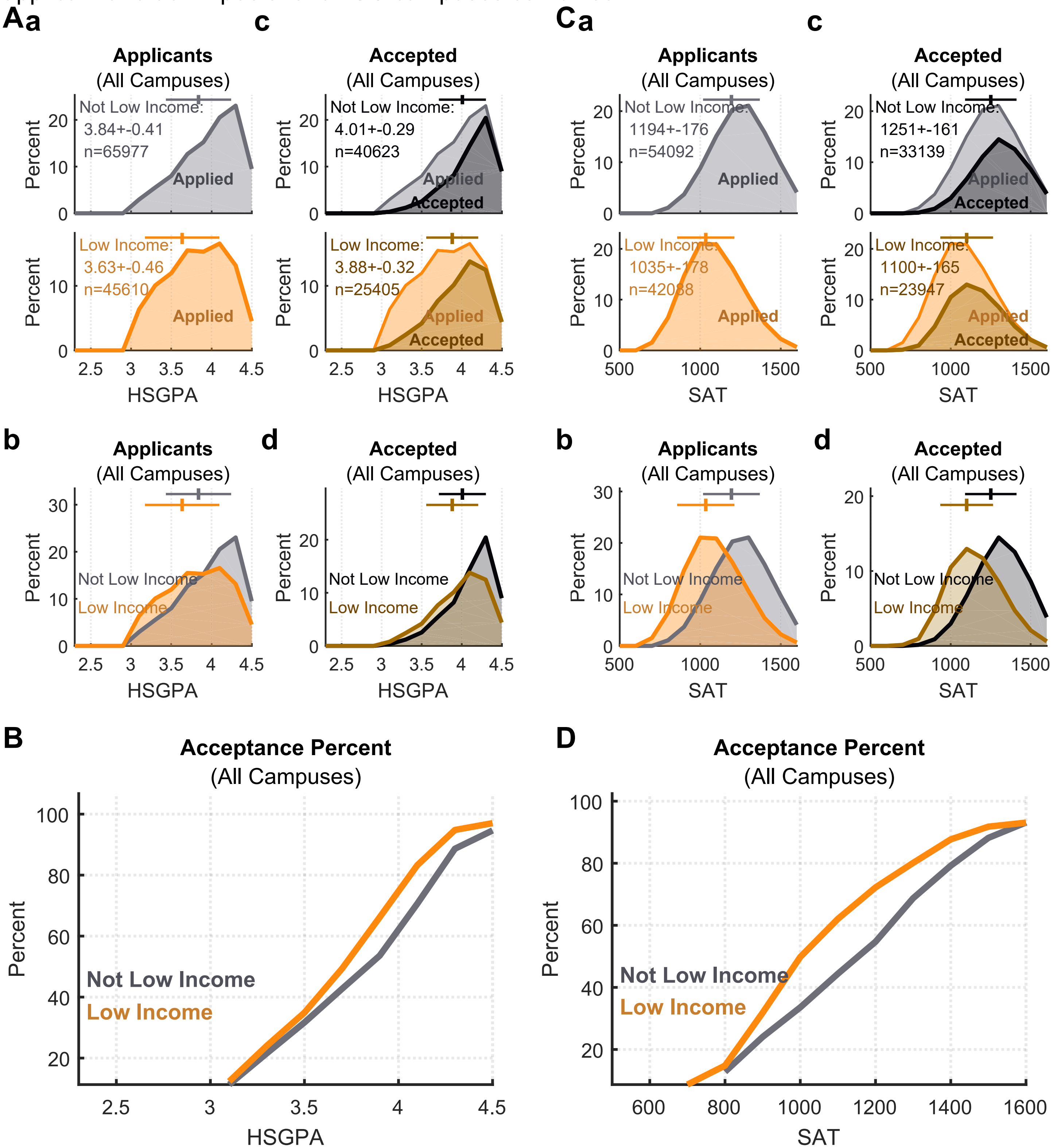
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 21. Family income, HSGPAs, and test scores in the applicant and admit pools for all UC campuses combined.



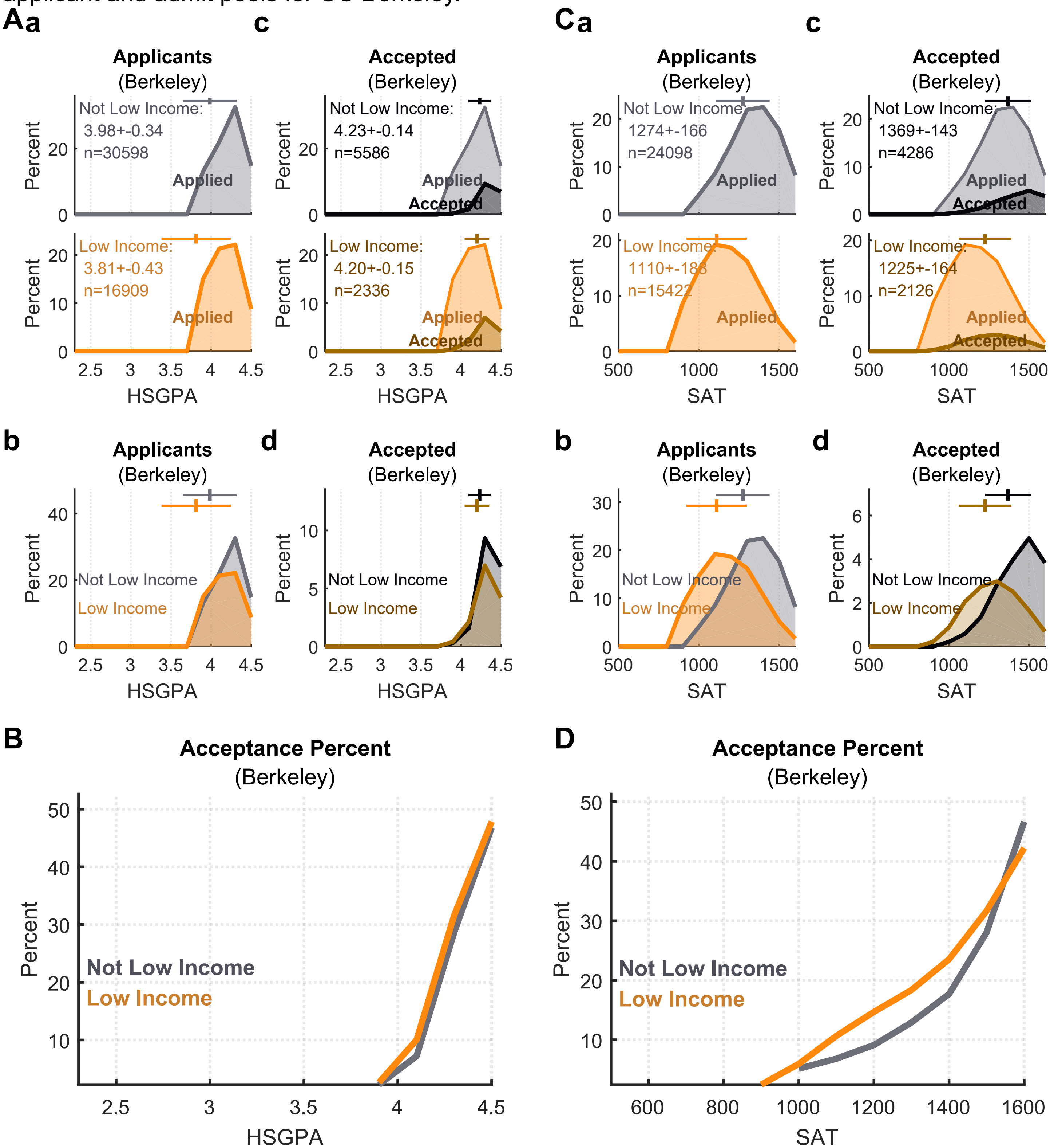
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 22. Family income, HSGPAs, and test scores in the applicant and admit pools for UC Berkeley.



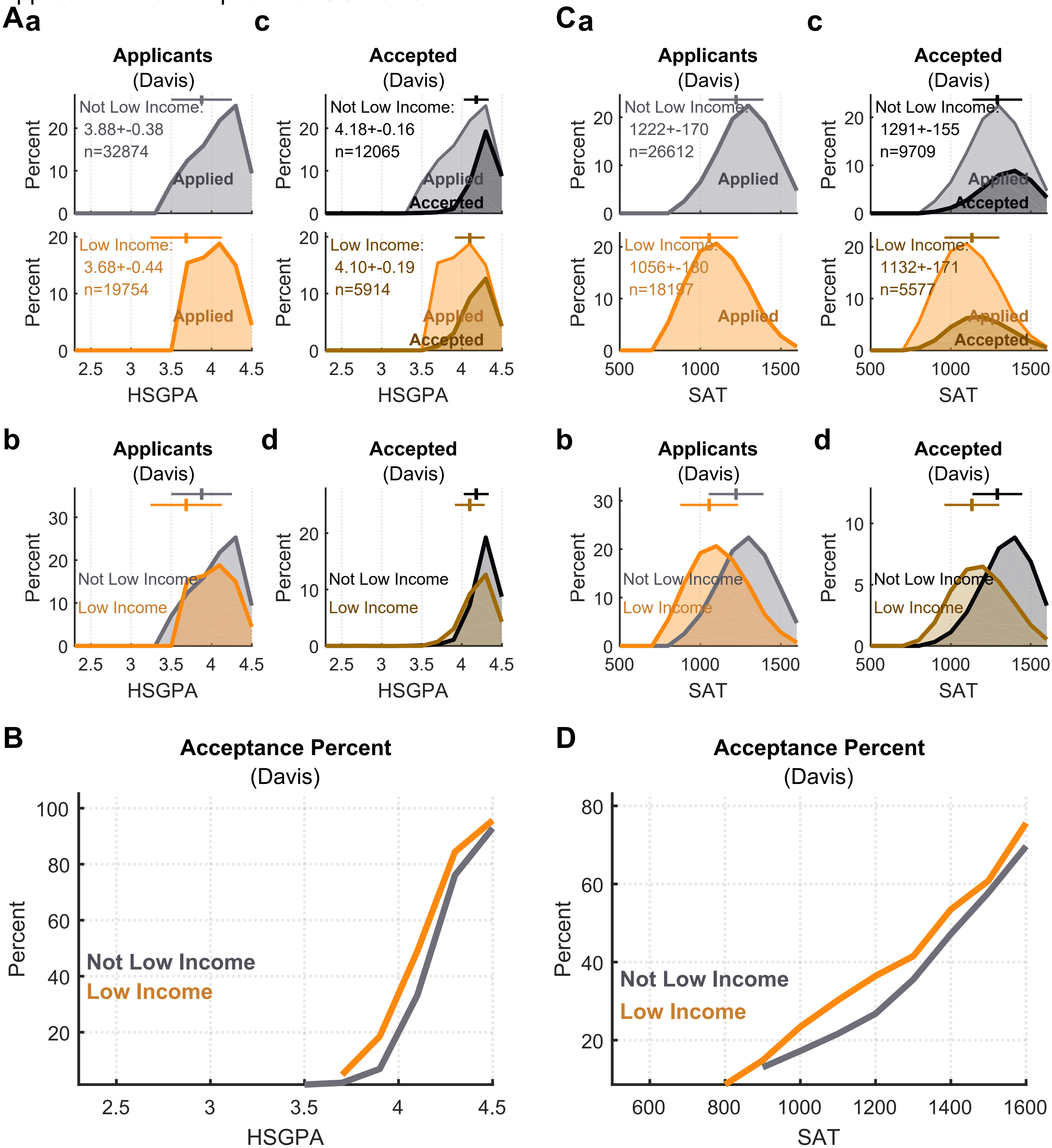
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 23. Family income, HSGPAs, and test scores in the applicant and admit pools for UC Davis.



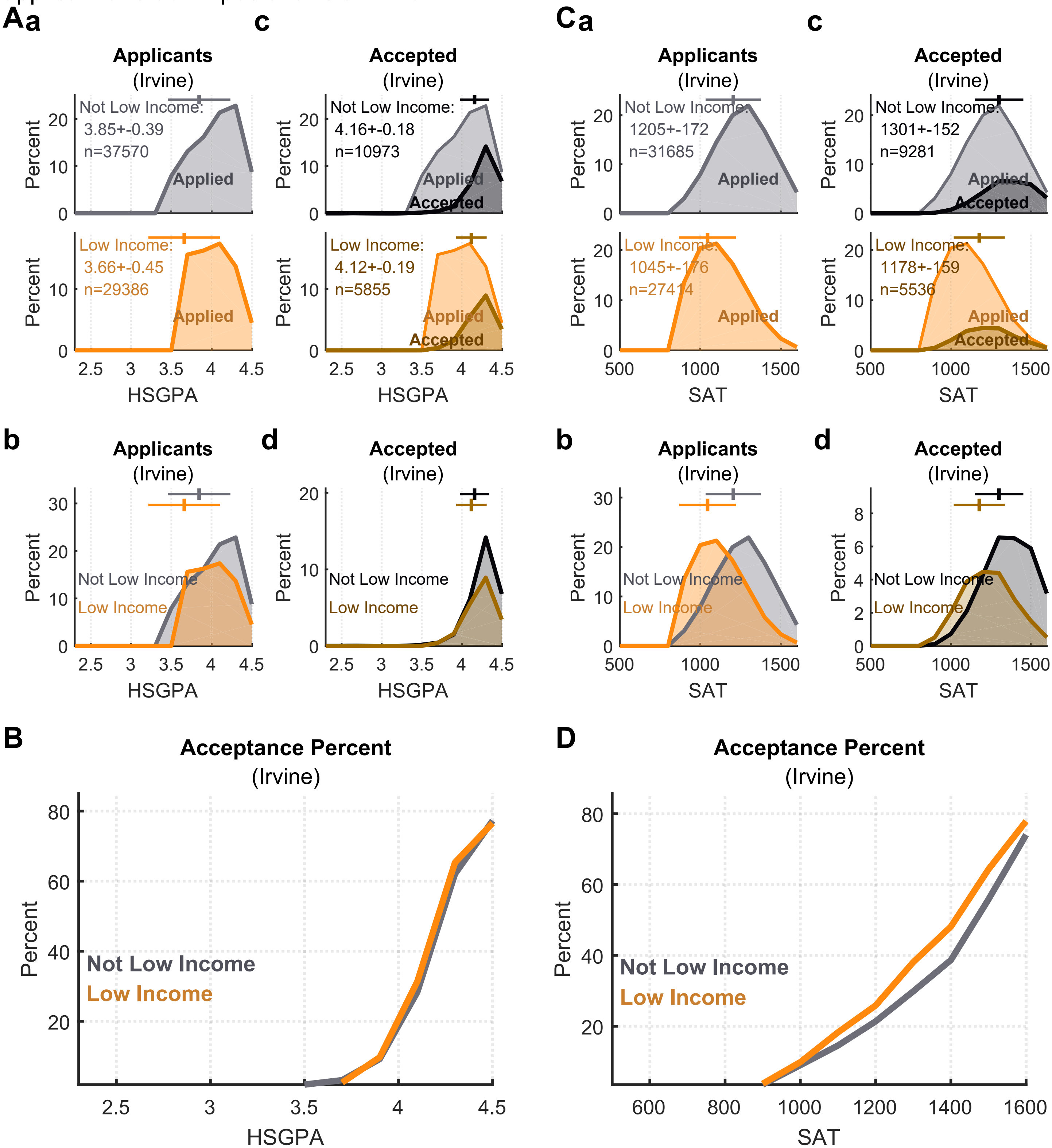
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 24. Family income, HSGPAs, and test scores in the applicant and admit pools for UC Irvine.



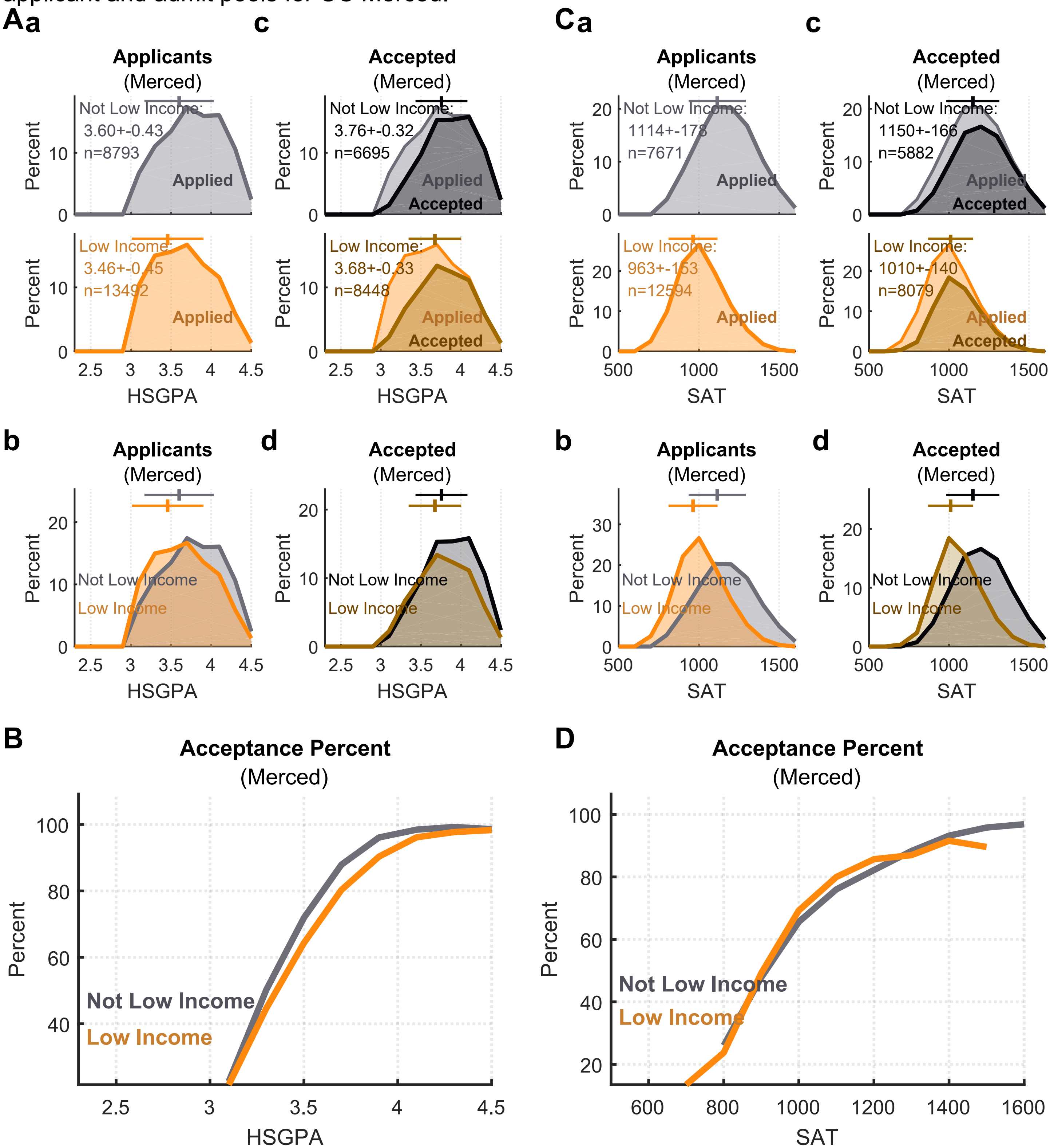
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 25. Family income, HSGPAs, and test scores in the applicant and admit pools for UC Merced.



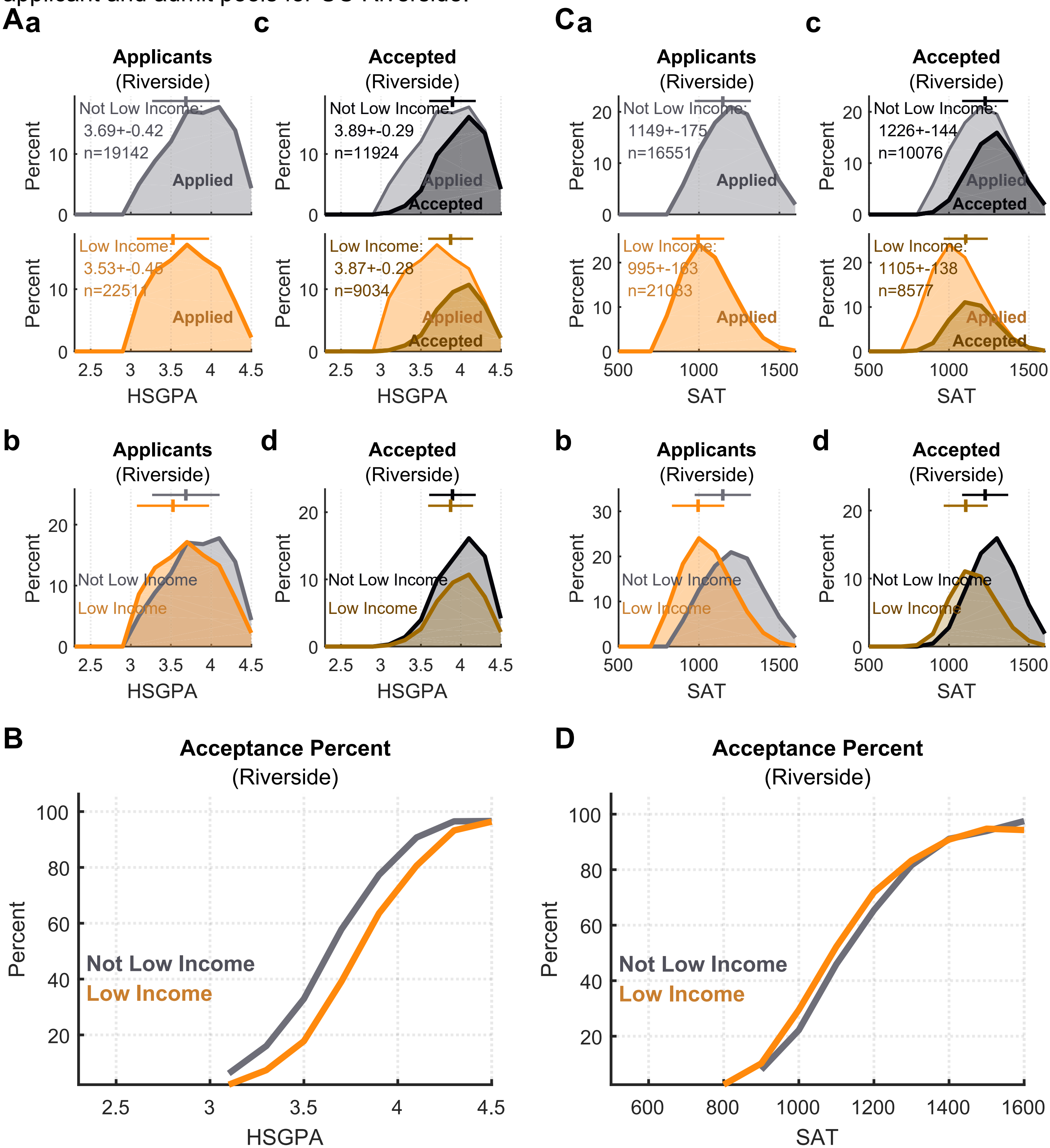
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 26. Family income, HSGPAs, and test scores in the applicant and admit pools for UC Riverside.



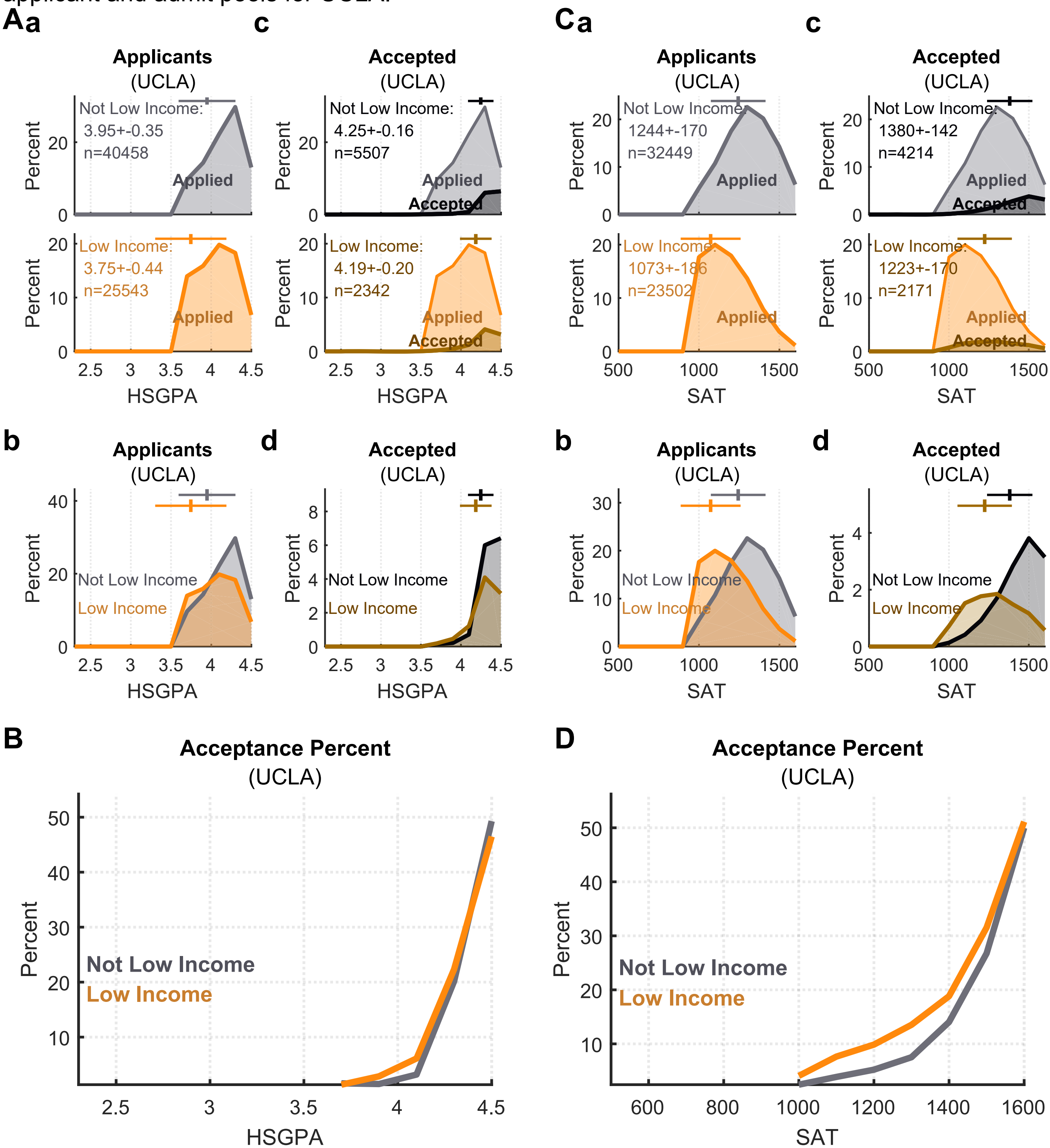
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 27. Family income, HSGPAs, and test scores in the applicant and admit pools for UCLA.



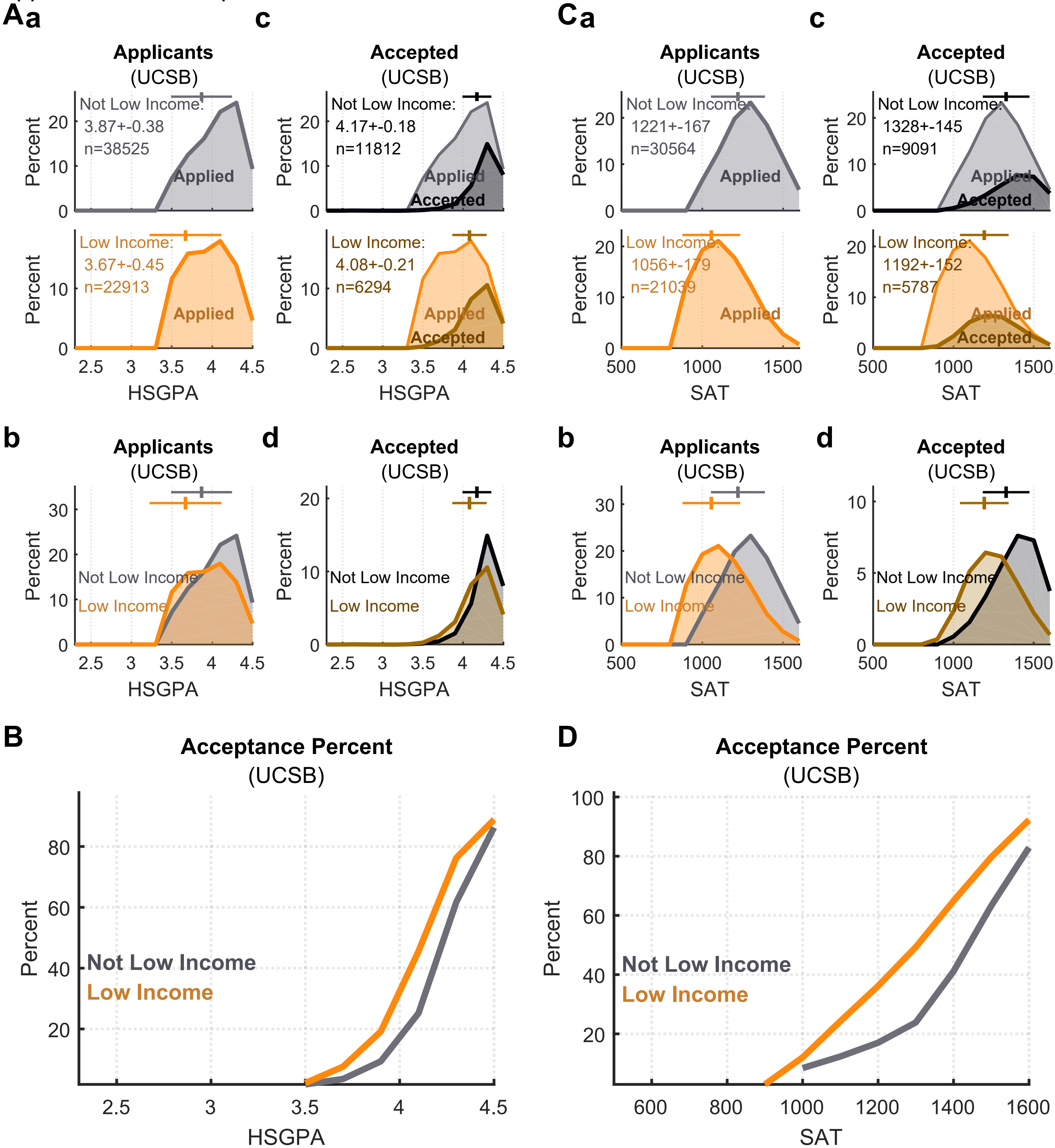
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 28. Family income, HSGPAs, and test scores in the applicant and admit pools for UC Santa Barbara.



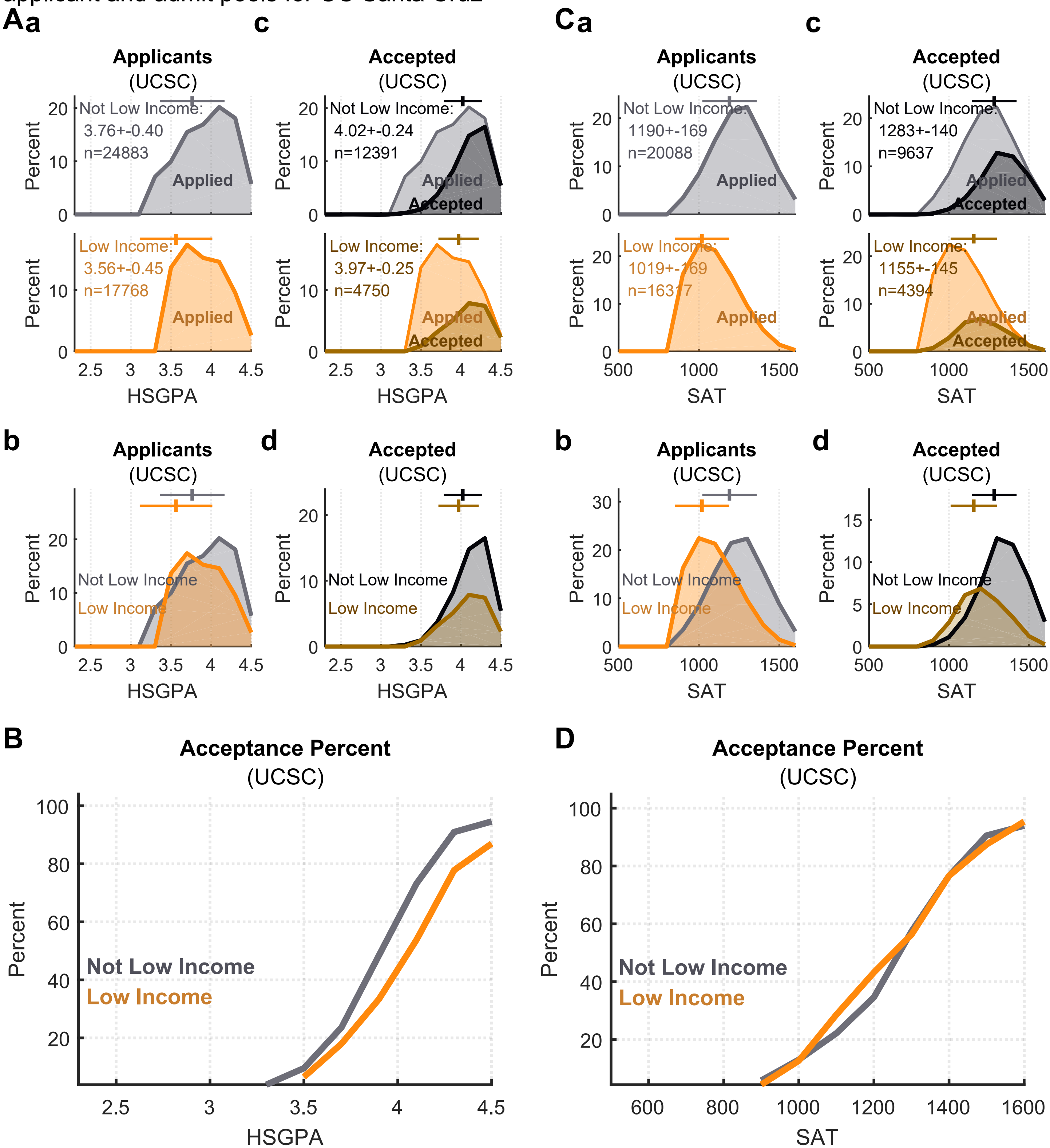
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 29. Family income, HSGPAs, and test scores in the applicant and admit pools for UC Santa Cruz



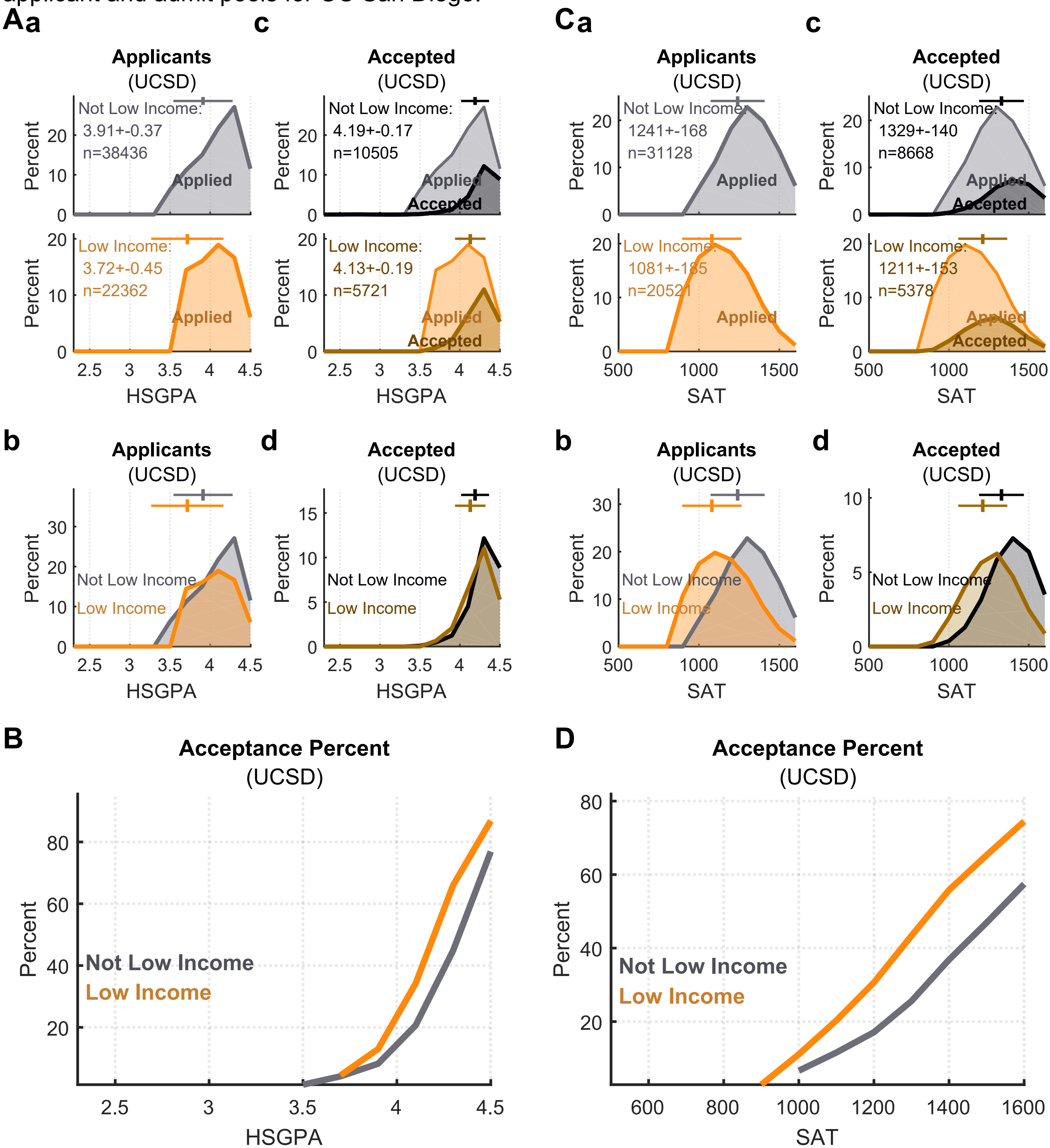
Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Figure 30. Family income, HSGPAs, and test scores in the applicant and admit pools for UC San Diego.



Aa: Rows show HSGPAs for **applicants** from different demographic groups. HSGPAs are divided into 13 equal bands from 2.1 to 4.5. Y-axis shows the percent of applicants in that demographic group who had a HSGPA in that range. Means and standard deviations are plotted above each group. Color corresponds to demographic group. **Ab:** HSGPAs for all applicant demographic groups are overlaid.

Ac: HSGPA distributions for students applying (lighter color) and accepted (darker color). Y axis is the percent of applicants in each demographic group who had HSGPA in that range (light color) or who had HSGPA in that range and were admitted (dark color). Means and standard deviations for HSGPAs of accepted students are plotted above each group. **Ad:** HSGPAs for **accepted** students from all demographic groups are overlaid.

B: Percent of applicants in each demographic group and HSGPA combination who were admitted.

C-D: As **A-B** but for SAT scores instead of HSGPA. To preserve student privacy, combinations with fewer than 30 accepted students are blanked out.

Appendix V. Table 1: HSGPA and Score Differences By Race/Ethnicity, Each Campus

All Campuses

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Af-Amer	3.55	3.86	0.31	0.7
Asian	3.85	3.99	0.14	0.32
Hispan	3.64	3.88	0.25	0.56
White	3.88	4.05	0.16	0.37

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Af-Amer	1043	1120	77	0.39
Asian	1221	1270	49	0.25
Hispan	1022	1080	57	0.29
White	1214	1262	48	0.24

Berkeley

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Af-Amer	3.67	4.13	0.46	1.19
Asian	3.99	4.24	0.25	0.64
Hispan	3.82	4.21	0.39	1.02
White	4.01	4.23	0.23	0.59

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Af-Amer	1088	1199	110	0.57
Asian	1294	1390	96	0.5
Hispan	1088	1195	107	0.55
White	1277	1362	85	0.44

Davis

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Af-Amer	3.59	4.08	0.49	1.2
Asian	3.87	4.18	0.31	0.75
Hispan	3.69	4.1	0.41	0.99
White	3.91	4.19	0.28	0.69

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Af-Amer	1055	1136	81	0.42
Asian	1236	1310	74	0.38
Hispan	1036	1108	72	0.37
White	1228	1290	62	0.32

UCLA

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Af-Amer	3.64	4.12	0.47	1.17
Asian	3.95	4.26	0.31	0.76
Hispan	3.75	4.19	0.44	1.09
White	3.98	4.24	0.26	0.65

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Af-Amer	1070	1200	130	0.66
Asian	1268	1411	143	0.72
Hispan	1056	1203	147	0.74
White	1250	1369	119	0.6

Appendix V. Table 1. Left: Each row shows, for one campus and one demographic group, the mean HSGPAs for applicants and for admitted students, the change in mean grades between the pool of applicants and the pool of accepted students, and the size of the change normalized to units of standard deviations (calculated across the entire applicant pool) – a measure of how strongly the campus selects based on that qualification. Normalized changes are color coded (red = strongest selection, white = no selection). **Right:** As Left, but for SAT. Note that 1) the HSGPAs of admitted students are typically substantially higher than the HSGPAs of applicants, 2) the SATs of admitted students are also typically higher than the SATs of applicants, 3) measured selectivity based on test scores is generally much smaller than the selectivity based on grades, across different demographic groups and campuses, and 4) for the groups most underrepresented at UC, the contribution of selection on HSGPA exceeds the contribution of selection on test scores.

Appendix V. Table 1 (continued): HSGPA and Score Differences By Race/Ethnicity, Each Campus

Riverside

HSGPA differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Af-Amer	3.44	3.85	0.4	0.91
Asian	3.71	3.88	0.17	0.39
Hispan	3.52	3.89	0.37	0.83
White	3.7	3.9	0.2	0.44

SAT differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Af-Amer	1006	1123	116	0.62
Asian	1183	1245	62	0.33
Hispan	977	1080	103	0.55
White	1153	1213	60	0.32

UCSD

HSGPA differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Af-Amer	3.61	4.12	0.51	1.23
Asian	3.91	4.19	0.27	0.66
Hispan	3.71	4.12	0.41	1
White	3.93	4.19	0.27	0.65

SAT differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Af-Amer	1067	1201	134	0.69
Asian	1262	1346	84	0.43
Hispan	1058	1177	120	0.62
White	1241	1324	82	0.42

UCSC

HSGPA differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Af-Amer	3.5	3.85	0.36	0.82
Asian	3.75	4.02	0.26	0.61
Hispan	3.56	3.98	0.41	0.95
White	3.81	4.05	0.24	0.54

SAT differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Af-Amer	1043	1146	103	0.54
Asian	1207	1301	94	0.49
Hispan	1007	1144	137	0.72
White	1205	1279	75	0.39

UCSB

HSGPA differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Af-Amer	3.59	4.06	0.47	1.14
Asian	3.87	4.16	0.3	0.71
Hispan	3.67	4.08	0.41	0.98
White	3.9	4.18	0.28	0.68

SAT differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Af-Amer	1064	1211	147	0.77
Asian	1248	1352	104	0.54
Hispan	1040	1159	119	0.62
White	1228	1324	96	0.5

Appendix V., Table 1 (continued). **Left:** Each row shows, for one campus and one demographic group, the mean HSGPAs for applicants and for admitted students, the change in mean grades between the pool of applicants and the pool of accepted students, and the size of the change normalized to units of standard deviations (calculated across the entire applicant pool) – a measure of how strongly the campus selects based on that qualification. Normalized changes are color coded (red = strongest selection, white = no selection). **Right:** As Left, but for SAT. Note that 1) the HSGPAs of admitted students are typically substantially higher than the HSGPAs of applicants, 2) the SATs of admitted students are also typically higher than the SATs of applicants, 3) measured selectivity based on test scores is generally much smaller than the selectivity based on grades, across different demographic groups and campuses, and 4) for the groups most underrepresented at UC, the contribution of selection on HSGPA exceeds the contribution of selection on test scores.

Appendix V. Table 1 (continued): HSGPA and Score Differences By Race/Ethnicity, Each Campus

Irvine

HSGPA differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Af-Amer	3.56	3.98	0.42	0.98
Asian	3.86	4.17	0.31	0.73
Hispan	3.65	4.12	0.47	1.1
White	3.86	4.17	0.31	0.73

SAT differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Af-Amer	1047	1178	131	0.68
Asian	1227	1325	98	0.51
Hispan	1022	1137	115	0.59
White	1207	1289	82	0.43

Merced

HSGPA differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Af-Amer	3.37	3.65	0.28	0.62
Asian	3.61	3.76	0.15	0.34
Hispan	3.46	3.67	0.21	0.47
White	3.63	3.78	0.15	0.34

SAT differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Af-Amer	978	1041	62	0.35
Asian	1140	1176	36	0.2
Hispan	954	999	45	0.25
White	1141	1168	27	0.15

Appendix V. Table 1 (continued). **Left:** Each row shows, for one campus and one demographic group, the mean HSGPAs for applicants and for admitted students, the change in mean grades between the pool of applicants and the pool of accepted students, and the size of the change normalized to units of standard deviations (calculated across the entire applicant pool) – a measure of how strongly the campus selects based on that qualification. Normalized changes are color coded (red = strongest selection, white = no selection). **Right:** As Left, but for SAT. Note that 1) the HSGPAs of admitted students are typically substantially higher than the HSGPAs of applicants, 2) the SATs of admitted students are also typically higher than the SATs of applicants, 3) measured selectivity based on test scores is generally much smaller than the selectivity based on grades, across different demographic groups and campuses, and 4) for the groups most underrepresented at UC, the contribution of selection on HSGPA exceeds the contribution of selection on test scores.

Appendix V. Table 2: HSGPA and Score Differences By Family Income, Each Campus

All Campuses

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	3.84	4.01	0.17	0.38
Low Income	3.63	3.88	0.25	0.56

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	1194	1251	57	0.29
Low Income	1035	1100	65	0.33

Berkeley

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	3.98	4.23	0.25	0.66
Low Income	3.81	4.2	0.39	1.01

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	1274	1369	95	0.49
Low Income	1110	1225	116	0.59

Davis

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	3.88	4.18	0.3	0.74
Low Income	3.68	4.1	0.41	1

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	1222	1291	68	0.35
Low Income	1056	1132	76	0.39

UCLA

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	3.95	4.25	0.3	0.74
Low Income	3.75	4.19	0.44	1.09

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	1244	1380	136	0.69
Low Income	1073	1223	151	0.76

Appendix V. Table 2. Left: Each row shows, for one campus and one demographic group, the mean HSGPAs for applicants and for admitted students, the change in mean grades between the pool of applicants and the pool of accepted students, and the size of the change normalized to units of standard deviations (calculated across the entire applicant pool) – a measure of how strongly the campus selects based on that qualification. Normalized changes are color coded (red = strongest selection, white = no selection). **Right:** As Left, but for SAT. Note that 1) the HSGPAs of admitted students are typically substantially higher than the HSGPAs of applicants, 2) the SATs of admitted students are also typically higher than the SATs of applicants, 3) measured selectivity based on test scores is generally much smaller than the selectivity based on grades, across different demographic groups and campuses, and 4) for the groups most underrepresented at UC, the contribution of selection on HSGPA exceeds the contribution of selection on test scores.

Appendix V. Table 2 (continued): HSGPA and Score Differences By Family Income, Each Campus

Riverside

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	3.69	3.89	0.21	0.47
Low Income	3.53	3.87	0.35	0.78

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	1149	1226	78	0.42
Low Income	995	1105	110	0.59

UCSD

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	3.91	4.19	0.28	0.68
Low Income	3.72	4.13	0.41	1

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	1241	1329	88	0.45
Low Income	1081	1211	131	0.68

UCSC

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	3.76	4.02	0.26	0.6
Low Income	3.56	3.97	0.41	0.94

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	1190	1283	93	0.49
Low Income	1019	1155	136	0.71

UCSB

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	3.87	4.17	0.3	0.73
Low Income	3.67	4.08	0.41	0.99

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not Low Income	1221	1328	107	0.56
Low Income	1056	1192	136	0.71

Appendix V. Table 2 (continued). **Left:** Each row shows, for one campus and one demographic group, the mean HSGPAs for applicants and for admitted students, the change in mean grades between the pool of applicants and the pool of accepted students, and the size of the change normalized to units of standard deviations (calculated across the entire applicant pool) – a measure of how strongly the campus selects based on that qualification. Normalized changes are color coded (red = strongest selection, white = no selection). **Right:** As Left, but for SAT. Note that 1) the HSGPAs of admitted students are typically substantially higher than the HSGPAs of applicants, 2) the SATs of admitted students are also typically higher than the SATs of applicants, 3) measured selectivity based on test scores is generally much smaller than the selectivity based on grades, across different demographic groups and campuses, and 4) for the groups most underrepresented at UC, the contribution of selection on HSGPA exceeds the contribution of selection on test scores.

Appendix V. Table 2 (continued): HSGPA and Score Differences By Family Income, Each Campus

Irvine

HSGPA differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not Low Income	3.85	4.16	0.31	0.74
Low Income	3.66	4.12	0.46	1.08

SAT differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not Low Income	1205	1301	96	0.5
Low Income	1045	1178	133	0.69

Merced

HSGPA differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not Low Income	3.6	3.76	0.16	0.35
Low Income	3.46	3.68	0.22	0.48

SAT differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not Low Income	1114	1150	36	0.2
Low Income	963	1010	47	0.26

Appendix V. Table 2 (continued). **Left:** Each row shows, for one campus and one demographic group, the mean HSGPAs for applicants and for admitted students, the change in mean grades between the pool of applicants and the pool of accepted students, and the size of the change normalized to units of standard deviations (calculated across the entire applicant pool) – a measure of how strongly the campus selects based on that qualification. Normalized changes are color coded (red = strongest selection, white = no selection). **Right:** As Left, but for SAT. Note that 1) the HSGPAs of admitted students are typically substantially higher than the HSGPAs of applicants, 2) the SATs of admitted students are also typically higher than the SATs of applicants, 3) measured selectivity based on test scores is generally much smaller than the selectivity based on grades, across different demographic groups and campuses, and 4) for the groups most underrepresented at UC, the contribution of selection on HSGPA exceeds the contribution of selection on test scores.

Appendix V. Table 3: HSGPA and Score Differences By Family Educational History, Each Campus

All Campuses

HSGPA differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not1stGen	3.87	4.02	0.16	0.36
1stGen	3.64	3.89	0.24	0.55

SAT differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not 1stGen	1226	1281	55	0.28
1stGen	1036	1097	62	0.31

Berkeley

HSGPA differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not1stGen	4	4.24	0.24	0.63
1stGen	3.82	4.2	0.38	1.01

SAT differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not 1stGen	1297	1390	93	0.48
1stGen	1103	1205	102	0.53

Davis

HSGPA differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not1stGen	3.89	4.19	0.3	0.72
1stGen	3.7	4.1	0.41	0.98

SAT differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not 1stGen	1250	1321	71	0.37
1stGen	1055	1127	72	0.37

UCLA

HSGPA differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not1stGen	3.96	4.25	0.29	0.72
1stGen	3.76	4.19	0.43	1.06

SAT differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not 1stGen	1268	1400	132	0.67
1stGen	1072	1209	137	0.69

Appendix V., Table 3. Left: Each row shows, for one campus and one demographic group, the mean HSGPAs for applicants and for admitted students, the change in mean grades between the pool of applicants and the pool of accepted students, and the size of the change normalized to units of standard deviations (calculated across the entire applicant pool) – a measure of how strongly the campus selects based on that qualification. Normalized changes are color coded (red = strongest selection, white = no selection). **Right:** As Left, but for SAT. Note that 1) the HSGPAs of admitted students are typically substantially higher than the HSGPAs of applicants, 2) the SATs of admitted students are also typically higher than the SATs of applicants, 3) measured selectivity based on test scores is generally much smaller than the selectivity based on grades, across different demographic groups and campuses, and 4) for the groups most underrepresented at UC, the contribution of selection on HSGPA exceeds the contribution of selection on test scores.

Appendix V. Table 3 (continued): HSGPA and Score Differences By Family Educational History, Each Campus

Riverside

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not1stGen	3.7	3.89	0.19	0.43
1stGen	3.53	3.88	0.34	0.77

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not 1stGen	1182	1251	69	0.37
1stGen	998	1105	108	0.58

UCSD

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not1stGen	3.92	4.2	0.27	0.66
1stGen	3.72	4.13	0.41	0.98

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not 1stGen	1264	1352	88	0.46
1stGen	1077	1205	128	0.66

UCSC

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not1stGen	3.79	4.03	0.24	0.55
1stGen	3.57	3.97	0.4	0.92

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not 1stGen	1219	1302	83	0.44
1stGen	1020	1147	127	0.67

UCSB

HSGPA differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not1stGen	3.89	4.18	0.3	0.72
1stGen	3.68	4.08	0.4	0.97

SAT differences between applicant and accepted pools				
Group	Appl.	Accepted	Change	Change in STD units
Not 1stGen	1246	1356	111	0.58
1stGen	1054	1184	130	0.68

Appendix V, Table 3. Left: Each row shows, for one campus and one demographic group, the mean HSGPAs for applicants and for admitted students, the change in mean grades between the pool of applicants and the pool of accepted students, and the size of the change normalized to units of standard deviations (calculated across the entire applicant pool) – a measure of how strongly the campus selects based on that qualification. Normalized changes are color coded (red = strongest selection, white = no selection). **Right:** As Left, but for SAT. Note that 1) the HSGPAs of admitted students are typically substantially higher than the HSGPAs of applicants, 2) the SATs of admitted students are also typically higher than the SATs of applicants, 3) measured selectivity based on test scores is generally much smaller than the selectivity based on grades, across different demographic groups and campuses, and 4) for the groups most underrepresented at UC, the contribution of selection on HSGPA exceeds the contribution of selection on test scores.

Appendix V. Table 3 (continued): HSGPA and Score Differences By Family Educational History, Each Campus

Irvine

HSGPA differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not1stGen	3.86	4.17	0.31	0.72
1stGen	3.67	4.12	0.45	1.05

SAT differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not 1stGen	1233	1333	99	0.52
1stGen	1047	1169	122	0.63

Merced

HSGPA differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not1stGen	3.61	3.77	0.16	0.36
1stGen	3.47	3.68	0.21	0.47

SAT differences between applicant and accepted pools				
Group	**Appl.**	**Accepted**	**Change**	**Change in STD units**
Not 1stGen	1157	1194	37	0.2
1stGen	967	1013	47	0.26

Appendix V, Table 3. **Left:** Each row shows, for one campus and one demographic group, the mean HSGPAs for applicants and for admitted students, the change in mean grades between the pool of applicants and the pool of accepted students, and the size of the change normalized to units of standard deviations (calculated across the entire applicant pool) – a measure of how strongly the campus selects based on that qualification. Normalized changes are color coded (red = strongest selection, white = no selection). **Right:** As Left, but for SAT. Note that 1) the HSGPAs of admitted students are typically substantially higher than the HSGPAs of applicants, 2) the SATs of admitted students are also typically higher than the SATs of applicants, 3) measured selectivity based on test scores is generally much smaller than the selectivity based on grades, across different demographic groups and campuses, and 4) for the groups most underrepresented at UC, the contribution of selection on HSGPA exceeds the contribution of selection on test scores.

Appendix V. Table 4: Uncompensated GPA and Score Differences By Race/Ethnicity, Each Campus

All Campuses

HSGPA: Appl. vs Asian / Accepted vs Asian					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Af-Amer	3.55	0.31	3.86	0.14	55%
Asian	3.85	---	3.99	---	---
Hispan	3.64	0.22	3.88	0.11	48%
White	3.88	-0.03	4.05	-0.05	-73%

SAT: Appl. vs Asian / Accepted vs Asian					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Af-Amer	1043	177	1120	150	16%
Asian	1221	---	1270	---	---
Hispan	1022	198	1080	190	4%
White	1214	6	1262	8	-23%

Berkeley

HSGPA: Appl. vs Asian / Accepted vs Asian					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Af-Amer	3.67	0.32	4.13	0.11	66%
Asian	3.99	---	4.24	---	---
Hispan	3.82	0.18	4.21	0.03	81%
White	4.01	-0.01	4.23	0.01	135%

SAT: Appl. vs Asian / Accepted vs Asian					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Af-Amer	1088	205	1199	191	7%
Asian	1294	---	1390	---	---
Hispan	1088	206	1195	195	5%
White	1277	16	1362	28	-70%

Davis

HSGPA: Appl. vs Asian / Accepted vs Asian					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Af-Amer	3.59	0.28	4.08	0.09	66%
Asian	3.87	---	4.18	---	---
Hispan	3.69	0.18	4.1	0.08	56%
White	3.91	-0.04	4.19	-0.01	67%

SAT: Appl. vs Asian / Accepted vs Asian					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Af-Amer	1055	182	1136	174	4%
Asian	1236	---	1310	---	---
Hispan	1036	200	1108	202	-1%
White	1228	9	1290	20	-136%

UCLA

HSGPA: Appl. vs Asian / Accepted vs Asian					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Af-Amer	3.64	0.31	4.12	0.14	54%
Asian	3.95	---	4.26	---	---
Hispan	3.75	0.20	4.19	0.07	67%
White	3.98	-0.02	4.24	0.02	179%

SAT: Appl. vs Asian / Accepted vs Asian					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Af-Amer	1070	199	1200	211	-6%
Asian	1268	---	1411	---	---
Hispan	1056	212	1203	208	2%
White	1250	18	1369	42	-129%

Appendix V. Table 4. **Left:** Each row shows, for one campus and one demographic group, mean HSGPAs for applicants, the difference in mean HSGPA (green) between applicants in that group and the group with the highest test scores, mean HSGPAs for admitted students, the difference in mean HSGPA between admitted students in that group and the group with the highest test scores (green), and the ratio of the Accepted difference vs the Applicant difference. If the difference in scores between two groups of accepted students is exactly the same as the difference in scores between applicants in those groups, (i.e., uncompensated difference of 0), that is consistent with the hypothesis that the UC admissions process effectively compensates for the mean differences between groups. **Right:** As Left, but for SAT (differences in blue). Note that 1) the remaining, uncompensated differences in HSGPAs are typically much larger than uncompensated differences in test scores (i.e., the left green columns are darker than the right green columns, but the left blue columns are similar to the right blue columns), 2) the uncompensated differences in HSGPA are largest for underrepresented groups.

Appendix V. Table 4 (continued): Uncompensated GPA and Score Differences By Race/ Ethnicity, Each Campus

Riverside

HSGPA: Appl. vs Asian / Accepted vs Asian					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Af-Amer	3.44	0.27	3.85	0.03	87%
Asian	3.71	---	3.88	---	---
Hispan	3.52	0.19	3.89	-0.01	104%
White	3.7	0.01	3.9	-0.02	280%

SAT: Appl. vs Asian / Accepted vs Asian					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Af-Amer	1006	177	1123	122	31%
Asian	1183	---	1245	---	---
Hispan	977	207	1080	165	20%
White	1153	30	1213	32	-5%

UCSD

HSGPA: Appl. vs Asian / Accepted vs Asian					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Af-Amer	3.61	0.30	4.12	0.07	78%
Asian	3.91	---	4.19	---	---
Hispan	3.71	0.20	4.12	0.06	69%
White	3.93	-0.01	4.19	-0.01	41%

SAT: Appl. vs Asian / Accepted vs Asian					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Af-Amer	1067	195	1201	144	26%
Asian	1262	---	1346	---	---
Hispan	1058	204	1177	168	18%
White	1241	21	1324	22	-7%

UCSC

HSGPA: Appl. vs Asian / Accepted vs Asian					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Af-Amer	3.5	0.26	3.85	0.17	36%
Asian	3.75	---	4.02	---	---
Hispan	3.56	0.19	3.98	0.04	77%
White	3.81	-0.06	4.05	-0.03	52%

SAT: Appl. vs Asian / Accepted vs Asian					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Af-Amer	1043	164	1146	156	5%
Asian	1207	---	1301	---	---
Hispan	1007	199	1144	157	21%
White	1205	2	1279	22	---

UCSB

HSGPA: Appl. vs Asian / Accepted vs Asian					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Af-Amer	3.59	0.28	4.06	0.10	64%
Asian	3.87	---	4.16	---	---
Hispan	3.67	0.19	4.08	0.08	57%
White	3.9	-0.03	4.18	-0.02	39%

SAT: Appl. vs Asian / Accepted vs Asian					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Af-Amer	1064	184	1211	140	24%
Asian	1248	---	1352	---	---
Hispan	1040	208	1159	193	7%
White	1228	20	1324	28	-38%

Appendix V. Table 4 (continued). **Left:** Each row shows, for one campus and one demographic group, mean HSGPAs for applicants, the difference in mean HSGPA (green) between applicants in that group and the group with the highest test scores, mean HSGPAs for admitted students, the difference in mean HSGPA between admitted students in that group and the group with the highest test scores (green), and the ratio of the Accepted difference vs the Applicant difference. If the difference in scores between two groups of accepted students is exactly the same as the difference in scores between applicants in those groups, (i.e., uncompensated difference of 0), that is consistent with the hypothesis that the UC admissions process effectively compensates for the mean differences between groups. **Right:** As Left, but for SAT (differences in blue). Note that 1) the remaining, uncompensated differences in HSGPAs are typically much larger than uncompensated differences in test scores (i.e., the left green columns are darker than the right green columns, but the left blue columns are similar to the right blue columns), 2) the uncompensated differences in HSGPA are largest for underrepresented groups.

Appendix V. Table 4 (continued): Uncompensated GPA and Score Differences By Race/ Ethnicity, Each Campus

Irvine

HSGPA: Appl. vs Asian / Accepted vs Asian					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Af-Amer	3.56	0.29	3.98	0.19	36%
Asian	3.86	---	4.17	---	---
Hispan	3.65	0.21	4.12	0.05	76%
White	3.86	---	4.17	---	42%

SAT: Appl. vs Asian / Accepted vs Asian					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Af-Amer	1047	180	1178	147	18%
Asian	1227	---	1325	---	---
Hispan	1022	205	1137	188	8%
White	1207	20	1289	36	-79%

Merced

HSGPA: Appl. vs Asian / Accepted vs Asian					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Af-Amer	3.37	0.24	3.65	0.12	52%
Asian	3.61	---	3.76	---	---
Hispan	3.46	0.15	3.67	0.09	40%
White	3.63	-0.01	3.78	-0.02	-19%

SAT: Appl. vs Asian / Accepted vs Asian					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Af-Amer	978	161	1041	136	16%
Asian	1140	---	1176	---	---
Hispan	954	186	999	177	4%
White	1141	-1	1168	8	---

Appendix V. Table 4 (continued). **Left:** Each row shows, for one campus and one demographic group, mean HSGPAs for applicants, the difference in mean HSGPA (green) between applicants in that group and the group with the highest test scores, mean HSGPAs for admitted students, the difference in mean HSGPA between admitted students in that group and the group with the highest test scores (green), and the ratio of the Accepted difference vs the Applicant difference. If the difference in scores between two groups of accepted students is exactly the same as the difference in scores between applicants in those groups, (i.e., uncompensated difference of 0), that is consistent with the hypothesis that the UC admissions process effectively compensates for the mean differences between groups. **Right:** As Left, but for SAT (differences in blue). Note that 1) the remaining, uncompensated differences in HSGPAs are typically much larger than uncompensated differences in test scores (i.e., the left green columns are darker than the right green columns, but the left blue columns are similar to the right blue columns), 2) the uncompensated differences in HSGPA are largest for underrepresented groups.

Appendix V. Table 5: Uncompensated HSGPA and Score Differences By Family Income, Each Campus

All Campuses

HSGPA: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Not Low Income	3.84	---	4.01	---	---
Low Income	3.63	0.20	3.88	0.13	38%

SAT: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Not Low Income	1194	---	1251	---	---
Low Income	1035	159	1100	151	6%

Berkeley

HSGPA: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Not Low Income	3.98	---	4.23	---	---
Low Income	3.81	0.17	4.2	0.03	80%

SAT: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Not Low Income	1274	---	1369	---	---
Low Income	1110	164	1225	144	12%

Davis

HSGPA: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Not Low Income	3.88	---	4.18	---	---
Low Income	3.68	0.19	4.1	0.08	58%

SAT: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Not Low Income	1222	---	1291	---	---
Low Income	1056	167	1132	159	5%

UCLA

HSGPA: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Not Low Income	3.95	---	4.25	---	---
Low Income	3.75	0.20	4.19	0.06	69%

SAT: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	Appl.	Diff	Accepted	Diff	Percent Uncompensated
Not Low Income	1244	---	1380	---	---
Low Income	1073	172	1223	157	9%

Appendix V. Table 5. Left: Each row shows, for one campus and one demographic group, mean HSGPAs for applicants, the difference in mean HSGPA (green) between applicants in that group and the group with the highest test scores, mean HSGPAs for admitted students, the difference in mean HSGPA between admitted students in that group and the group with the highest test scores (green), and the ratio of the Accepted difference vs the Applicant difference. If the difference in scores between two groups of accepted students is exactly the same as the difference in scores between applicants in those groups, (i.e., uncompensated difference of 0), that is consistent with the hypothesis that the UC admissions process effectively compensates for the mean differences between groups. **Right:** As Left, but for SAT (differences in blue). Note that 1) the remaining, uncompensated differences in HSGPAs are typically much larger than uncompensated differences in test scores (i.e., the left green columns are darker than the right green columns, but the left blue columns are similar to the right blue columns), 2) the uncompensated differences in HSGPA are largest for underrepresented groups.

Appendix V. Table 5 (continued): Uncompensated HSGPA and Score Differences By Family Income, Each Campus

Riverside

HSGPA: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not Low Income	3.69	---	3.89	---	---
Low Income	3.53	0.16	3.87	0.02	85%

SAT: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not Low Income	1149	---	1226	---	---
Low Income	995	154	1105	121	21%

UCSD

HSGPA: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not Low Income	3.91	---	4.19	---	---
Low Income	3.72	0.20	4.13	0.06	69%

SAT: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not Low Income	1241	---	1329	---	---
Low Income	1081	160	1211	117	27%

UCSC

HSGPA: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not Low Income	3.76	---	4.02	---	---
Low Income	3.56	0.20	3.97	0.05	74%

SAT: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not Low Income	1190	---	1283	---	---
Low Income	1019	171	1155	128	25%

UCSB

HSGPA: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not Low Income	3.87	---	4.17	---	---
Low Income	3.67	0.20	4.08	0.09	54%

SAT: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not Low Income	1221	---	1328	---	---
Low Income	1056	165	1192	136	18%

Appendix V. Table 5, continued. **Left:** Each row shows, for one campus and one demographic group, mean HSGPAs for applicants, the difference in mean HSGPA (green) between applicants in that group and the group with the highest test scores, mean HSGPAs for admitted students, the difference in mean HSGPA between admitted students in that group and the group with the highest test scores (green), and the ratio of the Accepted difference vs the Applicant difference. If the difference in scores between two groups of accepted students is exactly the same as the difference in scores between applicants in those groups, (i.e., uncompensated difference of 0), that is consistent with the hypothesis that the UC admissions process effectively compensates for the mean differences between groups. **Right:** As Left, but for SAT (differences in blue). Note that 1) the remaining, uncompensated differences in HSGPAs are typically much larger than uncompensated differences in test scores (i.e., the left green columns are darker than the right green columns, but the left blue columns are similar to the right blue columns), 2) the uncompensated differences in HSGPA are largest for underrepresented groups.

Appendix V. Table 5 (continued): Uncompensated HSGPA and Score Differences By Family Income, Each Campus

Irvine

HSGPA: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not Low Income	3.85	---	4.16	---	---
Low Income	3.66	0.18	4.12	0.04	78%

SAT: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not Low Income	1205	---	1301	---	---
Low Income	1045	160	1178	123	23%

Merced

HSGPA: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not Low Income	3.6	---	3.76	---	---
Low Income	3.46	0.14	3.68	0.08	43%

SAT: Appl. vs Not Low Income / Accepted vs Not Low Income					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not Low Income	1114	---	1150	---	---
Low Income	963	151	1010	139	8%

Appendix V. Table 5, continued. **Left**: Each row shows, for one campus and one demographic group, mean HSGPAs for applicants, the difference in mean HSGPA (green) between applicants in that group and the group with the highest test scores, mean HSGPAs for admitted students, the difference in mean HSGPA between admitted students in that group and the group with the highest test scores (green), and the ratio of the Accepted difference vs the Applicant difference. If the difference in scores between two groups of accepted students is exactly the same as the difference in scores between applicants in those groups, (i.e., uncompensated difference of 0), that is consistent with the hypothesis that the UC admissions process effectively compensates for the mean differences between groups. **Right:** As Left, but for SAT (differences in blue). Note that 1) the remaining, uncompensated differences in HSGPAs are typically much larger than uncompensated differences in test scores (i.e., the left green columns are darker than the right green columns, but the left blue columns are similar to the right blue columns), 2) the uncompensated differences in HSGPA are largest for underrepresented groups.

Appendix V. Table 6: Uncompensated GPA and Score Differences By Family Educational History, Each Campus

All Campuses

HSGPA: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	3.87	---	4.02	---	---
1stGen	3.64	0.22	3.89	0.14	38%

SAT: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	1226	---	1281	---	---
1stGen	1036	191	1097	184	3%

Berkeley

HSGPA: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	4	---	4.24	---	---
1stGen	3.82	0.18	4.2	0.04	80%

SAT: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	1297	---	1390	---	---
1stGen	1103	194	1205	185	5%

Davis

HSGPA: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	3.89	---	4.19	---	---
1stGen	3.7	0.20	4.1	0.09	56%

SAT: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	1250	---	1321	---	---
1stGen	1055	195	1127	194	1%

UCLA

HSGPA: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	3.96	---	4.25	---	---
1stGen	3.76	0.20	4.19	0.06	69%

SAT: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	1268	---	1400	---	---
1stGen	1072	196	1209	191	2%

Appendix V, Table 6. **Left:** Each row shows, for one campus and one demographic group, mean HSGPAs for applicants, the difference in mean HSGPA (green) between applicants in that group and the group with the highest test scores, mean HSGPAs for admitted students, the difference in mean HSGPA between admitted students in that group and the group with the highest test scores (green), and the ratio of the Accepted difference vs the Applicant difference. If the difference in scores between two groups of accepted students is exactly the same as the difference in scores between applicants in those groups, (i.e., uncompensated difference of 0), that is consistent with the hypothesis that the UC admissions process effectively compensates for the mean differences between groups. **Right:** As Left, but for SAT (differences in blue). Note that 1) the remaining, uncompensated differences in HSGPAs are typically much larger than uncompensated differences in test scores (i.e., the left green columns are darker than the right green columns, but the left blue columns are similar to the right blue columns), 2) the uncompensated differences in HSGPA are largest for underrepresented groups.

Appendix V. Table 6 (continued): Uncompensated GPA and Score Differences By Family Educational History, Each Campus

Riverside

HSGPA: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	3.7	---	3.89	---	---
1stGen	3.53	0.17	3.88	0.02	90%

SAT: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	1182	---	1251	---	---
1stGen	998	184	1105	145	21%

UCSD

HSGPA: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	3.92	---	4.2	---	---
1stGen	3.72	0.20	4.13	0.07	66%

SAT: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	1264	---	1352	---	---
1stGen	1077	186	1205	147	21%

UCSC

HSGPA: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	3.79	---	4.03	---	---
1stGen	3.57	0.22	3.97	0.06	73%

SAT: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	1219	---	1302	---	---
1stGen	1020	199	1147	155	22%

UCSB

HSGPA: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	3.89	---	4.18	---	---
1stGen	3.68	0.21	4.08	0.11	50%

SAT: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	1246	---	1356	---	---
1stGen	1054	192	1184	173	10%

Appendix V. Table 6, continued. Left: Each row shows, for one campus and one demographic group, mean HSGPAs for applicants, the difference in mean HSGPA (green) between applicants in that group and the group with the highest test scores, mean HSGPAs for admitted students, the difference in mean HSGPA between admitted students in that group and the group with the highest test scores (green), and the ratio of the Accepted difference vs the Applicant difference. If the difference in scores between two groups of accepted students is exactly the same as the difference in scores between applicants in those groups, (i.e., uncompensated difference of 0), that is consistent with the hypothesis that the UC admissions process effectively compensates for the mean differences between groups. **Right:** As Left, but for SAT (differences in blue). Note that 1) the remaining, uncompensated differences in HSGPAs are typically much larger than uncompensated differences in test scores (i.e., the left green columns are darker than the right green columns, but the left blue columns are similar to the right blue columns), 2) the uncompensated differences in HSGPA are largest for underrepresented groups.

Appendix V. Table 6 (continued): Uncompensated GPA and Score Differences By Family Educational History, Each Campus

Irvine

HSGPA: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	3.86	---	4.17	---	---
1stGen	3.67	0.19	4.12	0.05	73%

SAT: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	1233	---	1333	---	---
1stGen	1047	187	1169	164	12%

Merced

HSGPA: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	3.61	---	3.77	---	---
1stGen	3.47	0.13	3.68	0.08	38%

SAT: Appl. vs Not 1stGen / Accepted vs Not 1stGen					
Group	**Appl.**	**Diff**	**Accepted**	**Diff**	**Percent Uncompensated**
Not 1stGen	1157	---	1194	---	---
1stGen	967	191	1013	181	5%

Appendix V. Table 6, continued. Left: Each row shows, for one campus and one demographic group, mean HSGPAs for applicants, the difference in mean HSGPA (green) between applicants in that group and the group with the highest test scores, mean HSGPAs for admitted students, the difference in mean HSGPA between admitted students in that group and the group with the highest test scores (green), and the ratio of the Accepted difference vs the Applicant difference. If the difference in scores between two groups of accepted students is exactly the same as the difference in scores between applicants in those groups, (i.e., uncompensated difference of 0), that is consistent with the hypothesis that the UC admissions process effectively compensates for the mean differences between groups. **Right:** As Left, but for SAT (differences in blue). Note that 1) the remaining, uncompensated differences in HSGPAs are typically much larger than uncompensated differences in test scores (i.e., the left green columns are darker than the right green columns, but the left blue columns are similar to the right blue columns), 2) the uncompensated differences in HSGPA are largest for underrepresented groups.