

THE GROWING CORRELATION BETWEEN RACE AND SAT SCORES:
NEW FINDINGS FROM CALIFORNIA*

October 2015

Saul Geiser
Center for Studies in Higher Education
University of California, Berkeley

Copyright 2015 Saul Geiser, all rights reserved.

ABSTRACT

This paper presents new and surprising findings on the relationship between race and SAT scores. The findings are based on the population of California residents who applied for admission to the University of California from 1994 through 2011, a sample of over 1.1 million students. The UC data show that socioeconomic background factors – family income, parental education, and race/ethnicity – account for a large and growing share of the variance in students' SAT scores over the past twenty years. More than a third of the variance in SAT scores can now be predicted by factors known at students' birth, up from a quarter of the variance in 1994. Of those factors, moreover, race has become the strongest predictor. Rather than declining in salience, race and ethnicity are now more important than either family income or parental education in accounting for test score differences. It must be cautioned that these findings are preliminary, and more research is needed to determine whether the California data reflect a broader national trend. But if these findings are representative, they have important implications for the ongoing debate over both affirmative action and standardized testing in college admissions.

Keywords: Higher education, college admissions, standardized tests, race and ethnicity, affirmative action

Race, the SAT, and the University of California

The University of California has been ground zero for two of the most divisive issues in American college admissions: the use of race and SAT scores as selection criteria. The convergence of the two controversies at UC is no accident. UC is the largest single user of the SAT, and its decision to adopt that test in 1979 helped solidify the SAT as the most widely used national admissions test until it was overtaken by the ACT in 2012.

Early on, UC officials recognized the adverse impact of standardized tests on admission of California's growing Latino and African American populations, and this was a major consideration leading to adoption of affirmative action as a compensatory measure.¹ The US Supreme Court's decision in *Regents of the University of California v. Bakke* provided the legal rationale – the university's compelling interest in enrolling a diverse student body – which remains the principal justification for race-conscious admissions to this day.

Later, however, the SAT would play a key role in the backlash against affirmative action. The disparity in test scores between minority and non-minority admits led to claims of "reverse discrimination" that would provide the principal justification for dismantling affirmative action by vote of the UC Regents in 1995 and California's referendum on Proposition 209 in 1996.

Like two magnets that simultaneously attract and repel one another, the strangely symbiotic relationship between race and the SAT continued to play out at UC even after Proposition 209 took effect in 1998. Faced with plummeting minority enrollments at its flagship campuses, UC undertook a sweeping review of all of its admissions criteria. Based on that review, then-UC President Richard Atkinson called for eliminating the SAT in his landmark speech to the American Council on Education (Atkinson, 2001). In addition to its adverse effect on minority admissions, Atkinson observed, the SAT had only marginal validity in predicting

* The author thanks Bill Kidder, Pam Burdman, John Douglass, Juliet Popper Shaffer, and Mark Wilson for their helpful comments.

¹ Many have noted the peculiar connection between affirmative action and the SAT. See, among others, Fischer, et al. (1996), Bowen & Bok (1998), Sacks (2000), and Alon & Tienda, (2007).

student success at UC; curriculum-based tests like the SAT II Subject Tests were equal or better predictors and were more closely aligned with the curricula that students learn in school. In response, the College Board altered the SAT into a more curriculum-based test in 2006 and recently has announced changes that will move the test still further in this direction.

Atkinson's speech to ACE also marked the advent of a number of major admissions reforms at UC designed to diminish the role of the SAT. Among other changes, UC introduced its Top 4 Percent plan, extending eligibility to top students in each California high school, as well as holistic review, an admissions process that until then had been confined largely to elite, private colleges. The common denominator in these reforms was their de-emphasis on SAT scores as a selection criterion in favor of students' high-school record and indicators of socioeconomic disadvantage. The reforms have been closely watched for their national implications.

To both critics and supporters of affirmative action alike, the University of California's experience is a morality tale, if with two very different morals. For critics of affirmative action, it demonstrates what can be done "on the natural" (in the words of a former California governor) to expand minority enrollments by race-neutral means. For supporters, it illustrates not only the dire consequences of eliminating affirmative action but also the impossibility for American colleges and universities to keep pace with the growing diversity of the nation without taking account of race. For better or worse, UC has provided ample grist for both mills. The present study is the latest in a series of studies² employing the California data to explore these and related issues of equity and access in American higher education.

The Growing Correlation between Race and SAT Scores at UC

Figure 1 (next page) shows the multiple correlation between SAT scores and three measures of socioeconomic status -- family income, parental education, and race/ethnicity -- among 1,144,047 California high school graduates who applied for freshman admission at UC from 1994 through 2011. SAT scores are the sum of the verbal and mathematical components of the SAT I until that test was revised in 2006; the sum of the critical reading and math components is used thereafter. Family income is derived from information reported in the UC application and is adjusted for inflation. Following standard practice in economic research, this analysis uses the log of income, which takes into account the diminishing marginal effect of income at higher income levels (i.e., a \$10,000 increase in income is likely to have a greater effect for a student whose family earns \$35,000 annually than one from a family earning \$350,000). Data on parental education and race/ethnicity are also derived from information reported in the UC application. Parents' education is that of the applicant's highest-educated parent.³

The three socioeconomic variables were entered into a multiple regression to determine the extent to which they "explained" or statistically accounted for differences in SAT scores among UC applicants. To provide a point of comparison, the same regression was run with high-school grade point average as the dependent variable to determine the extent to which differences in students' HSGPA could be explained by these same factors.

The regression results show a marked increase since 1994 in the proportion of variance in SAT scores that can be predicted from socioeconomic background factors largely determined at students' birth. After falling slightly from 25% to 21% between 1994 and 1998, the proportion of explained variance increased each year thereafter, growing to 35% by 2011, the last year for which the author has obtained data. Remarkably, more than a third of the variance in SAT scores among UC applicants can now be predicted by family income, education, and race/ethnicity. This result contrasts sharply with that for high school GPA: Socioeconomic background factors accounted for only 7% of the variance in HSGPA in 1994 and 8% in 2011.⁴

This is not to suggest that socioeconomic factors "cause" SAT scores in any simple or direct fashion. As the College Board cautions in its annual reports on SAT performance, "[R]elationships between test scores and other factors such as educational background, gender, racial/ethnic background, parental education and household income are complex and interdependent. These factors do not directly affect test performance; rather, they are associated with educational experiences both on tests such as the SAT I and in schoolwork" (College Board, 2000, p. iii). Socioeconomic factors, in other words, are mediated by other,

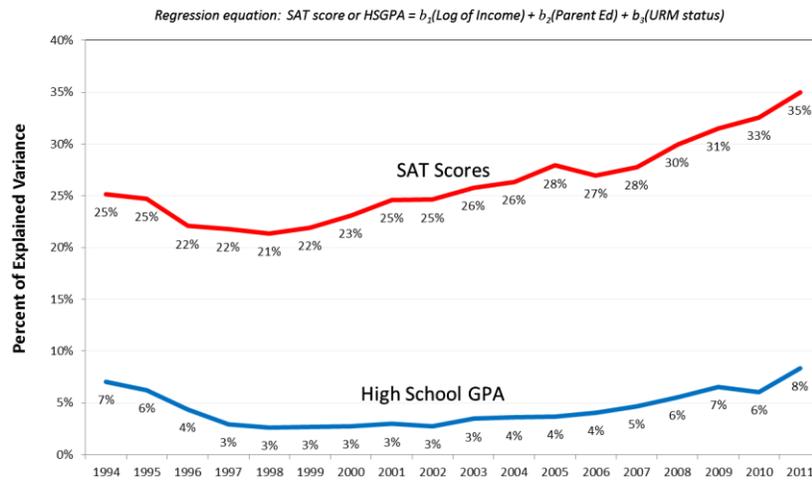
² Geiser & Studley, 2002; Geiser & Caspary, 2005; Geiser & Santelices, 2006; Geiser & Santelices, 2007; Geiser, 2009; Atkinson & Geiser, 2009; Geiser & Atkinson, 2013; Geiser, 2014.

³ See Appendix 1 for a descriptive summary of the variables employed in this study. SAT scores also include ACT-equivalent scores for a small percentage of UC applicants who took only the latter test. California is predominantly an "SAT state," and approximately half of all high school seniors take that test (see Figure 3, below). By comparison, for most of the period covered by this study, fewer than 15 percent took the ACT, although that percentage has been growing in recent years (ACT, 2015). Many UC applicants who take the ACT also take the SAT, in which case their SAT scores have been used in the following analysis. In any event, there is little, if any, difference between SAT and ACT scores with respect to either predictive validity or racial impact; see Bowen, Chingos, & McPherson, 2009, p.116.

⁴ See Appendix 2 for complete documentation of regression analysis and results.

more proximate experiences that do have an impact on test performance, such access to test prep services or the quality of the schools that students attend.

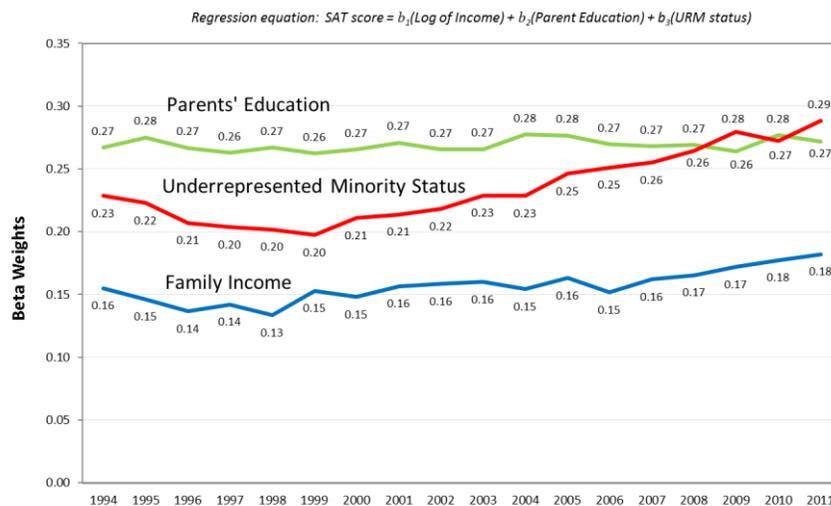
Figure 1
Percent of Variance in SAT Scores and High School Grades Explained by Family Income, Parents' Education, and Underrepresented Minority Status, UC Applicants, 1994 to 2011



Source: UC Corporate Student System data on California resident freshmen applicants, 1994 to 2011.

Nevertheless, even without being able to observe those intermediating experiences directly, regression analysis enables one to assess the relative importance of different socioeconomic factors in predicting test performance. Figure 2 provides standardized regression coefficients, or “beta weights,” for predicting SAT scores conditional on family income, parents’ education, and race/ethnicity. The coefficients show the predictive weight of each factor after controlling for the effects of the other two, thereby providing a measure of the unique contribution of each factor to the prediction.

Figure 2
Standardized Regression Coefficients for Predicting SAT Scores Conditional on Family Income, Parents' Education, and Underrepresented Minority Status: UC Applicants, 1994 to 2011



Source: UC Corporate Student System data on California resident freshmen applicants, 1994 to 2011.

In 1994, at the beginning of the period covered in this analysis, parental education was the strongest of the three socioeconomic predictors of test performance. (The standardized regression coefficient of 0.27 in that year means that, for each one standard

deviation increase in parental education, SAT scores increased by 0.27 of a standard deviation, when income and underrepresented minority status were held constant.) The predictive weight for parental education has remained about the same since then. The weight for family income has shown a small but steady increase from 0.13 in 1998 to 0.18 in 2011.

But the most important change has been the growing salience of race/ethnicity. By 2011, the predictive weight for underrepresented minority status, 0.29, was greater than that for either family income or parental education. When the regression results for the UC sample are pooled across applicant cohorts, race/ethnicity is the strongest predictor of SAT scores over the last four years.⁵

A key implication of this finding is that racial and ethnic group differences in SAT scores are not simply reducible to differences in family income and parental education. At least for the UC sample, there remains a large and growing residual effect of race/ethnicity *after* those factors are taken into account. Whatever mediating factors may be involved, it appears that their effects are different and more pronounced for students of color. If true, this conclusion has important implications about the efficacy of race-neutral policies for redressing racial disparities in college admissions.

The remainder of the paper is devoted to interrogating and interpreting these findings. A first question to be addressed is whether the findings reflect selection effects or other statistical anomalies. The following sections present data that discount that explanation and offer alternative hypotheses, based on extant research on the black-white “test score gap,” that may account for the observed trends. The paper then examines the differential impact of the SAT on admission of Latino and black applicants as well as the validity of that test in predicting student success at UC. The final sections of the paper explore the implications of the California findings, if confirmed, for the ongoing debate over the use of both race and the SAT as criteria for college admission.

Ruling Out Selection Effects

One possible explanation of the growing correlation between socioeconomic factors and SAT scores observed in the UC data is that it may reflect selection effects. That is, the sample of students who apply to UC may differ from the general population of Californians in some systematic but unknown way and thus bias the correlation. Selection effects may occur at several points in the admissions process. First is at the point of admission itself. When institutional admissions criteria are altered – UC introduced major changes in its admissions criteria between 1994 and 2011 -- the pool of admitted students may become increasingly different from the general population, thus creating selection bias.

Institutional selection effects of this kind cannot, however, account for the growing correlation between SAT scores and socioeconomic factors shown above, since those findings are based on the sample of all California applicants, whether or not admitted. In any case, the pattern of correlations among UC admits is almost identical to that for the UC applicant pool. For both UC admits and applicants, the proportion of variance in SAT scores explained by family income, parental education, and race/ethnicity increases sharply over time, and the pattern of beta weights is similar.^{6,7}

Still, even if institutional selection does not account for the observed patterns, it is possible that *self-selection* may be at work. That is, the composition of the pool who *apply* to UC may have changed, and the observed increase in the correlation between SAT scores and SES may reflect that change. To examine that possibility, it is necessary to compare changes in the UC applicant pool with changes in the broader pool of California SAT takers over the past two decades.

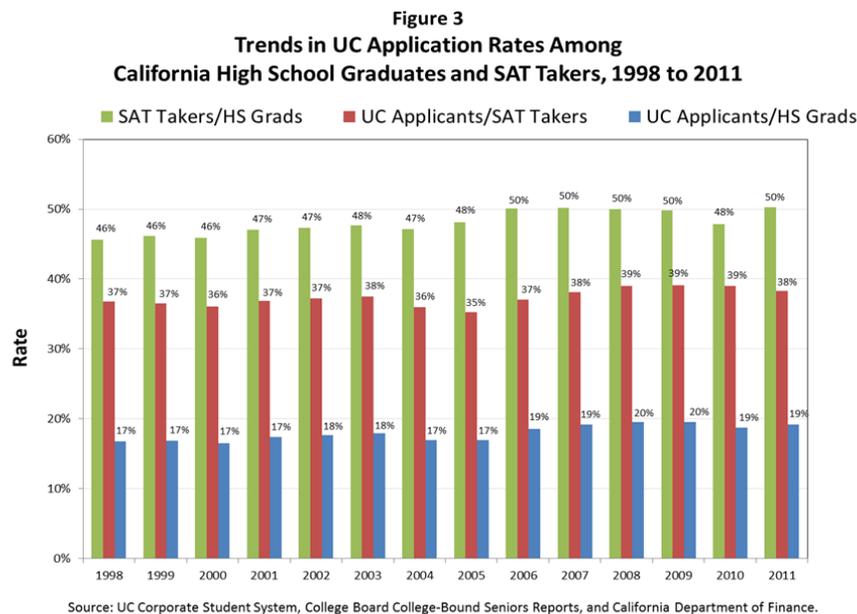
The only available historical data on California SAT takers are from the College Board’s *College-Bound Seniors Report*; the College Board has published annual summary data on SAT takers in each state since 1998. When students sit for the test, they are asked to fill out the SAT Questionnaire (formerly known as the Student Descriptive Questionnaire), which includes items on family income, parental education, and race/ethnicity. Although the quality of these data is uneven – due, among other problems, to the large proportion of students who decline to answer -- comparison of the statewide SAT data with the UC applicant pool reveals a very similar picture.

⁵ See Appendix 3. Due to very small sample sizes and attendant concerns about confidentiality of student records, the sample provided to the author by UC’s Office of the President does not allow separate identification of American Indian or Native American students. Thus, “underrepresented minorities” or “students of color,” as used in this paper, refer to black/African American and Chicano/Latino applicants.

⁶ See Appendix 4 for regression results for UC admits as against applicants.

⁷ For the technical reader, the data presented here were also checked carefully for both multicollinearity and interaction effects, two statistical problems that sometimes occur in regression analysis when predictor variables are related not only to the outcome variable of interest (in this case, SAT scores) but also to each other. Neither accounts for the results shown here. The one, non-trivial interaction effect observed in the UC data, between race and income, serves only to underscore the unique influence of race on SAT scores. See Appendix 5.

Because this conclusion forms the basis for much of the analysis that follows, it is important to examine the data in some detail. Figure 3 below provides an overview of changes since 1998 in the rate at which (1) California high school graduates take the SAT, (2) SAT takers apply to UC, and (3) high school graduates apply to UC. Although each of these rates has trended slowly upward during this period, the relationship among them has changed little. There is no indication that the proportion of SAT takers who apply to UC (red bars) has diverged markedly from the proportion of high school graduates who take the SAT (green bars) and apply to UC (blue bars).



Further evidence that selection effects are an unlikely explanation of the observed pattern of correlations in the UC data is shown in Figure 4 below, which compares mean and standard deviation SAT scores for UC applicants vs. all California SAT takers. For both groups, SAT scores have remained remarkably stable since 1998. To be sure, mean test scores for UC applicants are higher than for California SAT takers at large, as would be expected given UC's rigorous admissions standards. But the key point to note is that the relationship between the pools has changed little. Mean and standard deviation SAT scores for both UC applicants and California SAT takers in 2011 are virtually identical to 1998. There is no indication that the selectivity of the UC applicant pool has increased relative to the general population of California SAT takers during this period.

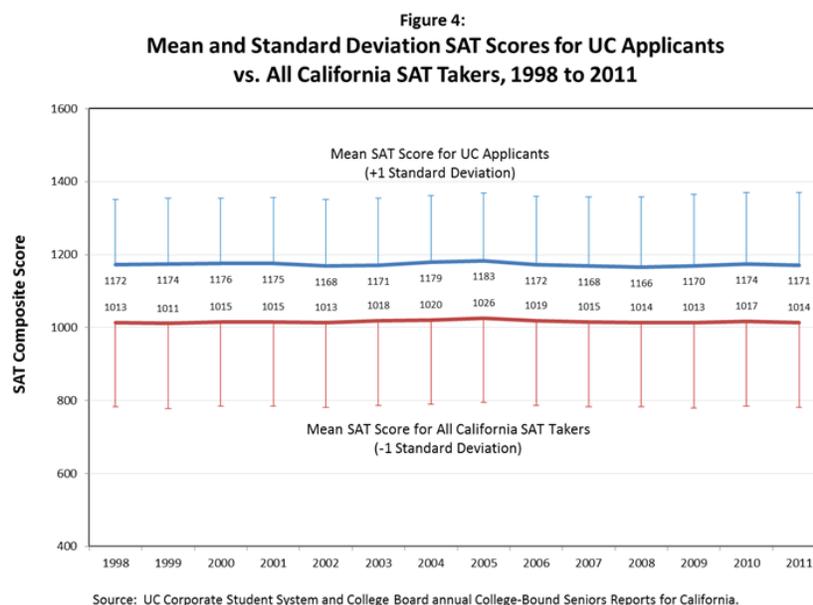
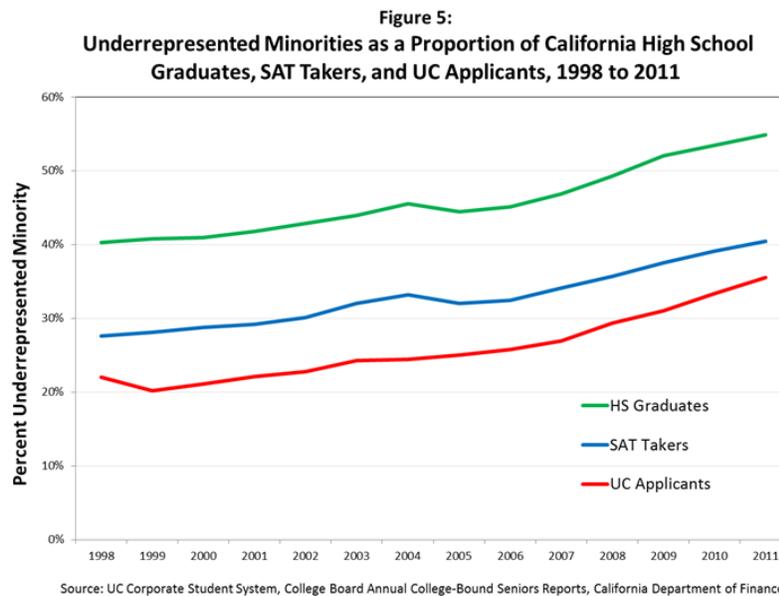


Figure 5 next examines the changing racial/ethnic composition of the UC applicant pool as compared with both California SAT takers and high school graduates over the same period. California's emergence as the first "majority-minority" state is reflected in the upward trend lines for each:



The gaps between the three lines are a stark reminder of the continuing educational disparities faced by students of color. The key point for purposes of the present discussion, however, is the consistent slope and relationship of the lines over time. Changes in the demographic composition of the UC applicant pool are entirely consistent with the changes that have occurred among both California SAT takers and high school graduates since 1998. Indeed, the gap between the proportion of underrepresented minorities in the UC applicant pool and their proportion among all California SAT takers actually narrowed after 2004, indicating that the two pools have grown more similar than different. Again, there is no indication that UC applicants have become less like the population of California SAT takers during the period covered by this analysis.⁸

Finally, with respect to family income and parental education, the available data also show no sign of selection effects. Unfortunately, however, the poor quality of the SAT Questionnaire data does not permit as definitive a comparison as one might like. Among other problems, the problem of missing data is most significant. As shown in Figure 6 (next page), the percentage of California SAT takers who declined to respond to SAT Questionnaire items on family income, parental education, or race/ethnicity grew sharply in the early 2000s, peaking in 2003. Non-response rates declined thereafter in the case of parental education and race/ethnicity. But non-response rates for income have continued to grow to the point where over half of all California SAT takers now decline to report family income on the SAT Questionnaire.^{9,10}

In addition to high non-response rates, the income data from the SAT Questionnaire exhibit several other problems that limit their reliability and usefulness. First, income is collected as a categorical rather than continuous variable; the student is asked to check one of eleven broad income categories, forcing researchers to use the midpoint of each category as an approximation of income. Second, research by the Educational Testing Service suggests that student responses to the SAT Questionnaire's income item are far from reliable; in one study, only a third of respondents accurately checked their true income category, as later verified by financial aid data (Freeberg, 1988). Finally, because the SAT Questionnaire's income categories are not

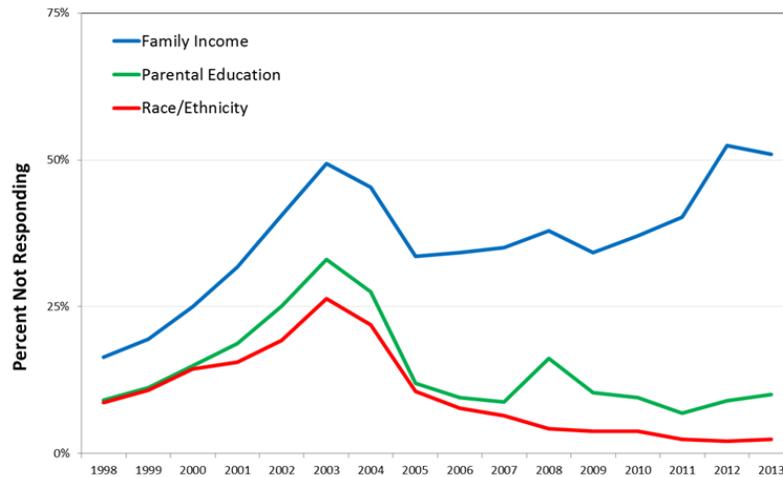
⁸ The main regression findings presented earlier were also examined to determine whether the growing weight of race/ethnicity as a determinant of SAT scores might be a linear function of the proportion of underrepresented minorities within the UC applicant pool. When entered into the main regression, the proportion of underrepresented minorities in each applicant cohort did add slightly to the overall explained variance but did not reduce the coefficient on race/ethnicity for individual applicants. See Appendix 6.

⁹ By comparison, 14 percent of UC applicants failed to report family income in 2011. Appendix 7 shows non-response rates for family income, parents' education, and race/ethnicity among UC applicants from 1994 to 2011. Cases with missing data were excluded from the analysis.

¹⁰ Researchers often find that non-response rates are greater among higher-scoring students, who are presumed to come disproportionately from upper-income families. However, the California SAT data show no consistent pattern in this regard: the proportion of non-respondents with higher SAT math and verbal scores rose significantly between 1998 and 2005, but fell sharply thereafter, and there is now little difference in mean test scores of respondents and non-respondents to the income item on the SAT Questionnaire; see Appendix 8.

adjusted for inflation, the proportion of students in the top categories eventually grew so large that the College Board was forced to revise the categories in 2008. The revisions have been helpful in reducing compaction at the top of the income distribution, but they introduce an obvious discontinuity when one attempts to compare historical income trends before and after 2008.

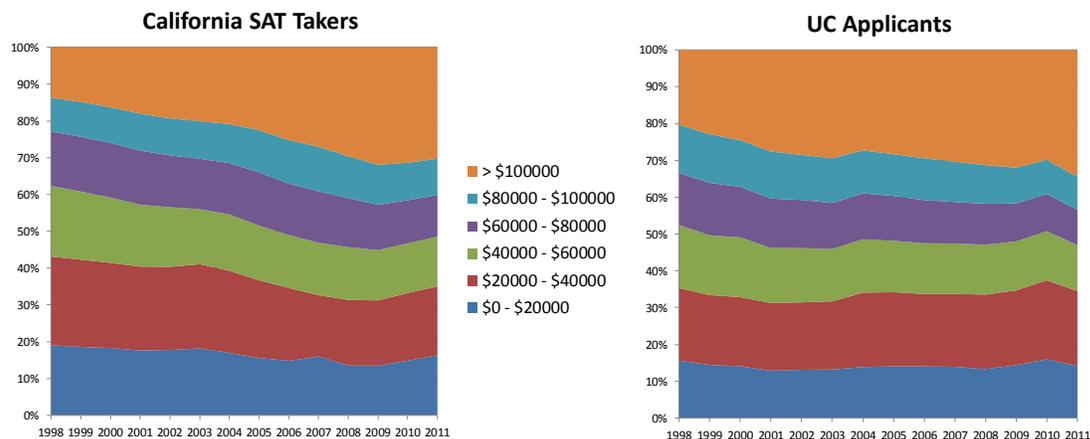
Figure 6
Percent of Non-Responses to SDQ Items on Family Income, Education, and Race/Ethnicity among California SAT Takers, 1998 to 2013



Source: College Board annual College-Bound Seniors Reports for California, 1998 to 2013.

For all of these reasons, the lack of reliable information on family income in the SAT Questionnaire prevents any definitive comparison of UC applicants and California SAT takers. The most that can be said is that, among California SAT takers who did respond to this item, the income distribution appears very similar to that of the UC applicant pool.¹¹

Figure 7
Distribution of Reported Family Income Among California SAT Takers vs. UC Applicants, 1998 to 2011



Source: UC Corporate Student System and College Board College-Bound Seniors Reports.

¹¹ In addition, the distribution of parental education among California SAT takers appears quite similar to that of the UC applicant pool, although exact comparison is not possible because the SAT Questionnaire categorizes parental education differently from the UC application; see Appendix 9.

To summarize: Comparison of the UC applicant pool with the pool of all California SAT takers provides no grounds to believe that selection effects account for the growing correlation between SAT scores socioeconomic background factors observed in the UC data. But neither does it definitively confirm the UC findings. The SAT Questionnaire data are inadequate to resolve the crucial questions: Are the UC data representative of a broader trend? Do socioeconomic background factors such as family income, parents' education, and race/ethnicity account for an increasing share of the variance in students' SAT scores? And does race have a large and growing *independent* effect after family income and education are taken into account?

A Call for Institutional Research

One might presume that the California findings could be readily confirmed or disconfirmed using national data on SAT takers. That presumption is mistaken, however, since the non-response rate for family income among national SAT takers is virtually identical to that among California SAT takers. In a recent year, 55 percent of all US high-school seniors who took the SAT declined to report their family income (College Board, 2012, Table 11). If one attempted to replicate the regression analyses presented here with either the national or California SAT Questionnaire data, the majority of cases would be eliminated.¹²

The best alternative source of data may be other state university systems that, like the University of California, receive applications from a sizeable proportion of SAT takers in their states. Because those institutions collect socioeconomic data on applicants for purposes of financial aid as well as admissions, the quality of the data they collect tends to be more complete and accurate than the SAT Questionnaire data. At UC, for example, information on family income is typically supplied by parents, and the data are subject to audit. Merging test-score data with admissions and financial aid information from institutional databases can thus provide a more reliable basis for examining trends over time in the relative effects of family income, parents' education, and race/ethnicity on test scores.

Accordingly, the author calls on institutional researchers in other state university systems to replicate and extend the above analyses with their own applicant data from the past 20 years. It is important that this work be conducted independently of the national testing agencies both because of the unreliability of the SAT Questionnaire data and also to avoid any potential conflict of interest, whether in perception or in fact.¹³ To that end, the National Association for College Admission Counseling's blue-ribbon commission on admissions testing has called for an independent national clearinghouse to coordinate research on admissions testing (NACAC, 2008). NACAC has offered its own services for this purpose, and the kind of institutional research proposed here seems ideally suited to take advantage of that offer.

Alternative Explanations of the "Test Score Gap"

If the California findings do reflect a broader trend, the question naturally arises: Why has the association between SAT scores and socioeconomic background factors, in particular race and ethnicity, apparently grown stronger over the past two decades?

The last 20 years have seen an extraordinary volume of research on the black-white "test score gap" beginning with Jencks and Phillips' *The Black-White Test Score Gap* in 1998 and Magnuson and Waldfogel's *Steady Gains and Stalled Progress* in 2008, among many others. Yet that research has produced no consensus about the underlying factors. Explanations that appeal to genetic differences cannot explain why the black-white gap narrowed significantly in the 1970s and 80s (Neal, 2006). Explanations that appeal to differences in income or SES cannot account for the large residual disparity that remains after controlling for those factors (Phillips, et al, 1998; Magnuson, Rosenbaum & Waldfogel, 2008). And explanations that emphasize the role of the schools founder on the fact that the test score gap appears before children enter school and persists even when black students attend the same schools as whites (Phillips, Crouse, & Ralph, 1998; Jencks and Phillips, 1998). "What explains these achievement gaps? It should come as no surprise that the issue has defied simple explanation" (Magnuson and Waldfogel, 2008, p.1).

No attempt will be made here to summarize this vast body of research, which is widely available elsewhere. With the benefit of hindsight on over two decades of work on this subject, instead the aim is to highlight possible explanations that seem most relevant to, and consistent with, the trends observed in the California data.

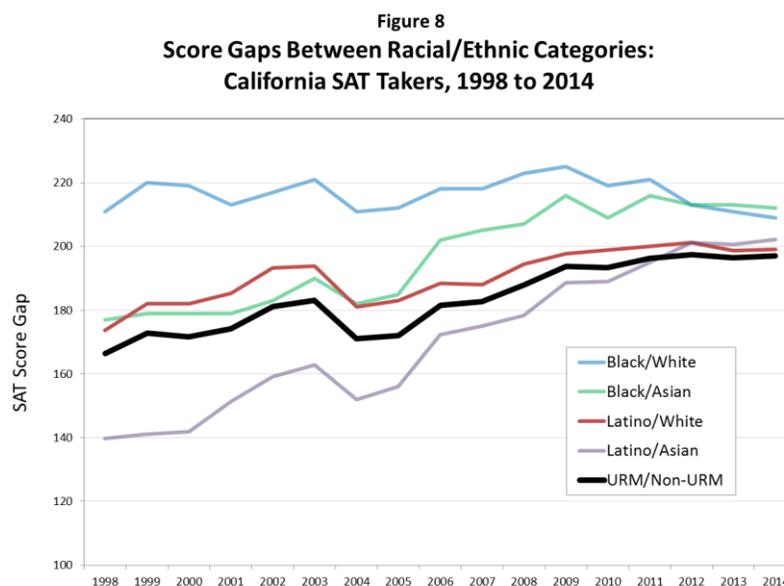
¹² While there are a number of methods for imputing missing data, such techniques are unlikely to be useful in the present case. Most imputation methods assume that missing values are "missing at random" (Little & Rubin, 2002), an assumption that is not tenable here. In addition, most imputation methods would require using other key variables of interest in the present study – e.g., SAT scores, high-school GPA, parent education – to derive missing income values, and so would compromise the regression results shown earlier in Figures 1 and 2, which estimate the relative weights of these same variables.

¹³ Not surprisingly, studies conducted by researchers with ties to the testing agencies have tended to downplay the effect of income and other socioeconomic factors on standardized test scores. See, for example, Sackett, et al. (2009), which relies on income data from the SAT Questionnaire. See also Zwick (2004).

Most research on the test score gap has focused on black-white differences and is thus not as directly relevant to the California experience as one might wish. The explosion of California's Chicano and Latino population together with continued growth among different Asian ethnic groups have overtaken that traditional racial dichotomy. Explanations or hypotheses based on black students' experience may not necessarily translate to Latino students. In addition, single-minded focus on trends on the black-white test score gap may miss the larger picture.

As shown in Figure 8, the test score gap in California is greatest between black and white SAT takers but has oscillated up and down and shows no consistent trend since 1998. If one were to draw inferences about racial and ethnic differences from the black-white gap alone, one might conclude that there has been little change in this respect.

But that conclusion would be wrong. For all other racial/ethnic comparisons, test score gaps between underrepresented minority and other students have been growing. The Black-Asian, Latino-White, and Latino-Asian test score gaps have increased almost every year since 1998. The glaring exception is 2003-04, when the gaps narrowed sharply in all of the comparisons (undoubtedly due to the spike in the number of students declining to state their race or ethnicity in that year; see Figure 6 earlier). Other than that one year, however, the overall gap between underrepresented minority and other students has steadily widened:¹⁴



Source: Author's calculation from College Board College-Bound Seniors Reports for California.

These trends suggest the need for caution in extrapolating from research on the black-white test score gap to the overall gap between students of color and others. The consensus among researchers is that the black-white gap narrowed substantially in the 1970s and early 1980s, plateaued in the late 1980s and 1990s, but resumed narrowing in the 2000s (Magnuson and Waldfogel, 2008; Magnuson, Rosenbaum, & Waldfogel, 2008).¹⁵ Indeed, some researchers argue that, as a result of growing economic inequality in American society, the black-white test score gap has become less significant than the income gap. Following the lead of Thomas Piketty, Reardon (2011) finds that the test score gap between students in the top and bottom income deciles has widened over the past 40 years to the point where it now exceeds the black-white gap. Again, however, all of these findings are limited to black-white comparisons.

In general, explanations of the test score gap can be divided into two types: Those that explain racial/ethnic gaps primarily by reference to general socioeconomic factors, such as differences in family wealth or income; and those that emphasize factors

¹⁴ Closer examination of test-score differences between specific ethnic groups in the UC applicant data indicates that most of the increase in the overall gap between underrepresented minority and other students reflects changes among Latinos. In a regression analysis predicting SAT scores conditional on family income, parents' education, and categorical variables for specific racial/ethnic groups, the beta weight for Latino applicants increased sharply from -0.14 (relative to white applicants) in 1997 to -0.22 in 2011. The beta weights for Asian applicants grew slightly from 0.04 to 0.06 and for black applicants from -0.14 to -0.16 (all relative to white applicants) over the same period.

¹⁵ Most research on the national black-white test score gap has relied on NAEP (National Assessment of Educational Progress) scores rather than SAT scores.

specifically associated with race, such as discrimination or segregation. The distinction is not perfect, nor are the two types of explanations mutually exclusive. But it is central to the debate over policies for ameliorating racial disparities in education, such as affirmative action.

A great deal of research on the black-white test score gap has favored the first type of explanation. Those studies have emphasized, among other influences, differences in family income (e.g., Magnuson & Duncan, 2006; Magnuson, Rosenbaum, & Waldfogel, 2008), parental education (e.g., Haveman & Wolfe, 1995; Cook & Evans, 2000), and quality of schools (e.g., Phillips, Crouse, & Ralph, 1998; Corcoran & Evans, 2008) as factors underlying black-white test score differences. As a group, black students are disproportionately affected by all of these factors. For example, black students have, on average, fewer resources in and out of the home, poorer health care, and less effective teachers, all which can have an impact on test scores (Haveman, et al., 2004; Meyers, et al., 2004; Phillips & Chin, 2004).¹⁶

Economic inequality, in particular, is most often cited in discussions of the black-white test score gap, and that explanation is plausibly consistent with the trends observed in the California data. The growing SAT-score gap between students of color and others parallels the marked increase in income inequality in California and the US over the last two decades (Congressional Budget Office, 2011).

The main difficulty with this explanation is that it does not account for the large and growing residual association between race and SAT scores observed in the California data after controlling for family income and parental education. Indeed, much the same problem is evident in extant research on the black-white test score gap: Though class differences account for a portion of the gap, they by no means eliminate it entirely. Conventional measures of socioeconomic status leave a large portion, even a majority of the gap, unexplained (Phillips, et al., 1998; Magnuson, Rosenbaum & Waldfogel, 2008).

Researchers who study the effects of economic inequality on the black-white test score gap argue, nevertheless, that this result may simply reflect deficiencies in available measures of SES. In place of family income, for example, wealth may be the more decisive factor. Measures of parental net worth not only provide a better indicator of current family resources but also capture family background effects that span more than one generation. Studies that have examined black-white wealth gap find that it is more pronounced than the income gap, and that it appears to have widened in recent decades (Oliver & Shapiro, 2006). Urging researchers to look more carefully at this and other family background factors, Wilson argues, “[M]easures of black-white family background differences are not only incomplete but probably underestimate the importance of such differences. ... [A]dditional measures of family environment might account for the remainder of the black-white gap” (1998, p. 505).

The problem, however, is that this explanatory framework knows no stopping point – there will always be additional variables that may be considered. While certainly plausible as an explanation of trends in the California test score data, the assumption that racial/ethnic differences are reducible, in principle, to general socioeconomic factors remains unproven.

A second type of explanatory framework looks to factors specifically associated with race and ethnicity, such as the experience of racial discrimination among students of color. The notion of “stereotype threat” among black students who take the SAT is one example of this type of explanation (Steele & Aronson, 1995). Another is the resurgence of racial segregation in schools and communities throughout California and the US.

Researchers who have studied the black-white test score gap have long noted the coincidence between trends in racial segregation and changes in the size of the gap. Following the US Supreme Court’s decision in *Brown v. Board of Education*, racial segregation in US schools decreased dramatically during the 1960s and 1970s. The black-white test score gap narrowed significantly during the same period. School desegregation stalled in the 1990s as the result of court decisions limiting busing and other integration measures. Progress in narrowing the test score gap stalled at the same time. The coincidence of these trends has provoked a great deal of research to determine if, and how, they may be causally linked (for a useful summary, see Vigdor & Ludwig, 2008).

Among the most rigorous efforts to determine whether a causal connection exists is Card and Rothstein’s (2007) influential study of SAT-score gaps in metropolitan areas across the US: “We find robust evidence that the black-white test score gap is higher in

¹⁶ Evans, Hout, & Mayer (2004) distinguish between two general kinds of effects that socioeconomic inequality may have on test score gaps. First are “mechanical” consequences; insofar as socioeconomic factors have a direct effect on individual student achievement, any increase in economic inequality, other factors held constant, will lead mechanically to an increase in the test score gap. Second are “externalities.” For example, growing economic inequality may affect state spending decisions for K-12 education or increase economic segregation within schools and neighborhoods, which may in turn affect student achievement.

more segregated cities” (p. 1). Their findings suggest that neighborhood segregation may have more causal importance than school segregation *per se*, although they were unable to account fully for the effects of the latter due to tracking within schools. Other researchers have found significant independent effects of both school and neighborhood segregation on black-white test score gaps (Vigdor, 2006; Rubiowitz & Rosenbaum, 2001; Guryan, 2004; Hanushek, et al., 2004; Hoxby, 2000).

One difficulty in this kind of research is differentiating the effects of school or neighborhood segregation from the effects of poverty, which often overlap. Even in the absence of segregation, substantial black-white test score differences can be expected for this reason. On the other hand, the frequent pattern of “double segregation” on the basis of both race and poverty (Orfield & Frankenberg, 2014) may produce an interaction effect that can accentuate such differences. As Rothstein (2014) has argued,

There are two aspects to this conclusion: First, social and economic disadvantage – not poverty itself, but a host of associated conditions – depresses student performance, and second, concentrating students with these disadvantages in racially and economically homogenous schools depresses it further (p. 1).

In short, racial segregation may amplify the effects of poverty and other kinds of social inequality on student achievement. The interaction of race and class may help explain why the effect of family income on SAT scores is non-linear and more than twice as large for black students (Dixon-Román, Everson & McArdle, 2013).¹⁷

If racial segregation does have an independent effect on student achievement, as much research on the black-white test score gap suggests, that explanation could account for observed trends in the California SAT data over the past two decades: why the test score gap has grown for Latino as well as black students, and why the gap remains after controlling for conventional measures of SES such as family income and parental education.

California’s schools have undergone a vast demographic transformation in the last 20 years, to an even greater extent than the nation’s schools. The most dramatic change has been among Latino students. Between 1993 and 2012, the number of Latinos enrolled in California schools increased by 68 percent, and their proportion increased from 25.5 to 52.7 percent, becoming a majority of public school enrollments. White and black enrollments declined in both absolute and relative terms, while the proportion of Asian enrollments held steady (Orfield & Ee, 2014, p. 28).

Rather than being dispersed throughout the state’s schools, however, the growing population of Chicano and Latino students has been concentrated in a relatively few, intensely segregated schools and school districts. Those schools also disproportionately enroll many black students, and a clear pattern of co-segregation of students of color has emerged (Orfield & Frankenberg, 2014). On a variety of measures, racial segregation among Latino and black students has increased sharply:

In 1993, more than two decades ago, about half of California’s schools were still majority white schools, and only one-seventh were intensely segregated (zero to 10 percent whites). Fewer than 5 percent were “apartheid” schools (99 percent to 100 percent students of color). By 2012, 71 percent of the state’s schools had a majority of students of color and fewer than 30 percent were majority white. The proportion of intensely segregated schools had doubled in just two decades, with one school in fourteen an apartheid school (Orfield & Ee, 2014, pp. 31-32).

Slightly more than half of the state’s Latino students now attend “intensely segregated” schools – 90 percent or more students of color -- and the comparable proportion for black students is 39 percent (Orfield & Ee, 2014, pp. 31-33). At the same time, the sharp rise in racial segregation has overlapped with rising poverty levels in those same schools:

In 1993, black and Latino students attended schools with 52 percent and 58 percent poor children, respectively, as measured by subsidized lunch eligibility (around 120 percent of the federal poverty level qualifies for free school lunch). By 2012, blacks on average attended schools whose populations were two-thirds poor children, and Latinos attended schools that were more than 70 percent poor (Orfield & Ee, 2014, p. 41).

This pattern of “double” or even “triple segregation” (when language differences are also considered) has created what has aptly been described as a “perfect storm” (Boger, 2003) of educational, social, and economic disadvantage in intensely segregated schools. That those multiple disadvantages may compound to exacerbate test score gaps would not be surprising and could account for the growing association between race and SAT scores observed in the UC data.

¹⁷ The UC applicant data also show a statistically significant, non-trivial interaction effect between underrepresented minority status and family income in more fully specified regression models predicting SAT scores; see Appendix 6.

In sum, both growing socioeconomic inequality and growing racial segregation offer plausible explanations of the California test score data, and it is likely that both factors are at work. The relative importance of these explanations still matters, however, since they have very different policy implications for the role of both race and SAT scores in college admissions, as discussed in the conclusion to this paper.

Adverse Impact of SAT Scores on Students of Color

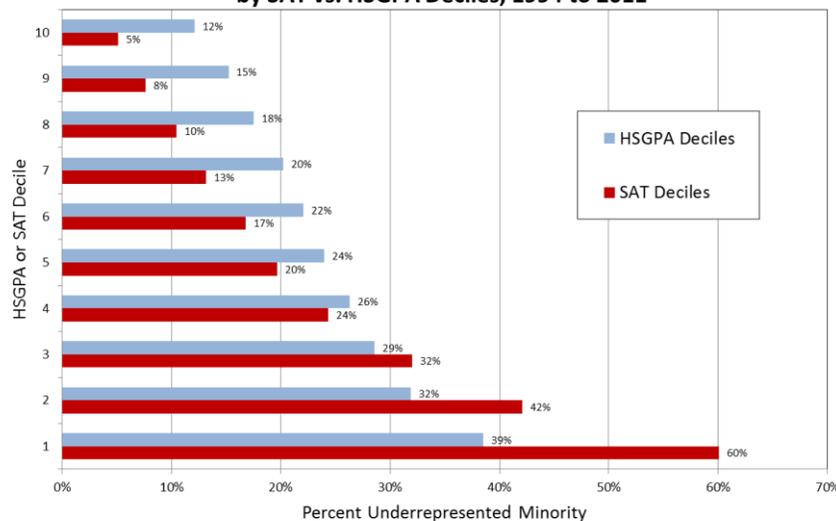
Before turning to those policy questions, however, it is important to consider two further aspects of the relationship between race and the SAT revealed in the UC data: (1) the adverse impact of SAT scores on admission of underrepresented minority applicants, and (2) the marginal validity of SAT scores in predicting college outcomes, especially for students of color.

Proponents of the SAT often emphasize that other measures of student achievement – high school grades, AP Exams, SAT Subject Tests, and other national assessments such as NAEP -- exhibit similar gaps between students of color and others. As a former president of the College Board famously remarked, “Getting rid of the SAT or any other standard is not going to change that fundamental fact. Instead of smashing the thermometer, why not address the conditions that are causing the fever?” (Stewart, 1998). Researchers with ties to the testing agencies emphasize the same point, if in less colorful language:

Results indicate that the score gaps that are observed in admission tests among members of different racial and ethnic groups and different socioeconomic groups are also found in other standardized tests and in other measures of educational attainment. It is hypothesized that these differences are a powerful illustration of an important social problem: inequitable access to high quality education (Camara & Schmidt, 1999, p.1). . . . The indisputable fact is that both high school grades and scores on admissions tests are reflections of the same education system, with all its flaws and inequities (Zwick, 1999, p. 323).

The problem with statements such as these is that they obscure a fundamental point: The *magnitude* of the gap between students of color and others is much greater for SAT scores than for admissions criteria such as high school grades. As a result, the SAT has a more adverse impact on Latino and black applicants’ chances of admission. Figure 9 illustrates the point. The figure is again based on the sample of 1.1 million California residents who applied for freshman admission from 1994 to 2011. In this case, applicants have been divided into ten equal groups, or deciles, based on their test scores. The red bars in the figure represent the percentage of applicants within each SAT decile who are Latino or black. The blue bars show the percentage of Latino and black applicants when the same students are ranked by high school GPA.

**Figure 9
Distribution of Underrepresented Minority Applicants
by SAT vs. HSGPA Deciles, 1994 to 2011**



Source: UC Corporate Student System data on CA resident freshman applicants, 1994 to 2011.

The difference is stark. Although SAT scores and high school GPA both have an adverse impact on students of color, the demographic footprint of the SAT is far more extreme. At the bottom of the applicant pool, Latino and black applicants comprise 60 percent of the lowest SAT decile but only 39 percent of the lowest HSGPA decile. Conversely, within the top decile – those

most likely to be admitted – Latinos and blacks comprise 12 percent of applicants when ranked by high school grades but just 5 percent when ranked by SAT scores. Using SAT scores to rank applicants produces more severe racial/ethnic stratification than high school GPA. The relative emphasis that colleges place on the two criteria can thus make a substantial difference in the demographic composition of the admitted pool.^{18,19}

Validity of SAT Scores in Predicting College Outcomes

A last set of findings from California concerns the predictive validity of the SAT. Predicting college success is the *raison d'être* for college admissions tests such as the SAT. As the UC data demonstrate, however, SAT scores are a relatively poor indicator of student success in college. This is especially true for students of color. While SAT scores have a strong adverse impact on admission of Latino and black students, they are relatively weak predictors of how those students actually perform in college.

The College Board conducts annual validity studies to provide empirical justification for use of the SAT as an admissions criterion. Those studies ordinarily focus on first-year college grades as an outcome measure, since that measure of college performance is most readily available. Typically, validity studies examine the multiple correlation among high school GPA, SAT scores, and first-year college grades for a given sample of college freshmen. They invariably find that combining SAT scores with high school GPA provides better prediction than either factor alone. While high school GPA is usually the best single predictor in most samples, inclusion of SAT scores in the correlation adds a statistically significant, if modest, increment to the prediction (see, for example, Patterson & Mattern, 2012). Such studies constitute the principal justification of the SAT's validity for use in college admissions under established national testing standards (AERA/APA/NCME, 2014).

The California data paint a rather different picture. Extending back over two decades, the UC database provides a longer-term and more significant measure of student performance than first-year grades: college completion. It also permits examination of how socioeconomic factors may affect the prediction. Because family background is correlated with *both* SAT scores and college outcomes, much of the apparent predictive power of the SAT actually reflects the proxy effects of socioeconomic status. Berkeley economist Jesse Rothstein (2004) conservatively estimates that traditional validity studies that omit socioeconomic variables overstate the predictive power of the SAT by 150 percent.

Figure 10 (next page) shows the relative weights of high school GPA and SAT scores in predicting 5-year graduation rates at UC. The sample includes all California residents who enrolled as UC freshmen from 1994 through 2005.²⁰ The coefficients are standardized and show the effect of a one standard-deviation change in each measure on the probability of graduating from UC, holding the other measure constant, thus permitting direct comparison of their relative weights. The weights are given both before and after controlling for SES (family income and parents' education), and for all entering freshmen vs. underrepresented minority freshmen.

Looking first at the weights for all UC freshmen, it is evident that, of the two admissions criteria, high school GPA is by far the stronger predictor of college completion. This finding replicates the results of an earlier, more comprehensive study of graduation rates at UC (Geiser & Santelices, 2007), subsequently confirmed in *Crossing the Finish Line*, the largest national study of college completion conducted to date. Based on a massive sample of entering freshmen at 21 state flagship universities and four state higher education systems, Bowen and his colleagues found that,

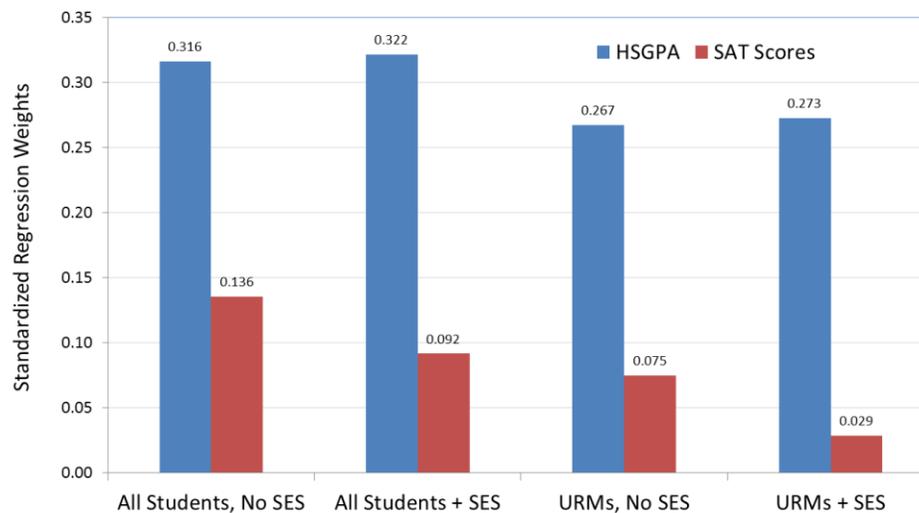
High school grades are a far better predictor of both four-year and six-year graduation rates than are SAT/ACT test scores ... The consistency of the results is extraordinary. *In all but one of these more than 50 public universities, high school GPA remains a highly significant predictor of six-year graduation rates after taking account of the effects of test scores.* . . . Test scores, on the other hand, routinely fail to pass standard tests of statistical significance when included with high school GPA in regressions predicting graduation rates . . . (Bowen, Chingos, and McPherson, 2009, pp. 113-115; italics in original).

¹⁸ The demographic footprint of the SAT has worsened relative to HSGPA during the period covered by this study, although the trend is partially obscured by the overall increase in the percentage of Latino and black applicants to UC, which rose from 21% to 34% during this time. In the bottom SAT quintile, the proportion of students of color increased from 43% in 1994 to 70% in 2011, as compared with an increase from 38% to 52% within the bottom HSGPA quintile. Conversely, at the top of the applicant pool, the proportion of students of color in the top SAT quintile rose from 5% in 1994 to 9% in 2011, while their proportion in the top HSGPA quintile grew by a much larger amount, from 9% to 19%.

¹⁹ SAT scores also correlate much more closely than HSGPA with family income and parents' education. In the overall UC sample, the correlation of SAT scores with family income was .36, and with parents' education, .45. For HSGPA, the correlations were .11 and .14, respectively. See Appendix 1.

²⁰ The modal time-to-degree for students who graduate from UC is about four years and a quarter. The analysis was limited to freshman cohorts entering UC before 2006 in order to allow sufficient lead time for modal time-to-degree. See Appendix 10 for full regression results.

Figure 10
Relative Weight of High School GPA and SAT Scores,
Before and After Controlling for SES, in Predicting 5-Year Graduation:
All UC Freshmen vs. Underrepresented Minorities, 1994 to 2005



Source: UC Corporate Student System data, 1994 to 2005. All estimates are statistically significant at .001 confidence level.

Moreover, when controls for family income and parental education are introduced, as shown in the second set of bars above, the limited predictive power of SAT scores becomes even more evident. While the weight for HSGPA remains steady or increases slightly, the weight for SAT scores declines by approximately a third. This result strongly suggests the proxy effects of socioeconomic background factors.

Finally, turning to the results for underrepresented minority students, the predictive validity of SAT scores is reduced still further, especially after controls are introduced for family income and parents' education. While SAT scores remain a "statistically significant" predictor of UC graduation for students of color, the effect size is very small indeed.²¹

Taken together with the previous findings on the adverse impact of SAT scores on admission of students of color, the predictive-validity findings shown here pose a fundamental question about costs and benefits: Is the adverse racial impact of the SAT justified in view of the relatively limited benefit it provides in predicting student performance in college? The concluding sections of the paper discuss the implications of the California findings for this and other policy issues.

Policy Implication #1: On "Workable Alternatives" to Race-Conscious Affirmative Action

Regents of the University of California v. Bakke decisively altered not only the legal rationale but the broader public debate over consideration of race in college admissions. Up to that time, the primary rationale for race-conscious admissions policies was remedial: overcoming racial disparities created by decades of segregation and other exclusionary laws. After *Bakke*, the policy narrative pivoted to emphasize the educational benefits of racial diversity. For better or worse, the diversity rationale has become the primary framework within which race-conscious admissions policies are now debated and litigated (Garces, 2015).

The California findings have important implications for several key aspects of the diversity rationale. The first concerns the notion of "workable alternatives" to race. In a series of decisions following *Bakke – Gratz v. Bollinger* (2003), *Grutter v. Bollinger* (2003), and *Fisher vs. University of Texas* (2013) -- the US Supreme Court has increasingly restricted the use of race-conscious admissions policies by applying a legal standard known as "strict scrutiny." That standard involves a two-pronged test requiring that, before taking race into account, college and universities first demonstrate a "compelling interest" in enrolling a diverse student body and second, that the admissions policies they adopt are "narrowly tailored" for this purpose.

One of the main tests of whether race-conscious admissions policies are narrowly tailored, in turn, is that colleges and universities have made good-faith efforts to explore workable, race-neutral alternatives to achieve diversity by other means. For example, some argue that "class-based affirmative action" – giving preference in university admissions to low-income or first-

²¹ Technical note: The difference in the validity coefficients for HSGPA and SAT scores is not attributable to range restriction; see Appendix 11.

generation college students – can yield a substantial improvement in racial diversity at the same time. A recent volume published by the Lumina and Century Foundations, *The Future of Affirmative Action*, proposes a number of different “proxies” for race, even to the use of zip codes, as a means of expanding admission of students of color through facially race-neutral means (Kahlenberg, 2014). Others are skeptical, noting that Latinos and African Americans represent a relatively small proportion of all low-income or first-generation college applicants. Simulations using national SAT data suggest that admissions policies that rely only on race-neutral measures such as class-based affirmative action or “percent plans” (i.e., admitting a fixed percentage of students from every high school), while helpful, cannot produce the same level of diversity as policies that give explicit consideration to race (Carnevale, Rose, & Strohl, 2014).

The California findings weigh decisively on the side of skepticism and demonstrate why class-based affirmative action is unlikely to provide a workable alternative to race-conscious admissions policies: After controlling for family income and parents’ education, race uniquely accounts for a substantial share of the variance in SAT scores.

The California data also have important implications for another element of the narrow-tailoring test, namely, that race-conscious admissions policies must be temporary and of limited duration. In her majority opinion in *Grutter v. Bollinger* (2003), then-Justice Sandra Day O’Connor famously declared, “We expect that 25 years from now, the use of racial preferences will no longer be necessary to further the interest approved today” (p. 31). Her opinion reflects a widely held view that racial and ethnic differences are receding in importance in contemporary society.

Yet the UC data show that, at least in California, race not only remains an important factor in explaining test score gaps but has grown in importance over the past two decades. Rather than declining in salience, race has now become more influential than either family income or parental education in accounting for test score differences. The California findings should give pause to those who assume that racial and ethnic disparities in educational opportunity will inevitably narrow over time.

Ultimately, the debate over workable alternatives to race-conscious affirmative action may turn on the question of causation: What accounts for the growing test score gap between students of color and others observed in the California data? Advocates of class-based affirmative action will point to growing socioeconomic inequality in California and the US as the likely culprit (Kahlenberg, 2014; Reardon, 2011). If true, this implies that ameliorating general disparities in wealth and income offers the best hope for closing the gap. But the resurgence of racial segregation in California’s schools suggests another plausible explanation. If true, that explanation implies that renewed efforts to integrate schools and neighborhoods are essential to remedy test score gaps among students of color.

Which, if either, of these two explanations is most dispositive is thus a matter of some consequence beyond the ivory tower. The conservative judicial trend limiting racial classification in college admissions has had significant repercussions for school integration efforts affecting a much larger number of students at the K-12 level.²² If racial segregation does account for a substantial part of the test score gap, beyond what can be explained by income or class differences alone, an overbroad Supreme Court decision ending race-conscious policies at elite colleges and universities could exacerbate underlying disparities in the nation’s schools.^{23,24}

Policy Implication #2: On the Disparate Impact of the SAT

At first glance, the California findings might seem to suggest a traditional remedy for racial disparities in college admissions – pursuing a claim of “disparate impact” against the SAT. The UC data provide two factual elements necessary to initiate a

²² “The comparison of affirmative action to school desegregation might strike some as odd. Most scholars considering both have limited their discussion to whether affirmative action is consistent or inconsistent with the *Brown v. Board of Education I* prohibition of assigning students according to their race (i.e., whether *Brown I* reflects an anticlassification or antisubordination principle). There is more to be said, however, because affirmative action and school desegregation cases are fundamentally quite similar. Both are Equal Protection Clause challenges to the racial activities of public schools that can result in federal court participation in local and state educational policy” (Parker, 2004, p. 1694).

²³ This paper is being written just as the US Supreme Court has agreed to rehear *Fisher v. University of Texas*, which will be argued in the 2015-16 term.

²⁴ Again, it is hoped that this paper may stimulate institutional researchers in other state public higher education systems to replicate and extend the analyses presented here. Researchers at those institutions will have access to all of the various data sources required: (1) longitudinal, student-level data for applicant cohorts going back to the early 1990s (when most institutions began keeping electronic records), (2) SAT or ACT scores for all applicants (who typically account for a sizeable fraction of test takers in each state), and (3) independent data on family income and parental education from admission and financial aid applications. In addition, for in-state students, most researchers will have access to (4) data on applicants’ high school of origin (including the racial/ethnic composition of those schools). Merging these various data sources will permit confirmation of whether the racial and SES trends in SAT scores observed in California are present in other states and also facilitate analysis of alternative explanations of such trends, if confirmed.

discrimination claim: (1) SAT scores have a demonstrable adverse statistical effect on admission of students of color, while at the same time, (2) they are relatively weak predictors of student outcomes such as college graduation, especially for Latino and black students. Together, these findings suggest that colleges' and universities' reliance on the SAT as an admissions tool unjustifiably disadvantages students of color.

Yet the likelihood of a successful challenge to the SAT is small, for a number of legal, political, and practical reasons outlined below. And even if such a challenge were successful, it is not clear that eliminating consideration of SAT or ACT scores in college admissions would have as great a remedial effect as preserving race-conscious affirmative action.

A first obstacle to mounting a disparate-impact claim against the SAT or ACT is that, because the College Board, ACT, and other test producers do not receive federal funding, they are not subject to such claims. Suing colleges and universities that *employ* the tests is the only available judicial remedy.²⁵ The potential liability of colleges and universities to disparate-impact claims may be one reason why the diversity rationale for affirmative action has proven more attractive to those institutions than the older, remedial rationale. Emphasizing the educational benefits of a diverse student body does not expose them to the liability they might face by acknowledging the adverse impact of the SAT on students of color.

Case law on disparate impact is unusually convoluted and, like that on affirmative action, has become increasingly constrained by the conservative trend of recent court decisions. Unless a college or university has a history of racial segregation, it is nearly impossible to challenge reliance on the SAT on an Equal Protection claim under the 14th Amendment of the U.S. Constitution, since courts have held that such a claim requires a showing that the institution was motivated by discriminatory intent. A showing that standardized tests have a discriminatory *effect* – as opposed to intent – is unlawful only if made so by statutes or regulations.

Of those statutes and regulations, Title VI of the 1964 Civil Right Act is most prominent. Although Title VI itself prohibits only intentional discrimination, federal agency regulations adopted to implement Title VI prohibit colleges and universities that receive federal funds from engaging in practices that have the “effect” of discriminating on the basis of race. Until 2001, the courts allowed plaintiffs to sue for enforcement of those regulations against colleges and universities, including loss of federal funding, based on a showing of disparate impact.

In its decision in *Alexander v. Sandoval* (2001), however, the U.S. Supreme Court ruled that there is no private right of action to bring a disparate impact suit to enforce Title VI regulations. Thus, the only remaining remedy for those who would challenge the adverse effects of standardized tests on students of color is to seek administrative enforcement through the U.S. Department of Education's Office for Civil Rights. Private individuals and advocacy organizations can file administrative complaints which may trigger OCR “compliance reviews” of colleges and universities when there is sufficient evidence that admissions criteria have disparate impact.

Even if a disparate-impact claim against the SAT were to proceed this far, moreover, there is no guarantee that it would succeed. Assessing a disparate-impact claim involves a three-step process, with a shifting burden of proof between the complainant and the institution employing the test in question. In the first step, the burden is on complainant to show that the test produces a significant racial disparity. The California data suggest strongly that this burden could be met. In the second step, the burden then shifts to the institution to show that the test is “educationally necessary,” that is, that it serves a “substantial legitimate” educational purpose. Then, in the third step, the burden shifts back to the complainant to demonstrate that there is an alternative test or practice that could serve the same educational purpose with less adverse impact on students of color.

It is in the latter steps that difficulties are likely to arise. In the first place, the term “educational necessity” is a misnomer because it suggests a stronger standard than case law actually requires. Rather than being essential or indispensable, the educational-necessity test requires only that the challenged practice “serves in some significant way” or “bears a manifest relationship” to the university's educational purpose (*Wards Cove*, 1989; University of California, 2008). While SAT scores are a relatively weak predictor of student outcomes such as first-year college grades, they do add a small predictive increment beyond that which is provided by high-school grades alone. To what extent this small increment is of practical significance is arguable (Crouse and Trusheim, 1988), but there is little question that the SAT meets the standard of statistical significance employed in most predictive-validity analyses, including those conducted at UC. Given the stakes involved, it is likely that the national testing agencies would mobilize all of their considerable analytical and political resources in support of the claim that test scores are valid indicators of student performance in college.

²⁵ The legal analysis here draws heavily on a paper by Kidder and Rosner (2002).

Second is the question of alternatives. While eliminating reliance on tests or going “test optional” is feasible at smaller, less selective colleges, it may be less feasible at large, highly selective universities that receive tens of thousands of applications. Admissions officers at those institutions often regard the SAT and ACT as indispensable tools for comparing applicants. Other than high school grades, test scores are often the only other metric that is available on all applicants, and admissions officers typically consider SAT or ACT scores in conjunction with high-school GPA in order to take account of variations in grading standards across different high schools.

To be sure, there are other *kinds* of tests that might be used for admissions purposes, but those tests tend to have similar impacts on students of color. On purely educational grounds, a strong case can be made for replacing norm-referenced tests such as the SAT or ACT with curriculum-based, subject-specific assessments such as the SAT II Subject Tests or Advanced Placement exams. Proponents of such assessments argue that they align more closely with materials that students study in the classroom, minimize the need for test prep, and better reinforce teaching and learning of a rigorous academic curriculum in the schools (Atkinson & Geiser, 2009). Yet such assessments do not provide any direct or immediate advantage to minority examinees, and score gaps are very similar to those found on the SAT or ACT. Purely with respect to disparate impact, there is little to choose between the two kinds of tests (Geiser & Studley, 2002; Kobrin, et al., 2006).²⁶

Thus, despite their adverse impact on admission of students of color, there is a strong probability that tests such as the SAT or ACT would survive a disparate-impact challenge, both in view of the statistically significant, if modest, correlation between test scores and first-year college grades and also because there are no readily available alternatives to the major national tests that have less adverse impact.²⁷

Finally, even if a disparate-action challenge were successful, its remedial impact would be limited the particular college or university against which the complaint was brought. And such an action would result only in elimination of the challenged test, leaving in place other admissions criteria that also have questionable impacts on students of color. By itself, a disparate-impact challenge to the SAT would likely have far less remedial impact than institutional policies that give explicit consideration to race in admissions decisions.

Policy Implication #3: Reclaiming the Remedial Rationale for Affirmative Action

Rather than eliminate the SAT, the more important lesson to draw from the California findings is the continuing relevance of the original, remedial rationale for affirmative action. While the courts may have rejected *historical* discrimination as a justification for race-conscious admissions policies, the UC data illustrate how SAT scores and other selection criteria pose an unwarranted barrier to admission of students of color in the present day, beyond what can be justified by their limited capacity to predict success in college. The SAT is not unique but only a more extreme example of the problem.

The notion that the adverse impact of standardized tests may provide a remedial rationale for race-conscious affirmative action is not new. In the U.S. Supreme Court’s decision in *Bakke*, Justice Powell opined that compensating for bias in standardized testing could conceivably justify race-sensitive admissions (*Regents of UC v. Bakke*, 1978, p. 306, note 50). More recently, the U.S. Sixth Circuit’s opinion in *Grutter* cited the strong correlation between race and LSAT scores as one factor in its decision to uphold the University of Michigan Law School’s affirmative action policy:

Diversity in education, at its base, is the desegregation of a historically segregated population and . . . *Bakke* and *Brown* must therefore be read together so as to allow a school to consider race or ethnicity in its admissions for many reasons, including to remedy past discrimination or present racial bias in the educational system. . . . Criticism of the use of numerical credentials such as LSAT scores is made in this opinion to support the law school’s use of other criteria in its admission policy - one of which is race or ethnicity (*Grutter v. Bollinger*, 2002).

The SAT is not, however, the only admissions criterion whose adverse racial impact is disproportionate to its measurement validity. Virtually all other academic criteria now employed in college admissions raise similar concerns, if to a lesser degree. The

²⁶ UC studies show that the predictive validity of curriculum-based tests such as the SAT II Subject Tests and AP exams is equal to or slightly better than the SAT or ACT, although the incremental gain in prediction is too small by itself to dictate adoption of one kind of test over the other (Geiser & Studley, 2002). The case for curriculum-based tests rests primarily on their educational value rather than their predictive validity or differential impact, both of which are similar to the SAT’s and ACT’s.

²⁷ Some testing experts have proposed a number of promising alternatives to the SAT or ACT aimed at assessment of “non-cognitive” factors such as achievement motivation and academic self-efficacy (Robbins, et al., 2004; Schmitt, et al., 2009; Sternberg, et al., 2012). However, such assessments raise issues of cost, scalability, and possible “gaming” of the tests by applicants and, in any case, are not yet widely available as alternatives to the established national admissions tests.

significance of the SAT is that it illustrates so clearly the costs and benefits involved. Yet eliminating the SAT, by itself, would do little to remedy the many other ways, large and small, that students of color are unjustifiably disadvantaged in contemporary college admissions.

“Strength of curriculum,” for example, is one of the top three criteria now employed in college admissions. Along with test scores and high school grades, that criterion is rated as of “considerable importance” by a substantial majority of admissions professionals at US colleges and universities (National Association for College Admission Counseling, 2015, pp. 27-35).

Strength of curriculum is most often measured by the number of Advanced Placement classes on an applicant’s transcript. Admissions officers typically interpret this as an indicator of an applicant’s willingness to challenge him or herself by taking more difficult, advanced-level coursework.

The problem is that AP classes are not equally available to all students, despite efforts to expand AP offerings in low-performing schools (Klopfenstein & Thomas, 2010). Even where they attend high schools with robust AP programs, minority students are often tracked out of such classes (Oakes, 1985; CSU Institute for Education Reform, 1999). For these reasons, employing AP as admissions criterion has an adverse impact on students of color (Oakes, 1990; Oakes, Gamoran, & Page, 1992).

A key question, therefore, is whether the adverse racial impact of this admissions criterion is justified by its utility in predicting student success in college. The California data indicate that the answer is no: The number of AP classes on a student’s transcript bears almost no relationship to how he or she performs at UC (Geiser & Santelices, 2004). While AP *exam scores* do correlate with college performance, merely taking AP classes, without passing the AP exams, is not a valid indicator of how students will perform in college. The key is not simply taking the classes but mastering the subject matter, as subsequent research has now confirmed (Dougherty, Mellor, & Jian, 2006; Klopfenstein & Thomas, 2009).

The same problem affects high school GPA itself. A survey conducted by the National Association for College Admission Counseling found that about 70 percent of all US high schools award “bonus points” for AP classes in calculating students’ GPAs (NACAC, 2004). For example, a grade of “B”, which would normally be worth 3 grade points, is counted as 4 points for an AP class. Bonus points are awarded just for taking the classes, whether or not students pass the AP exams. This creates an incentive for students to load their schedules with AP classes in order to raise their GPA and improve their college-admissions profile, but then skip the exams. It has also contributed mightily to grade inflation; the *average* high school GPA among applicants to UC Berkeley, for example, is now 4.2.

Few would advocate eliminating high school grades as an admissions criterion, nor is that the argument here. High school GPA remains the single strongest predictor of student performance in college. Yet the manner in which HSGPA is typically calculated – the so-called “weighted GPA,” with bonus points awarded for AP classes – unnecessarily disadvantages students of color while adding little, if any, improvement in predictive validity. In fact, *unweighted* grades are superior to weighted GPA in predicting student performance at UC (Geiser & Santelices, 2007).

The SAT is thus not alone in its cost-benefit calculus, but only a more extreme example. SAT scores have a far stronger racial imprint than high school grades but much weaker validity as an indicator of how students will perform in college. Nevertheless, under the existing legal standard for demonstrating educational necessity, SAT scores do exhibit a statistically significant, if modest, correlation with first-year college grades and are therefore likely to meet that standard.

It has been said that intelligence is the ability to hold simultaneously in mind two contradictory thoughts. The case for race-sensitive, remedial admissions policies follows from this contradiction: Though the SAT may meet the legal test of “educational necessity,” it has a substantial adverse effect on students of color beyond what can be reasonably justified by its predictive validity. Consideration of race is therefore necessary to compensate for that effect.

This conclusion is underscored by the standards for “fairness” in testing published jointly by the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. Standard 3.16 of the recently reissued *Standards for Educational and Psychological Testing* states,

When credible research indicates that test scores for some relevant subgroups are differentially affected by construct-irrelevant characteristics of the test or of the examinees, when legally permissible, test users should use the test only for those subgroups for which there is sufficient evidence of validity to support score interpretations for the intended uses (AERA/APA/NCME, 2014, p. 70).

Given the marginal validity of SAT scores in predicting college success, particularly for students of color,²⁸ consideration of race is essential to ensure fairness in college admissions.

Policy Implication #4: On Race, the SAT, and “Individualized Review”

A final issue concerns the role of race and the SAT in *individual* admissions decisions. Abstract notions of “academic merit,” “diversity,” or “college readiness” aside, college admissions ultimately comes down to review, comparison, and selection (or denial) of individual applicants. Collectively, of course, such decisions do produce group patterns, and colleges and universities often strive to expand access for groups that are underserved. But in the final analysis, admissions criteria must be perceived to be fair and make sense as applied to the individual applicant.

The Supreme Court has affirmed the importance of individualized review as part of its “narrow tailoring” doctrine in a series of affirmative action cases. In *Bakke*, the Court rejected mechanistic set-asides or quotas for minority applicants, but endorsed admissions policies that involve a flexible, individualized consideration of applicants so that race, while important, is only one of a number of factors considered. Subsequently in *Gratz* (2003), the Court rejected a point system which, though it considered race as only one of several factors, awarded undergraduate applicants to the University of Michigan a fixed number of points for underrepresented minority status. In its companion decision in *Grutter* (2003), however, the Court simultaneously affirmed UM’s admissions policy for graduate students, which also considered race as one of several factors but gave no fixed weight to minority status. A key difference between the two policies, in the Court’s opinion, was the individualized attention to race.

Whether or not one agrees with the overall direction of recent case law on affirmative action, it should be evident that a student’s race or ethnicity is an imperfect indicator for admissions purposes. Under a remedial rationale, for example, the assumption that underrepresented minority status is invariably an indicator of social disadvantage, regardless of other background circumstances in an applicant’s file, is sometimes unwarranted in individual cases. The assumption is usually but not always justified, as non-minority applicants from lower SES backgrounds may have faced similar challenges. Likewise under a diversity rationale, minority status alone, absent other social and cultural differences, is not invariably an indicator that an applicant will contribute more to the richness of the collective undergraduate experience than a non-minority candidate. In specific cases, determination of the extent to which race should count as a “plus factor” requires individualized review of the totality of information in an applicant’s file.

Yet the principle of individualized review is a double-edged sword. Like race, SAT scores are also an imperfect indicator of, in this case, academic merit or ability. Following the same logic, a strong argument can be made that mechanical use of the SAT to compare applicants is unjustified, and that SAT scores must also be subject to individualized review by admissions officers.

The UC data illustrate the extreme limitations of SAT scores as an admissions criterion when used to compare or rank individual applicants. Because SAT scores account for such a small fraction of the variance in student outcomes, the error bands around test-based predictions of college performance are quite large. Though test scores can be useful in predicting *average* performance for large groups of students, they are much less useful or accurate in individual cases.

The following example, based on the UC data, will illustrate the point. Consider two applicants who are identical in all other respects -- same high-school GPA, socioeconomic background, and so forth -- except for their SAT scores. The first student, with an SAT score of 1200, is predicted to earn a GPA of 3.0, or a B average, at UC. The second student, with an SAT score of 1300, is predicted to earn a GPA of 3.13. The choice may seem obvious. Test scores are sometimes used as a tie-breaker to choose between applicants who are otherwise equally qualified, and in this case the tie would go to the second student.

What this ignores, however, are the error bands around predicted performance. The error band around predicted GPA for the UC sample was plus or minus .81 grade points at the 95 percent confidence level, the standard normally employed in social-science research. The error bands, in other words, were over six times larger than the difference in outcomes -- .13 of a grade point -- predicted by the two applicants’ SAT scores. For both students, the most that can be said is that actual performance at UC is likely to fall somewhere in a broad range between an A- and a C+ average.

²⁸ See Figure 10 and accompanying text. Predictive-validity studies published by the College Board often emphasize that SAT scores tend to “overpredict” college grades for underrepresented minority students (see, for example, Mattern, et al, 2008). The implication is that the SAT may provide a small admissions advantage to minority students to that extent. Whatever small advantage SAT scores may provide with respect to predicted college grades, however, is more than offset by the severe racial stratification they create among applicants, illustrated in Figure 9 above.

Because the confidence intervals are so wide, reliance on SAT scores to compare individual applicants introduces a substantial degree of error in admissions decisions. Two types of errors are involved. First are “false positives,” that is, instances where an applicant admitted on the basis of higher test scores performs worse than an applicant denied admission would have performed. Second are “false negatives,” that is, students denied admission who would have performed better than some of those admitted. Both kinds of error are inevitable when the predictive power of tests is low. Nor are such errors randomly distributed across all groups. Because SAT scores produce such sharp racial stratification within the applicant pool, students of color are much more likely to fall among the false negatives than the false positives.

It follows that, just as consideration of race requires scrutiny, so consideration of SAT scores requires individualized scrutiny of the totality of circumstances in an applicant’s file. Especially when score differences are small – as is typically the case at selective universities, where almost all applicants have relatively high scores – mechanical use of test scores to compare or rank applicants is unjustified. Flexible, individualized review is essential since test scores often have different meaning in different circumstances.

The point is nicely illustrated by the Educational Testing Service’s ill-fated “Strivers” proposal (Carnevale & Haghghat, 1998). In 1999 ETS proposed a new measure derived by comparing a student’s actual SAT scores with their predicted scores based on socioeconomic and other factors, including race. Students whose actual score significantly exceeded their predicted score were deemed “strivers.” As an ETS official explained, “A combined score of 1000 on the SATs is not always a 1000. When you look at a Striver who gets a 1000, you’re looking at someone who really performs at a 1200. This is a way of measuring not just where students are, but how far they’ve come” (Marcus, 1999). The proposal was later withdrawn after it was opposed by conservative groups and rejected by the College Board.

This is not to suggest resurrection of the Strivers proposal or some similar model, since that proposal simply replaces one mechanical formula with another and is inconsistent with the principle of individualized review in that respect. But it does illustrate a larger point: SAT scores can have different meanings for different applicants, so that it is essential to consider test performance in light of individual context and circumstance. Race is a demonstrably relevant aspect of that context.

Conclusion

The continuing dominance of standardized admissions tests in American higher education is one of the most powerful arguments for affirmative action. Much of the original impetus for race-conscious policies grew out of recognition of the severe adverse impact of SAT scores on admission of students of color. Since then, that impact has not only continued but worsened, if the California data are any indication. Among the 1.1 million students who applied for freshman admission at the University of California between 1994 and 2011, socioeconomic background factors known at students’ birth – family income, parents’ education, and race/ethnicity – account for an increasing share, over a third, of the variance in SAT scores. Race now uniquely accounts for the largest share.

These findings underscore the continuing relevance of the original, remedial rationale for affirmative action. Rather than a remedy for historical discrimination, however, they show that race-conscious policies are essential to remedy unwarranted disparities in the present day. The adverse racial impact of SAT scores is far out of proportion with their limited capacity to predict how applicants will perform in college.

But eliminating the SAT is not necessarily the answer, for two reasons. First, other criteria commonly employed in college admissions exhibit the same questionable trade-offs between adverse impact and measurement validity. The SAT is not unique but only more extreme in the extent to which its negative cost for students of color is disproportionate to its marginal benefit in predicting college performance.

Second, a direct challenge to the SAT is unlikely to succeed insofar as the test may be deemed an “educational necessity” in both a legal and practical sense. As the UC data confirm, SAT scores add a small but statistically significant increment to predictions of college performance and are thus likely to meet the relatively modest standard of “educational necessity” that case law requires. An argument for the necessity of tests such as the SAT or ACT can also be made on practical grounds. Standardized tests have become virtually indispensable at major universities that receive tens of thousands of applications, and there are no readily available alternatives with less adverse impact.

Yet – and this is the key point – if the SAT is considered an educational necessity, then consideration of race is also necessary in order to remedy the substantial adverse impact of test scores, beyond what can be reasonably justified by their measurement validity.

Nor will proxies for race suffice. The UC data suggest that racial disparities in test performance cannot be reduced to differences in family income or parents' education -- race and ethnicity account for a large and growing proportion of the variance in test scores after controlling for those factors. This casts substantial doubt that class-based affirmative action can provide a "workable alternative" to race-conscious admissions policies.

Race and the SAT are oddly alike as admissions criteria. Both are imperfect indicators of what they purport to measure, yet both are arguably necessary in view of the realities of contemporary college admissions. Just as an applicant's race is not invariably an indicator of the social advantages or disadvantages he or she may have experienced, so SAT scores reflect broader social and economic influences in addition to individual ability or merit. At the same time, both racial stratification and standardized testing remain entrenched features of American society, and neither is likely to disappear anytime soon.

The only rational response to this complicated reality is a flexible, nuanced, and individualized admissions process that takes into account the totality of information in an applicant's file. The requirement for "strict scrutiny" of race must be extended to the SAT itself, so that test scores are not employed mechanically but viewed in appropriate context. An applicant's race or ethnicity is a uniquely relevant aspect of that context, as the California data illustrate. If the SAT is an educational necessity, then so is explicit consideration of race. In their curious symbiosis, the fate of affirmative action remains intertwined with the fate of the SAT.

REFERENCES

- ACT. (2015). National and state scores. Accessed on July 5, 2015 from <http://www.act.org/newsroom/act-national-and-state-scores/>.
- Alexander v. Sandoval*. (2001). 532 U.S. 275.
- Alon, S., & Tienda, M. (2007). Diversity, opportunity, and the shifting meritocracy in higher education. *American Sociological Review* 72(4), 487-511.
- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. Washington, DC: Author.
- Atkinson, R. (2001). Standardized tests and access to American universities. The 2001 Robert H. Atwell Distinguished Lecture. Washington, DC: American Council on Education.
- Atkinson, R., & Geiser, S. (2009). Reflections on a century of college admissions tests. *Educational Researcher*, 38(9), 665-676.
- Boger, J. (2003). Education's "perfect storm"? The effect of racial resegregation, high stakes testing, and school resource inequities on North Carolina's poor, minority students. *North Carolina Law Review* 81, 1375.
- Bowen, W., & Bok, D. (1998). *The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions*. Princeton, NJ: Princeton University Press.
- Bowen, W., Chingos, M., & McPherson, M. (2009). *Crossing the Finish Line: Completing College at America's Public Universities*. Princeton, NJ: Princeton University Press.
- CSU Institute for Education Reform. (2001). *Characteristics and Performance of Advanced Placement Classes in California*. Sacramento, CA: California State University.
- Camara, W. & Schmidt, A. (1999). Group differences in standardized testing and social stratification. College Board Report No. 99-5. Accessed on August 19, 2005 from <http://research.collegeboard.org/sites/default/files/publications/2012/7/researchreport-1999-5-group-differences-standardized-testing-social-stratification.pdf>.
- Card, D., & Rothstein, J. (2006). Racial segregation and the black-white score gap. NBER Working Paper 12078. Cambridge, MA: National Bureau of Economic Research.
- Carnegie, A., & Haggighat, E. (1998). Selecting the strivers: A report on the preliminary results of the ETS 'Educational Strivers' study. Pp. 122-128 *Hopwood, Bakke, and Beyond: Diversity on Our Nation's Campuses*, edited by D. Bakst. Washington, DC: American Association of Collegiate Registrars and Admissions Officers.
- Carnegie, A., Rose, S., & Strohl, J. (2014). Achieving racial and economic diversity with race-blind admissions policy. Chapter 15 in *The Future of Affirmative Action: New Paths to Higher Education after Fisher v. University of Texas*, edited by R. Kahlenberg. New York: Century Foundation Press.
- College Board. (2000). California College-Bound Seniors Report. Accessed on December 3, 2014 from <http://media.collegeboard.com/digitalServices/pdf/research/cb-seniors-2000-CA.PDF>.
- College Board. (2012). College-Bound Seniors: Total Group Profile Report. Accessed on August 2, 2015 from <http://media.collegeboard.com/digitalServices/pdf/research/TotalGroup-2012.pdf>.
- Congressional Budget Office. (2011). Trends in the distribution of household income between 1979 and 2007. Washington, DC: Author. Accessed on August 13, 2015 from <https://www.cbo.gov/publication/42729>.
- Cook, M., & Evans, W. (2000). Families or schools? Explaining the convergence in white and black academic performance. *Journal of Labor Economics* 18(4), 729-54.
- Corcoran, S., & Evans, W. (2008). The role of inequality in teacher quality. Chapter 6 in *Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap*, edited by K. Magnuson and J. Waldfogel. New York: Russell Sage Foundation.
- Crouse, J., & Trusheim, G. (1998). *The Case Against the SAT*. Chicago and London: University of Chicago Press.
- CSU Institute for Education Reform. (2001). Characteristics and performance of Advanced Placement classes in California. Sacramento: Author.

- Dixon-Román, E., Everson, H., & McArdle, J. (2013). Race, poverty and SAT Scores: Modeling the influences of family income on black and white high school students' SAT performance. *Teachers College Record* 115 (4), 1-33. Accessed on August 12, 2015 from <http://www.tcrecord.org/Content.asp?ContentId=16925>.
- Doughtery, C., Mellor, L., & Jian, S. (2006). The relationship between Advanced Placement and college graduation. Austin, TX: National Center for Educational Accountability.
- Duan, B., & Dunlap, W. (1997). The accuracy of different methods for estimating the standard error of correlations corrected for range restriction. *Educational and Psychological Measurement*, 57, 245-265.
- Evans, W., Hout, M., and Mayer, S. (2004). Assessing the effect of economic inequality. In *Social Inequality*, edited by K. Neckerman. New York: Russell Sage Foundation.
- Fischer, C., Hout, M., Sanchez Jankowski, M., Lucas, S., Swidler, A., & Voss, K. (1996). *Inequality by Design: Cracking the Bell Curve Myth*. Princeton, NJ: Princeton University Press.
- Fisher v. University of Texas*. (2013). 113 S.Ct. 2411.
- Freeberg, N. (1988). Analysis of the revised Student Descriptive Questionnaire: Phase I: Accuracy of student-reported information. College Board Report No. 88-5. New York: College Entrance Examination Board.
- Garces, L. (2015). The legal context and social science evidence in *Fisher v. University of Texas*. Chapter 1 in *Affirmative Action and Racial Equity: Considering the Fisher Case to Forge the Path Ahead*, edited by U. Jayakumar, L. Garces, & F. Fernandez. New York and London: Routledge.
- 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, 1-26.
- Geiser, S., & Caspary, K. (2005). 'No show' study: College destinations of University of California applicants and admits who did not enroll, 1997-2002. *Journal of Educational Policy*, 19(2), 396-417.
- Geiser, S., & Santelices, V. (2006). The role of Advanced Placement and honors courses in college admissions. Pp. 75-114 in *Expanding Opportunity in Higher Education: Leveraging Promise*, edited by P. Gandara, G. Orfield, and C. Horn. Albany, NY: State University of New York Press.
- Geiser, S., & Santelices, V. (2007). Validity of high school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes. Research and Occasional Paper Series CSHE.6.07. Center for Studies in Higher Education, University of California, Berkeley.
- Geiser, S. (2009). [Back to the basics: In defense of achievement \(and achievement tests\) in college admissions](#). *Change: The Magazine of Higher Learning* (January/February 2009), 41(1), 16-23.
- Geiser, S., & Atkinson, R. (2013). Beyond the Master Plan: The case for restructuring baccalaureate education in California. *California Journal of Politics and Policy*, 4(1), 67-123.
- Geiser, S. (2014). Back to the future: Freshman admissions at the University of California, 1994 to the present and beyond. Research and Occasional Paper Series No. 4.14 (April, 2014). Berkeley, CA: Center for Studies in Higher Education. Accessed January 5, 2015 from <http://www.cshe.berkeley.edu/publications/back-future-freshman-admissions-university-california-1994-present-and-beyond>.
- Gratz v. Bollinger*. (2003). 539 U.S. 244.
- Grutter v. Bollinger*. (2002). 137 F. Supp. 2d 821, 825 (E.D. Mich. 2001), reviewed *en banc* 288 F.3d 732 (6th Cir. 2002).
- Grutter v. Bollinger*. (2003). 539 U.S. 306.
- Guryan, J. (2004). Desegregation and black dropout rates. *American Economic Review* 94(4), 919-43.
- Hamilton, L. (1992). *Regression with Graphics: A Second Course in Applied Statistics*. Belmont, CA: Duxbury Press.
- Hanushek, E., Kain, J., & Rivkin, S. (2004). New evidence about Brown v. Board of Education: The complex effects of school racial composition on achievement. Hoover Institution Working Paper. Stanford, CA: Stanford University Press.
- Haveman, R., & Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature* 23(4):1829-78.
- Haveman, R., Sandefur, G., Wolfe, B., and Voyer, A. (2004). Trends in children's attainments and their determinants as family income inequality has increased. In *Social Inequality*, edited by K. Neckerman. New York: Russell Sage Foundation.
- Jencks, C., & Phillips, M. (1998). Introduction. In *The Black-White Test Score Gap*, edited by C. Jencks and M. Phillips. Washington, D.C.: Brookings Institution Press.
- Hoxby, C. (2000). Peer effects in the classroom: Learning from gender and race variation. NBER Working Paper 7867. Cambridge, MA: National Bureau of Economic Research.
- Jencks, C. & Phillips, M. (1998). *The Black-White Test Score Gap*. Washington, DC: Brookings Institution Press.
- Kahlenberg, R. (2014). Introduction. Chapter 1 in *The Future of Affirmative Action: New Paths to Higher Education after Fisher v. University of Texas*, edited by R. Kahlenberg. New York: Century Foundation Press.
- Kidder, W. (2001). Affirmative action in higher education: Recent developments in litigation, admissions and diversity research. *Berkeley La Raza Law Journal*, 12(2), 173-230.
- Kidder, W., & Rosner, J. (2002). How the SAT creates built-in-headwinds: An educational and legal analysis of disparate impact, 43, Santa Clara L. Rev. 131. Accessed on January 6, 2015 from <http://digitalcommons.law.scu.edu/lawreview/vol43/iss1/3>.
- Klopfenstein, K., & Thomas, S. (2009). The link between Advanced Placement experience and early college success." *Southern Economic Journal*, 75(3), 873-891.
- Klopfenstein, K., & Thomas, S. (2010). Advanced placement participation: Evaluating the policies of states and colleges. Chapter 9 in *AP: A Critical Examination of the Advanced Placement Program*, edited by P. Adler, G. Sonnert, R. Tai, and K. Klopfenstein. Cambridge, MA: Harvard University Press.
- Kobrin, J., Sathy, V., & Shaw, E. (2006). A historical view of subgroup performance differences on the SAT. College Board Research Report 2006-5. New York: The College Board.

- Lewis, E. (2015). An SAT validity primer. Online publication on College Board website. Accessed on June 30, 2015 from <http://research.collegeboard.org/sites/default/files/publications/2015/6/research-report-sat-validity-primer.pdf>.
- Little, R., & Rubin, D. (2002). *Statistical Analysis with Missing Data*. 2nd ed. Hoboken, NJ: Wiley.
- Magnuson, K., and Duncan, G. (2006). The role of family socioeconomic resources on racial test score gaps. *Developmental Review* 26(4), 365-399.
- Magnuson, K., & Waldfogel, J. (2008). Introduction. In *Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap*, edited by K. Magnuson and J. Waldfogel. New York: Russell Sage Foundation.
- Magnuson, K., Rosenbaum, D., & Waldfogel, J. (2008). Inequality and black-white achievement trends in the NAEP. Chapter 1 in *Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap*, edited by K. Magnuson and J. Waldfogel. New York: Russell Sage Foundation.
- Marcus, A. (1999). New weights can alter SAT scores: Family is factor in determining who's a 'striver'. *Wall Street Journal*, August 31, 1999, pp. B1, B8. Accessed on June 27, 2015 from <http://info.wsj.com/classroom/archive/wsjsce.99nov.casestudy.html>.
- Mattern, K., Patterson, B., Shaw, E., Kobrin, J. & Barbuti, S. (2008). Differential validity and prediction of the SAT. College Board Research Report No. 2008-4. Accessed on August 16, 2015 from <http://research.collegeboard.org/sites/default/files/publications/2012/7/researchreport-2008-4-differential-validity-prediction-sat.pdf>.
- Meyers, M., Rosenbaum, D., Ruhm, C., & Waldfogel, J. (2004). Inequality in early childhood education and care: What do we know? In *Social Inequality*, edited by K. Neckerman. New York: Russell Sage Foundation.
- National Association for College Admission Counseling. (2004). National school counselor survey. Alexandria, VA: Author.
- National Association for College Admission Counseling. (2008). Report of the commission on the use of standardized tests in college admission. Alexandria, VA: Author.
- National Association for College Admission Counseling. (2015). 2014 state of college admission report. Arlington, VA: Author. Accessed on June 30, 2015 from http://www.nxtbook.com/ygsreprints/NACAC/2014SoCA_nxtbk/.
- Neal, D. (2006). Why has black-white skill convergence stopped? In *Handbook of the Economics of Education*, edited by E. Hunushek and F. Welch. Amsterdam: North-Holland.
- Oakes, J. (1985). *Keeping Track: How Schools Structure Inequality*. New Haven, CT: Yale University Press.
- Oakes, J. (1990). Multiplying inequalities: Race, social class, and tracking students' opportunities to learn mathematics and science. Santa Monica, CA: RAND.
- Oakes, J., Gamoran, A., and Page, R.N. (1992). Curriculum differentiation: Opportunities, outcomes and meanings. In P. Jackson (ed.), *Handbook of Research on Curriculum*, pp. 570-608. New York: Macmillan.
- Oliver, M., & Shapiro, T. (1995). *Black Wealth/White Wealth: A New Perspective on Racial Inequality*. New York: Routledge.
- Orfield, G., & Ee, J. (2014). *Segregating California's Future: Inequality and its Alternative 60 Years after Brown v. Board of Education*. Los Angeles, CA: UCLA Civil Rights Project. Accessed on August 19, 2015 from <http://civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/segregating-california2019s-future-inequality-and-its-alternative-60-years-after-brown-v.-board-of-education>.
- Orfield, G., & Frankenberg, E. (2014). *Brown at 60: Great Progress, a Long Retreat and an Uncertain Future*. Los Angeles, CA: UCLA Civil Rights Project. Accessed on August 19, 2015 from <http://civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/brown-at-60-great-progress-a-long-retreat-and-an-uncertain-future/>.
- Parker, W. (2004). Connecting the dots: Grutter, school desegregation, and federalism. *William and Mary Law Review* 45(6), 1691-1768.
- Patterson, B., & Mattern, K. (2012). Validity of the SAT for predicting first-year grades: 2009 SAT validity sample. College Board Statistical Report No. 2012-2. Accessed on August 19, 2015 from <http://research.collegeboard.org/sites/default/files/publications/2012/7/statisticalreport-2012-2-validity-sat-1st-yr-gpa-2009-sample.pdf>.
- Phillips, M., and Chin, T. (2000). School inequality: What do we know? In *Social Inequality*, edited by K. Neckerman. New York: Russell Sage Foundation.
- Phillips, M., Crouse, J., & Ralph, J. (1998). Does the black-white test score gap widen after children enter school? Chapter 7 in *The Black-White Test Score Gap*, edited by C. Jencks and M. Phillips. Washington, DC: Brookings Institution Press.
- Reardon, S. (2011). The widening socioeconomic status achievement gap: New evidence and possible explanations. In *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, edited by R. Murnane and G. Duncan. New York: Russell Sage Foundation.
- Regents of the University of California v. Bakke*. (1978). 438 U.S. 265.
- Robbins, S., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological bulletin*, 130(2), 261-288.
- Rothstein, J. (2004). College performance predictions and the SAT. *Journal of Econometrics*, 121, 297-317.
- Rothstein, R. (2014). The racial achievement gap, segregated schools, and segregated neighborhoods – A constitutional insult. *Race and Social Problems* 6 (4). Accessed on August 19, 2015 from <http://www.epi.org/publication/the-racial-achievement-gap-segregated-schools-and-segregated-neighborhoods-a-constitutional-insult/>.
- Rubinowitz, L., & Rosenbaum, J. (2001). *Crossing the Class and Color Lines: From Public Housing to White Suburbia*. Chicago: University of Chicago Press.
- Sacks, P. (2000). *Standardized Minds: The High Price of America's Testing Culture and What We Can Do to Change It*. Cambridge, MA: Perseus Books.
- Sackett, P., Kuncel, N., Arneson, J., Cooper, S., & Waters, S. (2009) Socioeconomic status and the relationship between the SAT® and freshman GPA: An analysis of data from 41 colleges and universities. College Board Research Report No. 2009-1. New York: College Board. Accessed on August 3, 2015 from <http://research.collegeboard.org/sites/default/files/publications/2012/9/researchreport-2009-1-socioeconomic-status-sat-freshman-gpa-analysis-data.pdf>.

- Schmitt, N., Keeney, J., Oswald, F., Pleskac, T., Billington, A., Sinha, R., & Zorzie, M. (2009). Prediction of 4-year college student performance using cognitive and noncognitive predictors and the impact on demographic status of admitted students. *Journal of Applied Psychology, 94*(6), 1479-1497.
- Steele, C., & Aronson, J. (1998). Stereotype threat and the test performance of academically successful African Americans. Pp. 401-30 in *The Black-White Test Score Gap*, edited by C. Jencks and M. Phillips. Washington, DC: Brookings Institution Press.
- Sternberg, R. J., Bonney, C. R., Gabora, L., & Merrifield, M. (2012). WICS: A model for college and university admissions. *Educational Psychologist, 47*(1), 30-41.
- Stewart, D. (1998). Why Hispanic students need to take the SAT. *Chronicle of Higher Education*, January 30, 1998, p. A 48.
- Wards Cove Packing Co. v. Atonio*. (1989). [490 U.S. 642](#).
- Vigdor, J. (2006). Peer effects in neighborhoods and housing. In *Deviant Peer Influences in Programs for Youth: Problems and Solutions*, edited by K. Dodge, J. Dishion, and J. Lansford. New York: Guilford Press.
- Vigdor, J., & Ludwig, J. (2008). Segregation and the test score gap. Chapter 5 in *Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap*, edited by K. Magnuson and J. Waldfogel. New York: Russell Sage Foundation.
- Weber, D. (2001). Restriction of range: The truth about consequences and corrections. Paper presented at the Annual Meeting of the Southwest Educational Research Association (New Orleans, LA, February 1-3, 2001). Accessed on August 25, 2015 from <http://files.eric.ed.gov/fulltext/ED449213.pdf>.
- Wilson, W. (1998). The role of the environment in the black-white test score gap. Chapter 15 in *The Black-White Test Score Gap*, edited by C. Jencks and M. Phillips. Washington, D.C.: Brookings Institution Press.
- Zwick, R. (1999). Eliminating standardized tests in college admissions: The new affirmative action? *Phi Delta Kappan 81*.
- Zwick, R. (2004). Is the SAT a "wealth test"? The link between educational achievement and socioeconomic status. Pp. 203-216 in *Rethinking the SAT: The Future of Standardized Testing in University Admissions*, edited by R. Zwick. New York and London: RoutledgeFarmer.

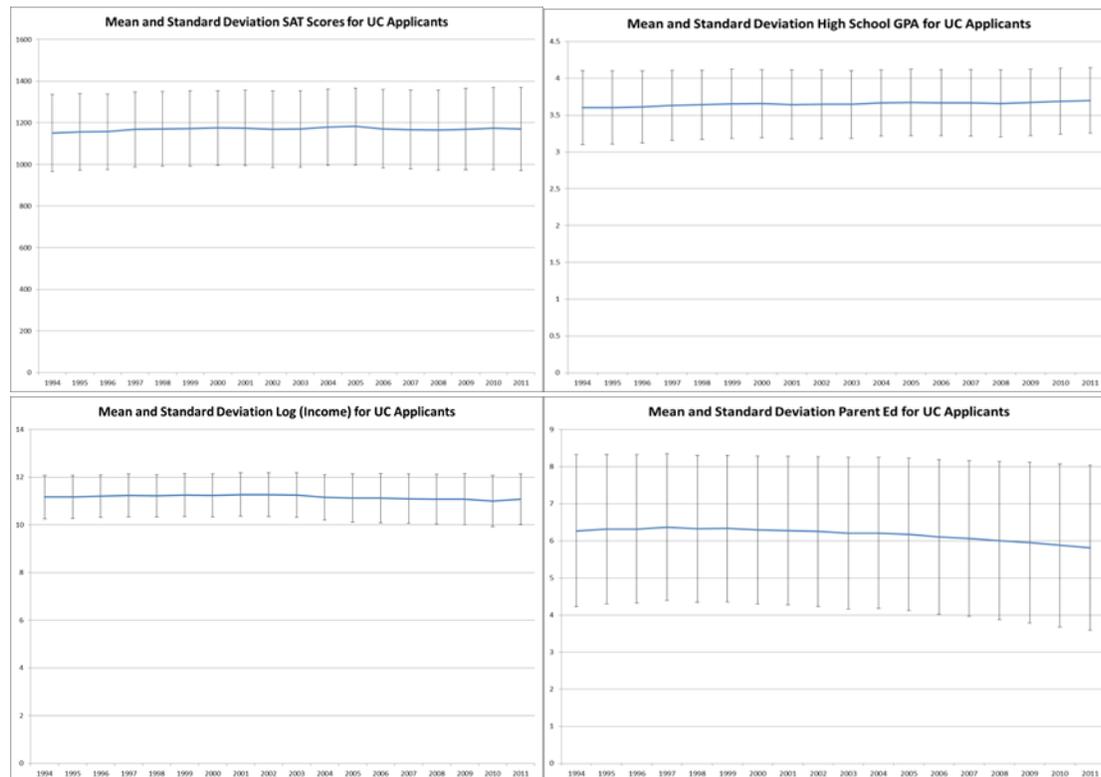
Appendix 1: Descriptive Information on Variables Employed in This Study

Descriptive Summary of Study Variables

Variable	Mean	Std. Dev.	Min	Max	Observations
Composite SAT Scores	1170	187	400	1600	1,135,175
High School GPA	3.66	0.46	0.07	4.88	1,135,113
Parent Education	6.14	2.08	2.00	8.00	1,088,936
Log Family Income (2013\$)	11.2	1.0	0.0	14.3	918,089
Underrepresented Minority Status	24%	N/A	0	1	1,144,047

Correlation Matrix of Study Variables

	SAT	HSGPA	Parent Ed	Log Inc	URM
Composite SAT Scores	1				
High School GPA	0.4859	1			
Parent Education	0.4451	0.1395	1		
Log Family Income (2013\$)	0.3646	0.1111	0.5361	1	
URM Status	-0.3824	-0.179	-0.3854	-0.2587	1



Appendix 2: Documentation for Figures 1 and 2 (STATA)

```
. ** Regress SAT scores on family income, parents' education and underrepresented minority status
. ** Sample: All CA residents applying to UC from 1994 through 2011
. ** "satc_aaa" = Composite SAT I score
. ** "loginc" = Log of family income in constant 2013 dollars
. ** "parent_ed" = Level of education of highest-educated parent
. ** "underrep" = Underrepresented minority status
. ** "yearapply" = Year applied to UC
. ** "beta" = Standardized regression coefficient
```

```
. by yearapply: regress satc_aaa loginc parent_ed underrep, beta
```

-> yearapply = 1994

Source	SS	df	MS	Number of obs =	36838
Model	313062024	3	104354008	F(3, 36834) =	4124.57
Residual	931920934	36834	25300.5629	Prob > F =	0.0000
				R-squared =	0.2515
				Adj R-squared =	0.2514
Total	1.2450e+09	36837	33797.0779	Root MSE =	159.06

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	31.428	1.080978	29.07	0.000	.1551103
parent_ed	23.92059	.4978586	48.05	0.000	.2669734
underrep	-102.2136	2.159817	-47.33	0.000	-.228583
_cons	667.5043	10.98656	60.76	0.000	.

-> yearapply = 1995

Source	SS	df	MS	Number of obs =	39222
Model	319088156	3	106362719	F(3, 39218) =	4287.69
Residual	972861862	39218	24806.5139	Prob > F =	0.0000
				R-squared =	0.2470
				Adj R-squared =	0.2469
Total	1.2920e+09	39221	32940.2621	Root MSE =	157.5

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	29.73947	1.056696	28.14	0.000	.1461782
parent_ed	24.75916	.4843433	51.12	0.000	.2751042
underrep	-98.25696	2.063139	-47.62	0.000	-.2227114
_cons	683.9914	10.73756	63.70	0.000	.

-> yearapply = 1996

Source	SS	df	MS	Number of obs =	42088
Model	298773214	3	99591071.4	F(3, 42084) =	3974.37
Residual	1.0546e+09	42084	25058.3133	Prob > F =	0.0000
				R-squared =	0.2208
				Adj R-squared =	0.2207
Total	1.3533e+09	42087	32155.4701	Root MSE =	158.3

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
----------	-------	-----------	---	------	------

loginc		27.7356	1.024661	27.07	0.000	.1367513
parent_ed		23.88504	.4686248	50.97	0.000	.2667493
underrep		-92.42571	2.046221	-45.17	0.000	-.2069254
_cons		710.9278	10.46813	67.91	0.000	.

-> yearapply = 1997

Source	SS	df	MS	Number of obs =
Model	294387372	3	98129124.1	42791
Residual	1.0562e+09	42787	24683.9955	F(3, 42787) = 3975.41
Total	1.3505e+09	42790	31562.082	Prob > F = 0.0000

R-squared = 0.2180
Adj R-squared = 0.2179
Root MSE = 157.11

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	28.07298	.9960659	28.18	0.000	.1420392
parent_ed	23.5667	.4669649	50.47	0.000	.262899
underrep	-92.24323	2.053326	-44.92	0.000	-.2036806
_cons	717.3678	10.14835	70.69	0.000	.

-> yearapply = 1998

Source	SS	df	MS	Number of obs =
Model	281898913	3	93966304.2	41558
Residual	1.0362e+09	41554	24935.1708	F(3, 41554) = 3768.42
Total	1.3181e+09	41557	31716.7986	Prob > F = 0.0000

R-squared = 0.2139
Adj R-squared = 0.2138
Root MSE = 157.91

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	26.86775	1.03526	25.95	0.000	.1336063
parent_ed	23.71946	.4721382	50.24	0.000	.267113
underrep	-90.56967	2.062524	-43.91	0.000	-.2018794
_cons	729.8235	10.53439	69.28	0.000	.

-> yearapply = 1999

Source	SS	df	MS	Number of obs =
Model	317980439	3	105993480	44916
Residual	1.1311e+09	44912	25185.4175	F(3, 44912) = 4208.53
Total	1.4491e+09	44915	32263.3399	Prob > F = 0.0000

R-squared = 0.2194
Adj R-squared = 0.2194
Root MSE = 158.7

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	30.6547	.9910894	30.93	0.000	.1529112
parent_ed	23.5564	.4581686	51.41	0.000	.2622893
underrep	-89.92766	2.003359	-44.89	0.000	-.1975023
_cons	688.0906	10.09069	68.19	0.000	.

-> yearapply = 2000

Source	SS	df	MS	Number of obs =	43694
Model	319286344	3	106428781	F(3, 43690) =	4369.21
Residual	1.0642e+09	43690	24358.7894	Prob > F =	0.0000
				R-squared =	0.2308
				Adj R-squared =	0.2307
Total	1.3835e+09	43693	31664.6111	Root MSE =	156.07

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	29.21157	.9877461	29.57	0.000	.1478992
parent_ed	23.33818	.4568376	51.09	0.000	.2652949
underrep	-93.40016	1.977715	-47.23	0.000	-.2108609
_cons	709.083	10.04516	70.59	0.000	.

-> yearapply = 2001

Source	SS	df	MS	Number of obs =	46656
Model	372111034	3	124037011	F(3, 46652) =	5069.52
Residual	1.1414e+09	46652	24467.2055	Prob > F =	0.0000
				R-squared =	0.2459
				Adj R-squared =	0.2458
Total	1.5136e+09	46655	32441.4341	Root MSE =	156.42

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	31.05748	.9534894	32.57	0.000	.1565821
parent_ed	23.91741	.4408529	54.25	0.000	.270795
underrep	-93.81399	1.886199	-49.74	0.000	-.213733
_cons	684.9693	9.741924	70.31	0.000	.

-> yearapply = 2002

Source	SS	df	MS	Number of obs =	48358
Model	393968867	3	131322956	F(3, 48354) =	5265.58
Residual	1.2059e+09	48354	24939.8817	Prob > F =	0.0000
				R-squared =	0.2462
				Adj R-squared =	0.2462
Total	1.5999e+09	48357	33085.4252	Root MSE =	157.92

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	31.4202	.9372394	33.52	0.000	.1582787
parent_ed	23.53504	.4347679	54.13	0.000	.265519
underrep	-95.63169	1.852779	-51.62	0.000	-.2181479
_cons	677.0037	9.575826	70.70	0.000	.

-> yearapply = 2003

Source	SS	df	MS	Number of obs =	50752
Model	435500148	3	145166716	F(3, 50748) =	5874.89
Residual	1.2540e+09	50748	24709.7093	Prob > F =	0.0000
				R-squared =	0.2578
				Adj R-squared =	0.2577
Total	1.6895e+09	50751	33289.3633	Root MSE =	157.19

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	31.39561	.8931883	35.15	0.000	.1600319
parent_ed	23.26681	.4157027	55.97	0.000	.2657537
underrep	-97.9853	1.76707	-55.45	0.000	-.2287652
_cons	684.3005	9.15951	74.71	0.000	.

-> yearapply = 2004

Source	SS	df	MS	Number of obs =	45900
Model	401321676	3	133773892	F(3, 45896) =	5468.34
Residual	1.1228e+09	45896	24463.3606	Prob > F	= 0.0000
Total	1.5241e+09	45899	33205.3438	R-squared	= 0.2633
				Adj R-squared	= 0.2633
				Root MSE	= 156.41

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	29.87872	.9198943	32.48	0.000	.1542903
parent_ed	23.83007	.4269397	55.82	0.000	.2773524
underrep	-95.14097	1.80153	-52.81	0.000	-.2284995
_cons	704.3812	9.389322	75.02	0.000	.

-> yearapply = 2005

Source	SS	df	MS	Number of obs =	49083
Model	462490793	3	154163598	F(3, 49079) =	6337.27
Residual	1.1939e+09	49079	24326.5025	Prob > F	= 0.0000
Total	1.6564e+09	49082	33747.8344	R-squared	= 0.2792
				Adj R-squared	= 0.2792
				Root MSE	= 155.97

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	29.84751	.8155734	36.60	0.000	.1630059
parent_ed	23.91343	.4035773	59.25	0.000	.2764876
underrep	-102.5001	1.724176	-59.45	0.000	-.2460963
_cons	713.6354	8.356165	85.40	0.000	.

-> yearapply = 2006

Source	SS	df	MS	Number of obs =	54321
Model	511937149	3	170645716	F(3, 54317) =	6688.20
Residual	1.3859e+09	54317	25514.4414	Prob > F	= 0.0000
Total	1.8978e+09	54320	34937.5011	R-squared	= 0.2698
				Adj R-squared	= 0.2697
				Root MSE	= 159.73

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	27.74383	.7811383	35.52	0.000	.1517353
parent_ed	23.64068	.3916924	60.36	0.000	.2696533
underrep	-105.773	1.668259	-63.40	0.000	-.2511913
_cons	730.5687	7.999464	91.33	0.000	.

-> yearapply = 2007

Source	SS	df	MS		
Model	555055758	3	185018586	Number of obs =	56154
Residual	1.4457e+09	56150	25746.3899	F(3, 56150) =	7186.20
				Prob > F	= 0.0000
				R-squared	= 0.2774
				Adj R-squared	= 0.2774
Total	2.0007e+09	56153	35629.7179	Root MSE	= 160.46

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	29.6596	.7627298	38.89	0.000	.1620382
parent_ed	23.508	.3829532	61.39	0.000	.2680527
underrep	-106.3577	1.614943	-65.86	0.000	-.2550162
_cons	707.8953	7.808712	90.65	0.000	.

-> yearapply = 2008

Source	SS	df	MS		
Model	664633075	3	221544358	Number of obs =	60890
Residual	1.5531e+09	60886	25508.0736	F(3, 60886) =	8685.26
				Prob > F	= 0.0000
				R-squared	= 0.2997
				Adj R-squared	= 0.2997
Total	2.2177e+09	60889	36422.3036	Root MSE	= 159.71

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	30.51947	.7361295	41.46	0.000	.1655113
parent_ed	23.54187	.3673215	64.09	0.000	.2691489
underrep	-109.0916	1.539362	-70.87	0.000	-.2646383
_cons	701.2469	7.556966	92.79	0.000	.

-> yearapply = 2009

Source	SS	df	MS		
Model	748806473	3	249602158	Number of obs =	62886
Residual	1.6288e+09	62882	25902.7166	F(3, 62882) =	9636.14
				Prob > F	= 0.0000
				R-squared	= 0.3149
				Adj R-squared	= 0.3149
Total	2.3776e+09	62885	37809.0339	Root MSE	= 160.94

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	31.42511	.7126935	44.09	0.000	.1718704
parent_ed	23.2623	.3627711	64.12	0.000	.2640872
underrep	-115.5396	1.507596	-76.64	0.000	-.2797483
_cons	703.8402	7.304336	96.36	0.000	.

-> yearapply = 2010

Source	SS	df	MS		
Model	795513980	3	265171327	Number of obs =	63965
Residual	1.6474e+09	63961	25755.7098	F(3, 63961) =	10295.63
				Prob > F	= 0.0000
				R-squared	= 0.3256
				Adj R-squared	= 0.3256
Total	2.4429e+09	63964	38191.4035	Root MSE	= 160.49

```
-----
```

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	32.36855	.7025186	46.08	0.000	.17739
parent_ed	24.18973	.3556824	68.01	0.000	.2769736
underrep	-111.8551	1.475295	-75.82	0.000	-.2722725
_cons	696.0479	7.142126	97.46	0.000	.

```
-----
```

-> yearapply = 2011

Source	SS	df	MS	Number of obs =	71833
Model	984336725	3	328112242	F(3, 71829) =	12901.77
Residual	1.8267e+09	71829	25431.5729	Prob > F =	0.0000
				R-squared =	0.3502
				Adj R-squared =	0.3501
Total	2.8111e+09	71832	39133.8286	Root MSE =	159.47

```
-----
```

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	34.23825	.6932287	49.39	0.000	.1819821
parent_ed	24.06674	.3449616	69.77	0.000	.2719899
underrep	-119.3687	1.395511	-85.54	0.000	-.2883279
_cons	690.2913	7.029099	98.20	0.000	.

```
-----
```

Appendix 3: Regression Results for UC Applicants, 2008 to 2011 (STATA)

```
. ** Regress SAT scores on family income, parents' education and underrepresented minority status
. ** Sample: All CA residents applying to UC from 2008 through 2011
. ** "satc_aaa" = Composite SAT I score
. ** "loginc" = Log of family income in constant 2013 dollars
. ** "parent_ed" = Level of education of highest-educated parent
. ** "underrep" = Underrepresented minority status
. ** "beta" = Standardized regression coefficient
```

```
. regress satc_aaa loginc parent_ed underrep if yearapply>2007, beta
```

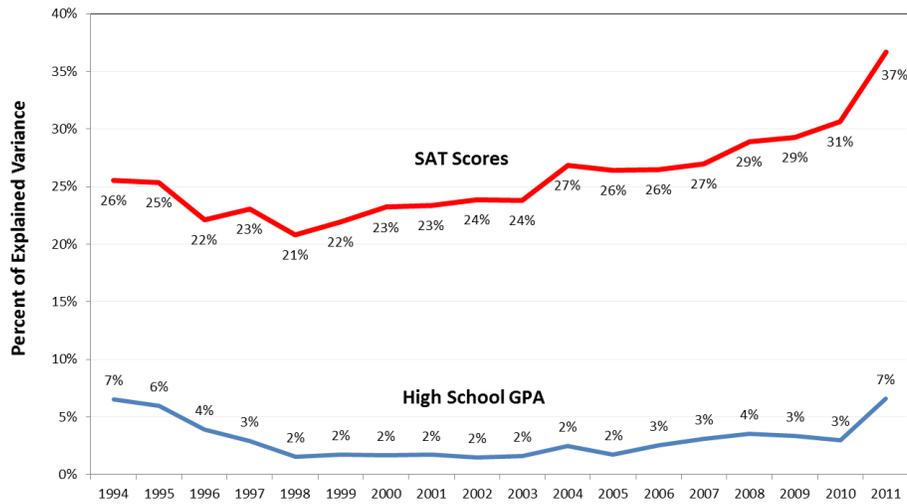
Source	SS	df	MS	
Model	3.1808e+09	3	1.0603e+09	Number of obs = 259574
Residual	6.6908e+09	259570	25776.3466	F(3,259570) =41133.74
Total	9.8716e+09	259573	38030.1465	Prob > F = 0.0000
				R-squared = 0.3222
				Adj R-squared = 0.3222
				Root MSE = 160.55

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
loginc	32.1855	.3560052	90.41	0.000	.1742724
parent_ed	23.87464	.1790776	133.32	0.000	.2714515
underrep	-113.3993	.7396411	-153.32	0.000	-.2747047
_cons	697.1142	3.631836	191.95	0.000	.

Appendix 4: Explained Variance and Standardized Regression Coefficients for UC Admits

Percent of Variance in SAT Scores and High School GPA Explained by Family Income, Parents' Education, and Underrepresented Minority Status: UC Admits, 1994 to 2011

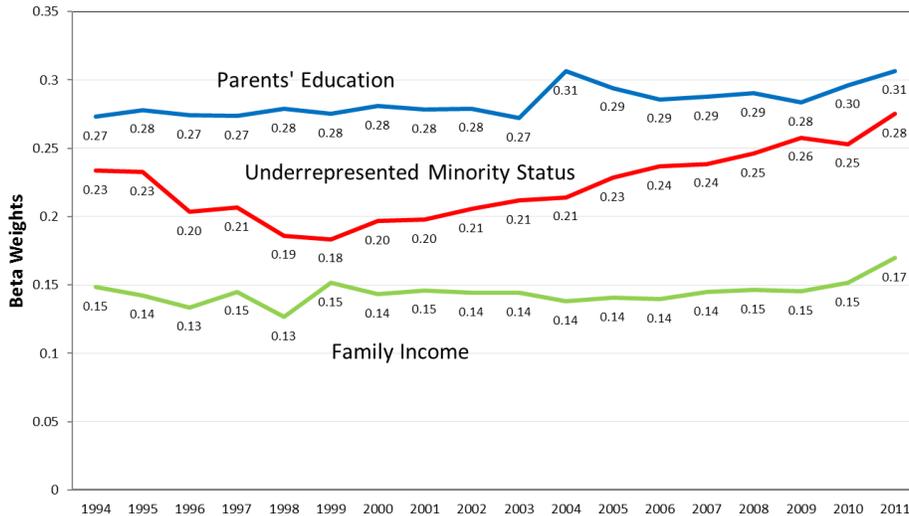
Regression equation: SAT score or HSGPA = $b_1(\text{Log of Income}) + b_2(\text{Parent Ed}) + b_3(\text{URM status})$



Source: UC Corporate Student System data on California resident freshmen admits, 1994 to 2011.

Standardized Regression Coefficients for Predicting SAT Scores Conditional on Family Income, Parents' Education, and Underrepresented Minority Status: UC Admits, 1994 to 2011

Regression equation: SAT score = $b_1(\text{Log of Income}) + b_2(\text{Parent Education}) + b_3(\text{URM status})$



Source: UC Corporate Student System data on California resident freshmen admits, 1994 to 2011.

Appendix 5: Checking for Multicollinearity and Interaction Effects

The present findings were checked carefully for both multicollinearity and interaction effects. The first problem occurs when predictor variables are related not only to the outcome variable of interest (in this case, SAT scores), but also to each other. When predictor variables are too closely related to one other, regression results can become unreliable. Given the well-known interrelationships among the three socioeconomic indicators considered in this analysis, it is therefore important to examine the UC data carefully to determine whether the findings reported here may reflect that problem.

**Correlations among Predictor Variables
and Tests for Multicollinearity**

	Correlations			Shared Variance and Tolerance	
	Family Income	Parents' Education	Underrep Minority	Shared Variance (R ²)	Tolerance (1 - R ²)
Family Income	1.000			0.290	0.711
Parents' Education	0.535	1.000		0.352	0.648
Underrep Minority	-0.261	-0.387	1.000	0.154	0.846

Source: UC Corporate Student System data on California resident freshman applicants, 1994 to 2011.

Although the three socioeconomic indicators are all correlated, as shown in the left-hand panel, none of the correlations are large enough to raise concerns about multicollinearity, as shown in the right-hand panel. "Shared variance" is the amount of variance each predictor variable shares with the other socioeconomic variables; it is derived by regressing each predictor against the other two. "Tolerance" is then derived by subtracting the shared variance from 1; that statistic shows the amount of variance *not* shared with the other predictors. As a rule of thumb, tolerances must be at least .1 or .2 to ensure stable regression results (Hamilton, 1992, pp.133-134). All of the tolerances shown here are well above what is necessary for reliable estimation of prediction weights.

Interaction effects refer to instances where the effect of a given predictor varies with other predictor variables. For example, it may be the case that the effect of family income on SAT scores increases with parents' education. If so, that might explain the direct or "main effect" of race/ethnicity on SAT scores observed earlier, since minority families fall disproportionately toward the lower end of both the income and education distributions.

Accordingly, all possible interactions among the three socioeconomic predictors – income by education, race by education, and race by income -- were introduced into the original regression analysis. Results are shown below:

**Predictive Weight* of Socioeconomic Factors
With and Without Interaction Effects:
UC Applicants, 1994 to 2011**

Regression equation: $SAT\ score = b_1(Log\ of\ Income) + b_2(Parent\ Ed) + b_3(URM\ status) + b_4(Income \times Education) + b_5(URM\ status \times Education) + b_6(URM\ status \times Income)$

	Main Effects Only	Main Effects with Interactions
Family Income	0.160	0.137
Parents' Education	0.268	0.282
URM status	-0.238	-0.232
Interaction terms:		
Income x Education	*	0.013
URM status x Education	*	0.019
URM status x Income	*	0.047
Number of students (N)	901,905	901,905
% Explained Variance (R ²)	26.8%	26.9%

* Weights shown in table are standardized coefficients.
All estimates are statistically significant at .001 confidence level.
Source: UC Corporate Student System data, 1994 to 2011.

The results do show the presence of small interaction effects, but rather than diminish the findings presented earlier, they underscore the importance of race as a determinant of SAT scores. While all three of the interaction effects are statistically significant, the first two – income x education and URM status x education – are very near zero; given the large size of the UC sample, even very small effects tend to attain statistical significance. On the other hand, the third interaction, between URM status and income, is larger and nontrivial. What this result implies is that the effect of income on SAT scores is intensified by URM status; family income has a greater effect on SAT scores for underrepresented minorities than for other students, other things being equal. This interaction or multiplier effect is in addition to the already strong direct effect of race/ethnicity on SAT scores.

Appendix 6: Effect of Changes in URM Proportion of Applicant Pool on Main Regression Findings (STATA)

```
. ** Regress SAT scores on parent education, family income, underrepresented minority (URM)
status, and URM percent of applicant cohort
. ** Sample: All CA residents applying to UC from 1994 through 2011
. ** "satc_aaa" = Composite SAT I score
. ** "loginc" = Log of family income in constant 2013 dollars
. ** "parent_ed" = Level of education of highest-educated parent
. ** "underrep" = Underrepresented minority status
. ** "urmpercent" = URM percent of applicant cohort in year applied
. ** "beta" = Standardized regression coefficient
```

```
. ** Main regression WITHOUT underrepresented minority percent of applicant cohort
```

```
. regress satc_aaa parent_ed loginc underrep, beta
```

Source	SS	df	MS	Number of obs =	901905
Model	8.3633e+09	3	2.7878e+09	F(3,901901) =	.
Residual	2.2812e+10	901901	25293.7119	Prob > F =	0.0000
Total	3.1176e+10	901904	34566.5712	R-squared =	0.2683
				Adj R-squared =	0.2683
				Root MSE =	159.04

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
parent_ed	23.49459	.0981201	239.45	0.000	.2678552
loginc	30.47913	.2037636	149.58	0.000	.1598298
underrep	-100.7651	.414734	-242.96	0.000	-.2376832
_cons	699.6704	2.079286	336.50	0.000	.

```
. ** Main regression WITH underrepresented minority percent of applicant cohort
```

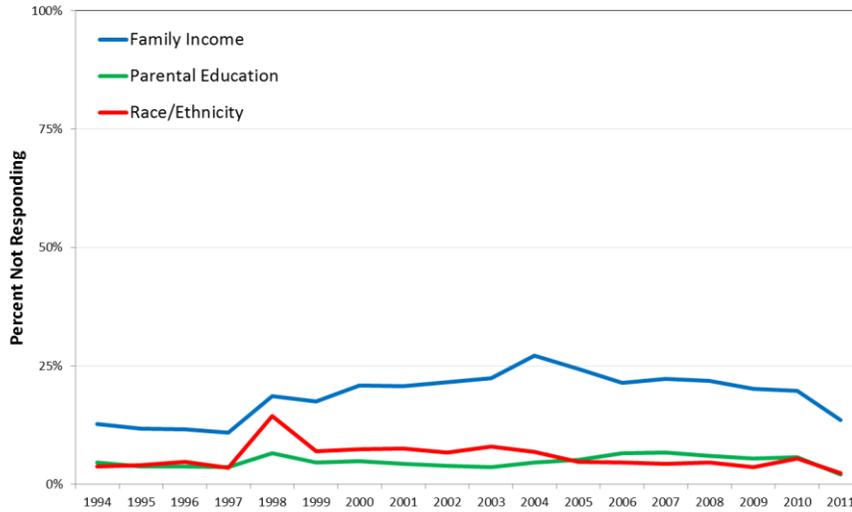
```
. regress satc_aaa parent_ed loginc underrep urmpercent, beta
```

Source	SS	df	MS	Number of obs =	901905
Model	8.4677e+09	4	2.1169e+09	F(4,901900) =	84078.34
Residual	2.2708e+10	901900	25177.9943	Prob > F =	0.0000
Total	3.1176e+10	901904	34566.5712	R-squared =	0.2716
				Adj R-squared =	0.2716
				Root MSE =	158.68

satc_aaa	Coef.	Std. Err.	t	P> t	Beta
parent_ed	23.7477	.0979743	242.39	0.000	.2707408
loginc	30.67432	.2033195	150.87	0.000	.1608534
underrep	-103.2915	.4156403	-248.51	0.000	-.2436426
urmpercent	223.6313	3.473054	64.39	0.000	.0584202
_cons	643.607	2.24983	286.07	0.000	.

Appendix 7: Non-Response Rates for SES among UC Applicants

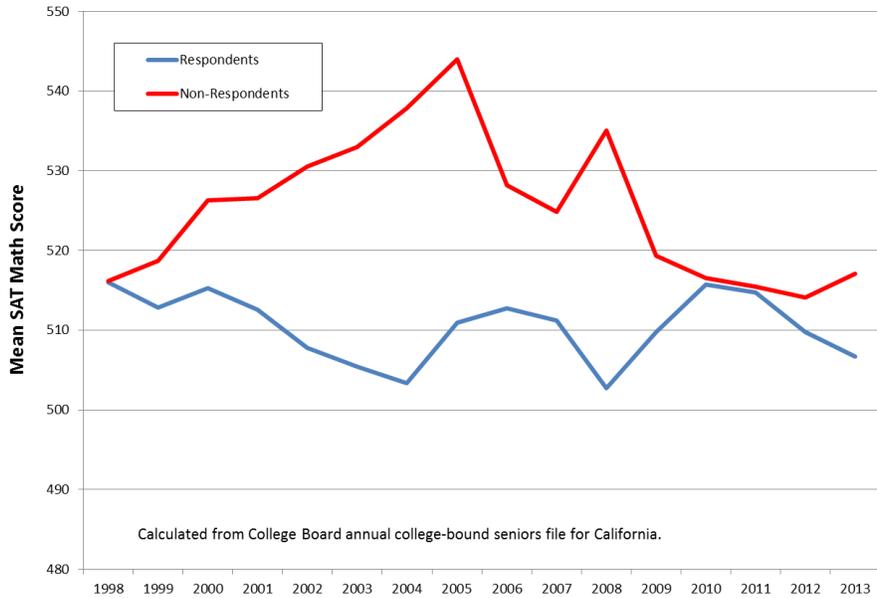
Percent of UC Applicants Not Reporting Family Income, Parents' Education, or Race/Ethnicity, 1994 to 2011



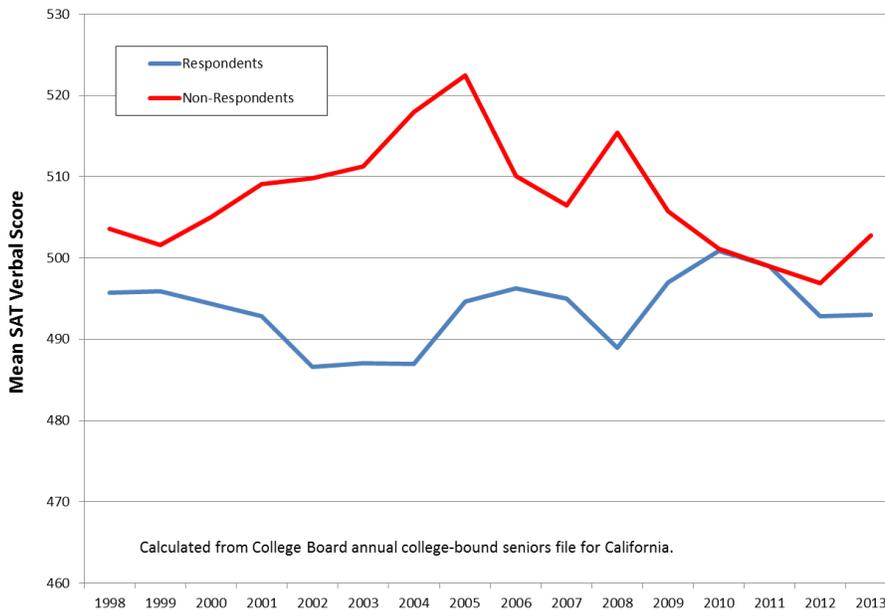
Source: UC Corporate Student System data for all California resident freshman applicants, 1994 to 2011.

Appendix 8: Mean Math and Verbal Scores for California SAT Takers Who Did Not Report Family Income, 1998 to 2011

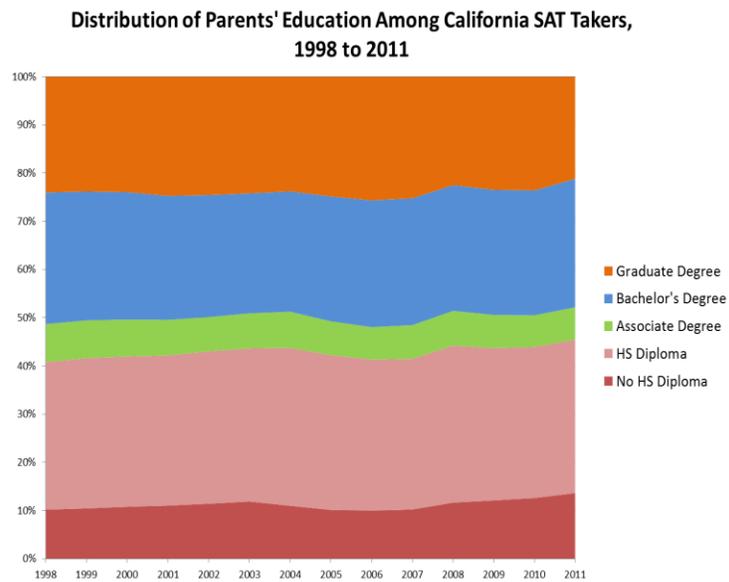
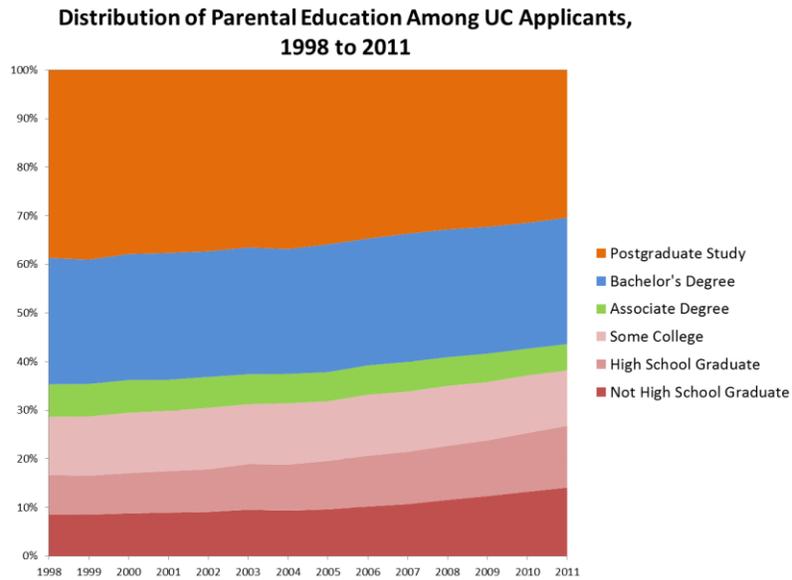
Mean SAT Math Scores for Respondents and Non-Respondents to SDQ Parental Income Item: California SAT Takers



Mean SAT Verbal Scores for Respondents and Non-Respondents to SDQ Parental Income Item: California SAT Takers



Appendix 9: Distribution of Parents' Education among UC Applicants and California SAT Takers, 1998 to 2011



Appendix 10: Documentation for Figure 10 (STATA)

```
. ** Sample: All CA residents enrolled as UC freshmen from 1994 through 2005
. ** "hasgrad500" = Graduated within 5 years
. ** "satc_aaa" = Composite SAT I score
. ** "hsgpa_calc" = High school grade-point average
. ** "log_inc" = Log of family income in constant 2013 dollars
. ** "parent_ed" = Level of education of highest-educated parent
. ** "underrep" = Underrepresented minority status
. ** "yearapply" = Year applied to UC
. ** "enr00" = Enrolled at UC
. ** "lstand" = Standardized regression coefficient for logistic regression
```

```
. . ** Predicted 5-year graduation for all entering freshmen without considering SES
```

```
. logistic hasgrad500 satc_aaa hsgpa_calc if enr00==1 & yearapply<2006
```

```
Logistic regression                               Number of obs   =    309391
                                                    LR chi2(2)      =   19129.64
                                                    Prob > chi2     =    0.0000
Log likelihood = -158907.82                       Pseudo R2      =    0.0568
```

```
-----+-----
```

hasgrad500	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
satc_aaa	1.001314	.0000286	45.95	0.000	1.001258	1.00137
hsgpa_calc	3.462845	.0414492	103.77	0.000	3.382551	3.545044

```
-----+-----
```

```
. lstand
```

Table of Predictor Estimates:
Standardized Coefficients and Partial Correlations

```
=====
```

No.	Var	Coef	OR	St.Coef	PartCorr	Prob(z)
0	Constant	-4.9546				
1	satc_aaa	0.0013	1.0013	0.1356	0.0791	0.000
2	hsgpa_calc	1.2421	3.4628	0.3164	0.1788	0.000

```
=====
```

```
. ** Predicted 5-year graduation for all entering freshmen controlling for SES
```

```
. logistic hasgrad500 satc_aaa hsgpa_calc parent_ed loginc if enr00==1 & yearapply<2006
```

```
Logistic regression                               Number of obs   =    254665
                                                    LR chi2(4)      =   16366.52
                                                    Prob > chi2     =    0.0000
Log likelihood = -131874.92                       Pseudo R2      =    0.0584
```

```
-----+-----
```

hasgrad500	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
satc_aaa	1.000888	.0000349	25.42	0.000	1.000819	1.000956
hsgpa_calc	3.537221	.0468009	95.48	0.000	3.446672	3.630148
parent_ed	1.027126	.0028685	9.58	0.000	1.021519	1.032763
loginc	1.142036	.0071318	21.27	0.000	1.128143	1.1561

```
-----+-----
```

```
. lstand
```


Standardized Coefficients and Partial Correlations

No.	Var	Coef	OR	St.Coeff	PartCorr	Prob(z)
0	Constant	-5.2763				
1	satc_aaa	0.0003	1.0003	0.0287	0.0138	0.000
2	hsgpa_calc	1.0713	2.9192	0.2729	0.1623	0.000
3	parent_ed	0.0141	1.0142	0.0161	0.0090	0.008
4	loginc	0.1617	1.1756	0.0874	0.0464	0.000

Appendix 11: A Note on Range Restriction

Studies conducted by the national testing agencies typically adjust the observed predictive-validity coefficients for SAT scores in order to account for range restriction among admitted students. Range restriction refers to the fact that, because applicants with higher scores are more likely to be admitted, there is a limited range of test scores with which to assess the utility of the SAT in predicting college outcomes. College Board researchers attempt to deal with the problem by estimating what the outcomes might have been if all SAT takers attended college, and their studies routinely report “corrected” regression coefficients much larger than those shown here (see, for example, Mattern, 2006).

The problem with such adjustments, however, is that they depend on statistical assumptions that cannot be directly verified. Those include the assumptions of *linearity* (the relationship between test scores and college grades is assumed to be linear and identical across the observed and unobserved ranges of the data) and *homoscedasticity* (the variance around the regression line is assumed to be the same across the observed and unobserved ranges). If either assumption is incorrect – if, for example, the true relationship between SAT scores and college outcomes is curvilinear – “corrected” correlation coefficients may actually be less accurate than observed coefficients (Duan & Dunlap, 1997; Weber, 2001).

In any case, range restriction does not account for the difference in the validity coefficients for HSGPA and SAT scores observed in the UC data. Because UC has traditionally put more emphasis on HSGPA than SAT scores in admissions decisions, HSGPA is significantly more range-restricted than SAT scores. In the UC sample, the standard deviation of HSGPA for admitted students was .38 as compared with .46 among applicants, a ratio of .83. For SAT scores, the standard deviation among admitted students was 175 as compared with 187 for applicants, a ratio of .93. Thus, if the regression coefficients for HSGPA and SAT scores in this study were both “corrected” for range restriction, the predictive superiority of high school grades would be even greater than shown in the text.