SAT/ACT SCORES, HIGH-SCHOOL GPA, AND THE PROBLEM OF OMITTED VARIABLE BIAS:
Why the UC Taskforce's Findings are Spurious

March 2020*

Saul Geiser
UC Berkeley

ABSTRACT
One of the major claims of the report of University of California's Task Force on Standardized Testing is that SAT and ACT scores are superior to high-school grades in predicting how students will perform at UC. This finding has been widely reported in the news media and cited in several editorials favoring UC's continued use of SAT/ACT scores in university admissions. But the claim is spurious, the statistical artifact of a classic methodological error: *omitted variable bias.* Compared to high-school grades, SAT/ACT scores are much more strongly correlated with student demographics like family income, parental education, and race/ethnicity. As a result, when researchers omit student demographics in their prediction models, the predictive value of the tests is artificially inflated. When student demographics are included in the model, the findings are reversed: High-school grades in college-preparatory courses are actually the stronger predictor of UC student outcomes. The Task Force should go back to the drawing board and provide the UC community with more realistic estimates of the true value-added by the tests.

Keywords: Higher education, college admissions, standardized tests, race and ethnicity, socioeconomic disadvantage

One of the major claims of the report of the University of California's Task Force on Standardized Testing is that SAT/ACT scores are strong predictors of academic performance in college and, in fact, are now equal to or better than high-school grades in predicting how students will perform at UC:

> At UC, test scores are currently better predictors of first-year GPA than high school grade point average (HSGPA), and about as good at predicting first-year retention, UGPA, and graduation. (UC Academic Senate: p. 3)

The Task Force's claim is conspicuous since it differs so markedly from an extensive body of published research that concludes the opposite: High-school grades in college-preparatory subjects are far superior to SAT/ACT scores in predicting how students perform in college, and this finding holds across all student-outcome measures, from first-year grades to college graduation rates.

For example, previous research has shown that high-school GPA is decisively superior to SAT/ACT scores in predicting first-year grades at UC (Geiser & Studley, 2002). Similarly, independent studies conducted at the vast majority of U.S. colleges and universities — including almost all institutions that have gone test-optional -— find that high school grades are consistently the best predictor of first-year performance at those institutions (Koretz & Langi, 2018; Hiss & Franks, 2014; Alon & Tienda, 2007). Even the College Board and ACT, Inc., which administer the tests and have a financial incentive to show them in the most favorable light, acknowledge that high-school grades in college-preparatory subjects are the better predictor of first-year college grades at the vast majority of institutions: "In the College Board analysis of the more than 150,000 students included in all 110 [institutional] studies, HSGPA was slightly more predictive than SAT scores" (College Board, 2008: p. 1).

The predictive superiority of high-school GPA over test scores becomes even clearer when one looks at longer-term outcomes like college graduation, which many would regard as more important. Still, the definitive examination of college-completion rates is

---

* Saul Geiser is a senior associate at the Center for Studies in Higher Education at UC Berkeley and former director of admissions research for the UC system.

Bowen, Chingos, and McPherson's (2009) study, *Crossing the Finish Line*, which compared the predictive validity of high-school GPA vs. SAT/ACT scores in a massive sample of students at 54 U.S. public universities:

> 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 (pp. 113-115).

Most recently, researchers at the University of Chicago have found that high school GPA is *five times stronger* than ACT scores at predicting who will graduate from college, according to January 2020 paper published in *Educational Researcher*, the lead journal of the American Educational Research Association (Allensworth & Clark, 2020).

Omitted Variable Bias

Why do the Task Force's findings differ so markedly from the vast majority of other studies? The Task Force findings illustrate a classic methodological error: *omitted variable bias*. Berkeley economist David Card describes the problem thusly:

> Omitted variable bias occurs whenever a regression model omits variables that (1) are correlated with the variable of interest and (2) affect the outcome variable. In that circumstance, the effect of the omitted variable on the outcome may incorrectly be attributed to the variable of interest (Card, 2017: p. 59).

In the present case, SAT/ACT scores and high-school grades are the variables of interest, and UC student performance (e.g., first-year UC grades, four-year graduation) is the outcome variable. From the very first table (Figure 3A-1 on p. 18) and throughout the report, the Task Force presents a dizzying array of correlations and regression findings intended to illustrate the supposedly stronger power of test scores than high-school grades in predicting UC student outcomes.

The problem, however, is that almost all of the many correlations and regressions presented in the report — with the exception of a few tables in the appendix, to which we shall return shortly — fail to include a key control variable within the prediction model. The omitted variable? Student demographics.

Omission of student demographics like family income or parental education from regression models can significantly overstate the predictive value of SAT/ACT scores. Family income, for example, is associated not only with the variable of interest, SAT/ACT scores, but is also strongly correlated with student outcomes like college completion (Bowen, Chingos, & McPherson, 2009: p. 29). As a result, when researchers fail to control for income in their prediction models, the predictive value of the tests is artificially inflated and appears much greater than is actually the case.

In a 2004 paper in the *Journal of Econometrics*, Berkeley economist Jesse Rothstein found that "sparse" prediction models — those that consider only test scores and high-school grades but omit student demographics — substantially overestimate the predictive value of the tests

> The results here indicate that the exclusion of student background characteristics from prediction models inflates the SAT's apparent validity, as the SAT score appears to be a more effective measure of the demographic characteristics that predict UC FGPA [freshman grade-point average] than it is of preparedness conditional on student background. … [A] conservative estimate is that traditional methods and sparse models overstate the SAT's importance to predictive accuracy by 150 percent (Rothstein, 2004: p. 297).

An indication that omitted-variable bias is at work in the Task Force findings can be found in the appendix to the report, as shown in Table 1.

As Table 1 demonstrates, the choice whether to include or omit student demographics from regression analyses can decisively affect the conclusions that one draws from the data. "Sparse" prediction models such as regression equation (1), which considers grades and test scores but omits student demographics, create the false impression that SAT scores are the stronger predictor than HSGPA with respect to first-year retention, first-year UCGPA, and graduation UCGPA, and are nearly as good as HSGPA in predicting four-year graduation. But when student demographics — family income, first-generation college status, high school Academic Performance Index quintile, campus enrolled, and major discipline — are entered in the prediction model, as shown in regression equation (2), the picture is reversed. HSGPA is actually the stronger predictor of first-year retention, the likelihood of graduating in four years, and graduation UCGPA, and are the same as SAT scores in predicting first-year UCGPA.

Table 1
Standardized Regression Coefficients
for High School Grade Point Average and SAT/ACT Scores
Before and After Controlling for Student Demographics

	Dependent Variables:			
	First-Year Retention	First-Year UCGPA	Four-Year Graduation	Graduation UCGPA
Regression Equations:				
HSGPA + SAT Total (SAT RM + SATW)				
SAT/ACT Scores	0.22	0.38	0.21	0.32
HSGPA	0.20	0.22	0.19	0.26
HSGPA + SAT Total + Student Demographics				
SAT/ACT Scores	0.14	0.34	0.16	0.32
HSGPA	0.19	0.34	0.25	0.37

Source: Derived from Appendix I, Tables 3, 8, and 9 of STTF Report.

The reason why omission of student demographics in prediction models has more of an effect on the apparent validity of test scores than high-school grades is straightforward. As the National Research Council's *Reference Guide on Multiple Regression* notes, "Other things being equal, the greater the correlation between the omitted variable and the variable of interest, the greater the bias caused by the omission" (Rubinfeld, 2011: p. 314). Compared to high-school grades, SAT/ACT scores are much more closely correlated with student demographics like family income, parental education, and race/ethnicity.

Table 2
Correlation of Demographic Factors with
High School GPA versus SAT/ACT Scores

	Family Income	Parents' Education	Underrepresented Minority
High School GPA	.11	.14	-.17
SAT/ACT Scores	.36	.45	-.38

Source: Geiser, 2017: p. 2. Based on UC Corporate Student System data on California residents who applied for admission from 1994 to 2011. N = 901,905.

As a result, failure to control for student demographics in prediction models inflates the apparent validity of test scores more than it affects high school grades, since test scores serve to a much greater extent as a proxy for demographics like family income and parental education. Inclusion of the omitted variables completely reverses the conclusions one draws about the relative contribution of high school grades vs. test scores to the prediction of UC student outcomes.[i]

Both the *Los Angeles Times* (2020) and the *San Diego Union-Tribune* (2020) have cited the Task Force findings in editorials urging UC to retain the SAT and ACT in admissions. Yet the findings are spurious, a statistical artifact of omitted variable bias. It is not surprising that the College Board and ACT, Inc. would omit student demographics from their prediction models, since that choice inflates the apparent predictive value of the tests:

> [E]ven though research has shown that other student level variables such as gender, ethnicity, and SES are related to retention, we intentionally excluded those variables as predictors in our model, knowing that omitted variable bias may be present to a limited extent (Mattern, Patterson & Wyatt: 2013: p. 6).

But it is surprising that the Task Force would make the same choice. Given the Task Force's charge to explore the impact of socioeconomic and racial disparities on admissions tests, it is curious that student demographics should be the omitted variable.

Value Added by SAT/ACT Scores

Yet the omitted-variable problem goes beyond student demographics. Many academic factors can also influence predictive-validity findings. Like family income, other academic indicators on the UC application, like class rank in high school or number of UC-required a-g courses taken, are also correlated with both SAT/ACT scores and UC outcomes. Omission of other academic information from regression models can thus also lead to overestimation of the predictive value of the tests.

In 2002, BOARS (Board of Admissions and Relations with Schools, the faculty committee charge with formulating UC admissions policy) issued what is believed to be the first statement of principles of any U.S. university on admissions testing. Among those principles was this: "An admissions test should be demonstrably useful in predicting student success at UC and provide information *beyond that which is contained in other parts of the application*" [italics added]. The Task Force report interprets BOARS' principle rather more narrowly: "Admissions tests [should] add substantially to UC's ability to predict student success beyond the predictive information in high school grades alone" (p. 19).

Yet the difference is crucial. The advent of holistic review in UC admissions has added substantially to the amount and quality of applicant data available to admissions officers. In real-world admissions, the question is not what SAT/ACT scores add to high-school grades alone, but what they add to the large body of other academic and socioeconomic information already available from the UC application.

An answer is provided in a 2008 analysis by Sam Agronow, former director of policy, planning, and analysis at Berkeley. In a regression model predicting first-year grades, Agronow entered all available data from the UC application. In addition to high-school GPA and SAT scores, these included: students' course totals in the UC-required "a-g" sequence, whether the student ranked in the top four percent of their class, scores on two SAT II Subject Tests, family income, parental education, language spoken in the home, participation in academic-preparation programs, and the rank of the student's high school on the state's Academic Performance Index.

Entering all these factors into the prediction model, the analysis found that they explained 21.7 percent of the variance in students' first-year grades at Berkeley. When SAT scores were eliminated from the model, thus isolating their effect, the explained variance fell to 19.8 percent. SAT scores, in other words, accounted for less than two percent of the variance in students' first-year grades at Berkeley. Across all UC undergraduate campuses, SAT scores contributed an increment of 1.6 percentage points (Agronow, 2008, cited in Geiser, 2017: p. 9).

These findings are now over a decade old and need to be replicated with more recent UC data. But they are strikingly similar to recent findings of other colleges and universities that have examined the incremental validity of SAT/ACT scores. Institutional validity studies conducted at Johns Hopkins found that test scores increased their explained variance, or R-square, by two percentage points, raising their model's explanatory punch from an R-square of 18 to 20. At the University of Georgia, SAT scores added one percentage point, raising their R-square from 30 to 31. And at DePaul, the ACT was found to contribute one percentage point, raising their R-square from 19 to 20 (Soares, 2012).

While predictive-validity studies examining only test scores and grades may have made sense at one time, when UC admissions relied almost exclusively on those two factors, this is no longer the case. Holistic review has substantially expanded the body of both academic and socioeconomic information considered in admissions decisions, with the result that SAT/ACT scores have become increasingly redundant. They uniquely predict only a very small percentage of the variance in UC student outcomes. The Task Force should go back to the drawing board and provide the UC community with more realistic estimates of the true value-added by the tests.

_____

REFERENCES

Agronow, S. (2008). Examining the predictive value of the SAT II subject exams in the prediction of first-year UC GPA – A report to BOARS. Appendix IV (pp. 88-132) in BOARS' revised "Proposal to reform UC's freshman eligibility policy." Oakland, CA:" University of California Academic Senate. Accessed on March 22, 2016 from http://senate.universityofcalifornia.edu/underreview/sw.rev.elgibility.02.08.pdf.

Allensworth, E. M., & Clark, K. (2020). High school GPAs and ACT scores as predictors of college completion: Examining assumptions about consistency across high schools. *Educational Researcher*, forthcoming, https://journals.sagepub.com/doi/pdf/10.3102/0013189X20902110

Alon, S., & Tienda, M. (2007). Diversity, opportunity, and the shifting meritocracy in higher education. *American Sociological Review*, 72(4), 487-511.

Bowen, W. G., Chingos, M., & McPherson, M. (2009). *Crossing the Finish Line: Completing college at America's public universities.* Princeton, NJ: Princeton University Press.

Card, D. (2017). *Expert witness report on Harvard admissions,* https://news.harvard.edu/gazette/story/2018/06/documents-released-in-admissions-lawsuit/

College Board. (2008). "SAT Validity Studies." Website content downloaded October 29, 2008. http://professionals.collegeboard.com/data-reports-research/sat/validity-studies.

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

Geiser, S. (2017). Norm-referenced tests and race-blind admissions: The case for eliminating the SAT and ACT at the University of California. Research and Occasional Paper Series CSHE.15.17. Berkeley: Center for Studies in Higher Education, University of California, Berkeley.

Hiss, W. C., & Franks, V. W. (2014). Defining promise: Optional standardized testing policies in American college and university admissions. The National Association for College Admission Counseling (NACAC) Conference, https://www.iacac.org/wp-content/uploads/2014/05/H59-Defining-Promise.pdf

Koretz, D., & Langi, M. (2018). Predicting freshman grade-point average from test scores: Effects of variation within and between high schools. Educational Measurement: Issues and Practice, 37(2), 9–19.

Los Angeles Times. (2020). Editorial: Despite complaints about bias, the University of California shouldn't dump the SAT and ACT. February 1, 2020. https://www.latimes.com/opinion/story/2020-02-21/fairness-success-university-of-california-sat-act

Mattern, K. D., Patterson, B. F., & Wyatt, J. N. (2013). How useful are traditional admission measures in predicting graduation within four years? College Board Research Report 2013-1.

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

Rubinfeld, D. (2011). Reference guide on multiple regression. In National Research Council. *Reference manual on scientific evidence (3rd edition).* National Academies Press.https://www.fjc.gov/sites/default/files/2015/SciMan3D08.pdf

San Diego Union-Tribune. (2020). Editorial: University of California system shouldn't drop the SAT or ACT for admission. March 9, 2020. https://www.sandiegouniontribune.com/opinion/editorials/story/2020-03-09/editorial-university-of-california-shouldnt-drop-sat-or-act-standardized-tests

Soares, J. (2012). *SAT Wars: The Case for Test-Optional College Admissions.* New York: Teachers College Press.

University of California Academic Senate. (2020). *Report of the UC Academic Council Standardized Testing Task Force .* Oakland, CA: Systemwide Academic Senate).

---

[i] The results here likely underestimate the predictive superiority of high school grades over SAT/ACT scores for several reasons, among them the fact that many other important student demographics are not included in the model (e.g., race/ethnicity, gender, disability status, English as a second language, high school of origin). I encourage UC Office of the President to make its data available for independent researchers to evaluate this and other of the Task Force's factual claims, which have not been subjected to normal academic peer review.