Predicting College Success: How Do Different High School Assessments Measure Up?

Michal Kurlaender
K.A. Kramer
Erika Jackson

March 2018

PACE
Policy Analysis for California Education

Predicting College Success: How Do Different High School Assessments Measure Up?

Michal Kurlaender, K.A. Kramer, Erika Jackson
*University of California, Davis**

In 2014, the state of California implemented the California Assessment of Student Performance and Progress (CAASPP) in order to align state assessment and accountability policies with the newly adopted Common Core State Standards (CCSS).[1] At the heart of the new performance and accountability system is the Smarter Balanced Assessment. The Smarter Balanced Assessment is designed to evaluate a student's full range of college- and career-readiness as defined by the CCSS. Three years after the implementation of the Smarter Balanced Assessment, we can now begin to analyze how well the assessments predict college readiness. In addition, we can investigate how the Smarter Balanced Assessments measure up to other commonly used assessments for predicting college success.

In this research brief, we describe early college outcomes for the 2014-15 cohort of California 11th graders, the first cohort of 11th grade Smarter Balanced Assessment takers. Specifically, we explain how well the Smarter Balanced Assessment, high school grade point average (HSGPA), and SAT predict first-year college outcomes for students enrolled in the California State University (CSU) system. We similarly explain outcomes for students enrolled in the University of California, Davis (UCD).[2] We also report how the relationship between these assessments and early college outcomes differ by key student subgroups (race/ethnicity and socioeconomic disadvantage).

*The research reported here was supported by Institute of Education Sciences, U.S. Department of Education, through Grant R305E150006, to the Regents of the University of California, and by The James Irvine Foundation. The opinions expressed are those of the authors and do not represent views of The Irvine Foundation, the Institute or the U.S. Department of Education, or of the agencies providing data. We thank the California Department of Education, the California State University Chancellor's Office, and the University of California Davis Office for Institutional Research for providing data access and expertise. We also thank the College Board for their assistance with restricted range adjustments. All errors are our own.

[1] California adopted the Common Core State Standards in August 2010 and subsequently aligned the state assessments in 2014.

[2] At the time of this analysis and report, we only have data for one University of California campus.

Data Description

Our primary analytic sample includes California 11th grade students who took both the Smarter Balanced Assessment and SAT, and who subsequently applied and enrolled as first time freshmen in 2016-17 at one of the 23 campuses of the CSU system. We use data from the California Department of Education (CDE) for all 11th graders who took the Smarter Balanced Assessment in the 2014-15 academic year. We match these individual-level 11th grade assessment data to application and enrollment data (which includes SAT scores) from the CSU Chancellor's Office.[3] In addition, we include individual-level demographic data (available from the 11th grade assessment files), HSGPA (available from the CSU application files), and high school characteristics from public-use data (available from the CDE).

We also have a smaller analytic sample of California 11th graders who took the Smarter Balanced Assessment in 2014-15, who also took the SAT, and who subsequently enrolled as first time freshmen in 2016-17 at the University of California, Davis (UCD). Similar to the CSU sample, we use data for all 11th graders who took Smarter Balanced Assessment in the 2014-15 academic year and match these individual-level 11th grade assessment data to application and enrollment data from UCD.[4] For this data set, we similarly include individual-level demographic data (available from the 11th grade assessment files), HSGPA (available from the UCD application files), and high school characteristics from public-use data (available from the CDE).

Measures

We focus on three early college outcomes: (1) first year college performance, measured by grade point average (GPA); (2) persistence to year two; and (3) total units accumulated in the first year. The primary predictors are High School Grade Point Average (HSGPA), 11th grade Smarter Balanced Assessment Consortium (SBAC) scores in math and English language arts/literacy (ELA); and SAT scores (verbal, math, and writing). We also include controls for key individual characteristics: gender (male/female); race/ethnicity (five mutually exclusive categories: Asian/Pacific Islander, Black/African American, Latino/Hispanic, White, and Other); socioeconomic disadvantage in high school (yes/no);[5] and English Learner (yes/no). In addition, our analysis includes high school quality as measured by the state's new College/Career Indicator (CCI) levels (Very Low, Low, Medium, High, Very High).[6] Finally, for the CSU analysis, we adjust

[3] Using student name, date of birth, gender, and high school attended, we are able to match 82 percent of Fall 2016 CSU applicants from California public high schools to the 11th grade assessment data.

[4] Using student name, date of birth, gender, and high school attended, we are able to match 82 percent of Fall 2016 UCD applicants from California public high schools, to the 11th grade assessment data.

[5] Students who meet the definition of socioeconomic disadvantaged (SED) either qualify for the free or reduced price school lunch program or do not have a parent who graduated from high school.

[6] The College/Career Indicator (CCI) is one of six indicators included in the new California School Dashboard. The CCI is first calculated at the student level, where students are labeled as *Prepared*, *Approaching Prepared*, or *Not Prepared* based on their grade 11th grade ELA and Math SBAC Scores, CTE Pathway Completion, AP and IB Exams, Dual Enrollment, and A-G Course Completion. A school-level indicator is then determined by considering the proportion of students in the graduating cohort that earned a *Prepared* status. (Continued on next page)

for campus differences to account for the host of other experiences students have depending on the CSU campus they choose to enroll in, and which may also influence our outcomes of interest. Descriptive statistics on the CSU analytic sample are presented in Table A1 and for the UCD sample in Table A2.[7]

Smarter Balanced Assessment

California first implemented the Smarter Balanced Assessment in Spring 2015. The Smarter Balanced Assessment includes three major components designed to improve teaching and learning: 1) an online library of formative assessments for use by teachers; 2) interim assessments for use by schools or districts to monitor student progress towards meeting standards; 3) a summative assessment administered annually to determine students' mastery of college and career readiness standards in ELA and math.[8,9] The annual summative assessment is the cornerstone of the Smarter Balanced Assessment and is administered to all students statewide in grades 3 through 8 and grade 11, with some exceptions.[10] The Smarter Balanced Assessment is delivered via computer and include both a computer-adaptive test and a performance task based on the CCSS for ELA and mathematics.[11] The computer-adaptive section includes a range of item types such as selected response, constructed response, table matching and fill-in, graphing, and drag and drop. The Performance tasks are extended activities that measure a student's ability to integrate knowledge and skills across multiple standards—a key component of college and career readiness. The estimated time for the 11th grade ELA and math tests combined is 7.5 hours.[12] The cohort included in this analysis is the first cohort in California to participate in the 11th grade Smarter Balanced Assessment.

SAT

The SAT is a college entrance exam accepted or required at nearly all four-year colleges and universities in the U.S. It measures students' college readiness and is validated as a reliable predictor of college outcomes, primarily first year GPA.[13] Most students take the SAT for the first time during spring of their junior year and a second time during the fall of their senior year. In

---

For example, schools with less than 10 percent of their graduating cohort earning *Prepared* receive a *Very Low* school-level indicator and schools with more than 70 percent of the graduating cohort receiving *Prepared* are considered *Very High* on the school-level indicator. For more detailed information on the College/Career Indicator (CCI) see the description and Technical Guide available from the CDE: https://www.cde.ca.gov/ta/ac/cm/cci.asp

[7] Univariate correlations are available in the Appendix, Table A3 and A4.

[8] https://www.cde.ca.gov/ta/tg/sa/

[9] Reliability coefficeints for 11th grade Smarter Balanced Assesment in 2014-2015 are .91 for English Language Arts and .89 for Math. https://www.cde.ca.gov/ta/tg/ca/documents/caaspp14techrpt.pdf (p. 281).

[10] Students with disabilities who participate in the alternate assessments are not required to participate in the Smarter Balanced Assessment. Additionally, English learners who are in 12 months of attending a school in the United States are not required to participate in the ELA test.

[11] https://www.cde.ca.gov/ta/tg/sa/sbacsummative.asp

[12]See: https://portal.smarterbalanced.org/library/en/estimated-testing-times.pdf

[13] See details for validity studies provided by college board: https://research.collegeboard.org/programs/sat/data/validity-studies

California, 86 percent of CSU applicants and 90 percent of UC applicants take the SAT. The format of the SAT for our sample consists of three sections: Mathematics, Critical Reading, and Writing (including an essay).[14] Each of the three respective sections is scored on a scale of 200-800, with a total score range of 600-2400. The test is administered on paper and is 3.75 hours long. Our analytical sample (the first 11th grade Smarter Balanced Asssessment takers) were the last cohort to take the SAT before its major redesign.

A redesigned SAT was launched in March 2016.[15] The new test has only two sections: Math and Evidence-Based Reading and Writing. The total score scale ranges from 400-1600, with an optional essay scored separately. The tests are still administered on paper and students are given 3 hours (+50 minutes for the optional essay) to complete the exam. Given the new format of the SAT, a follow up study, similar to the one presented here, is advised.

Data Analysis

We follow a standard validity approach to investigate the relationship between the 11th grade assessments and early college outcomes. Specifically, we fit a series of adjusted multiple correlations to examine the strength of the relationship between HSGPA, SBAC, and SAT, respectively, and our outcomes of interest, controlling for other variables. These multiple correlations tell us how well we can predict early college outcomes on the basis of high school performance indicators; for example, how well can we predict first year GPA on the basis of a test score. We compare the strength of these relationships across different specifications, using different predictor and control variables (e.g., HSGPA versus SBAC, or SBAC versus SAT, both with and without demographic control variables).[16]

Because applicants with higher HSGPAs or test scores are more likely to be admitted to college, our sample of college enrollees does not contain the full range of grades and test scores as the census of California 11th grade test takers. It is standard in validity studies to adjust for this "restricted range," which can result in a distortion of the true relationship between the predictors (e.g., high school assessments) and outcomes (e.g., college GPA).[17,18] We follow the standard practice of the College Board by adjusting for the restricted range of both of the assessments we investigate (SBAC and SAT). We provide additional information on our methodology in the Appendix. Finally, it is important to note that these analyses are entirely correlational, highlighting the association between each assessment and early college outcomes, and should not be interpreted as causal.

---

[14] Reliability coefficeints for the SAT are .92-.93 for Math, .91-.92 for Critical Reading, and .88-.91 for Writing. See: http://media.collegeboard.com/digitalServices/pdf/SAT-Test-Characteristics_of_SAT_2011.pdf

[15] See: https://collegereadiness.collegeboard.org/pdf/redesigned-sat-pilot-predictive-validity-study-first-look.pdf

[16] We also do this in a multiple regression framework to present R-squared statistics of the variance explained by these models.

[17] The narrowing of a score range by selection results in an underestimation of the true relationship between the predictor(s) and criterion (e.g., first year GPA).

[18] See: Shaw, Marini, Beard, Shmueli, Young & Ng, 2016; Mattern, Kobrin, Patterson, Shaw & Camara, 2009.

Results

California State University Results

Table 1 presents the adjusted (and raw) multiple correlation coefficients of the main predictors (HSGPA, SBAC, and SAT) and first year GPA (Panel A), 2nd year persistence rate (Panel B), and total credits accumulated in the first year (Panel C) for the CSU analysis. Each cell of the table reflects a separate model. Column 1 includes only the respective predictor variables without additional controls. Column 2 includes the respective predictor variables and controls for student demographics (i.e., gender, race/ethnicity, socioeconomic disadvantage, English Learner status). Column 3 includes the high school CCI indicator, while Column 4 also includes controls for enrollment at different CSU campuses.

This table presents several important findings. First, (as is expected in such validity studies), correlations adjusted for restricted range are higher than raw correlations. Second, HSGPA on its own is more strongly correlated to first-year college GPA than either SBAC or SAT scores by themselves. Specifically, we see in Column 1 of Panel A that the adjusted correlation coefficient of HSGPA and 1st year college GPA is .45, while the adjusted correlation of SAT and first-year GPA is .38 and SBAC and first-year GPA is .37. This finding is consistent with similar validity studies by Black et al. (2016) and Mattern and Patterson (2014). Third, the multiple correlations that include HSGPA show that SBAC and SAT are quite similar (if not identical) in their magnitude of association to first-year college GPA, (both .48 in column 1 with no additional controls, and both .51 in column 4 with all of the controls). Finally, we note that there is no meaningful benefit of including both SBAC and SAT. The multiple correlation of HSGPA, SAT, and SBAC with first-year GPA is .51, which is the same as the value when just SAT or SBAC are included with HSGPA.[19] The strength of the associations we find in this analysis is similar to that found in other validity studies; correlation coefficients—corrected for restricted range—between admission test scores and college GPA tend to be in the .40s and .50s (Shaw et al., 2016; Kuncel & Hezlett, 2007; Mattern & Patterson, 2014).

In looking at second year persistence rates (Panel B), magnitudes are much reduced relative to first year GPA, which is also typical of studies validating admissions tests such as the SAT (Mattern & Patterson, 2014). The overall patterns are similar to those found in Panel A; HSGPA on its own is more strongly correlated to second year persistence rates (.22) than either SBAC (.20) or SAT (.19) on their own. SBAC and SAT are similar in their magnitude of association to second year persistence rates when HSGPA and other controls are included, both at .25 with all of the additional controls (column 4). Including both SBAC and SAT in models predicting first year persistence does not offer any predictive benefit.

[19] F statistics are available from the authors by request.

Table 1: Multiple Correlation Coefficients, Adjusted (Raw) for CSU Analysis

Panel A: First Year GPA (N=36,519)

	1	2	3	4
HSGPA	.45 (.35)	.48 (.39)	.48 (.39)	.49 (.40)
SAT	.38 (.28)	.41 (.33)	.42 (.33)	.42 (.34)
SBAC	.37 (.28)	.41 (.33)	.41 (.33)	.42 (.34)
HSGPA & SAT	.48 (.38)	.50 (.41)	.50 (.41)	.51 (.42)
HSGPA & SBAC	.48 (.38)	.50 (.41)	.50 (.41)	.51 (.42)
HSGPA, SAT, & SBAC	.49 (.39)	.50 (.41)	.51 (.41)	.51 (.42)
Demographics	N	Y	Y	Y
HS CCI	N	N	Y	Y
Campus FE	N	N	N	Y

Panel B: Persistence To Second Year (N=43,791)

	1	2	3	4
HSGPA	.22 (.17)	.24 (.18)	.24 (.19)	.25 (.20)
SAT	.19 (.14)	.21 (.16)	.21 (.16)	.22 (.17)
SBAC	.20 (.15)	.21 (.17)	.22 (.17)	.22 (.18)
HSGPA & SAT	.24 (.18)	.24 (.19)	.25 (.20)	.25 (.20)
HSGPA & SBAC	.24 (.19)	.25 (.20)	.25 (.20)	.25 (.21)
HSGPA, SAT, & SBAC	.24 (.19)	.25 (.20)	.25 (.20)	.25 (.21)
Demographics	N	Y	Y	Y
HS CCI	N	N	Y	Y
Campus FE	N	N	N	Y

Panel C: Total Units, End of First Year (N=41,573)

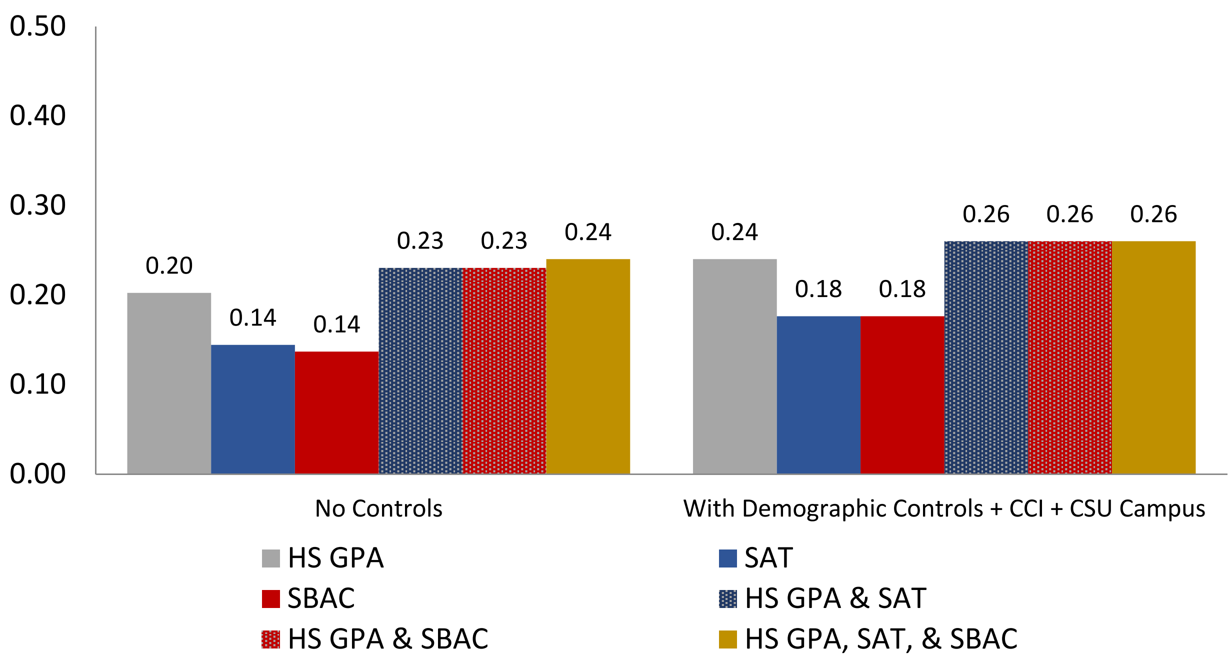
	1	2	3	4
HSGPA	.50 (.39)	.54 (.43)	.58 (.45)	.60 (.50)
SAT	.60 (.51)	.61 (.52)	.63 (.52)	.63 (.55)
SBAC	.55 (.46)	.56 (.47)	.59 (.47)	.60 (.52)
HSGPA & SAT	.64 (.55)	.64 (.55)	.66 (.55)	.66 (.58)
HSGPA & SBAC	.61 (.51)	.61 (.52)	.63 (.52)	.64 (.55)
HSGPA, SAT, & SBAC	.64 (.55)	.65 (.56)	.66 (.56)	.67 (.58)
Demographics	N	Y	Y	Y
HS CCI	N	N	Y	Y
Campus FE	N	N	N	Y

Notes: Sample sizes differ by outcome due to data constraints (first year GPA is only reported in the data if student returns for Year 2); Total Units include students enrolled in at least one term during the first year.

The last set of multiple correlations, where total units accumulated in the first year of college is the outcome of interest (Panel C), reveal a pattern similar to the other outcomes. With this outcome, however, we see a clearer advantage for the SAT in terms of the magnitude of the association. Specifically, in Panel C, column 1 the adjusted correlation of SAT and total units in year 1 is .60. The adjusted correlation between HSGPA and total units in Year 1 is .50, and between SBAC and total units is .55. When we look at each test score with HSGPA, we again note a slight advantage for SAT in predicting total units, at .66, compared to SBAC at .64 (column 4 with full controls). Overall, including both SBAC and SAT to predict total first-year units, offers no substantively meaningful contribution, albeit a statistically significant one, versus just including SAT and HSGPA.

In Figure 1 we present the R-squared values from multiple regression analyses based on the same models presented in Table 1 (Panel A: Columns 1 and 4), because presenting our results in terms of the percentage of variation in first year GPA explained by each of the predictors may be more familiar to some readers. For example, looking at the bars on the far left (No Controls), we see that HSGPA explains 20% of the variation in first year GPA among CSU freshmen, whereas SAT and SBAC each explain about 14% of the variation in first year GPA. The bars on the right include demographic controls, high school CCI levels, and CSU campus differences. As a result of these additional controls, more of the variance in first year GPA is being explained, but HSGPA remains the dominant predictor and SAT and SBAC explain virtually the same overall variation in college GPA. R-squared values for the other models presented in Table 1 can be obtained by squaring the multiple correlations presented in Table 1.

Figure 1: Proportion of Variation in First Year GPA Explained, for CSU Analysis



Exploring subgroup differences at CSU

Tables 2 and 3 present results from our fully specified models (those that include all controls as in Table 1, Column 4) by key student subgroups: race/ethnicity and socioeconomic disadvantage. Though the relationships between the assessments and early college outcomes for specific racial/ethnic subgroups (Latino students in particular) are weaker, our results overall reveal virtually the same pattern across all racial/ethnic groups. There are no meaningful differences between SAT and SBAC scores in predicting first year GPA and second year persistence (i.e., the coefficients never differ by more than .01). As with our main results, there are slightly larger differences between the two assessments in predicting total units in the first year; SAT is the strongest predictor of this outcome when compared to both HSGPA and SBAC for all racial/ethnic subgroups. Results by socioeconomic disadvantage similarly reveal no difference in the overall pattern of the predictive power of the assessments between groups. Here, however, we note more substantial differences in magnitudes, with consistently lower associations between high school assessments and first year outcomes for low income students than for more advantaged students.

Finally, the patterns of association controlling for school quality (as measured by the CCI) in Table 4 suggest similar relationships between each of the high school assessments and our outcomes of interest, regardless of the quality of the high school. HSGPA and both standardized tests are much stronger predictors of first-year college GPA and second-year persistence rates for students at the lowest CCI schools, when compared to higher CCI schools, which suggests that these assessments help identify college-ready students at some of the state's weakest high schools.

Table 2: Multiple Correlation Coefficients, Adjusted (Raw) for CSU Analysis, by Race/Ethnicity

Panel A: First Year GPA

	All (N= 36,519)	Race				
		Asian & PI (N= 7,163)	Black (N=1,725)	Latino (N=17,787)	White (N=8,754)	Other (N=1,090)
HS GPA	.49 (.40)	.47 (.38)	.45 (.37)	.41 (.32)	.51 (.45)	.49 (.42)
SAT	.42 (.34)	.41 (.32)	.38 (.32)	.34 (.27)	.38 (.33)	.44 (.38)
SBAC	.42 (.34)	.42 (.34)	.39 (.33)	.33 (.26)	.39 (.34)	.45 (.40)
HS GPA & SAT	.51 (.42)	.50 (.41)	.47 (.39)	.44 (.35)	.52 (.46)	.52 (.45)
HS GPA & SBAC	.51 (.42)	.50 (.42)	.47 (.40)	.43 (.34)	.52 (.46)	.53 (.46)
HS GPA, SAT, & SBAC	.51 (.42)	.51 (.42)	.48 (.41)	.44 (.35)	.52 (.46)	.53 (.46)

Panel B: Persistence to Second Year

	All (N= 43,791)	Race				
		Asian & PI (N= 8,093)	Black (N=2,189)	Latino (N=21,962)	White (N=10,240)	Other (N=1,307)
HS GPA	.25 (.20)	.23 (.19)	.22 (.18)	.23 (.18)	.23 (.21)	.22 (.21)
SAT	.22 (.17)	.21 (.17)	.21 (.17)	.20 (.15)	.19 (.17)	.20 (.19)
SBAC	.22 (.18)	.21 (.18)	.21 (.18)	.20 (.16)	.19 (.18)	.21 (.20)
HS GPA & SAT	.25 (.20)	.24 (.20)	.23 (.19)	.24 (.19)	.23 (.21)	.23 (.21)
HS GPA & SBAC	.25 (.21)	.24 (.20)	.23 (.20)	.24 (.19)	.24 (.21)	.23 (.22)
HS GPA, SAT, & SBAC	.25 (.21)	.24 (.20)	.23 (.20)	.24 (.19)	.24 (.21)	.24 (.22)

Panel C: Total Units, End of First Year

	All (N= 41,573)	Race				
		Asian & PI (N= 7,838)	Black (N=2,060)	Latino (N=20,695)	White (N=9,744)	Other (N=1,236)
HS GPA	.60 (.50)	.62 (.51)	.43 (.35)	.48 (.39)	.62 (.57)	.58 (.51)
SAT	.63 (.55)	.68 (.58)	.45 (.38)	.51 (.43)	.65 (.60)	.63 (.57)
SBAC	.60 (.52)	.65 (.55)	.43 (.37)	.48 (.40)	.61 (.57)	.60 (.55)
HS GPA & SAT	.66 (.58)	.70 (.61)	.48 (.41)	.55 (.47)	.69 (.64)	.66 (.59)
HS GPA & SBAC	.64 (.55)	.68 (.58)	.48 (.40)	.53 (.45)	.66 (.61)	.64 (.57)
HS GPA, SAT, & SBAC	.67 (.58)	.71 (.61)	.49 (.42)	.55 (.47)	.69 (.64)	.66 (.60)

Notes: All correlation coefficients presented include demographic controls (gender, socioeconomic disadvantage, English Learner status), CSU campus differences (fixed effects), and high school CCI levels.

Table 3: Multiple Correlation Coefficients, Adjusted (Raw) for CSU Analysis, by Socioeconomic Disadvantage

Panel A: First Year GPA

	All (N= 36,519)	Socio-Economic Disadvantaged Status	
		NOT SED (N= 17,402)	SED (N=19,117)
HS GPA	.49 (.40)	.51 (.43)	.43 (.34)
SAT	.42 (.34)	.42 (.34)	.37 (.29)
SBAC	.42 (.34)	.42 (.35)	.36 (.28)
HS GPA & SAT	.51 (.42)	.52 (.44)	.46 (.37)
HS GPA & SBAC	.51 (.42)	.53 (.44)	.45 (.36)
HS GPA, SAT, & SBAC	.51 (.42)	.53 (.45)	.46 (.37)

Panel B: Persistence to Second Year

	All (N= 43,791)	Socio-Economic Disadvantaged Status	
		NOT SED (N= 20,153)	SED (N=23,638)
HS GPA	.25 (.20)	.23 (.19)	.23 (.18)
SAT	.22 (.17)	.20 (.16)	.20 (.15)
SBAC	.22 (.18)	.21 (.17)	.21 (.16)
HS GPA & SAT	.25 (.20)	.24 (.20)	.24 (.19)
HS GPA & SBAC	.25 (.21)	.24 (.20)	.24 (.19)
HS GPA, SAT, & SBAC	.25 (.21)	.24 (.20)	.25 (.19)

Panel C: Total Units, End of First Year

	All (N= 41,573)	Socio-Economic Disadvantaged Status	
		NOT SED (N= 19,269)	SED (N=22,304)
HS GPA	.60 (.50)	.63 (.54)	.48 (.38)
SAT	.63 (.55)	.66 (.59)	.52 (.43)
SBAC	.60 (.52)	.63 (.55)	.49 (.40)
HS GPA & SAT	.66 (.58)	.70 (.62)	.56 (.46)
HS GPA & SBAC	.64 (.55)	.67 (.59)	.54 (.44)
HS GPA, SAT, & SBAC	.67 (.58)	.70 (.62)	.56 (.47)

Notes: All correlation coefficients presented include demographic controls (gender, race/ethnicity, English Learner status), CSU campus differences (fixed effects), and high school CCI levels.

Table 4: Multiple Correlation Coefficients, Adjusted (Raw) for CSU Analysis, School CCI Levels

Panel A: First Year GPA

	All	School CCI Levels				
	(N= 36,519)	Very Low (N=84)	Low (N=6,560)	Medium (N=17,281)	High (N=8,780)	Very High (N=3,683)
HS GPA	.49 (.40)	.67 (.56)	.47 (.36)	.46 (.37)	.50 (.42)	.52 (.45)
SAT	.42 (.34)	.59 (.52)	.37 (.29)	.39 (.32)	.42 (.34)	.44 (.37)
SBAC	.42 (.34)	.61 (.53)	.36 (.29)	.39 (.32)	.42 (.35)	.45 (.39)
HS GPA & SAT	.51 (.42)	.68 (.57)	.49 (.39)	.48 (.40)	.52 (.44)	.53 (.46)
HS GPA & SBAC	.51 (.42)	.68 (.57)	.48 (.38)	.47 (.39)	.52 (.44)	.53 (.47)
HS GPA, SAT, & SBAC	.51 (.42)	.70 (.59)	.49 (.39)	.48 (.4)	.52 (.44)	.53 (.47)

Panel B: Persistence to Second Year

	All	School CCI Levels				
	(N= 43,791)	Very Low (N=112)	Low (N=8,200)	Medium (N=20,899)	High (N=10,235)	Very High (N=4,160)
HS GPA	.25 (.20)	.48 (.47)	.24 (.19)	.23 (.19)	.25 (.21)	.23 (.21)
SAT	.22 (.17)	.50 (.48)	.21 (.16)	.20 (.16)	.22 (.18)	.22 (.20)
SBAC	.22 (.18)	.50 (.49)	.22 (.17)	.20 (.16)	.22 (.18)	.22 (.20)
HS GPA & SAT	.25 (.20)	.51 (.49)	.26 (.20)	.24 (.19)	.25 (.22)	.24 (.22)
HS GPA & SBAC	.25 (.21)	.51 (.49)	.26 (.21)	.24 (.19)	.25 (.22)	.24 (.22)
HS GPA, SAT, & SBAC	.25 (.21)	.55 (.52)	.26 (.21)	.24 (.20)	.26 (.22)	.24 (.22)

Panel C: Total Units, End of First Year

	All	School CCI Levels				
	(N= 41,573)	Very Low (N=105)	Low (N=7,695)	Medium (N=19,864)	High (N=9,735)	Very High (N=4,004)
HS GPA	.60 (.50)	.54 (.51)	.48 (.38)	.55 (.46)	.62 (.53)	.61 (.51)
SAT	.63 (.55)	.49 (.48)	.51 (.44)	.59 (.52)	.65 (.57)	.64 (.54)
SBAC	.60 (.52)	.49 (.49)	.49 (.42)	.55 (.48)	.62 (.54)	.61 (.51)
HS GPA & SAT	.66 (.58)	.56 (.52)	.56 (.47)	.63 (.55)	.68 (.61)	.67 (.58)
HS GPA & SBAC	.64 (.55)	.56 (.52)	.54 (.46)	.60 (.52)	.66 (.58)	.65 (.55)
HS GPA, SAT, & SBAC	.67 (.58)	.59 (.54)	.56 (.48)	.63 (.55)	.69 (.61)	.68 (.58)

Notes: All correlation coefficients presented include demographic controls (gender, race/ethnicity, socioeconomic disadvantage, English Learner status), CSU campus differences (fixed effects).

University of California, Davis Results

We conducted the same analyses with data from the University of California, Davis (UCD), where the student population is closely representative of the UC population as a whole. The percentage of underrepresented minority students (defined as Black/African American, Hispanic/Latino, or American Indian) in the Fall 2016 undergraduate cohort at UCD is 26 percent, compared to 30 percent across the UC system. The percentage of students who are first generation college-going is 44 percent at UCD and 43 percent for the UC overall. The percentage of students who receive Pell grants is the same at UCD as for the UC system as a whole. SAT Critical Reading and Math scores at UCD are also comparable to UC-wide rates among both applicants and enrollees.[20]

Results presented for UCD in Table 5 are set up in the same way as those presented for the CSU outcomes in Table 1. (There are three rather than four columns because we do not need to control for different campuses.) The UCD analysis overall reveals stronger associations between the predictors and first year GPA, relative to the CSU sample. We also note stronger associations between each of the assessments (SBAC and SAT) and first year GPA, relative to HSGPA. Specifically, in column 1 we note that the adjusted correlation between HSGPA and first-year college GPA is .54, whereas the adjusted correlation between SAT and first-year college GPA is .61 and SBAC and first-year college GPA is .56. We also note that the SAT has a stronger association with first year GPA than does the SBAC, even when controlling for HSGPA, but that the difference is reduced to .02 in models that include all the controls (column 3). There appears to be no benefit in predictive power to including both tests (SBAC and SAT). Specifically, the adjusted multiple correlation of HSGPA and SAT in predicting first-year college GPA is identical to the multiple correlation that includes SBAC as well (.67).

Results for second-year persistence (Panel B) reveal the same pattern as Panel A, but with weaker magnitudes overall, and with even greater similarities between SAT and SBAC. Specifically, we observe in Panel B column 1 that the adjusted correlation coefficient of HSGPA and second-year college persistence at UCD is .24, whereas the adjusted correlation of SBAC and second-year persistence rate at UCD is .28, and with SAT it is.29. In models with all the additional controls (column 3), SBAC and SAT are again very similar in their magnitude of association to second-year persistence rates at UCD, SBAC at .31 and SAT at .32. Including both SBAC and SAT does not offer an improvement in predicting second year persistence rates at UCD, versus just including one or the other test (SAT or SBAC).

[20] See sources: https://www.universityofcalifornia.edu/infocenter/fall-enrollment-glance; https://accountability.universityofcalifornia.edu/2016/chapters/chapter-1.html#1.3.4

Table 5: Multiple Correlation Coefficients, Adjusted (Raw) for UC Davis Analysis

Panel A: First Year GPA (N=3,544)

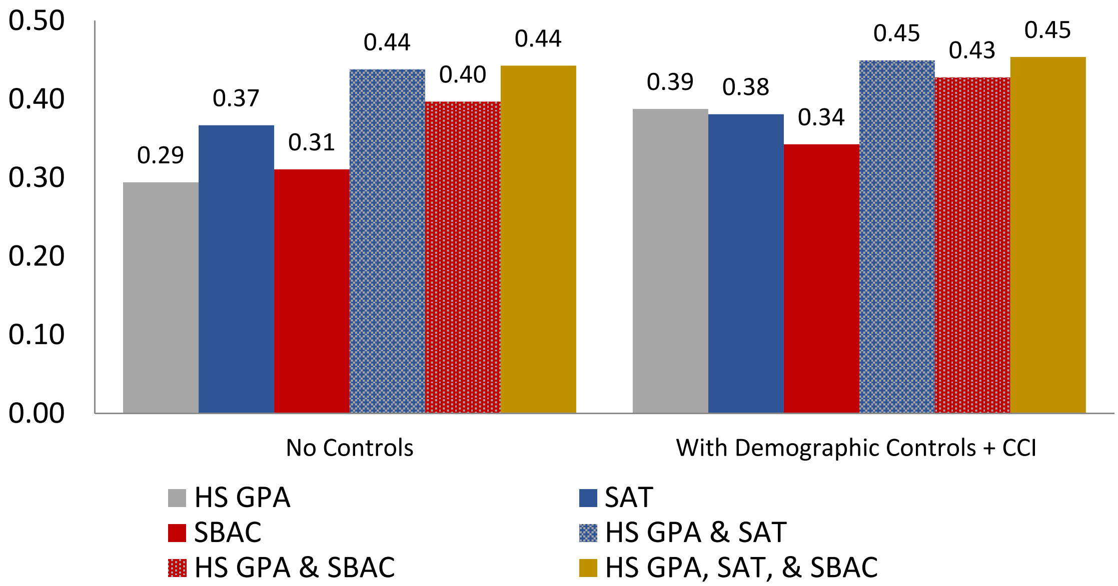
	1	2	3
HSGPA	.54 (.32)	.60 (.42)	.62 (.46)
SAT	.61 (.49)	.62 (.50)	.62 (.50)
SBAC	.56 (.44)	.58 (.47)	.59 (.48)
HSGPA & SAT	.66 (.52)	.67 (.53)	.67 (.54)
HSGPA & SBAC	.63 (.47)	.65 (.50)	.65 (.51)
HSGPA, SAT, & SBAC	.67 (.53)	.67 (.53)	.67 (.54)
Demographics	N	Y	Y
HS CCI	N	N	Y

Panel B: Persistence To Second Year (N=3,544)

	1	2	3
HSGPA	.24 (.13)	.27 (.17)	.29 (.20)
SAT	.29 (.22)	.29 (.22)	.30 (.23)
SBAC	.28 (.21)	.29 (.22)	.29 (.23)
HSGPA & SAT	.31 (.23)	.31 (.23)	.32 (.24)
HSGPA & SBAC	.30 (.22)	.31 (.23)	.31 (.23)
HSGPA, SAT, & SBAC	.31 (.23)	.32 (.24)	.32 (.24)
Demographics	N	Y	Y
HS CCI	N	N	Y

In Figure 2 we present the proportion of the variance explained in first year GPA among UCD freshmen on the basis of HSGPA, SAT, and SBAC (Table 5, Panel A: Columns 1 and 3). The bars on the far left, with no control variables, highlight the significantly greater proportion of the variation explained by the SAT, which is attenuated in models that include a second assessment, and also in the models on the right that include additional controls.

Figure 2: Proportion of Variation in First Year GPA Explained, for UCD Analysis



Exploring subgroup differences at UCD

We also conducted an analysis by subgroups for the UCD sample. Tables 6 through 8 provide results from the fully specified models (those that include all controls as in Table 5 Column 3) by key student subgroups: race/ethnicity (Table 6), socioeconomic disadvantage (Table 7), and high school CCI (Table 8). In general, results reveal similar patterns across all race/ethnicity and socioeconomic status groups. The SAT is the stronger predictor of first-year GPA, relative to SBAC, but that difference shrinks to a very small magnitude (or even closes) for some racial/ethnic subgroups in models that include the full set of controls. There is a less consistent pattern for predicting second-year persistence rates, with much lower magnitudes across all models and subgroups. Patterns by school quality (measured by the CCI) in Table 8 suggest a stronger SAT association (relative to SBAC) for students attending lower CCI high schools, but the pattern is not consistent for second-year persistence.

Table 6: Multiple Correlation Coefficients, Adjusted (Raw) for UCD Analysis, by Race/Ethnicity

Panel A: First Year GPA

	All (N= 3,544)	Race				
		Asian & PI (N= 1,322)	Black (N=125)	Latino (N=1,081)	White (N=822)	Other (N=194)
HS GPA	.62 (.46)	.63 (.46)	.64 (.45)	.46 (.31)	.62 (.45)	.62 (.48)
SAT	.62 (.50)	.63 (.51)	.49 (.37)	.48 (.38)	.57 (.46)	.60 (.50)
SBAC	.59 (.48)	.62 (.52)	.46 (.39)	.42 (.33)	.50 (.41)	.56 (.49)
HS GPA & SAT	.67 (.54)	.69 (.55)	.65 (.48)	.52 (.40)	.66 (.52)	.66 (.54)
HS GPA & SBAC	.65 (.51)	.69 (.55)	.65 (.47)	.49 (.35)	.64(.49)	.65 (.52)
HS GPA, SAT, & SBAC	.67 (.54)	.70 (.57)	.65 (.48)	.52 (.40)	.66 (.52)	.66 (.55)

Panel B: Persistence to Second Year

	All (N= 3,544)	Race				
		Asian & PI (N= 1,322)	Black (N=2,189)	Asian & PI (N= 1,322)	White (N=10,240)	Other (N=194)
HS GPA	.29 (.20)	.19 (.17)	.39 (.37)	.24 (.16)	.33 (.23)	.35 (.29)
SAT	.30 (.23)	.22 (.21)	.39 (.39)	.28 (.23)	.29 (.23)	.38 (.34)
SBAC	.29 (.23)	.25 (.24)	.38 (.38)	.24 (.19)	.29 (.23)	.32 (.30)
HS GPA & SAT	.32 (.24)	.23 (.21)	.42 (.40)	.30 (.23)	.35 (.26)	.40 (.34)
HS GPA & SBAC	.31 (.23)	.25 (.24)	.40 (.38)	.27 (.20)	.35 (.26)	.36 (.30)
HS GPA, SAT, & SBAC	.32 (.24)	.25 (.24)	.44 (.42)	.30 (.24)	.36 (.27)	.40 (.35)

Notes: All correlation coefficients presented include demographic controls (gender, socioeconomic disadvantage, English Learner status), and high school CCI levels.

Table 7: Multiple Correlation Coefficients, Adjusted (Raw) for UCD Analysis, by Socioeconomic Disadvantage

Panel A: First Year GPA

	All (N= 3,544)	Socio-Economic Disadvantaged Status	
		NOT SED (N= 1,954)	SED (N=1,590)
HS GPA	.62 (.46)	.59 (.41)	.54 (.37)
SAT	.62 (.50)	.58 (.45)	.54 (.44)
SBAC	.59 (.48)	.54 (.44)	.49 (.40)
HS GPA & SAT	.67 (.54)	.64 (.50)	.60 (.47)
HS GPA & SBAC	.65 (.51)	.63 (.48)	.58 (.43)
HS GPA, SAT, & SBAC	.67 (.54)	.65 (.51)	.60 (.47)

Panel B: Persistence to Second Year

	All (N= 3,544)	Socio-Economic Disadvantaged Status	
		NOT SED (N= 1,954)	SED (N=1,590)
HS GPA	.29 (.20)	.25 (.19)	.27 (.18)
SAT	.30 (.23)	.25 (.21)	.30 (.24)
SBAC	.29 (.23)	.26 (.21)	.27 (.22)
HS GPA & SAT	.32 (.24)	.27 (.22)	.32 (.25)
HS GPA & SBAC	.31 (.23)	.28 (.22)	.30 (.23)
HS GPA, SAT, & SBAC	.32 (.24)	.28 (.23)	.32 (.25)

Notes: All coefficients presented include demographic controls (gender, race/ethnicity, English Learner status), and high school CCI levels.

Table 8: Multiple Correlation Coefficients, Adjusted (Raw) for UCD Analysis, by School CCI Levels

Panel A: First Year GPA

	All (N= 3,544)	School CCI Levels: Very Low*	Low (N=531)	Medium (N=1,313)	High (N=860)	Very High (N=604)
HS GPA	.62 (.46)		.56 (.38)	.56 (.38)	.59 (.42)	.59 (.41)
SAT	.62 (.50)		.52 (.42)	.56 (.44)	.56 (.44)	.58 (.43)
SBAC	.59 (.48)		.46 (.37)	.51 (.40)	.55 (.44)	.58 (.45)
HS GPA & SAT	.67 (.54)		.61 (.47)	.63 (.48)	.64 (.49)	.63 (.48)
HS GPA & SBAC	.65 (.51)		.59 (.43)	.60 (.44)	.63 (.48)	.63 (.48)
HS GPA, SAT, & SBAC	.67 (.54)		.62 (.47)	.63 (.48)	.64 (.50)	.64 (.49)

Panel B: Persistence to Second Year

	All (N= 3,544)	School CCI Levels: Very Low*	Low (N=531)	Medium (N=1,313)	High (N=860)	Very High (N=604)
HS GPA	.29 (.20)		.29 (.20)	.25 (.16)	.33 (.21)	.16 (.15)
SAT	.30 (.23)		.30 (.24)	.28 (.21)	.28 (.19)	.19 (.17)
SBAC	.29 (.23)		.26 (.21)	.28 (.22)	.28 (.20)	.16 (.15)
HS GPA & SAT	.32 (.24)		.33 (.26)	.30 (.22)	.34 (.23)	.19 (.18)
HS GPA & SBAC	.31 (.23)		.31 (.23)	.30 (.23)	.34 (.23)	.16 (.15)
HS GPA, SAT, & SBAC	.32 (.24)		.33 (.26)	.31 (.24)	.34 (.24)	.19 (.18)

Notes: *Insufficient sample size to compute for very low CCI schools. All correlation coefficients presented include demographic controls (gender, race/ethnicity, socioeconomic disadvantage, English Learner status).

Discussion/Conclusion

Improving college readiness has been and remains a priority in California's education reform efforts and accountability agenda. Three years after adopting rigorous new college readiness assessments aligned with the Common Core State Standards, we examine how these new tests perform in predicting college success. More specifically, we investigate the first cohort of 11th grade Smarter Balanced Assessment takers and track them into college at the California State University campuses and at one University of California campus. We assess how three different high school assessments—GPA, SBAC, and SAT—fare in predicting students' early college outcomes—first year GPA, persistence to year two, and total units accumulated.

Our results for the CSU analysis reveal the following: (1) HSGPA is a stronger predictor than either standardized test score measure; (2) the SBAC does as well as the SAT in predicting college outcomes of CSU students; (3) none of these assessments is a strong predictor of college persistence; and (4) the overall pattern of results holds for different subgroups (race/ethnicity subgroups, socioeconomic disadvantage, and by high school quality).

For our analysis of UCD students, we find: (1) both standardized test scores are stronger predictors of college performance than of persistence; (2) the SAT is a stronger predictor of first year performance than SBAC, but the difference is quite small in models that account for other individual and school characteristics; (3) none of these assessments is a strong predictor of college persistence (probably because 92 percent of all UCD students persist to year two); and (4) the overall pattern of results holds for different subgroups (race/ethnicity, socioeconomic disadvantage, and by high school quality).

There are several factors that policymakers should take into account as they consider the use of the Smarter Balanced Assessment (or any other assessment) for predicting college outcomes. First, standardized tests explain a relatively small share of early college outcomes, especially when controlling for high school GPA and individual and institutional characteristics. In addition, our analysis was based on data from the first cohort of Smarter Balanced Assessment takers, who had had minimal exposure to the new standards. There have been large improvements in both curricular materials and test administration in the years since 2014-15. Third, the students in our analyses were the last cohort to take the SAT before its substantial redesign. Additional analyses with cohorts of students who have taken the new SAT, and also benefited more fully from the shift in curricular standards that followed the adoption of the Smarter Balanced Assessment, would provide an essential complement to the present study. Finally, these outcomes may be useful for understanding performance (as measured by grades) in college, but that is certainly not the only outcome policymakers (or society) cares about. In particular, we might care a lot about increasing access to college, particularly for underrepresented groups, or about the long-term impact of sorting students to different types of colleges or within colleges to majors. We might also be concerned about the need for developmental coursework once students enroll in college, college persistence to and through graduation, and longer-term employment and economic mobility.

Most colleges and universities nevertheless continue to use test scores on standardized tests to inform their admissions decisions and to assess the academic strengths of applicants who come from a broad range of secondary school experiences. Our results suggest that Smarter Balanced Assessment scores are no worse (or better) than SAT scores at predicting first year college GPA and persistence rates at CSU, and are only slightly weaker at predicting total first year units. Among UCD students Smarter Balanced Assessment scores do not predict first year GPA and second year persistence rates as well as SAT scores, but the difference in magnitude is trivial. One possible explanation for these fundamental similarities is that there is an overlap in the knowledge and skills assessed by the Smarter Balanced Assessment and the SAT; the two tests measure much of the same thing. Recent studies, however, also point out important differences in the content covered by the two assessments (Achieve, 2018; Assessment Solutions Group, 2018; HumRRO, 2016).[21] A second and closely related reason is that the Smarter Balanced Assessment has been designed to reflect the state's new effort to align K-12 standards to expectations for postsecondary success.

There are several important caveats to any effort to compare directly the Smarter Balanced Assessment to the SAT. First, there are important differences in test format between the SAT and Smarter Balanced Assessment, which were described earlier in this report. Second, the Smarter Balanced Assessment has—or at least is perceived to have—much lower stakes for individual students than the SAT; students taking the Smarter Balanced Assessment are not generally under the impression that their performance on the test will be considered when they apply to college.[22] Third, students may not retake the Smarter Balanced Assessment, which is common with the SAT. Finally, some students take preparatory courses to improve their scores on the SAT. For now it is unclear how and whether these differences (e.g., low stakes, retakes, and test preparation) between the Smarter Balanced Assessment and SAT might bias the results presented here, in either direction.

Our future work will continue to assess the use and effectiveness of various K-12 assessments in predicting a host of student outcomes, in college and after. As California seeks to strengthen alignment between K-12 and postsecondary schooling, our evaluation of the standards taught and tested in the K-12 years should provide key information to evaluate college and labor market readiness and success.

[21] While Smarter Balanced Assessment is specifically aligned to the Common Core State Standards for the purpose of measuring student achievement of the standards, recent studies suggest the SAT is not fully aligned with Common Core State Standards (Assessment Solutions Group, 2018; HumRRO, 2016).

[22] California was the first state to use the 11th grade assessment as an opportunity to give students a signal about their college readiness, which it began with the Early Assessment Program in 2004 and which it continues with the Smarter Balanced Assessment today. Thus, students' results on these tests do matter—at least for placement out of developmental coursework—though they are not perceived to matter as much as the SAT, which is more broadly used for admissions.

Appendix A: Data

Table A1: Summary Statistics, CSU Analytic Sample, Mean values [standard deviations]

	CSU Applicants (N=117,706)	CSU Enrollees (N=36,388)
Female	0.5805	0.5938
	[.4935]	[.4911]
Asian/PI	0.1998	0.1963
	[.3999]	[.3972]
Black	0.0531	0.0471
	[.2242]	[.2119]
Hispanic	0.4876	0.4868
	[.4998]	[.4998]
White	0.2282	0.2399
	[.4197]	[.427]
Other	0.0313	0.0299
	[.1741]	[.1703]
Limited English Proficient	0.0718	0.0636
	[.2581]	[.244]
SES Disadvantage	0.5339	0.5228
	[.4989]	[.4995]
CCI = Very Low	0.0027	0.0023
	[.0516]	[.048]
CCI = Low	0.187	0.1803
	[.3899]	[.3844]
CCI = Medium	0.4617	0.4749
	[.4985]	[.4994]
CCI = High	0.233	0.2413
	[.4227]	[.4279]
CCI = Very High	0.1156	0.1012
	[.3198]	[.3016]
HS GPA	3.358	3.416
	[.6235]	[.4799]
SAT Verbal	492.7	487.8
	[103.4]	[90.40]
SAT Math	505.9	502.2
	[110.2]	[95.7]
SAT Writing	488.4	481.6
	[102.4]	[86.90]
SBAC ELA	2651	2653
	[83.22]	[75.32]
SBAC Math	2631	2631
	[101.7]	[90.8]

Table A2: Summary Statistics, UC Davis Analytic Sample, Mean values [standard deviations]

	UCD Applicants (N=34,439)	UCD Enrollees (N=3,544)
Female	0.5697	0.6670
	[.4961]	[.4713]
Asian	0.3792	0.3730
	[.4852]	[.4837]
Black	0.0563	0.0353
	[.2305]	[.1845]
Hispanic	0.3302	0.3050
	[.4703]	[.4605]
White	0.2000	0.2319
	[.4000]	[.4221]
Other	0.0238	0.0203
	[.1523]	[.1411]
Limited English Proficient	0.0560	0.0615
	[.2298]	[.2403]
SES Disadvantage	0.4257	0.4486
	[.4944]	[.4974]
CCI = Very Low	0.0027	0.001693
	[.0522]	[.0411]
CCI = Low	0.1316	0.1498
	[.3381]	[.3570]
CCI = Medium	0.3486	0.3705
	[.4765]	[.4830]
CCI = High	0.2409	0.2427
	[.4277]	[.4288]
CCI = Very High	0.2211	0.1704
	[.4150]	[.3761]
HS GPA	3.770	4.028
	[.4206]	[.2209]
SAT Verbal	559.7	560.2
	[109.5]	[96.7]
SAT Math	583.5	585.5
	[116.3]	[100.5]
SAT Writing	561.3	563.3
	[112.4]	[95.8]
SBAC ELA	2694	2706
	[76.63]	[67.60]
SBAC Math	2697	2707
	[98.90]	[87.06]

Table A3: CSU Univariate Correlation Matrix for High School Assessment Measures

Panel A: CSU Applicants (N=118,210)

	HS GPA	SAT-Reading	SAT-Math	SAT-Writing	SBAC-ELA	SBAC-Math
HS GPA	1.0000					
SAT-Reading	.4618	1.0000				
SAT-Math	.4912	.7663	1.0000			
SAT-Writing	.4931	.8450	.7685	1.0000		
SBAC-ELA	.4512	.6997	.5999	.6614	1.0000	
SBAC-Math	.4987	.6568	.8204	.6498	.6598	1.0000

Panel B: CSU Enrollees (N=36,519)

	HS GPA	SAT-Reading	SAT-Math	SAT-Writing	SBAC-ELA	SBAC-Math
HS GPA	1.0000					
SAT-Reading	.3201	1.0000				
SAT-Math	.3624	.7030	1.0000			
SAT-Writing	.3584	.8034	.6924	1.0000		
SBAC-ELA	.3368	.6527	.5268	.6082	1.0000	
SBAC-Math	.3923	.5752	.7883	.5613	.5909	1.0000

Table A4: UCD Univariate Correlation Matrix for High School Assessment Measures

Panel A: UCD Applicants (N=34,439)

	HS GPA	SAT-Reading	SAT-Math	SAT-Writing	SBAC-ELA	SBAC-Math
HS GPA	1.0000					
SAT-Reading	0.4604	1.0000				
SAT-Math	0.4825	0.7495	1.0000			
SAT-Writing	0.4897	0.8342	0.7706	1.0000		
SBAC-ELA	0.4377	0.6669	0.5488	0.6212	1.0000	
SBAC-Math	0.5036	0.66	0.8253	0.6577	0.6359	1.0000

Panel B: UCD Enrollees (N=3,544)

	HS GPA	SAT-Reading	SAT-Math	SAT-Writing	SBAC-ELA	SBAC-Math
HS GPA	1.0000					
SAT-Reading	0.2564	1.0000				
SAT-Math	0.2934	0.6628	1.0000			
SAT-Writing	0.2552	0.7798	0.683	1.0000		
SBAC-ELA	0.2867	0.6372	0.48	0.5806	1.0000	
SBAC-Math	0.3616	0.5599	0.7824	0.5437	0.5754	1.0000

Appendix B: Methods

In this study, we estimate the multiple correlation coefficient, R, for assorted sets of explanatory variables and outcomes in order to observe the change in predictive power as different assessments of interest are included and excluded. Specifically, we calculate:

$$R = \sqrt{R^2} = \left(\mathrm{P}_{\boldsymbol{yx}}\mathrm{P}_{\boldsymbol{xx}}^{-1}\mathrm{P}_{\boldsymbol{xy}}\right)^{\frac{1}{2}}$$

where $\mathrm{P}_{\boldsymbol{yx}}$, $\mathrm{P}_{\boldsymbol{xx}}$, and $\mathrm{P}_{\boldsymbol{xy}}$ are partitions of the correlation matrix P such that:

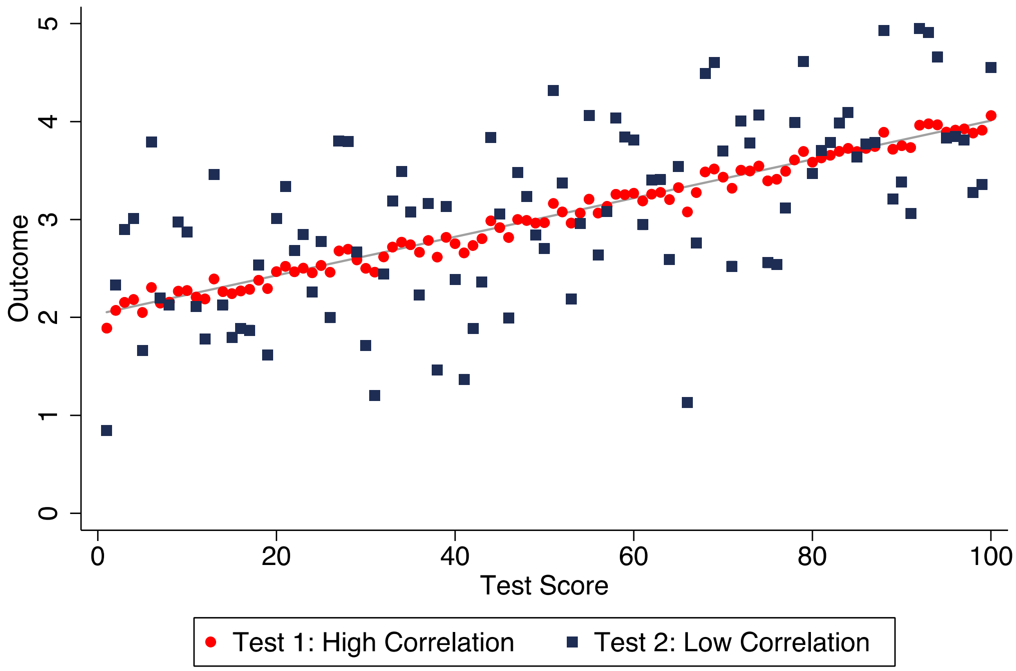
$$\mathrm{P} = \begin{bmatrix} \mathrm{P}_{\boldsymbol{xx}} & \mathrm{P}_{\boldsymbol{xy}} \\ \mathrm{P}_{\boldsymbol{yx}} & \mathrm{P}_{\boldsymbol{yy}} \end{bmatrix}$$

We calculate separate correlation matrices for each outcome (y) of interest, i.e. first year GPA, persistence to the second year, and units accumulated in the first year. Further, separate correlation matrices are calculated depending on what is included in the set of predictive (x) variables, the sparsest set of predictive variables being only high school GPA, while the fullest set being high school GPA, both SBAC scores, all three SAT scores, dummies for demographics, dummies for the College and Career Readiness level of a student’s high school, and dummies for CSU campuses (for the CSU analysis).

We can only observe outcomes for the students who are admitted and choose to attend the schools in our sample. As students are admitted to schools directly on the basis of high school grades and SAT scores and indirectly on the basis of SBAC scores (which are highly correlated with grades and SAT scores), our estimates will be biased due to selection. This is known as the “restriction of range” problem since we are much less likely to observe outcomes for the students with certain test scores and grades. For an extended discussion of why restriction of range causes problems in estimation see Gulliksen (1950). What follows is a brief illustration to develop intuition on the problem and proposed solution.

In Figure A1, we plot two tests hypothetical tests, Test 1 and Test 2, against an outcome. We note that both tests are positively correlated with the outcome, and that, on average, a student who receives some score on Test 1 would be expected to do the same as a student who receive the same score on Test 2. However, we can see that the outcome values for the students who took Test 1 are much more tightly clustered than the outcome values for students who took Test 2. This is the visual indication that Test 1 is much more highly correlated with the outcome than Test 2, (i.e. that knowing what a student receives on Test 1 is a better indication of their outcome, than knowing what a student receives on Test 2).

Figure B1: High vs. Low Correlation



Now imagine that we were to examine how closely related one test score was to an outcome, when we could observe the outcome for all students regardless of how they scored on the test. In the context of this study, this would be analogous to all high school students in California who took the SAT then attended a CSU. We can see what this might look like in Figure A2. If instead, we were only able to observe the outcomes of students who scored above a certain score on the test, because, for example, a student might need a certain SAT score in order to be admitted to college, then our data might look more like what is presented in Figure A3. Even though for tests scores over 50 these are the exact same points of data as in Figure A2, we now observe how much less tightly clustered the data points seem by excluding all outcomes for students with test scores below 50.

Figure B2: Outcomes Observed

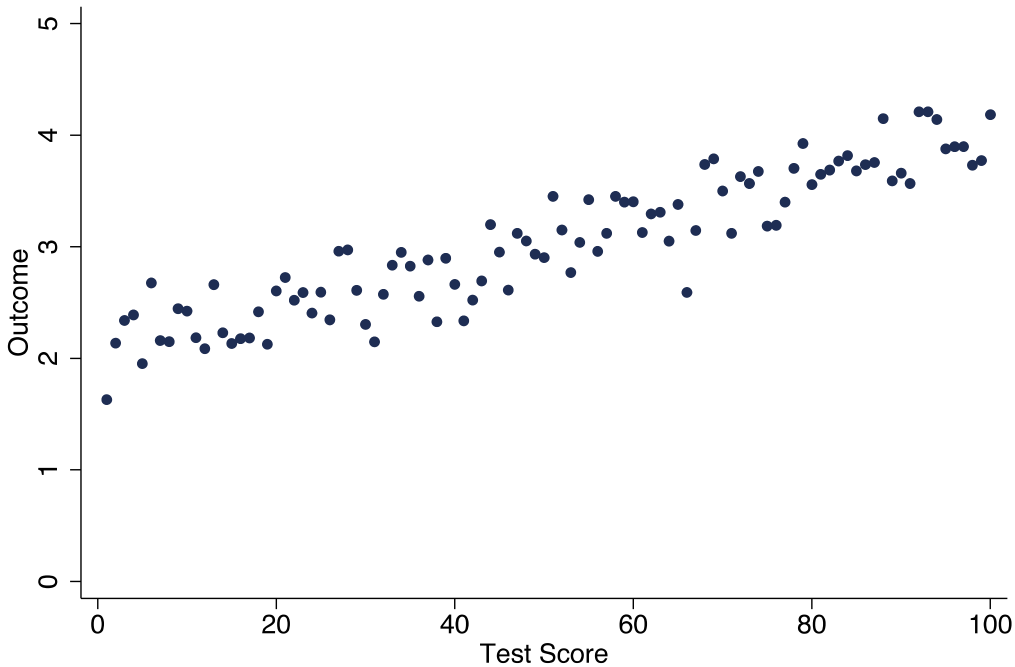
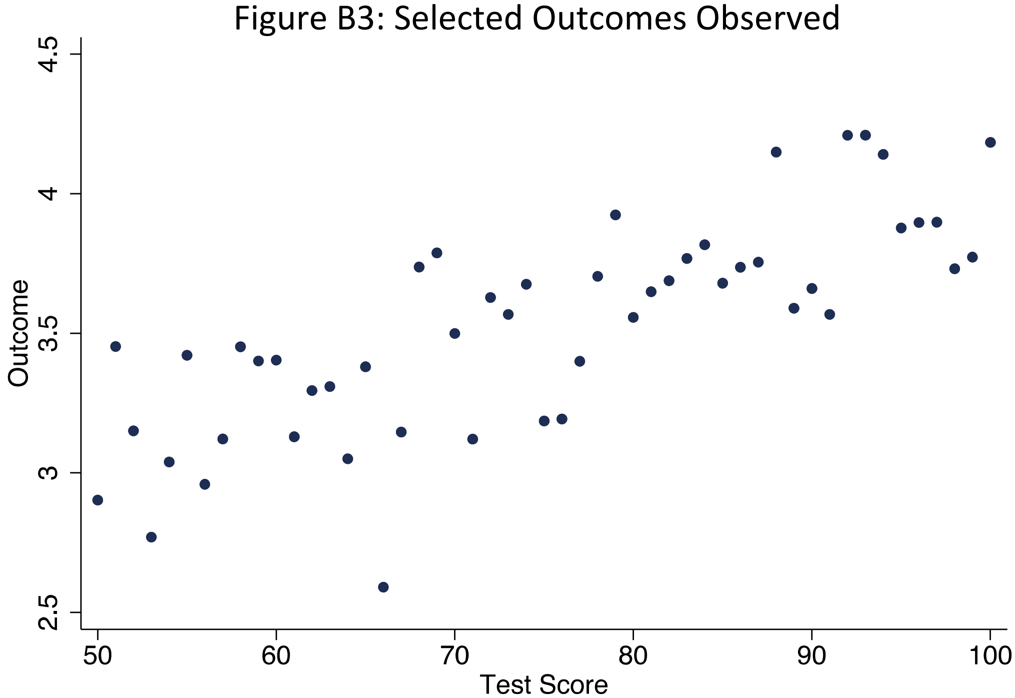


Figure B3: Selected Outcomes Observed



We correct for this bias due to this restriction of range using the method first proposed by Pearson (1903); developed in Lawley (1943); and then outlined in Gulliksen (1950), Lord and Novick (1968), and Lewis (2006). Following Lewis's notation, let **x** and **y** be vectors of random variables with covariance matrix $\sum$ for the full population[23]. Partition $\sum$ such that:

$$\Sigma = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}$$

Since **y** is not observed for the full population, $\sum_{xy}$, $\sum_{yx}$, and $\sum_{yy}$ cannot be estimated directly from the data. However, if we let **s** be a selection variable such that $s_i = 1$ if an individual from the full population is included in the selected population while $s_i = 0$ if the individual is not included, the covariance matrix $\sum$ for the selected population is:

$$\Sigma_s = \begin{bmatrix} \Sigma_{xx|s} & \Sigma_{xy|s} \\ \Sigma_{yx|s} & \Sigma_{yy|s} \end{bmatrix}$$

which we estimate using our selected sample.

Under the assumptions of linearity and homoscedasticity it can be shown that:

$$\Sigma = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xx}\Sigma_{xx|s}^{-1}\Sigma_{xy|s} \\ \Sigma_{yx|s}\Sigma_{xx|s}^{-1}\Sigma_{xx|s} & \Sigma_{yy|s} - \Sigma_{yx|s}\left(\Sigma_{xx|s}^{-1} - \Sigma_{xx|s}^{-1}\Sigma_{xx}\Sigma_{xx|s}^{-1}\right)\Sigma_{xy|s} \end{bmatrix}$$

In effect, this correction reduces the sample variance of the outcome to what we would theoretically observe if we were able to observe outcomes for the full population. As such, the full set of explanatory variables used in the most fully specified model should be included in $\sum_s$. We include all six test scores and dummies for demographics, CSU campus (where applicable), and CCI levels. Then in order to calculate the correlation matrix P, where

$$\mathrm{P} = \left(diag(\Sigma)\right)^{-\frac{1}{2}}\Sigma\left(diag(\Sigma)\right)^{-\frac{1}{2}}$$

we delete the rows and columns from $\sum$ which contain the variances and covariances for explanatory variables not included in the model being estimated. For additional derivations and further discussion, see Lawley (1943), Lord and Novick (1968), and Lewis (2006).

[23] Due to data considerations, we use the population of applicants to the CSUs (or UC Davis) as our full population.

References

Achieve. (2018, March). *What gets tested gets taught: Caustions for using college admissions tests in state accountability systems.* https://www.achieve.org/college-admissions-tests-accountability

Assessment Solutions Group (2018). Feasibility of the Use of ACT and SAT in Lieu of Florida Statewide Assessments. http://www.trbas.com/media/media/acrobat/2018-01/70109708365300-05065523.pdf

Black, S. E., Cortes, K. E., & Lincove, J. A. (2016). Efficacy versus equity: What happens when states tinker with college admissions in a race-blind era? *Educational Evaluation and Policy Analysis*, *38*(2), 336-363.

Gulliksen, H. (1950). *Theory of Mental Tests*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., Publishers.

HumRRO. (2016). *Delaware System of Student Assessment and Maine Comprehensive Assessment System: SAT Alignment to the Common Core State Standards*. https://www.doe.k12.de.us/cms/lib/DE01922744/Centricity/Domain/414/SATalignment.pdf

Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate students' success. *Science, 315*(5815), 1080-1081.

Lawley, D. N. (1943). A note on Karl Pearson's selection formulae. *Royal Society of Edinburgh, Proceedings, 62*(1), 28–30.

Lewis, C. (2006). Selected topics in classical test theory. In C.R. Rao, & S. Sinharay (Eds.), *Handbook of Statistics 26 Psychometrics* (pp.29-42). Amsterdam, NL: North Holland Publishing.

Lord, F.M., Novick, M.R., & Birnbaum, A. (1968). *Statistical Theories of Mental Test Scores*. Oxford, England: Addison-Wesley.

Mattern, K. D., Kobrin, J. L, Patterson, B. F., Shaw, E. J., & Camara, W. J. (2009). Validity is in the eye of the beholder: Conveying SAT research findings to the general public. In R.W. Lissitz (Ed.), *The concept of validity: Revisions, new directions, and applications* (213-240). Charlotte, NC: Information Age Publishing.

Mattern, K. D., & Patterson, B. F. (2014). *Synthesis of recent SAT validity findings: Trend data over time and cohorts* (College Board Research in Review 2014-1). New York: The College Board.

Pearson, K. (1903). Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. *Philosophical Transactions of the Royal Society of London, 200,* 1–66.

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

Shaw, E. J., Marini, J. P., Beard, J., Shmueli, D., Young, L., & Ng, H. (2016). The redesigned SAT pilot predictive validity study: A first look (College Board Research Report 2016-1). New York: The College Board.

About

Policy Analysis for California Education (PACE) is an independent, non-partisan research center based at Stanford University, the University of Southern California, and the University of California – Davis. PACE seeks to define and sustain a long-term strategy for comprehensive policy reform and continuous improvement in performance at all levels of California's education system, from early childhood to postsecondary education and training. PACE bridges the gap between research and policy, working with scholars from California's leading universities and with state and local policymakers to increase the impact of academic research on educational policy in California.

Founded in 1983, PACE

- Publishes policy briefs, research reports, and working papers that address key policy issues in California's education system
- Convenes seminars and briefings that make current research accessible to policy audiences throughout California
- Provides expert testimony on educational issues to legislative committees and other policy audiences
- Works with local school districts and professional associations on projects aimed at supporting policy innovation, data use, and rigorous evaluation

PACE
Policy Analysis for California Education

Stanford Graduate School of Education
520 Galvez Mall, CERAS 401
Stanford, CA 94305-3001
Phone: (650) 724-2832
Fax: (650) 723-9931

edpolicyinca.org