AEA Papers and Proceedings 2025, 115: 676–681
https://doi.org/10.1257/pandp.20251056

Standardized Test Scores and Academic Performance at Ivy Plus Colleges[†]

By John N. Friedman, Bruce Sacerdote, Douglas O. Staiger, and Michele Tine*

What is the appropriate role of standardized test scores in the admissions process at highly selective colleges? Proponents of their use argue that test scores provide valuable information on students' academic preparation that is crucial in admissions assessments at selective universities (Leonhardt 2024). Detractors argue that test scores are biased against students from less advantaged backgrounds, for instance because those students cannot devote the same resources to preparing for the test or relate less well to the cultural content of testing materials, and that alternative measures of academic preparation do not suffer from these problems (Lemann 2024). This debate has played out not only in the literature but also in the policies of colleges across the country. By early 2020, several hundred colleges had adopted test-optional admissions policies, in which applicants could choose to submit test scores or not with their application, with a smaller number of others (e.g., the University of California system) moving to exclude test scores entirely from the admissions process. Nearly all remaining colleges then adopted test-optional policies during the COVID-19 pandemic, although eight Ivy Plus schools have recently announced a return to test requirements.[1]

We contribute to this debate by analyzing the informational content in SAT and ACT scores for predicting academic performance during students' first year in college. We conduct this analysis in admissions data and academic records from multiple Ivy Plus colleges for those starting as first-time first-year students between 2017 and 2023. We present three key findings.

First, SAT and ACT scores have substantial predictive power in forecasting students' academic performance. Even when comparing students from the same racial or ethnic and socioeconomic background, students with the highest possible score (i.e., 1600 on the SAT or 36 on the ACT) achieve a first-year college GPA that is 0.43 higher on a 4.0 scale (equal to 1 standard deviation in the college GPA distribution) than students with an SAT score of 1200 or ACT score of 25 (equating to the seventy-fifth percentile of the national distribution of scores). These lower-scoring students are also 42 percentage points more likely to struggle academically during their first year (defined as receiving at least one grade of C+ or lower). We also find that students who do not submit test scores (as part of test-optional admissions policies at these schools in 2021–2023) achieve significantly lower levels of academic performance in college, equivalent to students who submitted SAT scores of roughly 1300 or ACT scores of 28.

Second, in contrast with standardized test scores, high school GPA has relatively little predictive power for academic success during a student's first year. Comparing students with a perfect 4.0 high school GPA to those with a 3.2 GPA—a gap of the same magnitude in the distribution of applicants as the test score gap discussed above—predicts a difference in first-year college GPA of less than 0.1. These findings align with those in Chetty, Deming, and Friedman (2023), who show that SAT and ACT scores similarly outperform high school GPA in predicting early-career outcomes such as attending a top-ranked graduate program or working at a prestigious firm for Ivy Plus students.

*Friedman: Brown University and NBER (email: john_friedman@brown.edu); Sacerdote: Dartmouth College and NBER (email: Bruce.I.Sacerdote@dartmouth.edu); Staiger: Dartmouth College and NBER (email: Douglas.O.Staiger@dartmouth.edu); Tine: Dartmouth College (email: Michele.T.Tine@dartmouth.edu). We thank Sian Beilock, Zachary Bleemer, Liz Cascio, Raj Chetty, Nathan Hendren, David Leonhardt, Adam Looney, Andrei Shleifer, and numerous seminar participants for helpful comments. This paper incorporates input from the work of Cascio et al. (2024). We are indebted to Hamidah Alatas and Arjun Shanmugam for their outstanding research assistance. This research was funded by the Gates Foundation (INV-069243), the Chan Zuckerberg Initiative (38382128), the JPB Foundation (2022-3516), and Overdeck Family Foundation (Opportunity Insights).

[†]Go to https://doi.org/10.1257/pandp.20251056 to visit the article page for additional materials and author disclosure statement(s).

[1]Ivy Plus includes the eight Ivy League colleges plus Chicago, Duke, MIT, and Stanford.

Third, SAT and ACT scores exhibit no calibration bias, in that students from different backgrounds but sharing the same test scores achieve similar average levels of academic success in college. This test follows the literature on algorithmic bias (e.g., Obermeyer et al. 2019) and notably differs from common uses of "bias" in discussion of test scores (that instead focus on differences in the average level of test scores). Because students from different backgrounds experience disparities in school quality, neighborhood exposure, and other environmental differences throughout childhood, each of which affects academic preparation, comparisons of the average levels of testing across students will not generally isolate calibration bias.

Collectively, these results suggest that standardized test scores provide important information to measure applicants' academic preparation that is not available elsewhere in the application file.

I. Data Description

We combine student-level data from the admissions process and the Office of Institutional Research for first-year first-time students from multiple Ivy Plus colleges. We include all students in the first-year classes who started in fall of 2017–2023 (excluding fall of 2020, due to the disruption from COVID-19) who reported high school GPA and completed the required minimum number of courses during their first full year.

Our key admissions variables are SAT/ACT test scores and high school GPA. We observe unweighted high school GPA on the standard 4.0 scale. We also measure a range of other student characteristics (such as student demographics and schooling background) from admissions records.

We define three outcome variables based on a student's grades in their first year of college. Our main outcome variable is first-year cumulative GPA (scaled in the standard way as above); the mean of first-year cumulative GPA across our colleges is 3.49 (on a 4.0 scale), with a (within–school year) standard deviation of 0.47. We additionally analyze two further measures of performance: percent of grades that are A or A− (63 percent on average in our sample) and an indicator for academic struggle, defined as having any grade of C+ or lower in the first year (25 percent on average in our sample). See Supplemental Appendix A for more details on our variable definitions.

II. Empirical Approach

We regress our three outcome measures on SAT/ACTs, high school GPA, and various sets of controls and interactions of the controls. Our specification is

$$(1) \quad Y_i = \beta_1 \times SAT_i + \beta_2 \times \mathbf{1}_{\{SATMISSING_i\}} + \beta_3 \times HSGPA_i + \mu_{s(i)t(i)} + \delta_{h(i)} + \boldsymbol{\gamma}'\mathbf{X}_i + \epsilon_i,$$

where Y_i is one of our academic outcomes for student i. Our key dependent variable is a student's standardized test score (SAT_i); we also include an indicator variable for students admitted without test scores during test-optional admissions cycles. For interpretability, we norm test scores to 0 for students with an SAT score of 1400 (ACT score of 31) and divide by 100; because we impute $SAT_i = 0$ for students' missing test scores, readers can interpret the coefficient β_2 as the extent to which students without test scores differ in their outcomes from students scoring 1400 on the SAT. All specifications include college-year fixed effects ($\mu_{s(i)t(i)}$). We then control for additional variables in certain specifications, including a vector of individual level characteristics $\mathbf{X}_i$ (including a student's gender, legacy status, early-decision applicant status, underrepresented minority (URM) status, athletic recruit status, first-generation college status, rural/urban home, US citizen or permanent resident, family income, and high school challenge index) and fixed effects for the high school from which they applied to college ($\delta_{h(i)}$).[2]

[2] Rothstein (2004) notes that equation (1) may produce biased estimates because we can only estimate it on students who attend rather than on all applicants. Intuitively, students admitted with particularly low SAT/ACT scores must have had some other countervailing factors; if these other factors positively (negatively) predict first-year college GPA, then the estimates would be negatively (positively) biased. If these other factors correlate in the same way with SAT/ACT scores as with high

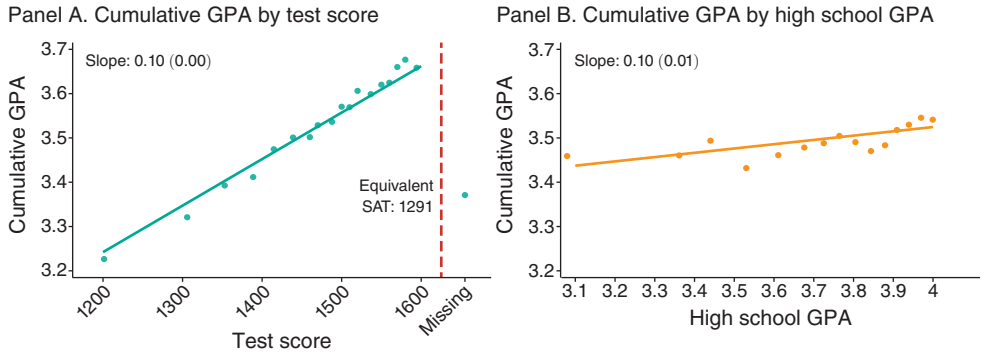


FIGURE 1. FIRST-YEAR STUDENTS' PERFORMANCE BY TEST SCORE AND HIGH SCHOOL GPA

Notes: Figure 1 presents binned scatterplots of first-year academic performance versus SAT/ACT test scores (panel A) or high school GPA (panel B) for students enrolled at selected Ivy Plus colleges controlling for some students' characteristics. Panel A presents binned scatterplots of academic outcomes versus SAT/ACT score in a specification that includes fixed effects for each school-by-year and for a student's race, gender, and parent income, as well as controls for other student demographics from admissions data. The specification in panel A also controls for high school GPA; the specification in panel B also controls for SAT/ACT score and the indicator for missing score. The rightmost dot in panel A is for students who do not submit a test score when applying under a test-optional admissions regime. These specifications match that in Supplemental Appendix Table 1, panel A, column 4. Slopes reported are for a 100 point change in SAT and a 1.0 point change in GPA.

In addition to our baseline specification in equation (1), we also test for calibration bias in test scores by separately estimating the relationship between test scores and academic performance in college for students from different subgroups. To do so, we augment equation (1) with two additional terms: interactions between an indicator G_i and each of our test score variables (SAT_i and $\mathbf{1}_{\{SATMISSING_i\}}$). We do so using four different definitions of G_i: students who attend a high school above the national twentieth percentile on our index of high school challenge indicators, first-generation students, students from families with incomes in the first tercile of distribution (below $91,800), and students from underrepresented racial and ethnic groups (URM students).

III. Results

A. *Association of Test Scores and Academic Performance*

Figure 1, panel A presents a nonparametric representation of the association between test scores and first-year college GPA from equation (1), including all control variables. More specifically, we estimate a version of equation (1) replacing the linear control for SAT with 17 indicators for quantiles of the test score distribution (since we observe test scores for roughly 85 percent of students). We then plot the value of these 17 coefficients (y-value) versus the average value of test score for students in each quantile (x-value), along with the coefficient for the remaining 15 percent of students missing a test score (on the right). The best-fit line plots the linear fit from the microdata.

There is a robust and linear relationship between test scores and first-year college GPA. This relationship is highly statistically significant and economically meaningful; moving from students with perfect scores to those with SAT scores of 1200 (or ACT scores of 25) predicts a 0.47 lower GPA, a shift that is 1 standard deviation in the distribution of first-year GPA. Students who do not submit test

school GPA, however, it would not affect the comparison between β_1 and β_3 and thus would not affect the qualitative conclusions from this analysis. Similarly, if these other factors correlate in the same way with SAT/ACT scores within different student subpopulations, it would not affect our test for bias.

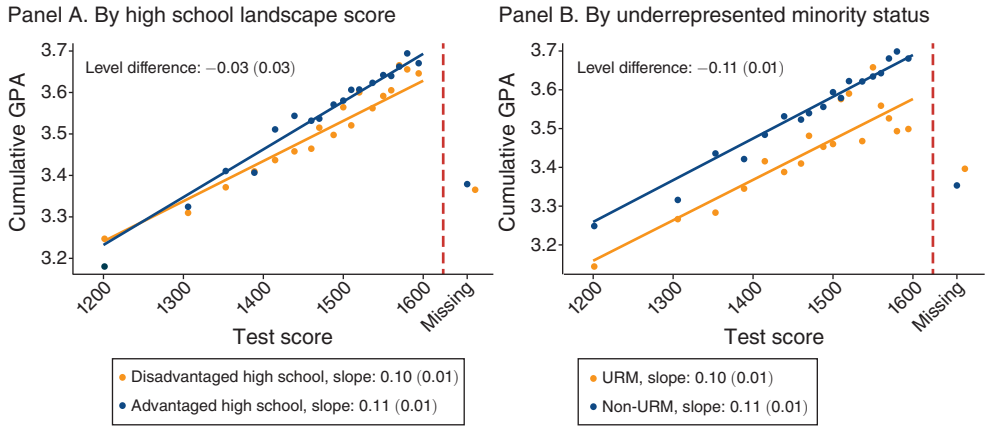


FIGURE 2. FIRST-YEAR CUMULATIVE GPA BY TEST SCORE AND STUDENTS' CHARACTERISTICS

Notes: Figure 2 replicates the binned scatterplot specification in Figure 1, panel A, except that students are split based on their characteristics into two different series. In panel A, we split students based on the challenge index for their high school; high schools with an index value above the twentieth national percentile—roughly the top quartile of student applicants—are "less advantaged." In panel B, we split students into those who are from historically underrepresented racial and ethnic groups and those who are not. The specifications in panels A and B match those in Supplemental Appendix Table 2, panel A, columns 1 and 4, respectively. Slopes reported are for a 100 point change in SAT.

scores (represented by the last dot on the right) achieve quite low first-year grades; on average, they receive the same first-year college GPA as students who reported an SAT score of just below 1300.

Figure 1, panel B presents the same nonparametric relationship but for high school GPA; in contrast with SAT/ACT scores, high school GPA has a relatively weak relationship with academic performance in college. Decreasing a student's high school GPA from 4.0 to 3.2 predicts a fall of less than 0.1 in a students' first-year college GPA.

Motivated by the strongly linear relationship, we estimate several versions of equation (1) with both SAT/ACT scores and high school GPA included linearly (see Supplemental Appendix Table 1, panel A). The results from Figure 1 are robust to a wide variety of controls. With the inclusion of high school fixed effects; the coefficient on test score falls somewhat, while the coefficient on high school GPA rises; still, the predicted change in academic outcomes from a 1 standard deviation increase in test scores is larger than for a 1 standard deviation increase in high school GPA. We also show similar results studying two different measures of academic performance: a student's fraction of A's during their first year (in Supplemental Appendix Table 1, panel B) and an indicator for academic struggle, defined as a student receiving at least one grade of C+ or below during their first year (Supplemental Appendix Table 1, panel C).

B. *Testing for Bias in Scores*

Having established that test scores include important and distinctive predictive power for students' future academic outcomes, we now test whether those predictions exhibit calibration bias. The intuition behind our test is straightforward: If test scores are biased against a certain group of students, then those students will have a higher level of underlying academic preparation as compared to others with the same score, leading them to outperform academically once in college and judged in a system without such bias. Such might be the case if, for instance, students from advantaged backgrounds had more resources with which to prepare specifically for the SAT or ACT, inflating their test scores relative to others with the same underlying level of academic preparation but lacking these extra resources.

Figure 2, panel A shows a nonparametric representation of our test for bias between students attending more versus less advantaged high schools. The relationship between test SAT/ACT scores

and first-year college GPA is quite similar for all students, no matter the advantage of the high school they attended. Although not statistically distinguishable, the point estimates suggest that students from less advantaged high schools slightly underperform their peers from more advantaged high schools with the same test scores. Figure 2, panel B replicates this test for bias between URM and non-URM students; we similarly find that students who are not from historically underrepresented racial and ethnic groups slightly outperform URM students with the same SAT/ACT scores, though the relationship (while statistically significant) is not large. We also find similar results in parametric regression specifications matching the nonparametric results in Figure 2, plus results comparing first-generation students and students from low-income backgrounds to others and using our alternative academic performance measures (see Supplemental Appendix Table 2). Similar results also obtain when including fixed effects for academic division (i.e., physical sciences) or course.

C. *Comparison with the Literature*

A long literature has found a positive relationship between standardized test scores and academic performance in college, college retention and completion, and early-career postbaccalaureate outcomes. Our findings align with this literature.

A much smaller body of work has compared the predictive power of SAT/ACT scores to that of high school GPA or other measures of academic preparedness in a multiple regression framework, and this literature finds a more mixed set of results. For instance, Allensworth and Clark (2020) show that ACT scores have only marginally statistically significant power to predict the six-year college graduation rates of students from Chicago Public Schools, while high school GPA has a much larger and more statistically significant effect. University of California, Office of the President (2020) shows that both standardized test (SAT and ACT) scores and high school GPA increase explanatory power in regressions that predict first-year GPA at University of California (UC) campuses, while Geiser and Santelices (2007) find that high school GPA outperforms standardized test scores in the UC system. Chetty, Deming, and Friedman (2023) find that SAT/ACT scores, but not high school GPA, predict early-career success.

One apparent pattern across these studies is that standardized test scores perform better in more selective college settings. Recent grade inflation could be eroding the information content of high school GPA most at the top, as more students are pushed up against the 4.0 cap. Goodman, Gurantz, and Smith (2020) show that retaking SAT tests increases scores more for students lower in the test score distribution. Finally, it may be that SAT/ACT tests and high school GPA simply capture different combinations of underlying student attributes that themselves are most relevant in different higher education situations; for instance, high school GPA in part captures variation in student attendance patterns, while test scores (by definition) measure scores for those students who are in attendance on the day of the test.

IV. Implications for Admissions Policy

These results show that SAT/ACT test scores possess important predictive power for students' academic performance in their first year of college at Ivy Plus institutions. While some of the raw correlation between these variables relates to students' demographic characteristics, much remains even when controlling for these variables. In contrast, high school GPA does not predict academic performance nearly as well. We also find no evidence of bias against students from less advantaged backgrounds, as students from such backgrounds do not outperform (and in some cases underperform) other students with the same test score.

It is important to acknowledge that students from low-income families and other less advantaged backgrounds have lower standardized test scores and are less likely to take the test than students from higher-income families. This fact is consistent with those presented above because of disparities experienced throughout childhood, including differences in school quality, neighborhood exposure (Chetty et al. 2020), and many other environmental conditions.

Collectively, these results suggest that standardized test scores provide important information to measure applicants' academic preparation that is not available elsewhere in the application file. In

practice, the current debate at most selective colleges is whether to maintain test-optional policies or return to required testing. In addition to depriving admissions officers of valuable information on students' academic preparation, test-optional policies also may affect students' choices whether and how to apply. Additional results from Sacerdote, Staiger, and Tine (2025) suggest that students from less advantaged backgrounds may be less well informed when making these decisions, for instance by choosing to withhold their test scores when those scores would actually help their chances of admissions. As a result, admissions policies that require applicants to submit test scores may benefit less-advantaged students in the application process.

REFERENCES

Allensworth, Elaine M., and Kallie Clark. 2020. "High School GPAs and ACT Scores as Predictors of College Completion: Examining Assumptions about Consistency across High Schools." *Educational Researcher* 49 (3): 198–211.

Cascio, Elizabeth, Bruce Sacerdote, Doug Staiger, and Michele Tine. 2024. *Report from Working Group on the Role of Standardized Test Scores in Undergraduate Admissions*. Communication to Sian Leah Beilock and Lee Coffin, January 30. https://home.dartmouth.edu/sites/home/files/2024-02/sat-undergrad-admissions.pdf.

Chetty, Raj, David J. Deming, and John N. Friedman. 2023. "Diversifying Society's Leaders? The Determinants and Causal Effects of Admission to Highly Selective Private Colleges." NBER Working Paper 31492.

Chetty, Raj, John N. Friedman, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2020. "The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility." NBER Working Paper 25147.

Geiser, Saul, and Maria Santelices. 2007. "Validity of High-School Grades in Predicting Student Success beyond the Freshman Year: High-School Record versus Standardized Tests as Indicators of Four-Year College Outcomes." University of California, Berkeley, Center for Studies in Higher Education Working Paper CSHE.6.07.

Goodman, Joshua, Oded Gurantz, and Jonathan Smith. 2020. "Take Two! SAT Retaking and College Enrollment Gaps." *American Economic Journal: Economic Policy* 12 (2): 115–58.

Lemann, Nicholas. 2024. *Higher Admissions: The Rise, Decline, and Return of Standardized Testing*. Princeton University Press.

Leonhardt, David. 2024. "The Misguided War on the SAT." *New York Times*, January 7. https://www.nytimes.com/2024/01/07/briefing/the-misguided-war-on-the-sat.html.

Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 2019. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." *Science* 366 (6464): 447–53.

Rothstein, Jesse M. 2004. "College Performance Predictions and the SAT." *Journal of Econometrics* 121 (1–2): 297–317.

Sacerdote, Bruce, Douglas O. Staiger, and Michele Tine. 2025. "How Test Optional Policies in College Admissions Disproportionately Harm High Achieving Applicants from Disadvantaged Backgrounds." NBER Working Paper 33389.

University of California, Office of the President. 2020. *Relationship of the SAT/ACT to College Performance at the University of California*. Institutional Research and Academic Planning.