High School Standardized Testing Participation in the Test-Optional Post-Pandemic Age

Audrey Liang ‘24
Data Science Major Capstone

Background & Research Question

In the wake of the COVID-19 pandemic that hit the U.S. in Spring 2020, there have been drastic changes in the college admissions landscape. This has been reflected especially in standardized testing for college admissions, with the number of test-optional post-secondary institutions soaring from 713 to 1,350 to adjust for the unique circumstances that made it difficult for students to take these tests. [1]

Fast forward to 2023, the number of undergraduate institutions adopting a test-optional policy has increased to 1,900 institutions, which represents 83% of four year U.S. post-secondary institutions. [2]

College admission offices are increasingly recognizing the negative consequences of mandating of standardized tests like the ACT or SAT, with a major concern with that policy being a barrier to entry to students of underprivileged backgrounds, especially impacting low-income students who many not be able to afford the fees associated with the tests. This has been discussed increasingly, especially with the 2023 Supreme Court decision to ban affirmative-action based admission policies, but by being test-optional, college may offset consequences of the Supreme Court decision by encouraging a more diverse applicant pool.

Thus, standardized test participation rates at high schools in 2020-2021 may reflect the socioeconomic status of students in the school. On the other hand, many students may have considered standardized testing as a factor that would benefit their college admission chances, so higher participation rates may reflect higher student intent to attend college and pursue top-ranking institutions.

This study uses four-year high school and student body statistics collected from the United States to fit multiple-linear regression model with the research questions outlined below:

(RQ1) What school and student body characteristics lead to higher standardized testing participation rates, especially in the test-optional age of the pandemic?

(RQ2) Are the identified predictors positively or negatively associated with SAT/ACT participation rate given other predictors in the model?

Data

The data was collected from the Civil Rights Data Collection Office for Civil Rights [3], which comes from the U.S. Department of Education. This dataset includes information relevant to students’ equal access to educational opportunities in the school year of 2020-2021 and was collected from all public schools in the 50 states, Washington D.C., and Puerto Rico.

The final dataset used for this analysis is a merged dataset of various csv files provided from the Civil Rights Data Collection Office for Civil Rights that contain various variables of interest related to school and student characteristics.

Cleaning

To conduct the analysis, 32 csv files with over 1000 potential predictor variables were merged to create a comprehensive dataset with over 90,000 observations.

Each of the files had from 30-170 variables. In total, the number of variables was in the thousands. All student-count related variables in the dataset are separated by sex, and all of the columns were even further separated into specific demographic categories, such as “Caucasian Female” or “Asian male” for every variable of interest, such as number of students participating in sports or number of students taking an AP.

For the purposes of this analysis, all aggregate female and male columns for each variable were summed to create a total number of students column for each variable that was related to total student number. These variables were then all divided by the total enrollment at the school to obtain a proportional ratio to account for schools of different sizes. For example, the columns with the number of female and male students taking biology were summed and divided by the sum of the columns with female and male total enrollment to create a new column indicating the proportion of all students in the school who were taking biology.

Similarly, many variables related to disciplinary punishments were more specifically divided into females with disabilities, females without disabilities, males with disabilities, and males without disabilities. For these variables, all columns were summed to include all students and divided by total student population to create a new column. Many other variables were collapsed, like nurses, psychologists, and social workers grouped as health professionals and types of physical attacks to the school collapsed into one variable as well.

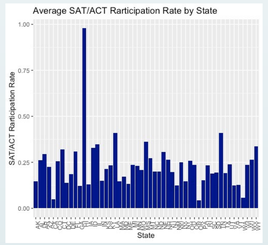
Additionally, the dataset was filtered for 4-year high schools by only including schools that offered years 9-12.Teacher, counselor, and health professional to student ratios as well as device per student ratios were calculated to better compare resource allocation across different school sizes. Any ratios with value over 1 were dropped from the dataset.

Missing values in the dataset are encoded as negative values, and according to the codebook this is because the surveyed question is non-applicable to the school. Thus, for many columns related to total student count, these values were encoded as zero. Rows with all zero values were dropped from the dataset, and rows with missing values for binary and categorical variables were dropped from the dataset.

To narrow down the variables and preserve interpretability of the model, each variable’s summary statistics were examined. Variables with high missingness (>30%) were omitted. Additionally, variables with third quartile zero values were omitted because of the rarity and specificity of the phenomenon.

After cleaning, the final dataset has 28 columns and ~10,000 rows. Our response variable of interest is SAT/ACT participation rate.

Data Exploration/Visualization



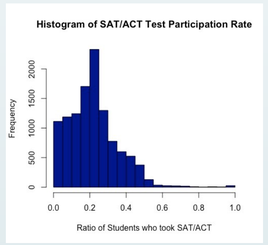


Figure I. Average SAT or ACT participation rate by state in 2020-2021
Figure II. Distribution of SAT or ACT participation rate in 2020-2021

In initial data exploration, it can be seen that state is a potentially important variable to include in the analysis because if its variation across different states, even though there are 50+ categorical levels. It is also apparent that the distribution of the participation rates across all high schools does not satisfy the normality assumption and has a right tail. This might be because only seniors and some juniors in high school traditionally take the exam for college admissions, so the average rate is close to 25% for most schools (senior class), but test-taking rates may also skew left due to the test-optional policy and the COVID-19 pandemic (or both).

First-order model

A multiple linear regression was ran with the response variable and 27 predictors to identify the school and student body related characteristics associated with standardized testing participation rate. In this full model, the adjusted R-squared value is 0.4496. Stepwise elimination was performed based on a threshold of 10 for the variance inflation factor to check for multicollinearity, but in the first iteration only the state identifier variable was over 10 at a value of 15. However, this is expected because state is a categorical variable representing over 50 states, so there is bound to be some overlap. In addition, removing the state variable greatly decreased the adjusted R-squared to 0.16, so it was not removed. Automatic selection was used to identify the best model, with AIC and BIC criterion. The adjusted R-squared for each model were 0.4498 and 0.4481 respectively, while the result of 10-fold cross-validation also yielded very similar results, with the CV score having a 0.0001 difference on the response variable scale from 0 to 1. In this case, the more parsimonious BIC model with 12 variables was chosen.

testParticipation= advancedMath_PER_ENR + APenroll_PER_ENR + algebraII_PER_ENR + biology_PER_ENR + calculus_PER_ENR + chemistry_PER_ENR + dual_PER_ENR + ESL program_PER_ENR + disabilities_PER_ENR + referals_PER_ENR + sportsrate + LEA_STATE

Higher-order model

Interaction terms were added to the first-order model using a AIC and BIC stepwise process. Again, the cross-validation scores for each resulting model were within a 0.0002 difference on the response variable scale. The AIC model yielded a higher adjusted R-squared score of 0.4938, while the BIC model yielded a 0.4549 adjusted R-squared. However, the AIC model had 34 predictors while the BIC model had 20 predictors. To preserve interpretability, the parsimonious BIC model with 20 predictors and a slightly higher adjusted R-squared was chosen as the best higher-order model.

testParticipation = advancedMath_PER_ENR + APenroll_PER_ENR + algebraII_PER_ENR + biology_PER_ENR + calculus_PER_ENR + chemistry_PER_ENR + dual_PER_ENR + ESL program_PER_ENR + disabilities_PER_ENR + TOT_REF_PER_ENR + sportsrate + LEA_STATE + calculus_PER_ENR:chemistry_PER_ENR + APenroll_PER_ENR:sportsrate + advancedMath_PER_ENR:disabilities_PER_ENR + calculus_PER_ENR:ESLprogram_PER_ENR + calculus_PER_ENR:disabilities_PER_ENR + APenroll_PER_ENR:calculus_PER_ENR + referals_PER_ENR :sportsrate + calculus_PER_ENR:dual_PER_ENR

Final model & Assumptions

To choose between the first-order and higher-order model, a partial F-test was conducted to compare the two models. The partial F-test yielded a small p-value of 2.2e-16, which indicated that the interaction terms in the higher-order model significantly improved the fit of the simple nested first-order model. Thus, the higher-order model is used, and assumptions are checked.

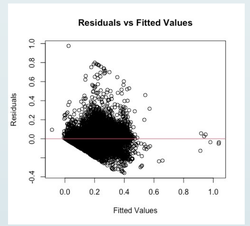


Figure III. Residuals vs. Fitted values plot for final model

As seen in figure II, the residuals vs. fitted values plot of the final model has uneven variance levels, with the values above the threshold zero values being more spread out than the ones below. A Box-cox transformation was performed on the model to satisfy this assumption for linear regression, and as result the response variable of SAT/ACT participation rate was square rooted to yield the final model:

√testParticipation= advancedMath_PER_ENR + APenroll_PER_ENR + algebraII_PER_ENR + biology_PER_ENR + calculus_PER_ENR + chemistry_PER_ENR + dual_PER_ENR + ESL program_PER_ENR + disabilities_PER_ENR + TOT_REF_PER_ENR + sportsrate + LEA_STATE + calculus_PER_ENR:chemistry_PER_ENR + APenroll_PER_ENR:sportsrate + advancedMath_PER_ENR:disabilities_PER_ENR + calculus_PER_ENR:ESLprogram_PER_ENR + calculus_PER_ENR:disabilities_PER_ENR + APenroll_PER_ENR:calculus_PER_ENR + referals_PER_ENR :sportsrate + calculus_PER_ENR:dual_PER_ENR

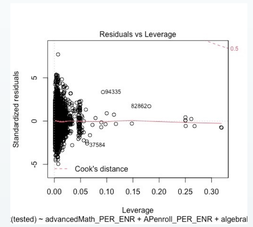


Figure IV. Cook’s Distance Residuals vs. Leverage Plot

As seen in Figure IV, there are many outliers that plotting Cook’s distance reveals. To adjust our model so that it does not overfit for influential outliers, the points that are identified as such are removed from the data, and the model is refitted again to yield final results.

After these adjustments, the distribution of the data points looks normal on the qqplot, with slight tails on either end. The final model with these adjustments leads to an adjusted R-squared of 0.6069 and a 10-fold cross-validation score of 0.003. This means on average, the predictions for SAT/ACT participation rate are 0.003 away from the actual participation rate on a scale from 0 to 1.

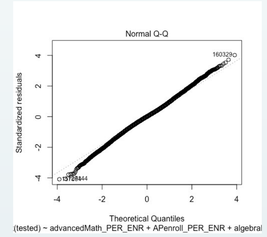


Figure V: Normality QQplot

Results & Conclusion

The model shows statistically significant positive associations for proportion of student taking advanced math, AP courses, and algebra II. There are also positive associations with proportion of students doing dual enrollment, participating in sports, and being referred to law enforcement. This model uses the baseline state as Alaska; Hawaii has a much higher test participation rate than Alaska at a statistically significant level.

The model shows statistically significant negative associations for proportion of students in biology class and ESL programs, as well as proportion of students with disabilities. California has a much lower test participation rate than Alaska at a statistically significant level.

Overall, it seems like schools where more students take fundamental classes, and participate in additional activities such as pre-college courses and sports have a higher rate of standardized test participation. Schools where bigger proportion of students are not native language speakers or have disabilities have challenges that might prevent them from taking the standardized test. However it is important to remember that this data is from when the pandemic first started and test-optionality first became a large-scale decision from institution, therefore it may not be generalizable to other time frames.

https://www.urban.org/research/publication/how-test-optional-college-admissions-expanded-during-covid-19-pandemic
https://www.forbes.com/sites/michaeltnietzel/2023/06/13/the-test-optional-college-admissions-movement-continues-to-grow/?sh=7d3ff92b1326
https://civilrightsdata.ed.gov/