Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 26-1581

Is Calculus a Gatekeeper in Selective Engineering Admissions? Insights from a Lab-in-the-Field Experiment with Admissions Officers

Bo-Kyung Byun
University of Michigan

Michael N. Bastedo
University of Michigan

Stephanie Carroll
University of Michigan

Julius DiLorenzo
University of Michigan

Jeremy Wright-Kim
University of Michigan

Lower-resourced high schools are far less likely to offer calculus, raising concerns that calculus availability may be a gatekeeper in selective college admissions. Using a lab-in-the-field experiment (292 admissions officers, 1768 total observations), we investigate whether engineering applicants are penalized for not taking calculus—even when it was unavailable to them. We also investigate whether Asian American applicants without calculus are more heavily penalized compared to similar White applicants. We find that not taking calculus decreases an applicant's probability of receiving an admit recommendation by 16 percentage points, but find no evidence that Asian American applicants without calculus are penalized more. Exploratory analyses using process-tracing data suggest the estimated calculus penalty varies by amount of time admissions officers spent examining the high school profile.

VERSION: September 2026

Suggested citation: Byun, Bo-Kyung, Michael N. Bastedo, Stephanie Carroll, Julius DiLorenzo, and Jeremy Wright-Kim. (2026). Is Calculus a Gatekeeper in Selective Engineering Admissions? Insights from a Lab-in-the-Field Experiment with Admissions Officers. (EdWorkingPaper: 26-1581). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/r5zn-e087

Is Calculus a Gatekeeper in Selective Engineering Admissions?

Insights from a Lab-in-the-Field Experiment with Admissions Officers

Nationwide, only 50% of public high schools offer calculus (U.S. Department of Education [ED], 2024), with lower-resourced schools significantly less likely to offer the course than their higher-resourced counterparts (U.S. Government Accountability Office [GAO], 2018). Students from high-income families are thus approximately four times more likely to complete calculus in high school compared to their peers from lower-income families (National Science Board [NSB], 2018). Similarly, only 35% of high schools with high Black and Latine student enrollment offer calculus (ED, 2024). Among high school graduates, only 9% of Black students and 15% of Latine students take calculus, compared to 22% of White students and 50% of Asian students (NSB, 2018).

Despite these disparities in access, calculus is widely perceived to carry significant weight in selective college admissions—particularly for students seeking admission to STEM programs (Anderson & Burdman, 2022). Within this context, how admissions offices evaluate otherwise high-achieving applicants who were unable to take calculus because it had not been offered at their high school remains a gray area in STEM admissions. In principle, institutions practicing holistic review should address such disparities by ensuring admissions officers evaluate student achievement in context of the opportunities that were available at their high school (Bastedo et al., 2018). Yet surveys and interviews with admissions officers across the country show discrepancies in admissions evaluation practices around calculus. While some admissions officers report they would not penalize an applicant who did not take calculus if the course was not offered at their high school, others dismiss or express hesitancy about STEM applicants without calculus—even when taking calculus is not a formal admissions requirement at their institution (Anderson & Burdman, 2022).

Using a lab-in-the-field experiment in which real-world admissions officers read simluated admissions materials, our study provides empirical evidence on whether otherwise high-achieving students who were unable to take calculus are penalized in selective engineering admissions. We also investigate whether Asian American engineering applicants who did not take high school calculus are more heavily penalized in admissions evaluations compared to otherwise similar White applicants, possibly due to stereotypes about Asian Americans and math proficiency in the U.S. (Biernat & Manis, 1994; Foschi, 2000). Finally, we conduct exploratory analyses of possible mechanisms driving findings from our experiment. In line with increasing calls for experimental researchers to pay more attention to potential mechanisms driving outcomes (Bertrand & Duflo, 2023; Macke et al., 2026), we collect process-tracing data—namely, the amount of time admissions officers spend reading the high school profile, which provides contextual information about applicants' educational opportunities (Nicola, 2021). These measures allow us to explore whether the weight placed on calculus varies with the amount of time admissions officers spend examining contextual information, a key element of holistic review.

Literature Review

Unequal access to calculus across socioeconomic and racial/ethnic lines is the result of accumulated structural barriers that disproportionately affect lower-resourced schools. These high schools often face immense pressure to raise student performance on state standardized tests, leading school administrators to prioritize core courses tested on state exams in budget allocation, course scheduling, and teacher assignments (Margolis et al., 2008). Advanced STEM-related coursework, on the other hand, is often considered peripheral and thus subject to eventual cancellation (Weis et al., 2015). Even where advanced courses are offered, access remains stratified along racial/ethnic lines both across and within schools: Black students are more likely than White students to attend high schools with fewer AP offerings, and less likely to enroll in AP coursework when attending the same school (Rodriguez & McGuire, 2019). This situation is exacerbated by the persistent challenge of out-of-field teaching: Lower-resourced schools frequently struggle to recruit and retain certified teachers capable of instructing advanced mathematics, limiting schools' capacity to offer courses like calculus even when student interest exists (Flores, 2007). Moreover, these challenges are compounded by course sequencing in mathematics, which often necessitates that students take Algebra I prior to high school to reach calculus coursework by senior year (Champion & Mesa, 2018). A curricular breakdown anywhere along the sequence can limit students' opportunity to pursue and complete high school calculus.

Questions also remain regarding the validity of calculus as a measure of college success (Giani et al., 2025). While a large body of literature examines the relationship between taking calculus in high school and college success and generally find a positive association between the two, these studies are limited by two key methodological constraints. Many of these studies rely on short-term college success such as grades received in introductory STEM courses taken during a student's first semester of college (Loehr et al., 2012; Tai et al., 2005), first-year college GPA (Patterson et al., 2011), and first-year retention (Mattern et al., 2009). Most critically, these studies tend to rely on correlational methods that fail to account for selection bias into high school calculus—making it difficult to disentangle whether observed advantages are due to having taken calculus in high school, or merely the underlying characteristics and resources of students who select into it.

Taken together, these two streams of research highlight a critical tension. Calculus is widely perceived to carry substantial weight in STEM admissions despite being systematically denied to many applicants from lower-resourced backgrounds, and despite limited evidence of its causal validity in predicting college success. Research has yet to empirically examine whether and to what extent admissions officers may penalize otherwise qualified applicants who were unable to take calculus. Our study addresses this gap by examining the following hypotheses:

Hypothesis 1a: Engineering applicants who were unable to take calculus because it was not offered at their high school will be less likely to receive an admit recommendation (0=deny/waitlist, 1=admit) compared to otherwise equivalent applicants who have taken calculus in high school.

Hypothesis 1b: Engineering applicants who were unable to take calculus because it was not offered at their high school will be evaluated more negatively in terms of admissions scores (1-10) compared to otherwise equivalent applicants who have taken calculus in high school.

We also investigate whether Asian American applicants face greater penalties than comparable White applicants if they did not take calculus due to lack of course offerings. Asian Americans are less likely to be admitted to the most selective universities than similarly qualified White applicants—a trend often attributed to application patterns and legacy, athletics, donor, private school or geographical preferences (Arcidiacono et al., 2022; Grossman et al., 2024; Kim & Bastedo, 2024; Park & Kim, 2020). We thus examine whether group-based stereotypes may also contribute to differential evaluation.

Asian Americans in the U.S. are stereotyped as being good at math (Cvencek et al., 2015; Shah, 2019), with 92% of Americans in a national survey agreeing with the statement that "in general, Asian Americans are good at math" (The Asian American Foundation, 2025, p. 11). These beliefs tie into the model minority stereotype, which in addition to perpetuating anti-Blackness and often pitting minoritized groups against one another, leads to assumptions that Asian Americans are generally successful and competent (Chou & Feagin, 2015; Poon et al., 2016). The Shifting Standards Model posits that evaluators may unconsciously adjust their standard of evaluation based on stereotypes associated with different groups, leading to different evaluations of similar qualifications (Biernat & Manis, 1994; Foschi, 2000). In the college admissions context, given Asian Americans as a group are often stereotyped as being good at math, they may be held to a higher standard for math coursework and penalized more severely for not meeting these expectations—even when such coursework was not available to them. This leads us to investigate the following hypotheses on how Asian American applicants are evaluated relative to White applicants in engineering admissions:

Hypothesis 2a: Asian American engineering applicants who were unable to take calculus because it was not offered at their high school will be less likely to receive an admit recommendation (0=deny/waitlist, 1=admit) compared to otherwise equivalent White applicants.

Hypothesis 2b: Asian American engineering applicants who were unable to take calculus because it was not offered at their high school will be evaluated more negatively in terms of admissions scores (1-10) compared to otherwise equivalent White applicants.

Data and Methods

Data for this study comes from a lab-in-the-field experiment employing a 2x2x2 full factorial design (Gerber & Green, 2012). The larger experiment involves eight combinations of applications that vary on three dimensions: (1) whether the applicant took calculus in high school or was not able to take calculus because it had not been offered, (2) whether the applicant is White or Asian American, and (3) whether the applicant submitted SAT scores. As outlined in our preregistration with OSF (link in Appendix A), this study focuses exclusively on the first two treatments, and controls for whether the applicant submitted SAT scores. Below, we first detail study participants, experimental stimuli, procedure, and analytical strategy used to investigate our study hypotheses. We then explain our use of process-tracing data—specifically, time admissions officers spent on each application, as well as each section of an application—to perform exploratory analyses of potential mechanisms driving our experimental findings.

Methods for Experiment

Participants. Study participants were college admissions officers working at four-year non-profit colleges or universities with an undergraduate admissions rate of 50% or lower based on 2019-2020 IPEDS data. We excluded admissions officers from institutions that did not offer STEM degrees at the undergraduate level, as well as admissions officers from institutions with specialized missions that made them meaningfully distinct from more general-purpose four-year colleges and universities that were the focus of our study (e.g., military academies, religious faith-based institutions). We also limited participants to admissions officers who reviewed applications regularly.

Participants were recruited through a partnership with the National Association for College Admissions Counseling (NACAC). A total of 295 admissions officers from 70 institutions participated in the experiment, but three participants did not respond to all questions surveying sociodemographic covariates to be included in our models. Given we did not have good data to impute these sociodemographic variables, we dropped the three participants with missing values from our sample. Our final analytical sample thus consists of 292 unique admissions officers from 69 institutions. Participants received USD $100 for completing the experiment. Appendix B provides a descriptive summary of institutions and admissions officer characteristics.

Experimental Stimuli. Our experiment involved eight combinations of applications that varied on three dimensions: whether the applicant took calculus in high school, whether the applicant is White or Asian American, and whether the applicant submitted SAT scores. All applications were embedded in an online Qualtrics survey, with application layout following that shown in Figure 1. Each application contained the same core components: an *applicant overview* section including student name, race/ethnicity, parental education, high school GPA, courses taken in current/most recent year of high school, and SAT/AP scores; a *high school profile* that clearly indicated calculus had not been offered at the student's high school; a *high school transcript* showing the student had taken all advanced math and science coursework available at their own high school; an *extracurriculars section*; and a *personal essay*. A header at the top of each application indicated the student was applying to engineering as an undeclared major, the *applicant overview* section indicated the student was a first-generation college applicant, and the *high school profile* section documented various indicators demonstrating the applicant came from a lower-resourced high school (e.g., high school-level college enrollment rates, average SAT/ACT scores, proportion of students on free/reduced price lunch). We conducted pilot testing and in-depth interviews with eight current and former admissions officers in our sampling frame prior to fielding our experiment to ensure all elements of the application unrelated to our treatment variables were comparable in quality across our eight combinations.

Because admissions standards differ by institutional selectivity, we adjusted our eight combinations of applications by institutional selectivity (highly selective, selective), which resulted in a total of 16 unique application files. This ensured admissions officers in different institutional selectivity tiers would read applications that were realistic for their own tier, and had a reasonable probability of acceptance or rejection (Bastedo & Bowman, 2017). Details about the creation of simulated admissions applications, pilot testing, and other data collection logistics are provided in Appendix A.

Procedure. Data collection occurred in two waves. In Wave 1, 292 admissions officers from 69 institutions each evaluated four applications tailored to the selectivity of the admissions officer's institution; the exact four combinations are listed in Section 1 of Table 1. In Wave 2, a randomly sampled subset of Wave 1 participants (150 admissions officers from 65 institutions) evaluated the remaining four application combinations (listed in Section 2 of Table 1). The four combinations included in Wave 1 (applications 1-4) consisted of applicants who had either taken calculus or submitted SAT scores but never both or neither. In contrast, Wave 2 applications (applications 5-8) consisted of applicants who had either both taken calculus and submitted SAT scores, or had neither taken calculus nor submitted SAT scores. Given lower anticipated variance in admissions evaluations for the Wave 2 combinations, the sample size for Wave 2 was set to roughly half of that allocated to Wave 1 combinations.

Aside from the combinations displayed, the format in which experimental stimuli were presented remained identical across waves. At the beginning of the experiment across both waves, admissions officers were informed that they would review simulated admissions applications for first-generation students applying to engineering in an undeclared major. Admissions officers were asked to use the same standards and criteria they would use when reading applications at their own institution. Each admissions officer then evaluated four files per wave using Qualtrics, with no opportunity to adjust evaluation of prior applications. Applications were counterbalanced to account for potential order effects.

After reading each application, admissions officers were asked to provide the admissions recommendation (deny, waitlist, accept) that they would make if that applicant had applied to their home institution. Admissions officers were also asked to rate the applicant on a scale of 1-10. Finally, following the experimental component, participants answered a brief survey that asked about their admissions office practices, as well as their own background and professional experience.

Empirical Strategy. We estimate a series of mixed-effects regression models with random intercepts at the admissions officer and institution level. We selected this approach over a fixed effects approach to maximize statistical efficiency and retain our full analytical sample while accounting for unobserved heterogeneity in baseline admissions tendencies across admissions officers and institutions. Because many admissions officers exhibited no within-admissions officer variation in their binary decisions (e.g., denied all applicants), a fixed effects logistic specification would result in the exclusion of 47% of our total sample (n=836) due to perfect prediction—reducing power and altering the effective estimand to rely only on the subset of admissions officers with within-variation. By employing a random intercepts approach, we use partial pooling to retain our complete sample (N=1,768) while still modeling unobserved heterogeneity. As a sensitivity check, we also estimate a series of regressions with admissions officer fixed effects, which yielded results consistent with our mixed-effects models in both direction and significance.

For hypotheses H1a and H2a with a binary dependent variable (admit recommendation: 0=deny/waitlist, 1=admit), we use mixed-effects logistic regression where:

$$\text{logit}\big[\Pr\big(Y_{ijk} = 1\big)\big] = \beta_0 + \beta_1(\mathit{TookCalc})_{ijk} + \beta_2(\mathit{Asian})_{ijk} + \beta_3(\mathit{TookCalc} \times \mathit{Asian})_{ijk} + X_{ij}\gamma + D_{ijk}\delta + Z_k\eta + v_k + u_{jk} + \varepsilon_{ijk} \quad (1)$$

Y_{ijk} is the binary admit recommendation (0=deny or waitlist, 1=accept) for each application *i* made by admissions officer *j* in institution *k*. This dependent variable was originally measured as 1=deny, 2=waitlist, 3=accept. However, many students who are placed on waitlists at selective institutions are never ultimately accepted (Clinedinst, 2019), so the accept decision is particularly important. In line with our preregistration, we collapse these categories into a binary outcome. $TookCalc$ is our treatment variable of interest, coded as 1 if the applicant took calculus in high school, and 0 if the applicant was unable to take calculus in high school because it had not been offered at their own school. Our experiment does not include a condition in which calculus was offered but the applicant chose not to take it. Estimates of the calculus penalty therefore reflect the extent to which admissions officers penalize applicants who lacked access to calculus, rather than the extent to which they penalize applicants who had access but chose not to enroll. We are also interested in the interaction term between $TookCalc$ and $Asian$ (0=White, 1=Asian American) to examine H2a.

X_{ij} is a vector for admissions officer-level covariates. These include admissions officer gender (1=male, 2=female, 3=non-binary/other), race/ethnicity (1=White, 2=Hispanic/Latine, 3=Black or African American, 4=Asian, 5=Two or more races, 6=Other), first-generation status (0=no, 1=yes), education level (1=BA, 2=MA, 3=Doctoral or other terminal degree), whether the admissions officer is working at their alma mater (0=no, 1=yes), and years of admissions officer experience on a 7-point scale (1=less than 1 year, 7=21 years or more). D_{ijk} is a vector for institution-level covariates, including institutional selectivity (1=highly selective, 2=selective), SAT/ACT policy (1=test optional, 2=testing required, 3=test free), whether the institution offers engineering degrees at the undergraduate level (0=no, 1=yes), and whether the institution is public (0=no, 1=yes). Average number of minutes spent reading each admissions file and the number of files read per week during busy times were measured through open-ended responses and then coded as continuous variables. A few participants reported ranges for these variables (e.g., 8-12 minutes per file); when this occurred, we used the mean. Z_k is a vector for covariates related to the design of our experiment, namely data collection wave (1=wave1, 2=wave2), the order an application was presented within each wave (1=first to 4=fourth), and whether the applicant submitted SAT scores (0=no, 1=yes). v_k refers to the institution-level random intercept, and u_{jk} refers to the admissions officer-level random intercept.

For H1b and H2b, which involve a continuous dependent variable (admissions score on a scale of 1-10), we use mixed-effects regression where:

$$Y_{ijk} = \beta_0 + \beta_1 (TookCalc)_{ijk} + \beta_2 (Asian)_{ijk} + \beta_3 (TookCalc \times Asian)_{ijk} + X_{ij}\gamma + D_{ijk}\delta + Z_k \eta + v_k + u_{jk} + \varepsilon_{ijk} \quad (2)$$

Appendix C shows raw coefficients from our mixed-effects regression models. To facilitate ease of interpretation, we present our findings below as Average Marginal Effects (AMEs). For our binary dependent variable, AMEs are reported as differences in predicted probabilities; For our continuous variable, AMEs are reported as differences in predicted percentage points. Finally, to account for the fact that we are testing multiple hypotheses, we set the false discovery rate at 0.05 using the Benjamini-Hochberg method (Benjamini & Hochberg, 1995).

Methods for Process-Tracing Data

To conduct exploratory analyses of potential mechanisms driving our experimental findings, we collected process-tracing data (e.g., where participants click, how much time they spend on different parts of provided stimuli) as admissions officers read each application. This data allows us to observe how participants allocate attention and search for information, thereby shedding light on cognitive processes underlying decision making (Daviss et al., 2026; Golder & Macy, 2014). Following Macke et al. (2026)'s approach to collecting digital trace data within online survey experiments, we adapted their publicly available JavaScript/CSS/HTML code to create static webpages corresponding to our eight application combinations. Our adaptation involved redesigning the layout from a labor market resume into a college admissions file resembling the Common App format (see Figure 1), and engineering the webpages to track process tracing measures relevant to our exploratory analyses of potential mechanisms. We presented the resulting admissions files as interactive webpages hosted on GitHub Pages, embedded into our Qualtrics surveys using an inline frame (iframe) of set pixel dimensions. This setup allowed us to collect two specific measures on how admissions officers allocated their time and attention when reading applications. First, we collected data on the total amount of time admissions officers spent interacting with each application. Second, we collected data on which sections of an application (high school profile, transcript, activities, personal essay) an admissions officer clicked to view, as well as how long they spent on each section. We were particularly interested in time spent on the high school profile, as this section provides contextual information necessary for admissions officers to interpret student achievement in light of available opportunities, a key condition of holistic review (Bastedo et al., 2018). To ensure more accurate measurement of time spent on particular sections, each admissions file webpage included several expandable sections that admissions officers had to click to view, and was designed so that only one section could be opened at a time.

To analyze our process-tracing data, we transform raw time (in seconds) spent on each application and section of an application into deciles, to account for the right-skewed distribution and aid interpretation. We then re-estimate analogous mixed-effects logistic regression models for admit recommendation (0=deny and/or waitlist, 1=admit) using time spent on each application section: high school profile, transcript, extracurriculars, and personal essay. For ease of presentation, the equation below shows the specification using time spent on the high school profile section.

$$\begin{aligned}\operatorname{logit}\left[\Pr\left(Y_{ijk}=1\right)\right] = \; & \beta_0 + \beta_1(\mathit{TookCalc})_{ijk} + \beta_2(\mathit{Asian})_{ijk} + \beta_3(\mathit{TookCalc} \times \mathit{Asian})_{ijk} \\ & + \beta_4(\mathit{Time\_HSProfile})_{ijk} + \beta_5(\mathit{TookCalc} \times \mathit{Time\_HSProfile})_{ijk} + \beta_6(\mathit{Time\_TotalApp})_{ijk} \\ & + X_{ij}\gamma + D_{ijk}\delta + Z_k\eta + v_k + u_{jk} + \varepsilon_{ijk}\end{aligned} \tag{3}$$

Equation (3) extends the baseline specification in Equation (1) in the section above by adding time spent on high school profile in deciles ($Time\_HSProfile$) and its interaction with $TookCalc$. This interaction tests whether the association between calculus and an admit recommendation is moderated by the amount of time admissions officers spent on the high school profile, which is our estimand of interest from the process-tracing data. We also include total time spent on an application in deciles ($Time\_TotalApp$) as a covariate, to address the possibility that this pattern reflects general thoroughness of reading rather than attention to context specifically. In line with how we report findings from our main experiment, we present these exploratory findings as differences in predicted probabilities derived from post-estimation margins.

Limitations

This study relies on a simulation of admissions decision making rather than a field experiment under real-world conditions. It is thus possible that admissions officers may not evaluate applicants as they would if their decisions had real consequences, although previous simulation studies in admissions have largely been confirmed by later real-world field experiments (e.g., Bastedo et al., 2022; Mabel et al., 2022). Additionally, to reduce cross-applicant comparisons, admissions officers in our experiment evaluated each application only once and were not allowed to revisit previous scores. The experimental setting thus diverges from real-world admissions, where admissions officers' application scoring and admit recommendations are often preliminary, with further review conducted by admissions committees and deans or directors of enrollment management before final decisions are rendered (Coleman & Keith, 2018). The admissions scores and admissions recommendations reported in study findings should therefore be interpreted as simulations of preliminary reader judgments rather than final admissions outcomes; Applicants who receive high ratings in our simulation may not ultimately secure admission in real-world settings.

Furthermore, this study limited its sample to White and Asian American applicants. This design choice allowed us to test whether Asian American students, who are often stereotyped as good at math, are penalized more heavily for not taking calculus than otherwise similar White applicants (Biernat & Manis, 1994; Foschi, 2000). A consequence of this design choice, however, is that our findings should not be assumed to generalize to racial/ethnic groups not represented in the study design. It is possible that admissions officers may interpret the absence of calculus differently for applicants from other racial/ethnic groups.

Finally, this study was conducted shortly after the Supreme Court's decision in *Students for Fair Admissions vs. Harvard*, which altered the legal landscape surrounding race-conscious admissions practices. Although applicants' racial/ethnic identity may emerge in real-world evaluations through personal essays or extracurricular involvement, participants in this study were shown applicants' racial/ethnic identities and instructed to evaluate applications according to the practices of their current institution. Consequently, it is possible that participants engaged with racial/ethnic information differently within the experimental setting than they would during admissions review processes in the current environment.

Findings

Main Experiment Findings

Hypotheses 1a and 1b examine whether students who were unable to take calculus in high school but maximized all other math and science curricular opportunities offered at their own school are penalized in engineering admissions, compared to otherwise identical applicants who took calculus. We find strong evidence confirming Hypothesis 1a. Holding all other factors constant, not taking calculus because it had not been offered at their high school decreases an engineering applicant's probability of receiving an admit recommendation by 16.4 percentage points on average (Table 2), even after taking all advanced math coursework available at their own high school. We observe similar trends when operationalizing the dependent variable as continuous (H1b), namely admissions score on a 1-10 point scale. All else equal, not taking calculus because it had not been offered at their high school decreased an engineering applicant's admissions score by 0.54 points on a 1-10 point scale—a 7.9% decline relative to the predicted admissions score of otherwise equivalent applicants who took calculus.

We find no evidence that Asian American engineering applicants who have not taken calculus are evaluated more negatively than equivalent White applicants, in terms of either admit recommendations (H2a) or admissions scores (H2b). We also find no evidence of heterogeneous effects for either H1a or H1b across admissions officer characteristics (e.g., whether the admissions officer is working at their own alma mater, admissions officer gender) or institution level characteristics (e.g., institution selectivity, type).

Process-Tracing Data Findings

Exploratory analyses of our process-tracing data suggest that the estimated admission penalty for not taking calculus varies by the amount of time admissions officers spent examining the high school profile ($OR = 0.89$, $p = .040$). As shown in Figure 3, while not taking calculus is associated with a lower probability of receiving an admit recommendation across all levels of time spent on the high school profile, the penalty for not taking calculus is smaller among reviews where admissions officers spent more time examining the high school profile. Among reviews where admissions officers spent the least time on the high school profile (lowest decile), not taking calculus is associated with a 22.1 percentage point lower probability of receiving an admit recommendation. However, the estimated penalty diminishes across deciles, reaching 11.0 percentage points in reviews where admissions officers spent the most time on the high school profile (highest decile)—effectively half the penalty estimated at the lowest decile. This moderation pattern remains statistically significant after controlling for time spent on total application, and is robust to alternative transformations of the time measures (e.g., log scale, alternative quantiles), as well as a specification using logistic regression with admissions officer fixed effects. Analogous models for the other application sections (high school transcript, extracurriculars, personal essay) yielded no statistically significant moderation effects on the calculus penalty.

Discussion and Implications

Our findings suggest that access to high school calculus functions as a critical gatekeeper for admission to selective engineering undergraduate programs. Given the socioeconomic and racial/ethnic disparities in access to calculus, these findings raise serious questions about admissions equity. The severe penalty for otherwise high-achieving engineering applicants who were unable to take calculus, even when the course had not been offered at an applicant's high school, indicates a disconnect between the goals of holistic review and what is implemented in practice. While holistic review intends to evaluate student achievement in light of opportunities that were available to them (Bastedo et al., 2018), our findings indicate that contextual information on course-taking opportunity may not be consistently incorporated when evaluating applicants' math preparation in engineering admissions. Without addressing this disconnect, calculus availability will continue to function as a barrier to otherwise high-achieving engineering applicants from lower-resourced backgrounds, perpetuating inequities in access to colleges and universities with the highest capacity to promote socioeconomic mobility.

Our exploratory process-tracing analyses provide further insight into this disconnect. The estimated calculus penalty identified in our experiment persisted across all levels of time spent on the high school profile, suggesting that engaging with contextual information does not fully eliminate the penalty associated with not taking calculus. At the same time, our process-tracing analyses show that this penalty is smaller among reviews where admissions officers spent more time on the high school profile. Although we cannot establish causality from our observational process-tracing measures, this pattern is consistent with the possibility that greater attention to contextual information could partially mitigate the penalty associated with not being able to take calculus. Admissions offices should thus explore ways to better integrate contextual information from high school profiles into their application reading platforms and increase the visibility of this information, especially regarding the most rigorous coursework available at each high school. Given the wide variance in the structure and quality of high school profiles (Nicola, 2021), it would also be beneficial for application platforms such as Common App to standardize the information required in the secondary school portion of the application to emphasize the information about course rigor that is most useful in the admission evaluation.

Additionally, admissions offices should consider ways to standardize their evaluation of curricular rigor so that context on calculus availability is considered in a systematic way (Rodriguez et al., 2025). This might be done through the incorporation of contextualized data points in the application reader, or through targeted training and norming around contextualized review that provides concrete tools for evaluating students without access to calculus. For example, admissions offices could study and discuss the characteristics of past students who have been successful at their institutions without having taken calculus in high school and create alternative data points by which to evaluate such applicants. Regardless of the mechanism used to support contextualization, it is critical that admissions officers are appropriately trained on implementing contextualized evaluative practices so that students are not penalized for the curricular limitations of their high school.

At the same time, we also must recognize that calculus is a necessary skill for many STEM majors and careers. Admitting students who have not taken calculus is only equitable if institutions provide the necessary academic infrastructure to ensure their success once they enroll in college. Investing in summer bridge programs and introductory calculus tracks designed for students with limited prior exposure should be viewed as essential infrastructure for equitable admissions. Some selective STEM-focused institutions, like Caltech, have also begun to accept alternative pathways to taking calculus in school by recognizing free certifications of calculus proficiency through Schoolhouse.world (Coffey, 2023).

Finally, our findings reinforce calls for admissions offices to provide greater transparency regarding the weight calculus holds in admissions. Admissions officers have widely claimed that lacking access to calculus should not—and does not—disadvantage potential applicants (Burdman et al., 2024; Williams, 2025). Institutions should clearly communicate their expectations around calculus in admissions materials and websites so that students can plan their coursework accordingly, or direct their applications to institutions that better align with their chosen curriculum.

Implications for Research

Our study highlights several avenues for future research. First, although calculus has received the most policy attention, it is but one example of how advanced math and science coursework tends to be unequally distributed along socioeconomic lines in the U.S. Advanced subjects like physics, chemistry, and computer science tend to be more difficult to access in lower-resourced schools, and admissions officers may inadequately adjust to this lack of access (Rodriguez et al., 2025). Future research should therefore investigate whether the penalty we observed for not being able to take calculus extends to these other subjects, thereby helping determine the full extent to which curricular inequities constrain postsecondary educational opportunities in STEM. Moreover, given the calculus penalty among engineering applicants in our study, future research should test whether a similar calculus penalty exists among applicants with other STEM majors, or even among non-STEM majors, at selective universities. While there continues to be debate around the importance of high school calculus for college STEM majors, a calculus penalty among students intending to major in English or Art History would be difficult to justify.

Second, our study contributes to growing work showing that pairing experiments with process-tracing measures can help open the “black box” of decision-making and evaluation (Bertrand & Duflo, 2023; Macke et al., 2026). In our study, process-tracing measures complement experimental estimates by exploring how admissions officers engage with the admissions file alongside their admit recommendations, helping identify promising directions for practice. Future research should continue to incorporate process-tracing tools into experiments to strengthen inference about the decision-making process underlying high-stakes evaluations.

Finally, we found no evidence that Asian American applicants who had not taken calculus because it was not offered at their high school were evaluated more negatively than similar White applicants. While the Shifting Standards model would suggest that Asian American applicants might be expected to demonstrate higher levels of math achievement (e.g., through taking higher level coursework), it is possible that the prevalence of stereotypes about Asian Americans and mathematics achievement (Cvencek et al., 2015; Shah, 2019) led to admission officers presuming mathematics competence for these applicants, despite the lack of calculus on their transcripts. The complicated relationship between race, mathematics achievement, and college admissions warrants further examination.

Conclusion

This research highlights an important issue that bridges the K-12 to postsecondary divide: how inequities in access to high school curriculum translate to inequities in postsecondary admissions. While our study demonstrates this in the specific case of access to high school calculus among engineering applicants, it is only one of the many ways that disparities in access to educational resources can impact postsecondary outcomes. However, if students from underrepresented socioeconomic and racial/ethnic backgrounds continue to be evaluated in selective STEM admission processes without consideration of these disparities, then we not only risk the perpetuation of social and educational inequality, but also the loss of significant talent from some of the fields that most need it.

References

Anderson, V., & Burdman, P. (2022). *A new calculus for college admissions: How policy, practice, and perceptions of high school math education limit equitable access to college*. Just Equations. https://justequations.org/resource/a-new-calculus-for-college-admissions-how-policy-practice-and-perceptions-of-high-school-math-education-limit-equitable-access-to-college

Arcidiacono, P., Kinsler, J., & Ransom, T. (2022). Legacy and athlete preferences at Harvard. *Journal of Labor Economics, 40*(1), 133–156. https://doi.org/10.1086/713744

Bastedo, M. N., Bell, D. W., Howell, J. S., Hsu, J., Hurwitz, M., Perfetto, G., & Welch, M. (2022). Admitting students in context: Field experiments on information dashboards in college admissions. *The Journal of Higher Education*, *93*(3), 327-374. https://doi.org/10.1080/00221546.2021.1971488

Bastedo, M. N., & Bowman, N. A. (2017). Improving admission of low-SES students at selective colleges: Results from an experimental simulation. *Educational Researcher*, *46*(2), 67-77. https://doi.org/10.3102/0013189X17699373

Bastedo, M. N., Bowman, N. A., Glasener, K. M., & Kelly, J. L. (2018). What are we talking about when we talk about holistic review? Selective college admissions and its effects on low-SES students. *The Journal of Higher Education*, *89*(5), 782-805. https://doi.org/10.1080/00221546.2018.1442633

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57*, 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bertrand, M., & Duflo, E. (2023). Field experiments: Correspondence studies. In A. Deshpande

(Ed.), *Handbook on Economics of Discrimination and Affirmative Action* (pp. 1-84). Springer.

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, *66*(1), 5-20. https://doi.org/10.1037/0022-3514.66.1.5

Burdman, P., Heller, R., & Baker, M. (2024). *The limits of calculus: Revisiting the role of math education in college admissions*. Just Equations. https://cdn.prod.website-files.com/61afa2b5ded66610900a0b97/674e10d56c6d7066cfaf20cd_JE_DEC%202024_THE%20LIMITS%20OF%20CALCULUS.pdf

Champion, J., & Mesa, V. (2018). Pathways to calculus in U.S. high schools. *PRIMUS*, *28*(6), 508–527. https://doi.org/10.1080/10511970.2017.1315473

Chou, R. S., & Feagin, J. R. (2015). *Myth of the model minority: Asian Americans facing racism* (2nd ed). Taylor and Francis. https://doi.org/10.4324/9781315636313

Clinedinst, M. (2019). *2019 state of college admission*. NACAC. https://nacacnet.org/wp-content/uploads/2022/10/soca2019_all.pdf

Coffey, L. (2023, September 21). No calculus? No problem at Caltech. *Inside Higher Ed*. https://www.insidehighered.com/news/admissions/traditional-age/2023/09/21/no-calculus-no-problem-caltechs-admissions

Coleman, A. L., & Keith, J. L. (2018). *Understanding holistic review in higher education admissions: Guiding principles and model illustrations*. The College Board and EducationCounsel. https://highered.collegeboard.org/media/pdf/understanding-holistic-review-he-admissions.pdf?utm

Cvencek, D., Nasir, N. S., O’Connor, K., Wischnia, S., & Meltzoff, A. N. (2015). The development of math-race stereotypes: “They say Chinese people are the best at math.”

Journal of Research on Adolescence, 25(4), 630–637. https://doi.org/10.1111/jora.12151

Daviss, C., Williams-Baron, E., & Macke, E. (2026). Hiring the ideal remote worker: The gendered implications of the rise of remote work. *Social Forces, 104*(4), 1-24. https://doi.org/10.1093/sf/soaf141

Flores, A. (2007). Examining disparities in mathematics education: Achievement gap or opportunity gap? *The High School Journal*, *91*(1), 29-42. https://www.jstor.org/stable/40367921

Foschi, M. (2000). Double standards for competence: Theory and research. *Annual Review of Sociology*, *26*(1), 21-42. https://www.jstor.org/stable/223435

Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. W. W. Norton & Company.

Giani, M., Lyra, F., & Tyner, A. (2025). *Calculus or statistics: Does it matter?* Thomas B. Fordham Institute. https://fordhaminstitute.org/national/research/calculus-or-statistics-does-it-matter

Golder, S. A., & Macy, M. W. (2014). Digital footprints: Opportunities and challenges for online social research. *Annual Review of Sociology, 40*(1), 129-152. https://doi.org/10.1146/annurev-soc-071913-043145

Grossman, J., Tomkins, S., Page, L., & Goel, S. (2024). The disparate impacts of college admissions policies on Asian American applicants. *Scientific Reports*, *14*(1), 4449. https://doi.org/10.1038/s41598-024-55119-0

Kim, S., & Bastedo, M. N. (2024). Who gets their first choice? Race and class differences in college admissions outcomes. *AERA Open*, *10*(1), 1-16. https://doi.org/10.1177/23328584241298951

Loehr, J. F., Almarode, J. T., Tai, R. H., & Sadler, P. M. (2012). High school and college biology: A multi-level model of the effects of high school courses on introductory course performance. *Journal of Biological Education, 46*(3), 165–172. https://doi.org/10.1080/00219266.2011.617767

Mabel, Z., Hurwitz, M. D., Howell, J., & Perfetto, G. (2022). Can standardizing applicant high school and neighborhood information help to diversify selective colleges? *Educational Evaluation and Policy Analysis*, *44*(3), 505-531. https://doi.org/10.3102/01623737221078849

Macke, E., Daviss, C., & Williams-Baron, E. (2026). Untapped potential: Designed digital trace data in online survey experiments. *Sociological Methods & Research, 55*(1), 299-339. https://doi.org/10.1177/00491241241268770

Margolis, J., Estrella, R., Goode, J., Holme, J. J., & Nao, K. (2008). *Stuck in the shallow end: Education, race, and computing*. MIT Press.

Mattern, K. D., Shaw, E. J., & Xiong, X. (2009). *The relationship between AP exam performance and college outcomes*. Research Report No. 2009-4. College Board. https://eric.ed.gov/?id=ED561021

National Science Board. (2018). *Science and engineering indicators 2018* (NSB-2018-1). National Science Foundation. https://www.nsf.gov/statistics/2018/nsb20181/assets/nsb20181.pdf

Nicola, T. P. (2021). Assessing applicants in context? School profiles and their implications for equity in the selective college admission process. *Journal of Diversity in Higher Education, 15*(6), 700-715. https://psycnet.apa.org/doi/10.1037/dhe0000318

Park, J. J., & Kim, S. (2020). Harvard's personal rating: The impact of private high school attendance. *Asian American Policy Review*, *30*(Spring), 2-4. https://www.proquest.com/openview/de368a656bec6995583a012898a6b092/1?pq-origsite=gscholar&cbl=436384

Patterson, B. F., Packman, S., & Kobrin, J. L. (2011). *Advanced Placement exam-taking and performance: Relationships with first-year subject area college grades*. Research Report No. 2011-4. College Board. https://eric.ed.gov/?id=ED561033

Poon, O., Squire, D., Kodama, C., Byrd, A., Chan, J., Manzano, L., Furr, S., & Bishundat, D. (2016). A critical review of the model minority myth in selected literature on Asian Americans and Pacific Islanders in higher education. *Review of Educational Research, 86*(2), 469–502. https://doi.org/10.3102/0034654315612205

Rodriguez, A., Lebioda, K., Rodríguez-Wilhelm, D., Arroyo, C., Skiles, J., McKenzie, B., & Wilson, N. J. (2025). "They're outside of the process": How admissions officers maintain inequality when evaluating course rigor through holistic admissions. *American Educational Research Journal*, *62*(5), 1019-1055. https://doi.org/10.3102/00028312251360905

Rodriguez, A., & McGuire, K. M. (2019). More classes, more access? Understanding the effects of course offerings on Black-White gaps in Advanced Placement course-taking. *The Review of Higher Education*, *42*(2), 641–679. https://doi.org/10.1353/rhe.2019.0010

Shah, N. (2019). "Asians are good at math" is not a compliment: STEM success as a threat to personhood. *Harvard Educational Review*, *89*(4), 661-686. https://doi.org/10.17763/1943-5045-89.4.661

Tai, R. H., Sadler, P. M., & Loehr, J. F. (2005). Factors influencing success in introductory college chemistry. *Journal of Research in Science Teaching, 42*(9), 987–1012.

https://doi.org/10.1002/tea.20082

The Asian American Foundation. (2025). *STAATUS Index 2025*. https://taaf-staatus-2025.vercel.app/STAATUS_Index_2025.pdf

U.S. Department of Education, Office for Civil Rights. (2024). *Student access to and enrollment in mathematics, science, and computer science courses and academic programs in U.S. public schools*. https://www.ed.gov/sites/ed/files/about/offices/list/ocr/docs/crdc-student-access-enrollment.pdf

U.S. Government Accountability Office. (2018). *K-12 education: Public high schools with more students in poverty and small schools provide fewer academic offerings to prepare for college* (Report to the Ranking Member, Committee on Education and the Workforce, House of Representatives GAO-19-8). https://www.gao.gov/assets/gao-19-8.pdf

Weis, L., Eisenhart, M., Cipollone, K., Stich, A. E., Nikischer, A. B., Hanson, J., Leibrandt, S. O., Allen, C. D., & Dominguez, R. (2015). In the guise of STEM education reform: Opportunity structures and outcomes in inclusive STEM-focused high schools. *American Educational Research Journal*, *52*(6), 1024-1059. https://doi.org/10.3102/0002831215604045

Williams, R. (2025). The inequitable race to calculus. *Journal of College Admission*, *265*, 21-25.

Tables

Table 1. Application Combinations by Wave

Section 1. Wave1 Data Collection

		Calculus Taken?	White vs Asian?	SAT Submitted?
Highly Selective	App1	No	White	Yes
	App2	No	Asian	Yes
	App3	Yes	Asian	No
	App4	Yes	White	No
Selective	App1	No	White	Yes
	App2	No	Asian	Yes
	App3	Yes	Asian	No
	App4	Yes	White	No

Section 2. Wave2 Data Collection

		Calculus Taken?	White vs Asian?	SAT Submitted?
Highly Selective	App5	Yes	White	Yes
	App6	Yes	Asian	Yes
	App7	No	Asian	No
	App8	No	White	No
Selective	App5	Yes	White	Yes
	App6	Yes	Asian	Yes
	App7	No	Asian	No
	App8	No	White	No

Table 2. AMEs of Calculus and Applicant Race on Admission Recommendation and Score

Panel A. Admit Recommendation (0=deny/waitlist, 1=admit)

	Predicted Probability (SE)	Average Marginal Effect (SE)
H1a: Main Effect of Taking Calculus		
Took Calculus (Baseline)	0.641 (0.029)	
Did Not Take Calculus	0.477 (0.030)	-0.164*** (0.019)
H2a: Effect of Applicant Race/Ethnicity (conditional on no calculus)		
White, Did not take calculus (Baseline)	0.462 (0.032)	
Asian American, Did not take calculus	0.491 (0.032)	0.029 (0.023)

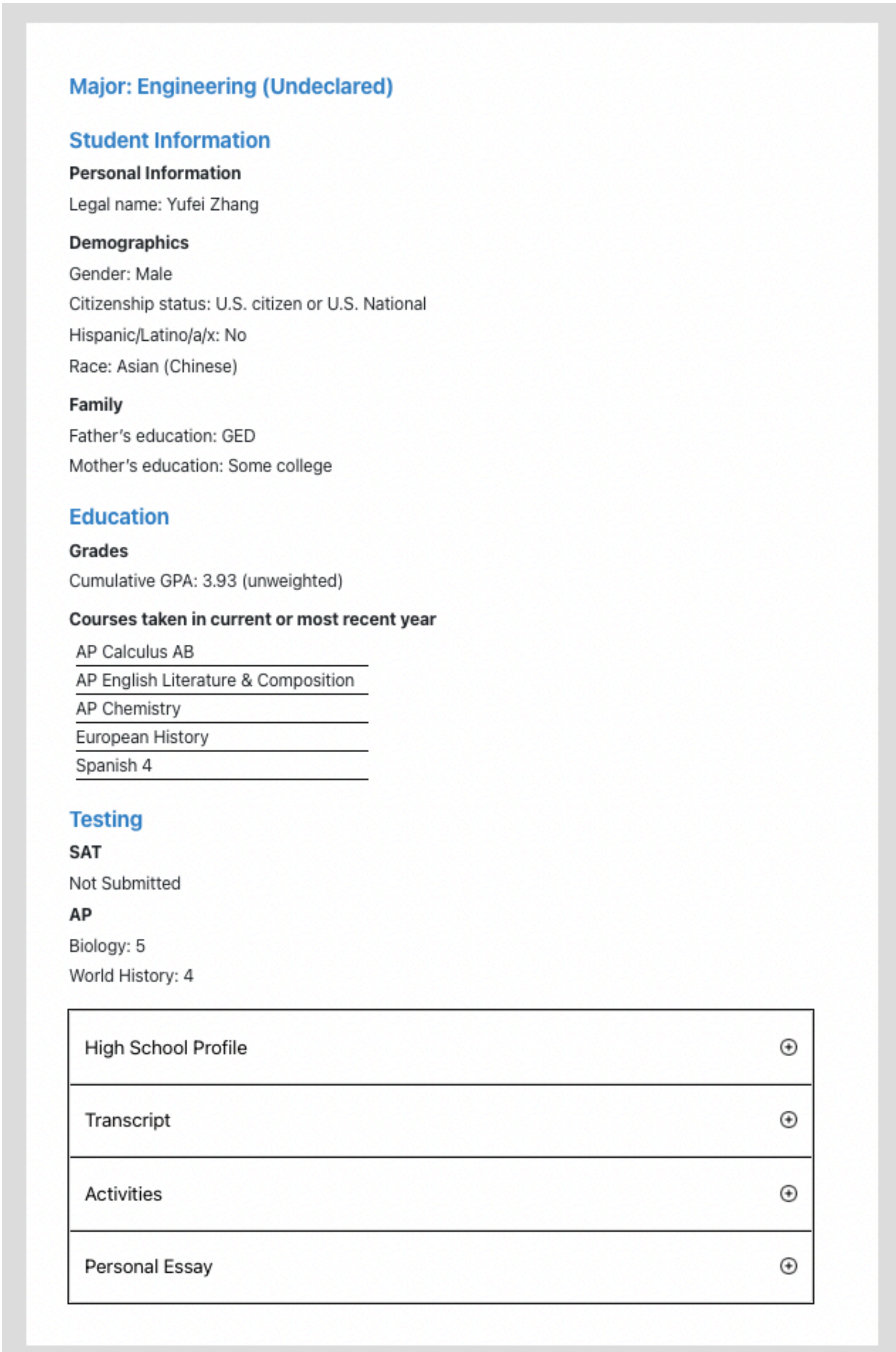
Panel B. Admission Score (continuous, 1–10)

Comparison	Predicted Percentage Points (SE)	Average Marginal Effect (SE)
H1b: Main Effect of Taking Calculus		
Took Calculus (Baseline)	6.848 (0.083)	
Did Not Take Calculus	6.305 (0.083)	-0.543*** (0.055)
H2b: Effect of Applicant Race/Ethnicity (conditional on no calculus)		
White, Did not take calculus (Baseline)	6.237 (0.091)	
Asian American, Did not take calculus	6.374 (0.091)	0.137 (0.076)

Note. Estimates are average marginal effects (AMEs) from a mixed-effects logistic regression model (Panel A) and a mixed-effects regression model (Panel B) with random intercepts at the admissions officer and institution level, admissions officer and institution covariates, and experimental design controls; N=1,768 application evaluations by 292 unique admissions officers. AMEs are defined as differences in predicted probabilities relative to the listed baseline (Panel A), and differences in predicted percentage points relative to the listed baseline (Panel B). The AME for "Did not take calculus" is averaged over all applicants; the AME for "Asian American, Did not take calculus" is conditional on applicants who did not take calculus. *** $p < .001$. AMEs for H1a and H1b remain statistically significant after Benjamini-Hochberg adjustment (False Discovery Rate=5%).

Figures

Figure 1. Example of application presented online through Qualtrics



Major: Engineering (Undeclared)

Student Information

Personal Information

Legal name: Yufei Zhang

Demographics

Gender: Male

Citizenship status: U.S. citizen or U.S. National

Hispanic/Latino/a/x: No

Race: Asian (Chinese)

Family

Father's education: GED

Mother's education: Some college

Education

Grades

Cumulative GPA: 3.93 (unweighted)

Courses taken in current or most recent year

AP Calculus AB
AP English Literature & Composition
AP Chemistry
European History
Spanish 4

Testing

SAT

Not Submitted

AP

Biology: 5

World History: 4

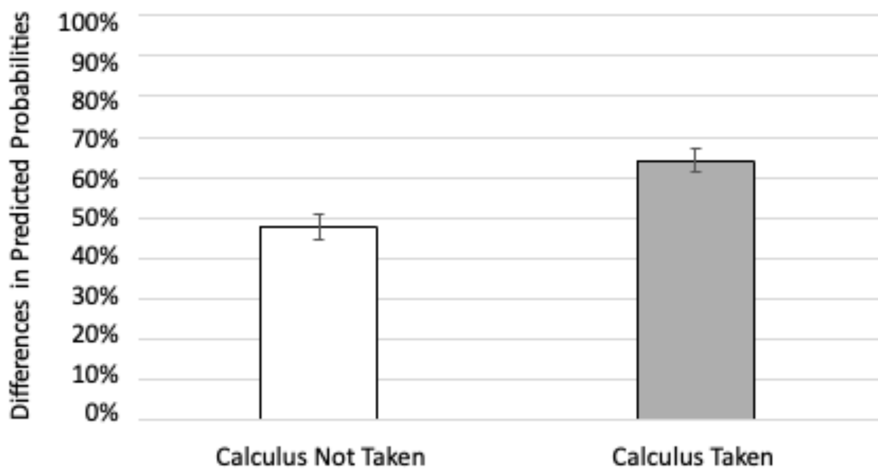
High School Profile

Transcript

Activities

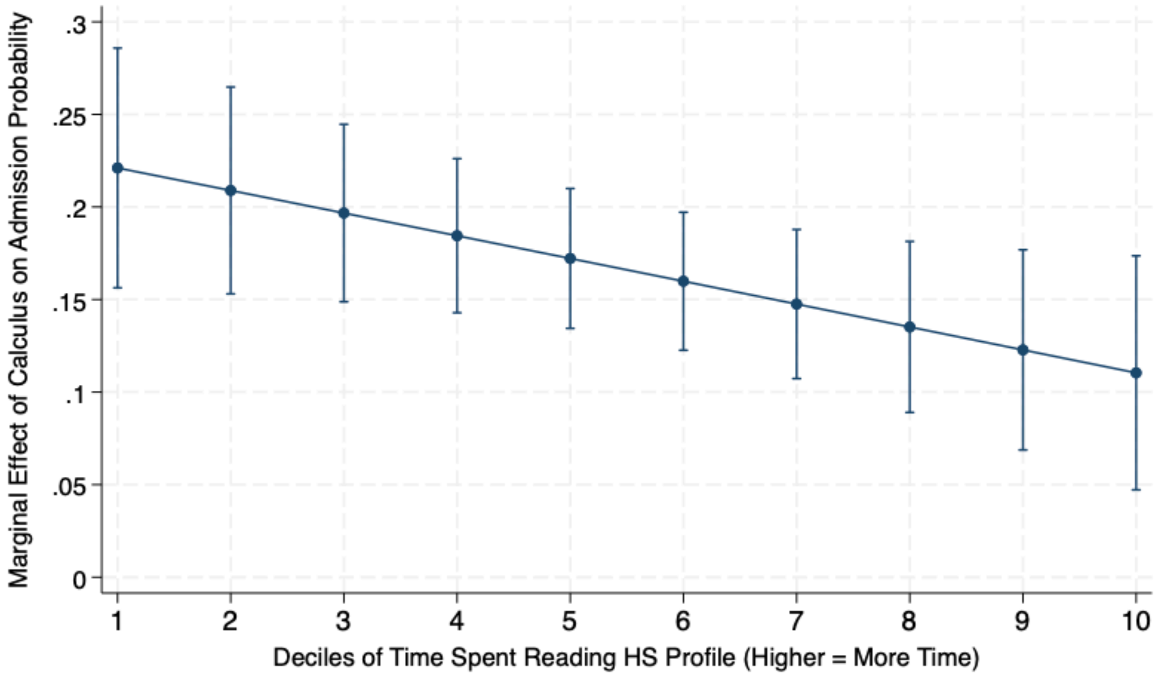
Personal Essay

Figure 2. Difference in Predicted Probability of Admit Recommendation, by Calculus



Note. Average marginal effects showing how the predicted probability of admission differs for applicants who have taken calculus versus those who have not. The outcome is a binary indicator of admission (Admit = 1, Waitlist/Deny = 0). Bars plot the estimated predicted probability of admission for each group, with standard errors shown as error bars.

Figure 3. Estimated Calculus Penalty by High School Profile Reading Time



Note. Estimates are based on admissions officer-application level evaluations. Time spent reading high school profile was observed, not experimentally assigned.

Appendix A

Development of Simulated Applications and Pilot Testing

The construction of simulated applications used in our experiment was guided by two critical goals: maximizing internal validity while ensuring external realism. Maximizing internal validity required that we clearly manipulate our key variables of interest (whether the applicant took calculus, and whether the applicant identified as White or Asian American) while ensuring the quality of all other application characteristics unrelated to our variables of interest remained constant across all experimental conditions. At the same time, we sought to achieve external realism by ensuring each set of eight applications created for each institutional selectivity tier (highly selective and selective) were appropriately tailored so that admissions officers in each tier would evaluate applications that reasonably reflected applications they regularly read for their own institutions in real-life settings. Below, we provide further details on how we systematically controlled for components of the application that did not have to do with our two variables of interest.

Academic Components

The process of developing our simulated applications began with benchmarking values for high school grades and coursework from simulated applications used in previous experimental research with admissions officers (e.g., Bastedo & Bowman, 2017; Bastedo et al. 2022). These benchmarked values were subsequently refined based on expert consultation with eight current and former admissions officers who had experience working at institutions within our target sampling frame. Specifically, we benchmarked values for high school grades and coursework from Bastedo and Bowman's (2017) "low-SES applicant" condition. This condition aimed to represent a high-performing student from a lower-SES background who had maximized all available curricular opportunities at their own high school. To assign appropriate high school grades and coursework for this condition, Bastedo and Bowman (2017) use the Education Longitudinal Study (ELS:2002), a federal dataset that tracks a nationally representative cohort of 10th-grade students starting in 2002, with follow-up data collected over the course of two, four, and ten years. While we acknowledge that ELS:2002 is older, it remains one of the most reliable sources of existing data documenting the high school grades and coursework for a nationally representative sample of students. This scope is vital given more recent data at the institution-level (e.g., IPEDS, Common Data Set) focus exclusively on the characteristics of admitted students, who likely differ from those of the overall applicant pool. Bastedo and Bowman (2017) assigned the low-SES applicant condition high grades (70th percentile for applicants applying to institutions with the highest tier of selectivity, and 60th percentile for applicants applying to institutions with the second to highest tier of selectivity), and set the number of AP courses taken to the 30th percentile for both selectivity tiers. Given the literature documents trends of high school grade inflation since the time ELS data was collected (Hurwitz & Lee, 2018), we consulted current/former admissions officers with experience working at institutions included in our study sampling frame. Based on their input, we slightly adjusted upward applicants' high school GPA.

The calibration of SAT scores for applications in the "test score submitted" condition required integrating data from multiple sources to ensure assigned scores were both current and realistic for our established two selectivity tiers. We collected SAT percentile data from each institution's incoming class as reported in the 2019-2020 IPEDS database, as this represented the most recent data year where the large majority of applicants submitted test scores prior to a wider array of institutions adopting test-optional policies. For institutions in our sampling frame that possessed missing values for this variable, we supplemented the dataset by drawing upon information from the corresponding year of the Common Data Set (CDS). After this process, none of the institutions in our sampling frame had missing values for SAT percentile. We then assigned applications in the "test score submitted" condition specific SAT scores tailored to the two institutional selectivity tiers. To be more specific, we calculated—across institutions within each selectivity tier—the 25th percentile EBRW score and 75th percentile math score among enrolled students. The asymmetry was deliberate, given engineering applicants typically have stronger math than EBRW scores relative to the general applicant pool. Moreover, this approach helped produce composite test scores that were neither too high to make the file an easy admit, nor too low to make the file an easy deny, thus situating applicants near the margin of admission. This procedure yielded composite SAT scores of 1470 for the highly selective tier (with subscores of either 790 for math and 680 for EBRW, or 780 for math and 690 for EBRW), and composite scores of 1360 for institutions in the selective tier (with subscores of either 740 for math and 620 for EBRW, or 730 for math and 630 for EBRW). Consultations with current and former admissions officers confirmed that these assigned SAT scores were competitive for engineering applicants from lower-resourced backgrounds when using standards they would regularly use to evaluate similar applications at their own institutions.

Non-Academic Components

To ensure that non-academic components in the application (extracurricular activities, personal essay) did not confound estimation of effects of our variables of interest, these additional components were specifically designed to be similar in quality and quantity across all our simulated applications. For extracurricular activities, we maintained consistency across both institutional selectivity tiers: Each applicant participated in one STEM-related club, one non-

STEM related club, was a member of NHS, participated in one volunteer activity, and worked a paid job. In line with consultations with current and former admissions officers, the time commitment and continuity of these activities (including hours per week, weeks per year, and total years participated) were adjusted to be largely consistent across all applications.

Similarly, we paid particular attention to controlling for the quality of personal essays, which were identical across institutional selectivity tiers. A study author with former admissions officer work experience wrote the initial essay drafts. All essays expressed a general interest in engineering, and did not talk about any form of experience that would provide insight into the applicant's socioeconomic status (e.g., overcoming dearth of resources, financial obstacles). Consultations with current/former admissions officers confirmed essays were similar in overall quality. In response to feedback received from admissions officers, the length of each essay was adjusted to be roughly equivalent.

Data Collection Logistics and Procedure

We recruited study participants through a partnership with the National Association for College Admissions Counseling (NACAC). Invitations were distributed to admissions officers who met our selection criteria (admissions officers who worked at institutions in our sampling frame and regularly reviewed applications as part of their job) to participate in our study in-person at the NACAC annual meeting in Los Angeles. Given our in-person recruitment yielded a sample size lower than our target specified in our preregistration, we subsequently recruited additional admissions officers who met our inclusion criteria to participate online after the NACAC annual meeting. Both in-person and fully online participants completed identical computer-based surveys, which were administered using Qualtrics survey software. All participants who completed our survey received $100 gift cards as compensation for their time.

For in-person data collection, participants signed up for time slots, were seated at one of six computers in a dedicated room at the conference center, where the survey link was loaded onto the computer for them. For fully online data collection, participants were emailed a survey link. Aside from this difference, the procedure for both in-person and fully online participants was identical: Both viewed and completed a consent form on the first page of the online survey prior to participating in our experiment, read four simulated applications, and answered a brief survey that asked about their admissions office practices as well as their own background and professional experience. To mitigate the risk of potential contamination effects, participants were explicitly instructed at the conclusion of the experiment to refrain from discussing its content with any other individuals. Further post-hoc analyses confirmed that our study findings were not moderated by whether participation occurred in-person or entirely online. We pre-registered our planned analyses at https://osf.io/xrg54/overview?view_only=4c31b6006eb54803ae3fe2635344a326.

Supplementary References

Bastedo, M. N., & Bowman, N. A. (2017). Improving admission of low-SES students at selective colleges: Results from an experimental simulation. *Educational Researcher*, *46*(2), 67–77. https://doi.org/10.3102/0013189X17699373

Bastedo, M. N., Glasener, K. M., Deane, K. C., & Bowman, N. A. (2022). Contextualizing the SAT: Experimental evidence on college admission recommendations for low-SES applicants. *Educational Policy*, *36*(2), 282-311.

Hurwitz, M. & Lee, J. (2018). Grade inflation and the role of standardized testing. In J. Buckley, L. Letukas, & B. Wildavsky (Eds.), *Measuring success: Testing, grades, and the future of college admissions* (pp. 64-93). Johns Hopkins University Press.

Appendix B

Table B-1. Admissions Officer- and Institution-Level Descriptive Statistics

	Participants (N=292)	Institutions (N=69)
Institution-Level Covariates		
Selectivity		
Highly Selective	45%	45%
Selective	55%	55%
Engineering Degree Offered		
Yes	78%	75%
No	22%	25%
Control		
Public	27%	25%
Private	73%	75%
Testing Status		
Test-Required	12%	10%
Test-Optional	77%	78%
Test-Free/Blind	11%	12%
Admissions Officer-Level Covariates		
Gender		
Female	64%	
Male	34%	
Non-Binary/Other	2%	
Race/Ethnicity		
White	56%	
Hispanic/Latine	14%	
Black/African American	12%	
Asian	9%	
Two or more races	7%	
Other	1%	
Years Work Experience		
1-2 Years	16%	
3-5 Years	26%	
6-10 Years	22%	
11-15 Years	17%	
16-20 Years	9%	
21+ Years	10%	
Education Level		
Bachelors' Degree	37%	
Masters' Degree	56%	
Doctoral/Terminal Degree	7%	
Works at Alma Mater	29%	
First Generation	34%	

Note. Total observations are N=1768, as each of the 292 participants reviewed 4-8 files.

Appendix C

Table C-1. Mixed Effects Logistic Regression Estimates (DV: admit recommendation)

	Odds Ratio	Standard Errors
Variables of Interest		
Calculus Taken *(vs. not taken)*	5.300***	(1.198)
East Asian American Applicant *(vs. White)*	1.141	(0.318)
Calculus x East Asian American Applicant	0.696	(0.212)
Admissions Officer-Level Covariates		
Gender *(vs. Female)*		
Male	0.853	(0.288)
Non-Binary/Other	1.186	(1.319)
Race/Ethnicity *(vs. White)*		
Hispanic/Latine	0.570	(0.274)
Black/African American	0.868	(0.440)
Asian	0.846	(0.485)
Two or more races	0.707	(0.451)
Other	0.204	(0.287)
Employee Characteristics		
Years Work Experience	0.842	(0.095)
Works at Alma Mater *(vs. Non-Alum)*	0.693	(0.243)
First Generation *(vs. Continuing Gen,)*	1.926	(0.684)
Education Level *(vs. Bachelors)*		
Masters' Degree	0.924	(0.323)
Doctoral/Terminal Degree	0.395	(0.266)
Institution-Level Covariates		
Selective *(vs. Highly Selective)*	6.288**	(3.640)
Engineering Degree Offered *(vs. Not Offered)*	0.801	(0.538)
Public *(vs. Private)*	2.191	(1.491)
Testing Status *(vs. Test-Optional)*		
Test-Required	0.138*	(0.115)
Test-Free/Blind	1.093	(0.925)
Reading behaviors		
Minutes per File	0.907	(0.058)
Files per Week	0.988**	(0.004)
Design-Related Covariates		
Wave 2 *(vs. Wave 1)*	1.106	(0.185)
App Presented Order *(vs. Order 1)*		
2	0.832	(0.168)
3	0.943	(0.193)
4	0.882	(0.179)
SAT Status *(vs. SAT Not Submitted)*		
SAT Submitted	1.772**	(0.385)
SAT x East Asian American Applicant	1.284	(0.389)

Note: N=1768 unique evaluations by 292 admission officers at 69 unique institutions. Standard errors in parentheses. * $p < .05$. ** $p < .01$. *** $p < .001$

Table C-2. Mixed Effects Regression Estimates (DV: admissions score)

	Coefficients	Standard Errors
Variables of Interest		
Calculus Taken *(vs. not taken)*	0.645***	(0.078)
East Asian American Applicant *(vs. White)*	0.138	(0.104)
Calculus x East Asian American Applicant	-0.206	(0.111)
Admissions Officer-Level Covariates		
Gender *(vs. Female)*		
Male	0.257	(0.133)
Non-Binary/Other	-0.091	(0.415)
Race/Ethnicity *(vs. White)*		
Hispanic/Latine	0.032	(0.194)
Black/African American	0.088	(0.202)
Asian	-0.029	(0.228)
Two or more races	0.238	(0.252)
Other	-0.722	(0.541)
Employee Characteristics		
Years Work Experience	-0.098*	(0.045)
Works at Alma Mater *(vs. Non-Alum)*	-0.104	(0.140)
First Generation *(vs. Continuing Gen,)*	0.149	(0.140)
Education Level *(vs. Bachelors)*		
Masters' Degree	-0.073	(0.140)
Doctoral/Terminal Degree	-0.119	(0.264)
Institution-Level Covariates		
Selective *(vs. Highly Selective)*	0.376*	(0.179)
Engineering Degree Offered *(vs. Not Offered)*	-0.038	(0.213)
Public *(vs. Private)*	0.707**	(0.212)
Testing Status *(vs. Test-Optional)*		
Test-Required	-1.043***	(0.263)
Test-Free/Blind	0.062	(0.275)
Reading behaviors		
Minutes per File	-0.016	(0.022)
Files per Week	-0.003**	(0.001)
Design-Related Covariates		
Wave 2 *(vs. Wave 1)*	-0.075	(0.062)
App Presented Order *(vs. Order 1)*		
2	-0.040	(0.074)
3	-0.016	(0.074)
4	-0.021	(0.074)
SAT Status *(vs. SAT Not Submitted)*		
SAT Submitted	0.183*	(0.078)
SAT x East Asian American Applicant	-0.003	(0.111)

Note: N=1768 unique evaluations by 292 admission officers at 69 unique institutions. Standard errors in parentheses. * $p < .05$. ** $p < .01$. *** $p < .001$