The Impact of Race-Blind and Test-Optional Admissions on Racial Diversity and Merit*

Allen Sirolly† Hongyao Ma‡ Yash Kanoria§

March 14, 2025

Abstract

How significant was the role of racial preferences in U.S. college admissions before the Supreme Court's 2023 decision to ban race-based affirmative action? How much might test-optional admission policies impact racial diversity and academic merit? In this work, we estimate a simple model of college admissions decisions from 2012–2020, leveraging a novel dataset of applicant profiles and admissions outcomes across the full spectrum of college selectivity. We find that, broadly, the impact of race and testing policies on diversity and merit of admits decreases by college selectivity. For America's less selective colleges that collectively enroll over three-quarters of students, fully eliminating preferences for race (and for unobserved factors which are correlated with race) has little impact on the proportion of underrepresented minorities (URM) and on the average SAT score of admitted students. In contrast, for the 34 most selective colleges accounting for 3 percent of total enrollment, our estimates suggest that admissions going "race blind"—absent any compensating changes in admissions criteria—could reduce URM admission by one-third while increasing the average SAT score of admits by no more than 10 points. We also estimate that universal test-optional admission may lead to a small increase in the proportion of URMs at the most selective colleges, while decreasing the average SAT score by up to 10 points. At less selective institutions, the effects are estimated to be negligible.

1 Introduction

In June 2023, the U.S. Supreme Court banned the consideration of race in undergraduate admissions in the landmark case *Students for Fair Admissions (SFFA) v. Harvard*. In the wake of the decision, highly selective colleges rallied to reaffirm their commitments to racial diversity on campus, demonstrating the significant role that race had played in college admissions.[1] In a parallel

*The authors thank Itai Ashlagi, Nick Arnosti, Omar Besbes, Yeon-Koo Che, Wouter Dessein, Laura Doval, Nikhil Garg, Sergey Gitlin, Navin Kartik, Tzuo Hann Law, Irene Lo, Jake Marcinek, Doron Nissim, Jonah Rockoff, Nicola Rosaia, Peng Shi, Nachum Sicherman, Fanyin Zheng, as well as participants in the Columbia DRO Brown Bag Seminar, INFORMS 2023 and 2024, and the 2023 Market Design Workshop in Santiago, for helpful comments and discussion. AS and YK gratefully acknowledge the support of the National Science Foundation through grant CMMI-1653477.

†Decision, Risk, and Operations Division, Columbia Business School, ASirolly26@gsb.columbia.edu

‡Decision, Risk, and Operations Division, Columbia Business School, hongyao.ma@columbia.edu

§Decision, Risk, and Operations Division, Columbia Business School, ykanoria@columbia.edu

[1] President Christopher L. Eisgruber of Princeton University, for example, announced shortly after the ruling "Princeton will pursue [diversity] with energy, persistence, and a determination to succeed despite the restrictions imposed by the Supreme Court". https://www.princeton.edu/news/2023/06/29/president-eisgrubers-statement-supreme-court-affirmative-action-decision, accessed February 6, 2024.

development, many colleges have relaxed standardized testing requirements for admission, and no longer require their applicants to submit SAT or ACT scores (see Figures 11 and 12 in Appendix A). Some colleges, including colleges in the University of California and California State University systems, have taken this one step further and no longer take test scores into consideration at all.[2] The adoption of such policies preceded but was greatly accelerated by the COVID-19 pandemic, which limited testing access. Yet, as of 2024, after pandemic-related public health concerns have receded, most colleges—across the entire spectrum of selectivity—have elected to keep these policies in place.

While the motivation for not reinstating test requirements may be multifaceted (e.g., boosting the size of the applicant pool and improving ranking), going *test-optional* is now viewed by many as part of colleges' toolbox to promote racial diversity in the new *race-blind* regime. The rationale is that underrepresented minorities (URM) score substantially lower on standardized tests than non-URMs on average, and that the gap has scarcely narrowed over the past two decades.[3] Notably, *Strategies for Increasing Diversity and Opportunity in Education*, a government report published at the behest of the Biden administration, identifies "reassessing the use of entrance exams" as one possible strategy for improving diversity that institutions are encouraged to consider (Office of the Under Secretary, 2023).[4]

Meanwhile, a sizable minority of highly selective colleges, following Dartmouth and MIT, have opted to reinstate testing requirements.[5] Dartmouth's announcement cited two recent studies of Ivy Plus colleges (Chetty, Deming, and Friedman (2023) and Friedman, Sacerdote, and Tine (2024)), both indicating that standard test scores more accurately predict various measures of college success in comparison to high school GPA.[6] MIT's internal research similarly also concluded that "our ability to accurately predict student academic success at MIT is significantly improved by considering standardized testing—especially in mathematics—alongside other factors".[7] In the wake of *SFFA v. Harvard*, deliberating between test-optional and test-mandatory policies necessitates a careful assessment of the former role of racial preferences, as well as the potential impact of testing policies on both the racial diversity and the average SAT scores of admitted students.

A number of recent studies focus on quantifying the potential impact of race and testing policies in college admissions. For example, Arcidiacono, Kinsler, and Ransom (2023) estimate that "at Harvard, the admit rates for typical African American applicants are, on average, over four times higher than if they had been treated as white". As another example, Lee et al. (2024) analyze data from a 'highly selective, engineering-focused' college—where the average SAT score of admitted students is around 1500—and show that complying with the *SFFA v. Harvard* policy changes may

[2] https://www.universityofcalifornia.edu/press-room/university-california-board-regents-unanimously-approved-changes-standardized-testing, accessed February 6, 2024.

[3] Using data from the College Board, Smith and Reeves (2020) and Reeves and Halikias (2017) illustrate a 70-100 point difference in average SAT math scores between URM and Caucasian students. We observe similar disparities in GPA and standardized test scores across racial groups over the years. See Figures 16 and 18 in Appendix A.

[4] The report points to evidence that racial disparities in test scores may be driven in part by differences in test preparation and retake rates that are reflective of family income (Goodman, Gurantz, and Smith, 2020; Dixon-Román, Everson, and Mcardle, 2013), and that high school GPA may have stronger relationship with college graduation than ACT scores (Allensworth and Clark, 2020).

[5] As of this writing, these institutions include: Harvard, Brown, Yale, Dartmouth, MIT, Stanford, Georgetown, Purdue, Caltech, and UT–Austin. Other test-mandatory colleges include U. Florida, Georgia Tech, U. Georgia, and Florida State. See https://www.nytimes.com/2024/06/08/us/stanford-standardized-tests.html, accessed September 12, 2024, and https://www.usnews.com/education/best-colleges/the-short-list-college/articles/top-colleges-that-still-require-test-scores, accessed October 31, 2024.

[6] https://president.dartmouth.edu/news/2024/02/reactivating-satact-requirement-dartmouth-undergraduate-admissions, accessed February 5, 2024.

[7] https://mitadmissions.org/blogs/entry/we-are-reinstating-our-sat-act-requirement-for-future-admissions-cycles/, accessed February 5, 2024.

lead to a reduction of over 60% in the share of URM applicants and increase the average SAT score by around 10 points. We provide a more extensive review of the literature in Section 1.2, but these recent studies focus on a small number of such elite institutions, to the best of our knowledge.

For the less selective colleges that collectively enroll the vast majority of college attendees, the datasets used in the latest studies are outdated or otherwise significantly limited. For example, Long (2004) uses responses from the 1994 vintage of the National Educational Longitudinal Study (NELS). Reber, Goodman, and Nagashima (2023), while more recent, relies on a coarse self-report from colleges on the significance of race in their admission decisions (instead of analyzing admissions data). This lack of data availability is perhaps not surprising, given that the U.S. admissions system is decentralized and opaque. Individual colleges typically do not have comprehensive information on their applicants' other applications, admissions, and enrollment decisions. On the other hand, application portals such as the Common App and organizations like the College Board may not have access to individual colleges' admission decisions.

1.1 This Work

In this study, we collect a novel dataset that includes application profiles and admissions outcomes for over 650,000 applicants across all levels of college selectivity. We estimate a logistic mixed-effects model for colleges' admission decisions, and simulate the *potential* impact of race-blind and test-optional admissions policies on the proportion of underrepresented minorities and the average SAT score of admitted students. We find that racial preferences are negligible at less selective colleges, but for the most selective institutions, eliminating considerations of race could reduce URM admission by one-third while increasing the average SAT score by up to 10 points. Moreover, universal adoption of test-optional admission may slightly increase URM representation, but it is unlikely to offset much of the decline due to race-blind admissions for the most selective colleges.

Data We collect and analyze a dataset containing self-reported application and admission information for over 650,000 applicants to over 1,300 colleges from 2012–2022. For each applicant, we observe demographic and academic characteristics: gender, ethnicity, home state, class rank, GPA, and SAT/ACT scores. We also observe the list of colleges they applied to, the corresponding admission outcomes, and the enrollment decision for a subset of the applicants. While the number of covariates we observe is modest, observing multiple admissions outcomes for each applicant allows us to leverage random effects to partially account for applicant-level characteristics which are unobservable to us, such as extracurriculars and recommendation letters. The application portfolios and enrollment decisions also allow us to partially account for spillover effects between colleges, such that we are able to study the composition of the set of *unique* admits and enrollees to each tier of colleges (we discuss this in more detail below).

Model Taking advantage of the breadth of our data, we group colleges into five selectivity tiers to allow the importance of various factors—including demographics and academic credentials—to vary with college selectivity. Our tier assignments are based on the average SAT scores of enrolled students: 1400+, 1300–1399, 1200–1299, 1100–1199, and 400–1099. The most selective tier ("T1") comprises 37 colleges enrolling 3 percent of students, while the least selective ("T5") comprises 901 colleges enrolling 50 percent of students; in between, enrollment shares double in successive tiers. We estimate a logistic mixed-effects model of admission outcomes which depends on applicant and college characteristics, as well as their interactions (to capture e.g., the relative advantage given to in-state applicants at public colleges). We also include college fixed effects, which can be interpreted as admission *cutoffs*, to allow for college-specific differences in selectivity within each tier.

Main Results Our main results are based on the subset of nearly 550,000 applicants in 2012–2020, so as to exclude the influence of near-universal test-optional admission following the COVID-19 pandemic. We pool applications from these admission cycles and estimate one model, allowing only colleges' admission cutoffs to vary by year. To compare the importance of race and other covariates across tiers, we compute tier-specific average marginal effects: specifically, we find that racial preferences are very large at the most selective colleges and diminish in magnitude as selectivity decreases. Moreover, our simulations indicate that if all T1 applicants were treated as Caucasian—what we call "race-blind" preferences—our model implies that the share of African-Americans and Hispanics among unique T1 admits declines by 38% and 27%, respectively.[8] At the same time, the mean SAT score of unique admits—which we take as a proxy of "academic merit"—increases by 10 points in T1 (the standard deviation of the mean enrollee SAT scores of T1 colleges is about 35 points). The estimated impact for Tiers 2 and 3 are smaller, and in Tiers 4 and 5 we do not detect meaningful racial preferences.

We additionally analyze the implications of universal adoption of test-optional policies. We assume that colleges treat applicants who do not report test scores as if they had the mean SAT score (among each college's applicant pool) conditioned on their GPA. Moreover, an applicant reports her score if and only if it exceeds this imputed value.[9] Under status quo racial preferences—where *status quo* refers to our estimated model—we find a small (5-10%) increase in the share of URM admits in Tiers 1–3, and even smaller increases in Tiers 4 and 5. The reason for the increase in selective tiers is that conditioned on GPA, URM applicants have lower SAT scores on average, and therefore benefit more from our score imputation rule. As a consequence of applicants' masking of below-average scores, we observe a small decrease in the mean SAT score of all admitted students in Tiers 1–3. Furthermore, our simulations suggest that the effects of implementing race-blind and test-optional policies are approximately additive when implemented simultaneously. Under such a "race-blind, test-optional" regime, the shares of African-American and Hispanic admits decrease by 34% and 21% in T1, with progressively smaller impact as selectivity decreases. The net effect on academic merit is negligible in T1, while a small decrease is observed in T2 and T3.

One should bear in mind three caveats when interpreting our results. First, our results are affected by selection bias, due to the nature of our observational data set in which applicants self-select to report their information. We make detailed comparisons between our data and more authoritative external data sets in Section 2. Second, due to the limited set of demographic features we observe, our estimates of racial preferences include preferences for unobserved correlates of race, such as family income, neighborhood, and other socioeconomic markers. We discuss the implications in Section 3. Third, we ignore general equilibrium effects that may accompany the policy changes under consideration. In particular, we hold fixed applicants' application sets and assume that colleges maintain the weights on other features in their admissions scoring rules. As such, we do not claim to predict the actual impact of the affirmative action ban. To better understand the overall allocative effect on the market, we take the additional step of accounting for the enrollment decisions of admits (also assumed to be policy-invariant). This allows us to present our estimates in terms of the change in tier-level enrollment, which we find to be very similar to that for unique admits. We further discuss these points in Section 4.3.

The remainder of the paper is organized as follows. Section 1.2 presents an overview of related literature. Section 2 discusses the data, and Section 3 details the model and caveats. The main

[8]We choose to focus our discussion of URMs on African-Americans and Hispanics, as Native Americans represent a small ($< 1\%$) share of applicants in our data.

[9]The language that many colleges use to describe their test-optional policies suggests that they intentionally do not account for *adverse selection* in reporting. See Section 4.2.2 for an explanation that if this adverse selection is accounted for, all applicants will choose to report their test scores in equilibrium.

empirical results are presented and discussed in Section 4. Further details and figures are provided in the appendices.

1.2 Related Work

Group Preferences (Elite Colleges) There is an extensive literature on group preferences in college admission, including affirmative action for underrepresented racial minorities. A number of studies use detailed internal data from admissions departments, and are consequently able to examine only a small number of (typically elite) colleges. Most relevant to this work are Arcidiacono, Kinsler, and Ransom (2023) and Lee et al. (2024), and our results for the most selective colleges are largely aligned with their findings. The former focuses on Harvard and UNC admissions, using detailed data from court documents in the Students for Fair Admissions cases. With access to a rich set of applicant covariates, the authors estimate that removing racial preferences—and appropriately adjusting for capacity constraints—would lead to a 72% (51%) decrease in the African-American (Hispanic) share of admits at Harvard; an 88% (60%) decrease among out-of-state UNC admits; and a 36% (18%) decrease among in-state UNC admits (Table 9 in their paper). Lee et al. (2024) uses internal data from one elite college and simulates the impact of a race-blind policy on the priority order of applicants during the application review process, when a test-optional policy is in place. The authors estimate that going race-blind decreases the share of URM applicants ranked in the top decile by 62%, while increasing the average SAT score by no more than 10 points.

Also focusing on highly selective colleges, Espenshade, Chung, and Walling (2004) and Espenshade and Chung (2005) use admissions data in the 1980s and 1990s, finding significant preferences for racial minorities, athletes, and legacies. Hurwitz (2011) estimates the extent of legacy preferences at 30 elite colleges, leveraging multiple applications to control for unobserved applicant qualities. Using the same data as Arcidiacono, Kinsler, and Ransom (2023), Arcidiacono, Kinsler, and Ransom (2022a) presents evidence that Harvard's "personal rating" may be a vehicle for discrimination against Asian American applicants, and Arcidiacono, Kinsler, and Ransom (2022b) finds that Harvard's preferences for ALDCs (athletes, legacies, dean's interest list, and children of faculty and staff) significantly tilts the racial distribution of admitted students toward whites. Using application lists and enrollment proxies from the Common App, Grossman et al. (2024) examines disparities in the *enrollment rates* (instead of admission outcomes) of Asian and white applicants to eleven "Ivy-Plus" colleges. They find that (i) white applicants enroll at higher rates than similarly qualified Asian applicants; (ii) geography and legacy status explain much of the disparity, as Asians are less likely to be legacies and more likely to reside in states with low admit rates; and (iii) these disparities are greater for South Asian than East Asian applicants.[10] Chetty, Deming, and Friedman (2023) merge tax records with admissions data from (an unspecified subset of) 33 highly selective colleges, which the authors use to quantify disparities in application, admission, and enrollment rates by family income, after controlling for SAT/ACT scores.

Group Preferences (All Colleges) The latest existing studies that estimate preferences over the *full spectrum of college selectivity* use data from surveys of colleges or applicants. Reber, Goodman, and Nagashima (2023) use colleges' self-reported admissions considerations, based on responses to the Common Data Set (an annual survey of colleges), and find that less selective colleges generally do not take applicants' race into consideration. Arcidiacono (2005) analyzes data from the National Longitudinal Study of the Class of 1972 (NLS72), a representative survey of

[10] We find evidence that South Asian applicants to top-tier colleges may face an admissions penalty relative to East Asian applicants (see Figure 3). At the same time, South Asian applicants admitted to at least one top-tier college may have a lower propensity to enroll in a top tier college than East Asian admits (see Appendix D.1).

high school seniors. The author finds that equalizing admission rules for blacks and whites reduces the enrollment share of black males at top colleges by over 40%, but does not meaningfully change overall college enrollment rates or average future earnings. Kane (1998) and Long (2004) use survey data collected in 1982 and 1994, respectively. Both find that racial preferences are large only at highly selective colleges ("the most academically selective fifth of all four-year institutions, where SAT scores averaged 1100 or more" in the case of Kane, 1998). Our paper suggests that things have not changed much over the last 30-40 years, prior to the SFFA decision. However, our larger sample size allows us to make sharper inferences; for example, Long (2004) does not observe differences in preferences for Blacks and Hispanics, or for Whites and Asians, but we do.

Race-Neutral Alternatives to Affirmative Action A number of the aforementioned studies also evaluate the effectiveness (in terms of recovering the level of racial diversity) of "race-neutral alternatives" to race-based affirmative action. Kane (1998) argues that replacing race-based with income-based preferences cannot recover a comparable diversity level, as URMs represent only a small fraction of high-achieving, low-income students ("roughly one out of six" in the author's 1992 data). Levine and Reber (2023) examines the possibility of replacing racial preferences with preferences for low-income and first-generation students at the 153 most-selective colleges representing 15% of total enrollment. They find that maintaining current diversity levels by giving preferences to Pell grant recipients would require an $8 billion per year (80%) increase in aggregate financial aid expenditure. Long (2004) estimates the joint impact of the elimination of racial preferences and universal adoption of top percent policies—which grant automatic admission to students in the top x% of their high school classes—on racial diversity, finding that "x% policies are unable to replace traditional affirmative action", as most minority applicants who qualify under these policies would have been accepted even without them.[11] Bleemer (2023), using a difference-in-differences design, finds that "top 4%" (later "top 9%") and holistic review policies lifted the share of URM and low-income enrollees in the University of California system by no more than about 4 percent. The author also finds that affirmative action had lifted URM enrollment by 20 percent—and by 60 percent at the more-selective campuses—relative to the period after the ban. We contribute to this body of work by demonstrating that, in a race-blind world, universal test-optional admission also has limited efficacy in this regard.

Responses to Affirmative Action Bans Several reports and studies examine the applicant-side response to affirmative action bans. Kim et al. (2024), using comprehensive Common App data, does not find a meaningful change in application behavior or in the racial composition of the applicant pool in the 2023-24 cycle, relative to pre-existing trends (prior to *SFFA v. Harvard*). This is consistent with earlier findings from Card and Krueger (2005) and Antonovics and Backes (2013), which examine the effects of the California affirmative action ban (Prop 209); using SAT score reports as a proxy for application intent, neither finds evidence for a material "chilling effect" on competitive URM applications.[12] Together, these empirical findings lend support to the simplicity

[11] Top percent policies are notably implemented in California, Texas, and Florida, where they are arguably most effective due to high racial segregation across high schools as of 1996. For a universal policy, the author's simulations indicate diminishing returns to relaxing the threshold: relative to the drop after banning racial preferences, a "top 10%" policy yields an 18% rebound in the overall URM share (38% at the most selective decile of colleges); the effect is "generally much smaller" under a "top 4%" policy, while a "top 20%" policy yields a 19% rebound.

[12] Antonovics and Backes (2013) finds that "the relative decline in URM score-sending rates... was small and concentrated at Berkeley and UCLA among URMs who experienced the largest relative drop" in predicted admission probability. The authors conclude that, overall, there is "stability of URM application behavior in the face of substantial declines in their admission rates."

of our approach, in particular our assumption of no application response to an affirmative action ban. There is ample evidence, however, that colleges may change their recruitment strategies or admissions criteria. For example, Chan and Eyster (2003) and Antonovics and Backes (2014) document that some UC schools de-emphasized SAT scores (among other changes) after California's Prop 209; Long and Tienda (2008) find similar changes in Texas following the *Hopwood v. Texas* ruling (in which a ban on affirmative action was effected by court order). We focus our main analysis on the direct effect absent endogenous policy responses, but discuss the possibility of a policy response in Section 4.3.

Informativeness of Test Scores As colleges now contemplate whether to reinstate standardized tests, several recent studies aim to understand the informativeness of GPA and test scores as signals for success in college. These studies have found that test scores are better predictors of first-year college GPA than high-school GPA (Friedman, Sacerdote, and Tine, 2024), and aid in the identification of high-potential applicants from nontraditional backgrounds (Sacerdote, Staiger, and Tine, 2025). This is especially true considering widespread grade inflation which is more prevalent among well-resourced high schools, a phenomenon documented in Hurwitz and Lee (2018). In our data we observe, from 2012-2022, a slight decrease in the average SAT score and a substantial increase in the average GPA (Figure 16), adding to the descriptive evidence for the overall trend.

Test-Optional Admission Several recent works study the impact of test-optional admission on the representation of target groups on campus, from both theoretical and empirical perspectives. In the model of Dessein, Frankel, and Kartik (2023), colleges may voluntarily deprive themselves of test information in order to achieve a desired tradeoff between admission outcomes and "disagreement costs" with society when preferences over whom to admit differ. Garg, Li, and Monachou (2023) posits a theoretical model featuring (i) differential informativeness of both test scores and non-test score features across groups, and (ii) access barriers to standardized tests, and identifies conditions under which eliminating testing requirements can improve both diversity and merit.

There are existing studies which empirically analyze the impact of testing policies among a subset of colleges. Our test-optional results are consistent with those of Kelly (2022), which uses data from one moderately selective college and finds that switching from test-optional to test-mandatory does not have an appreciable impact on URM shares and average test scores.[13] The college has an existing test-optional policy, under which some students sent scores but elected to not have them considered. As a result, the author only needs to assume that, under a test-mandatory regime, non-reporters would be treated like reporters were under the test-optional regime. By contrast, Bennett (2022), using a matched difference-in-differences design, finds that early adopters of test-optional policies (primarily liberal arts colleges) modestly increased the enrollment of Pell Grant recipients, URMs, and women relative to otherwise comparable late adopters.

Lastly, our stylized assumptions on applicants' score-reporting behavior under test-optional policies are informed by McManus, Howell, and Hurwitz (2023), which uses data from 50 colleges and the College Board and finds that applicants are strategic when deciding whether to disclose or withhold their test scores (specifically, they condition their choice on their other characteristics).

[13]The author considers various assumptions which support partial non-disclosure of test scores: (i) "cursedness" in the sense of Eyster and Rabin (2005), in which applicants (or colleges) do not account for the dependence of strategies on information; (ii) differential weights placed on the non-SAT components of the score W_{ij} (see Section 3); and (iii) colleges' desire to improve their rankings, which depends on the average SAT scores of enrollees who reported them.

Methodology Many studies adopt an empirical approach based on "policy simulations", and do not account for potential general equilibrium effects such as application response and changes to other aspects of colleges' admissions policies. For example, starting from a probit regression estimating the probability of admission, Long (2004) simulates the elimination of racial preferences by setting the coefficients corresponding to race equal to zero. Chetty, Deming, and Friedman (2023) similarly eliminate preferences given to Ivy-Plus applicants from high-income families in the form of legacy advantage, athletic recruitment, and stronger non-academic evaluations. We take a similar approach in this work.

Other studies postulate a structural model, which allows for an equilibrium approach and counterfactual analysis. Arcidiacono (2005), mentioned earlier, takes a partial equilibrium approach to policy simulation, allowing for changes in application portfolios (in response to changes to admissions or financial aid rules) but ignoring, e.g., colleges' capacity constraints. Fu (2014) builds a full equilibrium model using the 1997 National Longitudinal Survey of Youth (NLSY97), allowing both sides to respond in their counterfactual simulations. Borghesan (2023), adapting the framework of Fu (2014), is perhaps the most similar to our paper in aim. Using the Educational Longitudinal Study of 2002 (ELS 2002), the author investigates the differential impact (across college selectivity tiers) of standardized testing requirements, while either allowing or banning race-based affirmative action. Under the author's model, URM enrollment is largely unresponsive to testing policy, regardless of whether affirmative action is permitted, which aligns with our findings. On the other hand, the model estimates a large (19%) decline in the overall URM share among enrollees at any college under an affirmative action ban, which differs from our result that the decline is concentrated among the most selective colleges. Borghesan (2023) explicitly models an applicant's decision of whether to take the SAT and apply to college—dependent on factors such as testing access and application costs—as well as the effect of admissions policy on incentives for human capital accumulation, in the form of high school study effort. Notwithstanding the richness of the model, there are limitations such as the assumption that all colleges in the same tier share the same admission cutoff. We further compare Borghesan (2023)'s approach and substantive findings to ours in Appendix D.5.

2 Data

We obtain a large college application data set ("the data") from an online service which mediates the exchange of academic credentials such as transcripts and diplomas between students, high schools, and colleges. In addition, this service provides college applicants ("users") an optional tool that estimates an admission probability for each college in their portfolios, based on the users' self-reported demographic information (ethnicity, gender, state of residence, year entering college) and academic performance (GPA, class rank, standardized test scores). We observe this information on each user's public profile page, which may also include the user's self-reported college list (or "portfolio") and admission outcomes. See Figure 1 for an example applicant profile.

Academics Users may report their unweighted—and, optionally, weighted—high school GPA, along with class rank and test scores for the SAT and/or ACT. In this study, we use the unweighted GPA only, which does not adjust for course difficulty. Class rank is reportable as a categorical variable, with levels shown in Table 8. The SAT and ACT are the standardized tests which are most commonly accepted or required by colleges, with the ACT being more popular among students in the Midwest. We use concordance tables by The College Board to convert ACT scores into an SAT-equivalent score, which we henceforth refer to as "SAT". (See Section 2.2 for more details

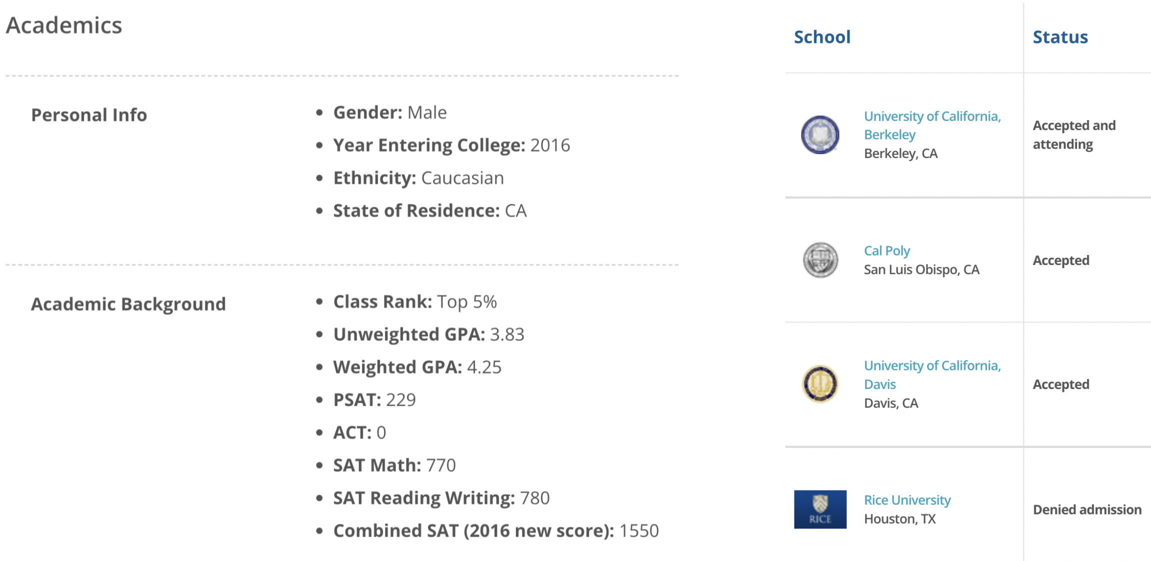

Figure 1: Applicant profile example, personal and academic information (left) and the list of colleges with the corresponding admission outcomes (right; truncated).

regarding data preparation.) Note that we do not observe whether a user ultimately disclosed her SAT score to colleges in her application portfolio, if some colleges were test-optional.

College List A user may indicate interest in a college by adding it to her list, in which case it is given the default status “Applying”. She may update the status later to indicate that she has applied to that college (“Applied”), and to report the admissions outcome and enrollment decision: Denied admission, Wait-listed, Accepted, or Accepted and attending.

Our Sample We consider users who reported entering college between 2012 and 2020, and on Title IV participating four-year colleges in the *Integrated Postsecondary Education Data System* (IPEDS).[14] Table 1 reports the total number of applications, applicants, and colleges in different subsets of the data. The first column includes applicants who reported (a) all available demographic information, i.e., gender, year entering college, ethnicity, and state of residence; (b) all academic credentials used in our analysis, i.e., class rank, GPA, and SAT scores; and (c) a list of colleges, with or without admission outcomes reported.[15] The yearly count (also including 2021–22) is shown in the left panel of Figure 2.

Our main analysis focuses on applicants who reported at least one admission outcome to a 4-year college represented in IPEDS, corresponding to the middle column of Table 1. This sample retains 535,099 of the original 1.15 million applicants, with the yearly count shown in the right

[14]IPEDS is a U.S. Department of Education data repository which contains detailed information about educational institutions that participate in federal student aid programs under Title IV (virtually all reputable colleges and universities). There are 1,992 colleges satisfying the “Title IV participating”, “U.S. only”, and “Institutional category: Degree-granting, primarily baccalaureate or above” filters in the 2019 data collection year, 1,470 of which report quartiles of enrollee SAT scores to IPEDS, and 1,322 of which are also represented in our data.

[15]Some high schools do not rank their students, and users may select “Not Ranked” accordingly; in our data, 47% of applicants indicate this, and we include them in our analysis. Notably, these no-rank policies are not adopted randomly by high schools: Hurwitz and Lee (2018) document trends in rank suppression which—along with GPA inflation—are concentrated at high-socioeconomic-status high schools.

Number of:	All applications	Applications with outcome	Applicants with all outcomes
Applications	7,171,185	1,786,596	733,713
Applications (to 4-year colleges)	6,519,152	1,631,844	683,785
Applicants (to 4-year colleges)	1,151,966	535,099	192,425
4-year colleges	1,322	1,318	1,308

Table 1: Total number of applications, applicants, and colleges over all years 2012–2020, for different subsets of the data. Four-year colleges include only those which reported SAT information for enrollees in IPEDS, and thus can be assigned to a tier according to our classification scheme. Additional filters described in Section 2.2 further reduce the number of colleges in the data; see Table 4 for an accounting of all filtering steps.

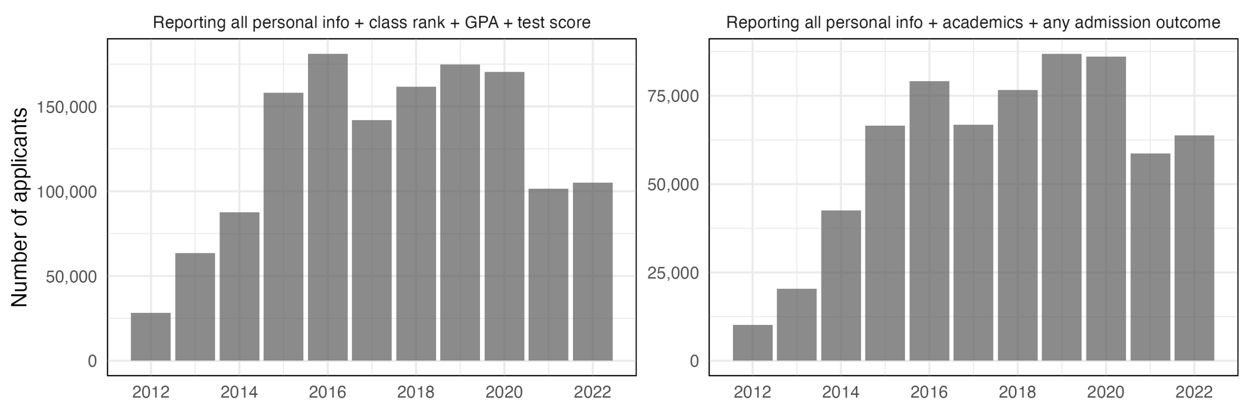

Figure 2: Number of applicants in the data by year entering college. Our analysis focuses on the subset of applicants who reported all demographic information, academic credentials (class rank, GPA, standard test scores), and at least one admission outcome among 4-year colleges in their portfolio.

panel of Figure 2.[16] We henceforth term this sample "The Data". The corresponding applications represent 1,318 colleges; the average number of applications per applicant is 5.65, while the average number of reported outcomes is 3.06. Fifty percent of the applicants in the data reported the college at which they ultimately enrolled, and 46% reported enrolling at a 4-year college in IPEDS.

2.1 Summary Statistics

We group colleges into five tiers based on their selectivity, measured by the average SAT score of enrollees reported to IPEDS in 2015, as shown in Table 2 (see Appendix D for how composite SAT averages are estimated from lower and upper quartiles of math and verbal component scores). The SAT ranges are 1400+, 1300–1399, 1200–1299, 1100–1199, and 400–1099 (see Table 9 for a selection of colleges belonging to each tier). The columns in the middle of the table compare the shares of applications, admissions, and enrollment among these tiers. We find a reasonable degree of alignment between our data and IPEDS in terms of these shares, and note that (i) the IPEDS enrollment share roughly doubles in successive tiers, and (ii) the most selective tiers receive a disproportionate number of applications, relative to enrollment.

The tier-level admit rates, shown in the rightmost pair of columns, are considerably higher in the data than in IPEDS. We flag this discrepancy as possibly the most significant selection bias

[16] We repeat our analysis for the set of roughly 200,000 applicants who had reported all admissions outcomes (last column in Table 1), and do not find a meaningful qualitative difference in our results. See Appendix B.5.

			Applications Share		Admissions Share		Enrollment Share		Admit Rate	
Tier	SAT range	Colleges	Data	IPEDS	Data	IPEDS	Data	IPEDS	Data	IPEDS
T1	1400+	37	7.8%	8.0%	2.2%	1.8%	3.0%	3.1%	22.8%	12.8%
T2	1300–1399	55	10.6%	11.1%	7.0%	6.0%	10.2%	6.7%	52.4%	30.3%
T3	1200–1299	102	12.4%	13.7%	11.6%	12.2%	14.0%	12.0%	75.0%	50.3%
T4	1100–1199	223	26.2%	23.2%	29.0%	26.8%	27.2%	25.7%	88.5%	65.1%
T5	400–1099	901	43.1%	44.0%	50.1%	53.2%	45.6%	52.5%	92.6%	68.0%

Table 2: The share of applications, admissions and enrollment, and admit rate, of different tiers of colleges across 2012–2020, grouped according to average SAT score of enrollees in 2015, in IPEDS and the data.

in our data.[17] A possible explanation for this discrepancy is a reporting bias among users, with a higher propensity to report applications and outcomes which were successful. We nevertheless find in Appendix C that such reporting bias is not significantly associated with race: for applications with reported outcomes, the average of the *overall* admit rates of colleges applied to is higher than that of applications with unreported outcomes, but the magnitude of this gap does not vary significantly by race. We also note that applicants represented in the data are virtually all domestic U.S. students, while the data in IPEDS includes international students. The latter are separately identified in enrollment numbers, but not in application or admission statistics.

In Appendix A, we make a number of distributional comparisons with known population-level data obtained from IPEDS and other definitive sources such as The College Board and The Common App.[18] These comparisons indicate that our data is broadly representative of the population with respect to observable metrics, including the distribution of the number of applications per applicant (Figure 20) and the college-level SAT score distributions of enrollees (Figure 21). Figure 17 shows that the racial composition of enrollees is comparable to that in IPEDS (Asians are slightly overrepresented in the data in tiers 1, 2, and 4, and Caucasians in tiers 3 and 5, while IPEDS has a larger fraction that are not categorized as Asian, Caucasian, African-American, or Hispanic).[19] An exception to the general level of agreement is the overrepresentation of residents of states in the Midwest such as Michigan, Kentucky, and Indiana, as seen in Table 6. This may be a result of a partnership between the platform and the Midwestern Higher Education Compact (MHEC).

Table 3 shows various summary statistics for each race and tier, averaged over all years in the sample, for applications and acceptances (i.e., the averages are weighted by each applicant's number of applications or acceptances in the tier). For all races, there is a pattern of decreasing acceptance rates and increasing GPA and SAT in tier selectivity (i.e., going from T5 to T1). In T1, African-Americans and Hispanics have the highest unconditional acceptance rates (32.8% and 28.0%), while South Asians have the lowest (19.0%). At the same time, African-Americans have the lowest average GPA and SAT, followed by Hispanics—a pattern which holds in all tiers, for both applications and acceptances (see Figure 16 for average GPA and SAT by year.) In Tiers

[17] We note that representative survey data exhibit a comparable bias: for example, Borghesan (2023), using data from the Education Longitudinal Study of 2002, finds a difference in admit rates of "eleven to thirteen percentage points (pp) for private colleges, and between six and eight pp for public colleges", compared to IPEDS. Nevertheless, our results are broadly aligned with those from similar studies, for example Arcidiacono, Kinsler, and Ransom (2023) which examines the impact of race-blind admissions at Harvard and the University of North Carolina.

[18] The College Board is the organization which, *inter alia*, designs and administers the SAT. The Common App is a popular application platform for students and colleges.

[19] Note that these overall averages conceal significant within-tier heterogeneity, especially in less selective tiers: In the IPEDS data, the enrollment-weighted standard deviation of college-level URM shares is 3.5 percentage points (pp) in T1; 4.4pp in T2; 6.9pp in T3; 10.3pp in T4; and 22.5pp in T5.

2–5, Caucasians have the highest unconditional acceptance rates. Moreover, for both applications and acceptances, the fraction that are in-state is generally decreasing in tier selectivity, while the fraction that are male is increasing. Finally, we note that the average number of applications per applicant is higher among enrollees in more selective tiers: those who reported enrolling in a T1 college have 8.6 applications (5.2 outcomes) on average; in a T2 college, 7 applications (4.6 outcomes); and in a T5 college, 4.9 applications (2.5 outcomes).

2.2 Data Preparation

We take several data processing steps before estimating our model. As mentioned above, we convert ACT scores into SAT-equivalent scores using a conversion table from The College Board. If a user reports test scores for both the SAT and the ACT, we take the maximum of the former and the SAT-equivalent score of the latter.[20] We standardize GPA and SAT, and recode `class_rank` and `ethnicity` by folding together categories with few observations (see Table 8). The retained `class_rank` categories are Top 1%, Top 5%, Top 10%, Top 25%, Top 50%, Bottom 50%, and Unranked (NR), and the retained `ethnicity` categories are African-American, Hispanic, Caucasian, East Asian, South Asian, and Other.

We exclude 511 colleges whose median enrollee SAT score in the data falls outside the interquartile range of scores reported in IPEDS, or vice versa (including 145 for which there are no self-reported enrollees in the data). Lastly, we drop 15 colleges for which we observe only acceptances or only rejections, as the corresponding fixed effect is not identified. These last two steps together remove 9% of applications (and 5% of applicants) from the data. The entire sequence of data filtering steps is shown in Table 4. The final numbers of colleges represented in Tiers 1–5 are 34, 44, 77, 164, 474. The number of application outcomes observed is 1,483,297 (1,870 per college on average, and 208 per college-year).

3 Model

In this section, we propose a logistic mixed-effects model of admission outcomes. In our model, colleges assign a scalar-valued *score* to each applicant and admit those whose score exceeds a cutoff. Consider an applicant i who has applied to a college j. We assume that i's overall score consists of a systematic, tier-specific component (which we label W_{ij}) and an idiosyncratic, college-specific component (which we label ξ_{ij}). Applicant i receives an offer of admission from college j, coded as $Y_{ij} = 1$, if her overall score exceeds the cutoff c_j, that is,

$$Y_{ij} = \mathbb{1}\{W_{ij} + \xi_{ij} \geq c_j\}. \tag{1}$$

We model the systematic part W_{ij} as a linear function of observed applicant-college characteristics X_{ij} and an unobserved factor η_i . The observed X_{ij} include factors such as SAT score, GPA, class rank, and whether the applicant is in-state for the college (more details below). The unobserved factor η_i captures common essays, recommendation letters, and extracurriculars—information accessible to the colleges but not available in our data. Recognizing that colleges of varying selectivity may evaluate these factors differently, we specify W_{ij} using tier-specific coefficients:

$$W_{ij} = \beta_{T(j)}^{\top} X_{ij} + \delta_{T(j)} \eta_i, \tag{2}$$

[20] Students who take these exams more than once are often permitted on applications to report a "superscore" which combines the best within-section scores across exam attempts. We do not observe whether the score reported on an applicant's profile is a superscore or a single-exam score, but assume it is the score reported to colleges.

	African-American		Hispanic		Caucasian		East Asian		South Asian	
	App.	Acc.	App.	Acc.	App.	Acc.	App.	Acc.	App.	Acc.
Pr(Accepted)										
T1	0.328	1	0.280	1	0.216	1	0.212	1	0.190	1
T2	0.488	1	0.444	1	0.561	1	0.506	1	0.509	1
T3	0.654	1	0.640	1	0.789	1	0.709	1	0.722	1
T4	0.801	1	0.789	1	0.917	1	0.844	1	0.873	1
T5	0.879	1	0.860	1	0.952	1	0.906	1	0.911	1
GPA										
T1	0.453	0.596	0.661	0.784	0.795	0.808	0.828	0.885	0.777	0.866
T2	0.268	0.486	0.463	0.673	0.642	0.757	0.643	0.785	0.591	0.751
T3	0.092	0.292	0.328	0.511	0.520	0.629	0.487	0.617	0.427	0.567
T4	-0.199	-0.053	0.093	0.206	0.299	0.362	0.350	0.426	0.266	0.332
T5	-0.500	-0.389	-0.148	-0.051	0.077	0.126	0.159	0.220	0.097	0.159
SAT										
T1	0.666	0.999	0.946	1.256	1.397	1.435	1.683	1.788	1.615	1.744
T2	0.288	0.575	0.458	0.783	1.056	1.207	1.324	1.497	1.259	1.454
T3	0.038	0.246	0.289	0.505	0.780	0.882	1.049	1.164	1.056	1.181
T4	-0.324	-0.192	-0.020	0.112	0.429	0.485	0.707	0.771	0.672	0.744
T5	-0.711	-0.643	-0.458	-0.387	-0.018	0.014	0.290	0.338	0.186	0.234
Male										
T1	0.404	0.422	0.499	0.533	0.510	0.523	0.518	0.479	0.572	0.581
T2	0.360	0.357	0.424	0.432	0.453	0.456	0.481	0.472	0.499	0.502
T3	0.384	0.381	0.431	0.426	0.451	0.443	0.476	0.457	0.494	0.477
T4	0.342	0.338	0.386	0.387	0.398	0.397	0.437	0.429	0.437	0.437
T5	0.311	0.306	0.339	0.332	0.352	0.350	0.411	0.401	0.394	0.389
In State										
T1	0.122	0.128	0.166	0.143	0.109	0.133	0.142	0.135	0.148	0.171
T2	0.495	0.507	0.622	0.562	0.445	0.469	0.562	0.552	0.539	0.542
T3	0.519	0.521	0.665	0.630	0.524	0.526	0.600	0.575	0.543	0.521
T4	0.590	0.590	0.657	0.616	0.561	0.553	0.690	0.665	0.696	0.677
T5	0.677	0.675	0.847	0.836	0.782	0.781	0.858	0.854	0.852	0.847
Count										
T1	8,730	2,861	12,114	3,389	57,043	12,335	19,350	4,108	12,376	2,348
T2	11,405	5,570	19,818	8,793	84,635	47,494	25,803	13,048	16,635	8,475
T3	13,048	8,539	21,133	13,518	117,215	92,483	17,653	12,511	14,203	10,257
T4	43,190	34,585	43,599	34,416	253,993	232,956	21,866	18,459	20,273	17,690
T5	108,820	95,693	81,662	70,258	346,376	329,666	20,034	18,146	16,657	15,167

Table 3: Summary statistics by race and tier, for applications with reported outcomes (App.) and acceptances (Acc.), pooled across all years 2012–2020. SAT and GPA reported as z-scores (the number of standard deviations above the mean). Race category "Other" not shown.

where $T(j)$ denotes the tier to which college j belongs. We assume that the unobserved factor η_i is a *random effect* drawn from a normal distribution with zero mean and unit variance. We make the natural identifying assumption that η_i is uncorrelated with $\beta_{T(j)}^\top X_{ij}$, and remark that the $\beta_{T(j)}^\top X_{ij}$ may be interpreted as the part of W_{ij} which is *predictable* from observed covariates, and $\delta_{T(j)}\eta_i$ the part which is not predictable.

The idiosyncratic component ξ_{ij} can be interpreted as the heterogeneous preferences of different colleges, as well as the noise of the admissions processes. We assume that ξ_{ij} is logistically distributed with zero mean and unit scale parameter, and is uncorrelated across i's multiple applications to different colleges (this effectively ignores endogeneity in ξ_{ij} which may arise because applicant i may be more likely to apply to college j where they have a high ξ_{ij}). Together, our assumptions yield the following *generalized linear mixed model* (GLMM):

$$p_{ij} = \sigma\left(-c_j + \beta_{T(j)}^\top X_{ij} + \delta_{T(j)}\eta_i\right) \tag{3a}$$

$$Y_{ij} \mid p_{ij} \sim \text{Bernoulli}(p_{ij}) \tag{3b}$$

$$\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \tag{3c}$$

where $\sigma(\cdot) := \exp(\cdot)/[1+\exp(\cdot)]$ is the logistic (sigmoid) function.[21] In the statistics and applied sciences literatures, such a model is typically referred to as a *hierarchical* or *random effect* model, as each η_i may be viewed as a clustered error at the applicant level, or as a random coefficient which multiplies a dummy variable representing applicant i. In the discrete choice literature, these random coefficients also manifest in the mixed-logit family of models as a way of representing unobserved choice heterogeneity among individual consumers (see Train (2009)). From a Bayesian perspective, our normality assumption represents a prior distribution on these random effects.

In Appendix D.6, we provide details regarding our maximum likelihood estimation approach, which requires integration over the distribution of η_i. Given the scale of the likelihood computation, we use an adaptive gradient descent method while taking advantage of both sparsity (applicants apply to few colleges) and parallelizability (admissions are independent across applicants). Finally, we derive the posterior distribution for an applicant's η_i, which we use to calculate model predictions in subsequent sections.

Model Specification As discussed in Section 2.1, we group colleges into $K = 5$ tiers based on the average SAT scores of enrollees. The covariates included in the observable characteristics X_{ij} are class rank, GPA, SAT, the GPA × SAT interaction, race, an indicator of whether the applicant is male, and indicators `IN_STATE` and `IN_STATE_PUBLIC`. `IN_STATE` takes value 1 if the application is to a college in the applicant's state of residence, and `IN_STATE_PUBLIC` further requires that the college is public. Also note that male/female are the only selectable gender options on the platform.

We present model fit statistics for our main specification in Section B.3. Our main model achieves an in-sample pseudo-R^2 of 0.479, compared to 0.441 for a model which does not include the applicant random effects $\boldsymbol{\eta}$.[22] Figure 6, Figure 24, and Figure 25 suggest that the model is well calibrated across race, tier, and relevant ranges of SAT and GPA. To allow more flexibility and non-linearity in the admission scoring rules, we consider in Section B.3 a number of alternative specifications that include higher order terms in SAT and GPA and their interactions with race.

[21] With K tiers, an alternative representation of (3) specifies a random effect vector $\eta_i = (\eta_{i1}, \ldots, \eta_{iK})$ for applicant i, with unknown covariance parameters $\delta = (\delta_1, \ldots, \delta_K)$. Under this standard form, our random effects assumption corresponds to $\eta_i \sim \mathcal{N}(0, \delta\delta^\top)$, which implies a common $\bar{\eta}_i = \eta_{ik}/\delta_k$ for all $k = 1, \ldots, K$.

[22] As a point of comparison, Arcidiacono, Kinsler, and Ransom (2023) report pseudo-R^2 values of 0.260, 0.420, and 0.588 for Harvard, UNC Out-of-State, and UNC In-State applicants using a "demographics + academic" specification roughly comparable to ours, and higher values of 0.556, 0.588, and 0.727 using their "preferred" specification with many more controls. See their Table 7.

This does not yield a material improvement in model fit, in terms of either pseudo-R^2 or in-sample AUROC. Moreover, the aggregate estimates from our policy simulations—which we describe in Section 4.2—are largely unaffected under these alternative models, including the model which excludes the random effects $\boldsymbol{\eta}$ (see Appendix B.3.3). Thus, we focus our subsequent discussion on the simple model described above.

Unobserved Correlates of Race Colleges often have preferences for characteristics other than race that are nonetheless correlated with race. For example, *need-aware* colleges "will examine a student's financial need at the time of admission".[23] Colleges may also have diversity goals beyond race, as noted by MIT admissions dean Stu Schmill: "Given the clear educational benefits, we still [after *SFFA*] consider many kinds of diversity: prospective fields of study and areas of research, extracurricular activities and accomplishments, as well as economic, geographic, and educational background—just not race."[24] Our model coefficients on race covariates capture not only racial preferences but also preferences for such correlates of race that are unobserved in our data. These factors can impact our estimated importance of race in both directions. For example, when preferences are given to Pell Grant-eligible and first-generation college students, our model is likely to overestimate the importance of race, as URM applicants are over-represented in these groups.[25] On the other hand, when colleges prioritize applicants with lesser financial need, this may lead to an underestimation of the importance of race. We further discuss the implications for our policy simulations in Section 4.2.

4 Empirical Results

In this section, we present the estimated model coefficients; the average marginal effects of GPA, SAT, and race; and the race-blind and test-optional policy simulations. For the main results presented below, we pool admission cycles and estimate one model using applications in years 2012–2020, allowing the college cutoffs to vary by year.[26] Figure 3 displays the estimated coefficients $(\beta_k, \delta_k)_{k=1}^K$ (see Appendix B.1 for a tabular format). The signs of the coefficients reveal which factors contribute positively or negatively to the admission probability; a more informative quantification of their contribution is given by average marginal effects, which we discuss in Section 4.1.

We first observe that selective tiers generally have stronger preferences for applicants with higher class rank. Tiers 1 and 2 exhibit preferences for "Top 1%" applicants compared to "Top 5%" applicants, while Tier 3 appears to not make a distinction between these categories, but prefers both to applicants who rank "Top 10%" or lower. Interestingly, we also observe that Tiers 4 and 5 show preferences for lower-ranked applicants compared to those in "Top 1%." This is consistent with the practice of "yield protection", in which colleges reject high-ranked applicants who are unlikely to accept their offers.[27] Note that some high schools deliberately do not rank their

[23] `https://finaid.brown.edu/basics/financial-need-eligibility/need-blind`, accessed February 13, 2025.

[24] `https://news.mit.edu/2024/qa-undergraduate-admissions-in-wake-of-supreme-court-ruling-0821`, accessed September 12, 2024.

[25] `https://www.brookings.edu/articles/can-colleges-afford-class-based-affirmative-action/`, accessed October 4, 2024.

[26] During these admission cycles, most colleges represented in our data required the submission of test scores and were allowed to take race into consideration. In our data, 15% of applications (with reported outcomes) are to colleges which did not require test score submission (6% in T1; 9% in T2; 12% in T3; 18% in T4; and 16% in T5). And 28% of applications are to colleges which were affected by race-based affirmative action bans in twelve states including Arizona, California, Florida, and Texas (0% in T1; 41% in T2; 24% in T3; 29% in T4; and 31% in T5). In Appendix B.4 we repeat our analysis excluding the affected colleges and find qualitatively similar results.

[27] See, for example, `https://www.wsj.com/articles/SB991083160294634500`, accessed June 21, 2024.

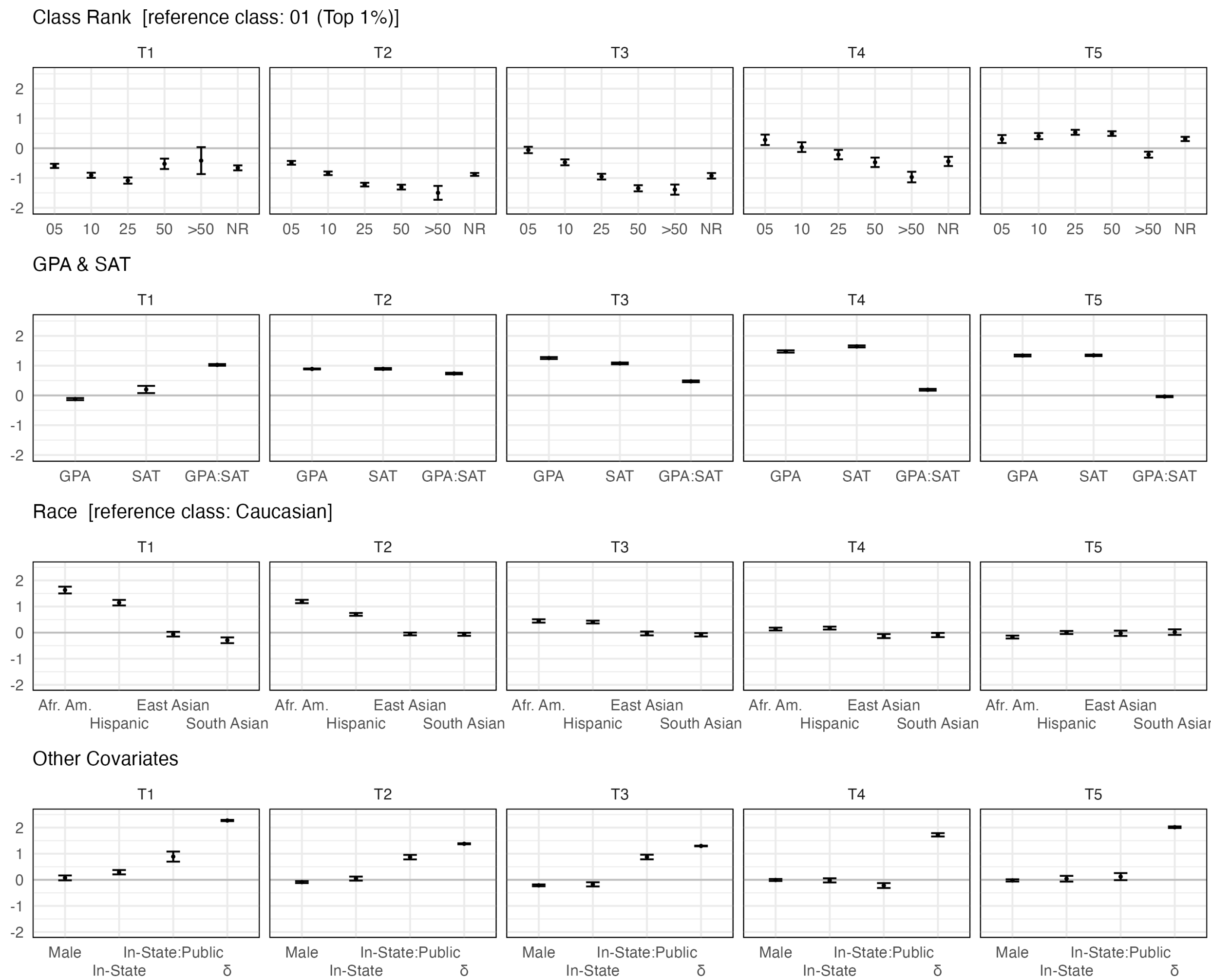

Figure 3: Estimated model coefficients (with 95% confidence intervals). Interactions are represented with a colon (e.g., GPA:SAT). See Table 10 in Appendix B.1 for a tabular format.

students (Buckley, Letukas, and Wildavsky, 2018, p. 83); in our data, 47% of applicants affirm that they are "Not Ranked" (NR). This is different from the scenario where an applicant did not submit any class rank information to the platform, in which case the applicant is excluded from our analysis. We find that such applicants are accorded preferences slightly lower than those for the average applicant in each tier.

Second, the more selective the tier, the larger the GPA × SAT interaction. Colleges in less selective tiers may admit students on the basis of a high GPA or a high SAT, while selective colleges prefer applicants who score highly on both. This may be reflective of the adage that highly selective colleges look for a "reason to reject" applicants rather than a "reason to admit".

Third, preferences for URMs (relative to the Caucasian reference class) are large and significant in the most selective tier and diminish in magnitude as selectivity decreases. As previously discussed, these coefficients reflect not only direct preferences based on race but also preferences for factors such as economic, geographic, and educational background that are unobserved in our data.[28] Additionally, we observe statistically significant negative coefficients for South Asian appli-

[28]In the less selective tiers, the small magnitude of these coefficients does not necessarily contradict Table 3, which

cants in Tiers 1–4, and for East Asian applicants in T1 and T4, though these are small in magnitude compared to the preferences for URMs. The differential cutoffs for South and East Asians applicants are aligned with findings in Grossman et al. (2024), as we discussed in Section 1.2.

Finally, the overall in-state effect is positive and significant in T1, while the public college in-state effect in positive and significant in Tiers 1–3. We also observe that the coefficient (δ) on the applicant-specific effect—which reflects the variance of applicants along unobservable qualities such as course difficulty, extracurriculars, letters of recommendation, need for financial aid, etc.—is U-shaped in selectivity, achieving the largest magnitude in T1 and T5.

In Appendix B.2, we estimate a separate model for each admission cycle from 2012–2022, allowing for the possibility of temporal changes in the importance of different factors. We find that the coefficients are generally stable during 2012–2020, and comparable to those of our pooled-year model. For 2021 and 2022, we observe that the SAT coefficients are roughly half their pre-COVID magnitudes in Tiers 2–5, which comports with the near-universal switch to test-optional policies during this period. As mentioned earlier, we also consider in Appendix B.3 a number of alternative model specifications (which include, for example, quadratic terms in SAT and GPA and a Race $\times$ SAT interaction), and observe very similar results in our policy simulations.

4.1 Marginal Effects

We provide in this section the importance of SAT, GPA and race at the margin. Intuitively, for a continuous-valued covariate (SAT and GPA) measured in standardized units, a marginal effect of Δ means that for small $\varepsilon > 0$, an ε-standard deviation increase in this covariate translates (on average) into a $\Delta \cdot \varepsilon$ increase in the admission probability predicted by our model.[29] The marginal effect of race (indicators) is the difference in the estimated probability of admission compared to the scenario where the applicants are treated as Caucasian (the reference class).

Figure 4 illustrates the average marginal effects of SAT and GPA over (a) the set of applications received by colleges in each tier; (b) the set of acceptances for each tier in the data; and (c) the set of *borderline* applications with predicted admission probability between 0.48 and 0.52. With the exception of T1, there is a pattern of declining marginal importance of GPA and SAT (among all who applied or were accepted) as one progresses from more selective to less selective tiers, for the full set of applications. The seemingly smaller marginal effects for T1 applicants are reflective of relatively low baseline admission probabilities: In our data, the overall admit rates in T1 and T2 are 22.8% and 52.4%, respectively, such that the displayed absolute (percentage-point) average marginal effects for applications ("applied") correspond to significantly larger relative (percent) effects of 37% and 29% for GPA, and 28% and 24% for SAT. We remark that borderline applications are subject to the largest possible marginal effect (as the factor $p_i(1-p_i)$ in Footnote 29 is maximal when p_i is close to 0.5). We observe less of a trend across tiers on this metric, for both GPA and

indicates that the URM admits have substantially lower average GPA and SAT scores than the non-URM admits. As Figure 14 shows, within each tier we still observe differences in the average admit rate of colleges applied to: in less selective tiers, we find that African-Americans and Caucasians apply to colleges with higher admit rates compared to Asian and Hispanic applicants.

[29] Technically, for an application with covariates X, the *marginal effect* of a continuous covariate x_ℓ is the derivative $\Delta_\ell := \mathbb{E}_g[\frac{\partial p(X,\eta)}{\partial x_\ell}]$, where $\mathbb{E}_g[\cdot]$ denotes expectation over the posterior distribution of the applicant's η. In the estimated logit model, the marginal effects of GPA and SAT for an application i with admission probability $p_i(\eta_i)$ are of the form (here, β is for the relevant tier)

$$\Delta_{i\mathrm{GPA}} = \mathbb{E}_{g_i}\left[p_i(1-p_i)\right] \times \left[\beta_{\mathrm{GPA}} + \beta_{\mathrm{GPA}\times\mathrm{SAT}}\mathrm{SAT}_i\right]$$
$$\Delta_{i\mathrm{SAT}} = \mathbb{E}_{g_i}\left[p_i(1-p_i)\right] \times \left[\beta_{\mathrm{SAT}} + \beta_{\mathrm{GPA}\times\mathrm{SAT}}\mathrm{GPA}_i\right].$$

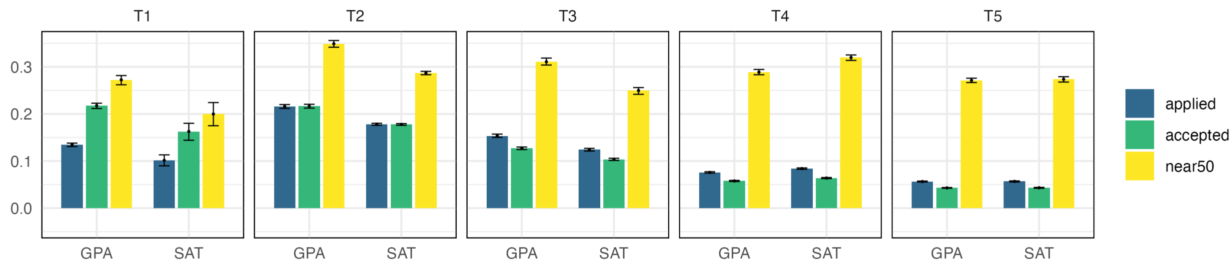

Figure 4: Average marginal effects for GPA and SAT, for (i) all applications; (ii) all acceptances; and (iii) applications with (predicted) admission probability between 0.48 and 0.52. The covariates are measured in standardized units.

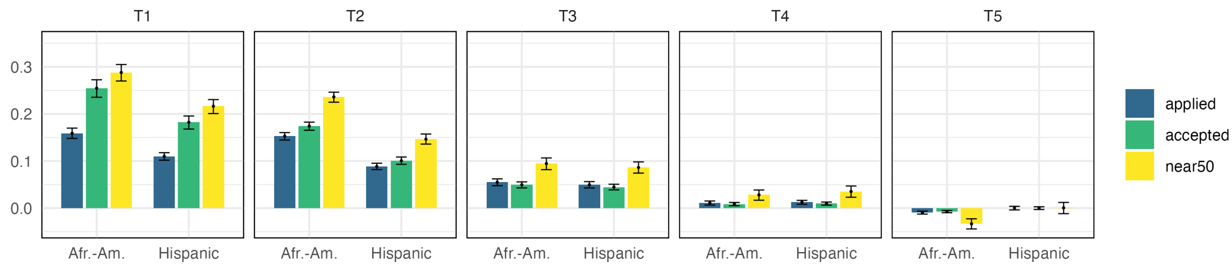

Figure 5: Average marginal effects of racial preferences for African-Americans and Hispanics, for (i) all applications; (ii) all acceptances; and (iii) all applications with (predicted) admission probability between 0.48 and 0.52.

SAT. Finally, we note that the marginal effect of GPA is on average higher than that of SAT in Tiers 1–3, but the two are comparable in Tiers 4–5.

Figure 5 presents the average marginal effects of race for African-American and Hispanic applications, acceptances, and borderline applications. We can see that African-American applicants received, on average, a 16 percentage point (pp) increase (or a 93% relative increase) in admissions probability in T1; a 15pp [46%] increase in T2; and a 6pp [9%] increase in T3, due to race-conscious admissions (again, encompassing preferences for unobserved factors correlated with race, as we discussed in the previous section). Hispanic applicants received smaller increases of 11pp [63%], 9pp [25%], and 5pp [8%] in Tiers 1–3. The percentage-point increase in T1 is larger among admitted students: 26pp [64%] for African-Americans and 18pp [42%] for Hispanics. In less selective tiers, URMs who were admitted received, on average, a comparable increase to those who applied.[30]

4.2 Alternative Admissions Policies

We now discuss our empirical exercise in which we replace our estimated model of admissions (which we refer to as the *status quo*) with several alternatives: (a) universal race-blind admissions; (b) universal test-optional admissions; and (c) universal race-blind, test-optional admissions. Relative to the status quo, we examine the impact of alternative policies on the racial composition and mean SAT score of unique admitted students within each tier. We refer to the exercises as “simulations” rather than counterfactual predictions as we do not account for general equilibrium effects, a point that we revisit in Section 4.3. As a matter of curiosity, we also report in Appendix D.3 results for test-blind admissions, in which standardized test scores are not taken into consideration at all.

[30] Arcidiacono, Kinsler, and Ransom (2023) reports the average marginal effects of race for Harvard (classes of 2014–19) and UNC (classes of 2016–21) URM applicants in Table 9. While the institution-years and baseline admit rates are not exactly comparable, the results in percentage terms are broadly similar to ours. UNC is in our T2.

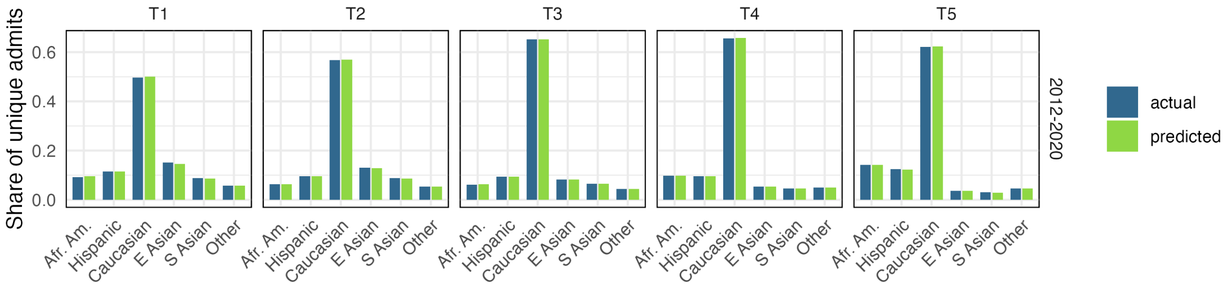

Figure 6: Race shares of unique admitted students, observed in the data versus predicted under the baseline model. We note that our estimated model satisfies certain moment conditions which equalize observed and predicted shares of acceptances for each tier, but not necessarily of unique admits (see Appendix D.6 for details).

In each simulation, we modify the admission model in accordance with the considered policy and compute a new score (W_{ij}) for each application.[31] To respect capacity constraints approximately, we adjust each college's cutoff for each year such that its expected number of *admitted* students remains unchanged from the status quo. We then calculate the new admission probability for each application, and the probability that each applicant is admitted to at least one college in her application set in each tier.[32] By summing these probabilities, we calculate for each tier the expected number of unique admitted students by race, and the corresponding shares, as well as the expected SAT score of unique admits. The predicted shares under the status quo model are shown in Figure 6; they are close to the observed shares, showing that our model is well-calibrated along this dimension.[33] Figures 24 and 25 in Appendix B.1 further demonstrate that the model is well-calibrated in terms of predicted admit rates as a function of SAT and GPA, respectively.

4.2.1 Race-Blind Admissions

To simulate race-blind admissions, we set all race coefficients in the model to zero, effectively treating all applicants as if they were Caucasian (the reference class). Mechanically, this is the same procedure as in Long (2004) and Arcidiacono, Kinsler, and Ransom (2023). As discussed earlier, this corresponds to elimination not only of direct racial preferences, but also of preferences for unobserved characteristics which are correlated with race.

The simulated impact is shown in absolute terms in Figure 7, and as a relative change in the top row of Figure 8. In T1, the African-American and Hispanic shares are estimated to decline by 38% and 27% respectively, while the Caucasian, East Asian, and South Asian shares are estimated to increase by 8%, 9%, and 19%. In T2, the African-American and Hispanic shares are estimated to decline by 25% and 13%; in T3, by 6% and 5%. The negative (relative) effect on the URM admitted share decreases as one progresses from higher to lower tiers, corresponding to the diminishing magnitude of the race coefficients as shown in Figure 3. The mean SAT score increases by 10

[31] We take into consideration all applications with reported outcomes (the same set of data used to fit our model). We find virtually identical results when also including the applications which do not have reported outcomes (by applicants with at least one reported outcome).

[32] Given the independence of the ξ_{ij}'s, the probability that applicant i is admitted to at least one college in tier k is $\mathbb{E}_{g_i}[1 - \prod_{j \in \mathcal{J}_i; T(j)=k}(1 - p_{ij}(\eta_i))]$, where $\mathcal{J}_i$ is applicant i's application set, $T(j)$ is the tier to which college j belongs, and $\mathbb{E}_{g_i}[\cdot]$ denotes expectation over the posterior distribution of η_i (see Section D.6 for more details on g_i).

[33] The way we adjust the cutoffs by college-year does not guarantee that the expected number of unique admits in each tier remains unchanged from the status quo. For each alternative policy, we find small deviations of less than 3 percent in T1; less than 2 percent in T2; and less than 1 percent in Tiers 3–5; see Figure 23.

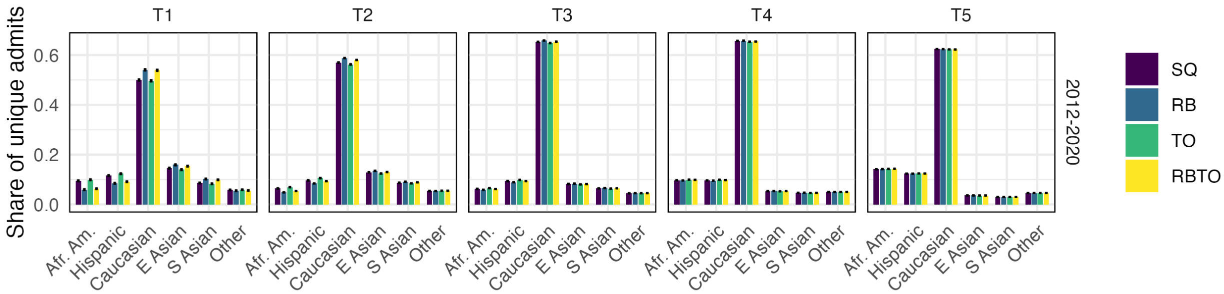

Figure 7: The simulated impact on race shares of unique admitted students in each tier under the status quo (SQ); a "race-blind" (RB) policy which sets race coefficients equal to zero; a "test-optional" (TO) policy in which colleges impute the mean SAT score conditioned on GPA and applicants withhold scores below this imputed value; and a "race-blind test-optional" policy (RBTO) which combines these elements.

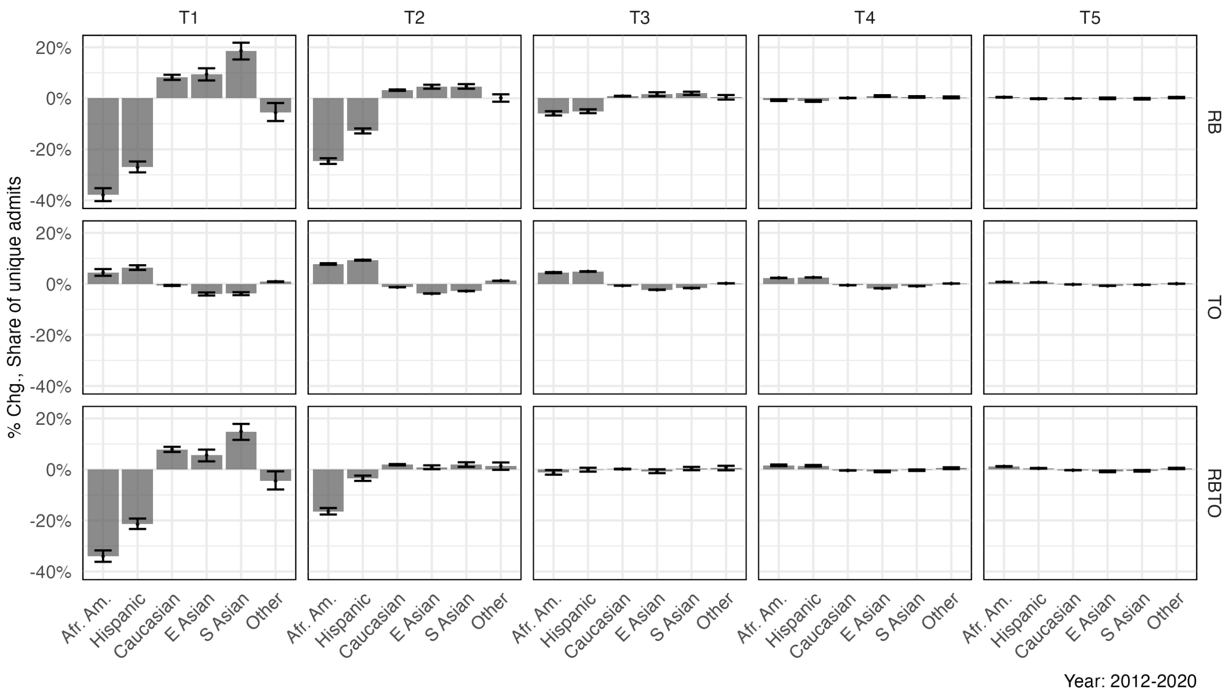

Figure 8: The simulated impact on race shares of unique admitted students in each tier (relative to the predicted shares under the baseline model) under a "race-blind" (RB) policy which sets race coefficients equal to zero; a "test-optional" (TO) policy in which colleges impute the mean SAT score conditioned on GPA and applicants withhold scores below this imputed value; and a "race-blind test-optional" policy (RBTO) which combines these elements.

points in T1, and by less than 5 points in the less selective tiers.[34]

[34] As a point of comparison, Lee et al. (2024) analyze admissions data from a "highly selective, engineering focused American university" (which is in our T1). They estimate that removing race as a predictor—in a model that predicts past admission decisions—leads to a 62% decline in the share of URM applicants in the top quintile of admission probability, and a 10-point increase in the mean SAT score.

4.2.2 Test-Optional Admissions

Test-optional admission requires modeling assumptions regarding which applicants withhold scores and how colleges treat these applicants. We consider a policy in which colleges impute an SAT score for non-reporters, but otherwise apply the same admissions scoring rule to both reporters and non-reporters. One subtle concern is the information with which applicants or colleges make decisions: If a college selects some imputation rule—and if applicants know this rule and respond rationally by withholding scores lower than their would-be imputed scores—then the imputations will have been too optimistic. If colleges account for this adverse selection, what follows is an "unraveling" whereby no applicants withhold scores in equilibrium. We direct the reader to Appendix D.7 (and to Kelly (2022), Dessein, Frankel, and Kartik (2023) and Liu and Garg (2021)) for a more detailed equilibrium argument, and state here our assumptions under which some applicants will choose to not report their test scores.

We assume that colleges do *not* condition an applicant's imputed SAT score on their decision to withhold her score. This corresponds to the "no adverse inference" assumption of Dessein, Frankel, and Kartik (2023), or the "cursed equilibrium" notion of Eyster and Rabin (2005).[35] We otherwise allow the imputation to condition on specific observable characteristics, excluding race. For simplicity, we consider the applicant's GPA ventile among all of college j's applicants, denoted by $v_j(\text{GPA}_i)$, as the only conditioning variable. Accordingly, we define $\widetilde{\text{SAT}}_{ij}$ as the sample average of the SAT scores of applicants to college j who are in the same GPA ventile as i:

$$\widetilde{\text{SAT}}_{ij} = \mathbb{E}[\text{SAT} \mid v_j(\text{GPA}_i)]\,.$$

We further assume that applicants withhold their scores if and only if the score is lower than the imputed value. This strategy is optimal as long as W_{ij} is monotonically increasing in SAT, which is the case in our estimated model even for T5 where the coefficient on GPA:SAT is slightly negative. As discussed in Section 1.2, strategic behavior by applicants is supported by an empirical examination of score-reporting behavior in McManus, Howell, and Hurwitz (2023): conditional on true SAT score, applicants with weaker GPA are more likely to disclose their score. Combining assumptions on both sides, we simulate the test-optional outcome by replacing applicants' SAT by

$$\widehat{\text{SAT}}_{ij} = \begin{cases} \widetilde{\text{SAT}}_{ij} & \text{if } \text{SAT}_i \le \widetilde{\text{SAT}}_{ij}\,, \\ \text{SAT}_i & \text{otherwise.} \end{cases}$$

In Appendix D, we find similar results when assuming that applicants make reporting decisions based on their position in the distribution of SAT scores among each college's enrollees. This information is available as colleges often publish the quartiles of their enrollees' scores.

The second row of Figure 8 shows the impact on race shares. We find that the policy as conceived leads to a 4% increase in the African-American share, and a 6% increase in the Hispanic share, relative to the status quo in T1, and a less than 5% decrease in the Caucasian, East Asian, and South Asian shares. One can explain this result by observing that, since the imputation rule does not condition on race and URM applicants tend to have lower SAT scores conditioned on GPA (see Figure 16), these applicants disproportionately benefit from the imputation. However, the increase in URM representation is modest, especially for the less selective colleges, since URM applicants also tend to have lower GPAs. The increase may be larger if colleges do not condition their imputations on GPA.

[35] This assumption is motivated by statements by colleges which declare that non-submitting applicants will not be penalized or disadvantaged in the admissions process. See, for example, https://web.archive.org/web/20240000000000*/https://undergrad.admissions.columbia.edu/apply/process/testing.

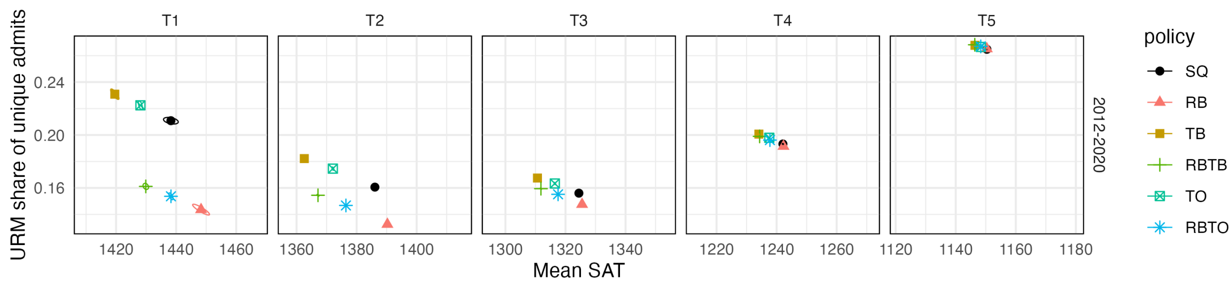

Figure 9: The predicted "diversity" (URM share of unique admitted students) and "merit" (average SAT score of unique admitted students) levels associated with each policy under consideration, with a 95% confidence region. The policies are abbreviated as SQ (status quo), RB (race blind), TB (test blind), and TO (test optional).

We also observe that among accepted T1 applications, African-Americans and Hispanics report their scores only 35% and 42% of the time under our assumed rule, while Caucasians, East Asians, and South Asians report scores 62%, 81%, and 80% of the time. As a consequence of applicants' masking of below-average scores, we find that the mean *true* SAT score of all admitted students (regardless of whether they reported their scores) decreases by 10 points in T1; by 14 points in T2; by 8 points in T3; and by less than 5 points in Tiers 4 and 5.

4.2.3 Race-Blind, Test-Optional Admissions

Our third simulation estimates the impact of simultaneously implementing the race-blind and test-optional policies, with the results shown in the third row of Figure 8. The effects are approximately additive, and the outcomes are more similar to race-blind alone than to test-optional alone: In Tiers 1 and 2, we find a 34% and 16% decrease in the African-American share relative to the status quo, respectively. We find a 21% decrease in the Hispanic share in T1, but a negligible impact on the Hispanic share in Tiers 2–5. This suggests that for colleges in T1 and T2, eliminating requirements for standardized tests is not sufficient to offset the decrease in URM shares due to race-blind admission. We also observe that the mean *true* SAT score of unique admits (including those who choose not to report) decreases by 10 points in T2; by 7 points in T3; and by less than 5 points in all other tiers including T1. This is consistent with the fact that the decrease in SAT scores in Tiers 2–3 under test-optional is larger than the corresponding increase under the race-blind policy.

4.2.4 Diversity and Merit

We now compare the predicted impact of different policies on both the URM share and the average SAT score of unique admits to each tier. See Figure 9, in which we also include, for reference, the test-blind and race-blind, test-blind policies which are described in Appendix D. We use the two metrics as operational definitions of "racial diversity" and "academic merit", respectively. Our narrow definitions of these terms are not a normative claim about how they should be measured, but merely serve to advance our analysis within the limitations of the data and our model.

Relative to the status quo under our estimated model, the increase in racial diversity in T1 under the test-optional and test-blind policies is small compared to the decline under race-blind policies. Moreover, due to the additivity mentioned above, the race-blind, test-optional (RBTO) outcome is weakly dominated by the status quo outcome in both diversity and merit in Tiers 1–3, while the impacts on Tiers 4 and 5 are substantially smaller. In less selective tiers, the choice of policy does not appear to substantially affect either the racial diversity or academic merit metrics.

4.3 Discussion

As we have discussed earlier, while interpreting our results one should bear in mind that our estimated racial preferences may partly reflect preferences for unobserved characteristics correlated with race (such as family income, high school background, and intended major). Moreover, rather than making precise predictions, our policy simulations quantify the magnitude of these preferences during the 2012–2020 admission cycles, in terms of the extent to which they might have affected admission outcomes. When substantial policy changes take effect, we may expect shifts in students' application decisions, and adjustments in colleges' admission scoring processes. We briefly discuss market equilibration below as well as our additional findings presented in the appendices which supplement our main analysis.

Enrollees Our preceding analysis is focused on changes in the racial composition of the set of *unique admitted students* in each tier. Since applicants commonly receive offers from colleges in more than one tier—and only enrollee statistics are reported in public data sources such as IPEDS—colleges, policymakers and the public may also be concerned with the impact on *enrollment*. To bridge this gap, we estimate in Appendix D.1 a simple model of tier-level enrollment choice based on reports of enrollment in the data. In our model, the probability that an admit enrolls in a tier is based on the number of offers he has in each tier, including an outside option, with race-specific weights which we estimate from the data (the more offers an admit has from a given tier, the more likely he is to enroll in that tier). Assuming that the weight parameters are policy-invariant, we calculate the percent change in each race's share of *enrollees* relative to that under the status quo, and compare the results in Figure 41 with those presented above for unique admitted students. We see that the impact on URM enrollment is virtually the same for T1 and T2, while tiers 3-5 observe a slightly smaller decrease or larger increase in URM enrollment under race-blind admission policies. This is intuitive, since the students who are no longer admitted to (or received fewer offers from) top tier colleges may now decide to enroll in the less selective tiers.

Changes in Application Decisions In response to changes in admissions policies, applicants may change their application patterns, for example by shifting their applications toward or away from colleges where they are now less likely to be accepted. In Appendix D.4, we consider two scenarios in which non-URM applications are held fixed, while URM applicants to T1 and T2 colleges send one or two additional applications (this may be effected, for example, by targeted outreach by colleges or by URM applicants' own self-directed efforts). We find that extra applications has a large moderating effect on the decline in the URM share of unique admits under our race-blind simulation, and substantially increases the URM share under our test-optional simulation. Given that the typical T1 (or T2) URM applicant applies to fewer than three T1 colleges (or two T2 colleges) in our data, two additional applications represents a large application response. We anticipate that, at least in the short run, there will not be such an increase in URM application volume at selective institutions. In fact, past studies have focused on whether URM applications had declined in response to decreased admission probabilities after state-level affirmative action bans; as we discuss in Section 1.2, these studies have found little to no impact on applications among competitive URM applicants (Antonovics and Backes, 2013; Card and Krueger, 2005). Addressing the *SFFA* ban directly, Kim et al. (2024), using comprehensive Common App data from the 2023-24 admissions cycle, does not find stark deviations in (i) the fraction of URM applicants applying to at least one "most-selective" college (those with admit rates below 25%); (ii) the average admit rate of colleges in URM applicants' portfolios; (iii) the number of applications per URM applicant, or (iv) the racial composition of most-selective colleges' applicant pools, relative to preexisting trends.

Changes in Admissions Criteria Concerning the possibility of colleges modifying their admissions practices, Antonovics and Backes (2014) presents evidence that University of California (UC) campuses adjusted their admissions criteria—namely by decreasing the effective weight given to SAT scores and increasing the effective weights given to GPA and family background—in such a manner as to partially offset the decline in URM admit rates that would have obtained absent the adjustments.[36] The results are similar to those found by Long and Tienda (2008) in a study of the effects of the affirmative action ban in Texas. Additionally, there are studies which allow for endogenous changes in admissions policies under a structural modeling approach. For example, as we discuss in Appendix D.5, Borghesan (2023) assumes that colleges estimate ability from noisy measurements of academic performance (GPA, test scores), as well as demographics and endogenous study effort; under a no-SAT counterfactual, the set of measurements is reduced, hence the mapping from inputs to admissions decisions changes. Notwithstanding the richness of such structural models, survey data limitations and estimation challenges necessitate assumptions regarding, e.g., the granularity of college differentiation (through admission cutoffs) and the structure of applicants' portfolio choices. For this reason, we believe the extent to which colleges will change their practices in response to the *SFFA* ban is an empirical question best answered *ex-post*. We do not consider it in our analysis but include preliminary observations from the 2023-24 admissions cycle in our concluding remarks in Section 5.

5 Concluding Remarks

Treating our estimated model of admissions outcomes in 2012–2020 as a *status quo* baseline, we simulate the potential impact of universal race-blind and test-optional admissions on the URM share ("racial diversity") and mean SAT scores ("academic merit") of colleges. We find that the effects of these policy changes are concentrated in more selective tiers of colleges, where we find racial preferences to be more pronounced: In T1, and to a lesser extent in T2, racial diversity declines and academic merit increases under the race-blind outcome, relative to the status quo. The reverse is true of test-optional policies, albeit with substantially smaller magnitudes. These tiers represent elite and highly selective colleges, respectively, and collectively enroll less than 10% of college students. In contrast, we find that the effects are minimal in the less selective tiers which account for the remaining 90% of enrollment in four-year colleges. Moreover, we find that when race-blind and test-optional policies are adopted simultaneously, the outcome closely resembles that under race-blind alone. In this respect, our study complements the literature showing that various race-neutral measures—such as top percent policies and class-based preferences—fall far short of matching the impact of race-based affirmative action.

Class of 2028 As of this writing, a number of selective colleges have released enrollment statistics for the Class of 2028, the first cohort of entering freshmen affected by the universal affirmative action ban; we present the college-level percent change in URM enrollment shares in Figure 10.[37] The numbers reveal stark heterogeneity in the realized impact on racial composition, even among peer institutions: At MIT, the African-American share of enrollees decreased to 5% from 15% the previous year, while the Hispanic share decreased to 11% from 16% and the Asian share increased

[36] We call these *effective* weights as they correspond to the authors' model of admissions, rather than to the true admission rule employed by the colleges. The authors find a "rebound" of approximately 20 percent of the predicted drop in URM admit rates at UC–Berkeley, and larger rebounds at less selective UC campuses, relative to an admissions rule which sets the URM coefficient equal to zero.

[37] These numbers are obtained from https://edreformnow.org/2024/09/09/tracking-the-impact-of-the-sffa-decision-on-college-admissions/, accessed September 12, 2024.

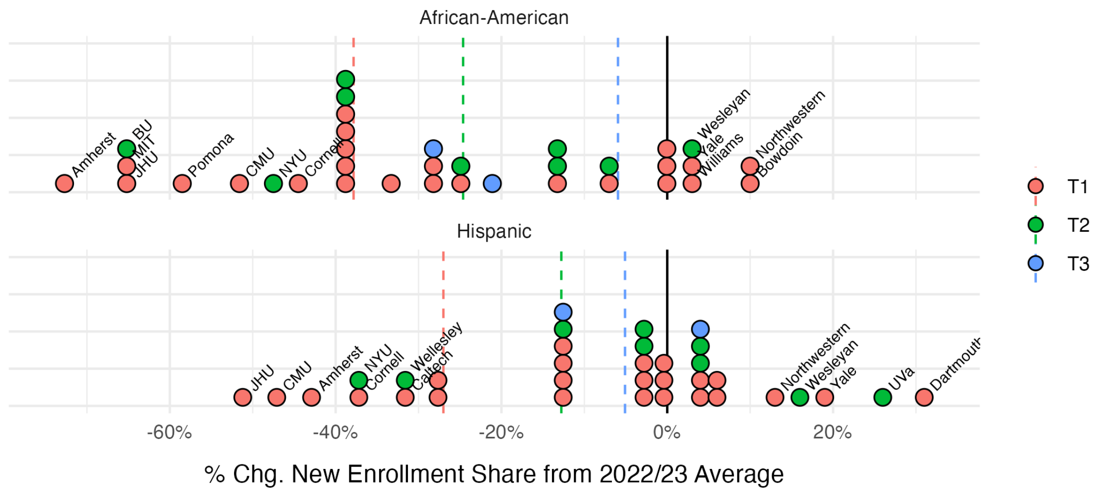

Figure 10: The realized percent change in the African-American and Hispanic shares of enrollees at selected colleges, in 2024 compared to the average enrollment share in 2022-23. Numbers obtained from https://edreformnow.org/2024/09/09/tracking-the-impact-of-the-sffa-decision-on-college-admissions/, accessed October 28, 2024. Note that the included colleges are limited to those which were previously race-aware, and which reported data on the racial composition of enrollees in sufficient detail. Moreover, the numbers are not necessarily reported consistently across institutions; see the figure notes in the link. The dashed vertical lines correspond to the results of our "race-blind" simulation for enrollees in Tiers 1–3; see Section D.1 for details.

to 47% from 40%.[38] At Yale, by contrast, the African-American share remained unchanged at 13%, while the Hispanic share increased to 19% from 18%, and the Asian share decreased to 24% from 30%.[39] Other colleges including Harvard and UNC saw intermediate declines in their URM enrollment shares, relative to the previous year.[40]

We cannot yet present a comprehensive comparison with our tier-level policy simulation results, as enrollment statistics for the Class of 2028 are not yet published by IPEDS. However, we note that at least several elite universities have adopted strategies which represent deviations from our modeling assumptions, namely the absence of changes in admissions practices or application behavior. Some expanded recruitment efforts targeting high-potential students from non-traditional backgrounds, or modified financial aid policies to increase affordability. For example, Duke—which, like Yale, did not see a substantial change in the enrollment share of URMs—reported that "admissions staff visited more high schools to spread the word about the new tuition grant program for students from the Carolinas and reviewed and updated admissions processes to cast the broadest net possible."[41] Similarly, MIT reported that "After the [Supreme Court] decision, we responded with expanded recruitment and financial aid initiatives designed to improve access to MIT for students from all backgrounds. These efforts include a new targeted outreach program to identify and encourage students in rural America to apply to MIT. They also include a new policy under which most families earning less than $75,000 a year pay nothing to attend [...]"[42] Notably, the

[38] https://www.nytimes.com/2024/08/21/us/mit-black-latino-enrollment-affirmative-action.html, accessed September 12, 2024.

[39] https://admissions.yale.edu/sites/default/files/classprofile2028web.pdf and https://admissions.yale.edu/sites/default/files/2027classprofileweb.pdf, accessed September 12, 2024.

[40] https://www.nytimes.com/2024/09/11/us/harvard-affirmative-action-diversity-admissions.html, and https://www.nytimes.com/2024/09/05/us/unc-affirmative-action-black-enrollment.html, accessed September 12, 2024.

[41] https://today.duke.edu/2024/09/dukes-newest-undergraduates-numbers, accessed September 11, 2024.

[42] https://news.mit.edu/2024/qa-undergraduate-admissions-in-wake-of-supreme-court-ruling-0821, ac-

fraction of new Yale enrollees who are Pell Grant recipients—an indicator of low-income status—increased to 25%, from 22% the previous year and 20% two years prior.[43] Our analysis does not account for the potential impact of changing income-based preferences and financial aid policies on applications, admissions, and enrollment.

Long-term Impact Moreover, there may be differential short-term and long-term impacts of the ban, as applicants and colleges adapt to the new policy environment. As we have noted, colleges may continue to express preferences for characteristics which are correlated with race, such as family income, high school background, and intended major (they are also still permitted to consider "an applicant's discussion of how race affects his or her life, be it through discrimination, inspiration, or otherwise"[44]). In a recent legal case involving Thomas Jefferson High School—a selective public high school in Virginia—the Supreme Court hinted that it may refrain from delineating boundaries around admissions policies, as long as such policies do not make explicit racial considerations. In particular, following the *SFFA* decision, the Court declined to review a ruling from the Fourth Circuit Court of Appeals which upheld a revised set of admissions criteria—criteria which were ostensibly race-neutral but had the notable effect of "quadrupling" the number of offers to Black and Hispanic students, and reducing offers to Asian students.[45] Thus, over time, colleges may choose to modify their admissions criteria, as was the case in the University of California (UC) system after the passage of Prop 209, the state-level affirmative action ban in California. As argued in Antonovics and Backes (2014), UC campuses which initially experienced large declines in URM enrollment modified their admissions policies so as to produce a rebound in the following years. Chan and Eyster (2003) documents specific changes in admissions criteria which led to rebounds at UC–Berkeley, Berkeley Law School, and the University of Texas at Austin (which was similarly affected by a state-level ban), including the overhaul of the Academic Index Score (AIS) at Berkeley and the enactment of the "Top 10 Percent" law in Texas.

Informativeness of Admission Determinants It remains to be seen whether a substantial fraction of colleges will follow the small number of elite institutions which have reinstated testing requirements (see Footnote 5). While the focus of this study is the interaction between racial preferences and testing policies, there are other trends which are pertinent to the decision of which testing policy to adopt, in particular the erosion of the signal value of traditional determinants of admission. The advent of large language models (LLMs) has provided a convenient technology to help applicants craft their application essays. Steady high school GPA inflation has made it more difficult for students to distinguish themselves academically in the absence of standardized tests. (An analysis by Brown University notes that "standardized test scores are a much better predictor of academic success than high school grades, which are exceptional for the vast majority of Brown applicants but also carry the complication of being increasingly subject to grade inflation".[46]) And, starting in 2024—and likely in response to the proliferation of test-optional policies—the SAT itself has been relaunched as a digital-only exam, with exam time reduced from 3 hours to

cessed September 11, 2024.

[43] Note that this coincides with a change in eligibility criteria leading to a 10% overall increase in Pell Grant recipients; see https://www.insidehighered.com/news/admissions/traditional-age/2024/10/29/fafsa-change-driving-colleges-surging-pell-numbers, accessed February 1, 2025.

[44] *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College*, 600 U.S. 181 (2023).

[45] https://www.scotusblog.com/2024/02/justices-decline-to-intervene-in-another-dispute-over-race-and-school-admissions/, accessed July 15, 2024.

[46] https://www.brownalumnimagazine.com/articles/2023-06-20/to-test-or-not-to-test, accessed February 11, 2024. We provide evidence of grade inflation in our data set in Figure 16.

2 hours and 15 minutes, and featuring shorter reading passages.[47] Together, these developments may necessitate more discretionary judgments by admissions officers and incentivize applicants to turn toward alternative signals. Access to standardized tests is often cited as a reason for going test-optional, but access to these alternatives is not necessarily more equitable.[48]

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. URL https://doi.org/10.3102/0013189X20902110.

Antonovics, Kate and Ben Backes. 2013. "Were Minority Students Discouraged from Applying to University of California Campuses after the Affirmative Action Ban?" *Education Finance and Policy* 8 (2):208–250. URL https://doi.org/10.1162/EDFP_a_00090.

———. 2014. "The Effect of Banning Affirmative Action on College Admissions Policies and Student Quality." *The Journal of Human Resources* 49 (2):295–322. URL http://www.jstor.org/stable/23799086.

Arcidiacono, Peter. 2005. "Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?" *Econometrica* 73 (5):1477–1524. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0262.2005.00627.x.

Arcidiacono, Peter, Josh Kinsler, and Tyler Ransom. 2022a. "Asian American Discrimination in Harvard Admissions." *European Economic Review* 144:104079. URL https://www.sciencedirect.com/science/article/pii/S0014292122000290.

———. 2022b. "Legacy and Athlete Preferences at Harvard." *Journal of Labor Economics* 40 (1):133–156. URL https://doi.org/10.1086/713744.

———. 2023. "What the Students for Fair Admissions Cases Reveal about Racial Preferences." *Journal of Political Economy Microeconomics* 1 (4):615–668. URL https://doi.org/10.1086/725336.

Bennett, Christopher T. 2022. "Untested Admissions: Examining Changes in Application Behaviors and Student Demographics Under Test-Optional Policies." *American Educational Research Journal* 59 (1):180–216. URL https://doi.org/10.3102/00028312211003526.

Bleemer, Zachary. 2023. "Affirmative action and its race-neutral alternatives." *Journal of Public Economics* 220:104839. URL https://doi.org/10.1016/j.jpubeco.2023.104839.

Borghesan, Emilio. 2023. "The Heterogeneous Effects of Changing SAT Requirements in Admissions: An Equilibrium Evaluation." Working paper. URL https://www.emilioborghesan.com/s/JMP_Borghesan12_2023.pdf.

[47] https://newsroom.collegeboard.org/digital-sat-launches-across-country-completing-transition-digital-and-providing-simpler-testing, accessed September 11, 2024.

[48] As one example, there is a burgeoning industry providing students with opportunities to publish in "peer-reviewed" high-school research journals, with most families paying "between $2,500 and $10,000" for research mentorship. See https://www.propublica.org/article/college-high-school-research-peer-review-publications, accessed September 12, 2024.

Bradley, Ralph A. and John J. Gart. 1962. "The asymptotic properties of ML estimators when sampling from associated populations." *Biometrika* 49 (1-2):205–214. URL https://doi.org/10.1093/biomet/49.1-2.205.

Buckley, Ed., Jack, Ed. Letukas, Lynn, and Ed. Wildavsky, Ben. 2018. *Measuring Success: Testing, Grades, and the Future of College Admissions*. Johns Hopkins University Press. URL https://eric.ed.gov/?id=ED598409.

Card, David and Alan B. Krueger. 2005. "Would the Elimination of Affirmative Action Affect Highly Qualified Minority Applicants? Evidence from California and Texas." *ILR Review* 58 (3):416–434. URL https://doi.org/10.1177/001979390505800306.

Chan, Jimmy and Erik Eyster. 2003. "Does Banning Affirmative Action Lower College Student Quality?" *American Economic Review* 93 (3):858–872. URL https://www.aeaweb.org/articles?id=10.1257/000282803322157124.

Chetty, Raj, David J Deming, and John N Friedman. 2023. "Diversifying Society's Leaders? The Causal Effects of Admission to Highly Selective Private Colleges." Working Paper 31492, National Bureau of Economic Research. URL http://www.nber.org/papers/w31492.

Dessein, Wouter, Alexander P. Frankel, and Navin Kartik. 2023. "Test-Optional Admissions." Working Paper 4422766, Columbia Business School. URL https://ssrn.com/abstract=4422766.

Dixon-Román, Ezekiel J., Howard T. Everson, and John J. Mcardle. 2013. "Race, Poverty and SAT Scores: Modeling the Influences of Family Income on Black and White High School Students' SAT Performance." *Teachers College Record* 115 (4):1–33. URL https://doi.org/10.1177/016146811311500406.

Espenshade, Thomas J. and Chang Y. Chung. 2005. "The Opportunity Cost of Admission Preferences at Elite Universities." *Social Science Quarterly* 86 (2):293–305. URL http://www.jstor.org/stable/42956064.

Espenshade, Thomas J., Chang Y. Chung, and Joan L. Walling. 2004. "Admission Preferences for Minority Students, Athletes, and Legacies at Elite Universities." *Social Science Quarterly* 85 (5):1422–1446. URL https://doi.org/10.1111/j.0038-4941.2004.00284.x.

Eyster, Erik and Matthew Rabin. 2005. "Cursed Equilibrium." *Econometrica* 73 (5):1623–1672. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0262.2005.00631.x.

Friedman, John N, Bruce Sacerdote, and Michele Tine. 2024. "Standardized Test Scores and Academic Performance at Ivy-Plus Colleges." URL https://opportunityinsights.org/wp-content/uploads/2024/01/SAT_ACT_on_Grades.pdf.

Fu, Chao. 2014. "Equilibrium Tuition, Applications, Admissions, and Enrollment in the College Market." *Journal of Political Economy* 122 (2):225–281. URL https://doi.org/10.1086/675503.

Garg, Nikhil, Hannah Li, and Faidra Monachou. 2023. "Dropping Standardized Testing for Admissions Trades Off Information and Access." Tech. rep., ArXiv. URL https://arxiv.org/abs/2010.04396.

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. URL https://www.aeaweb.org/articles?id=10.1257/pol.20170503.

Grossman, Joshua, Sabina Tomkins, Lindsay C Page, and Sharad Goel. 2024. "The Disparate Impacts of College Admissions Policies on Asian American Applicants." *Scientific Reports* 14 (4449). URL https://www.nature.com/articles/s41598-024-55119-0.

Hurwitz, Michael. 2011. "The impact of legacy status on undergraduate admissions at elite colleges and universities." *Economics of Education Review* 30 (3):480–492. URL https://www.sciencedirect.com/science/article/pii/S0272775710001676.

Hurwitz, Michael and Jason Lee. 2018. "Grade Inflation and the Role of Standardized Testing." In *Measuring Success: Testing, Grades, and the Future of College Admissions*, edited by Jack Buckley, Lynn Letukas, and Ben Wildavsky, chap. 3. Johns Hopkins University Press, 64–83. URL https://eric.ed.gov/?id=ED598409.

Kane, Thomas J. 1998. "Racial and Ethnic Preferences in College Admissions." *Ohio State Law Journal* 59:971–996. URL https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.828.5851&rep=rep1&type=pdf.

Kelly, Jack. 2022. "Who Benefits from Multiple Choice(s)?: The Equilibrium Impacts of Test-Optional College Admission." Senior thesis, Yale University. URL https://economics.yale.edu/sites/default/files/Forms/jack_kelly_senior_essay.pdf.

Kim, Brian, Elyse Armstrong, Mark Freeman, Rodney Hughes, Sarah Nolan, Tara Nicola, and Trent Kajikawa. 2024. "Application trends following the end of race-conscious admissions." Research brief, The Common Application. URL https://www.commonapp.org/files/Common-App-Race-Ethnicity-SCOTUS-2024.pdf.

Kim, Brian, Mark Freeman, Trent Kajikawa, Honeiah Karimi, and Preston Magouirk. 2022. "First-year applications per applicant: Patterns of high-volume application activity at Common App." Research brief, The Common Application. URL https://s3.us-west-2.amazonaws.com/ca.research.publish/Research_Briefs_2022/2022_12_09_Apps_Per_Applicant_ResearchBrief.pdf.

Lee, Jinsook, Emma Harvey, Joyce Zhou, Nikhil Garg, Thorsten Joachims, and René F Kizilcec. 2024. "Algorithms for College Admissions Decision Support: Impacts of Policy Change and Inherent Variability." URL osf.io/preprints/socarxiv/hds5g.

Levine, Phillip and Sarah Reber. 2023. "Can colleges afford class-based affirmative action?" Tech. rep., The Brookings Institution. URL https://www.brookings.edu/articles/can-colleges-afford-class-based-affirmative-action/.

Liu, Qing and Donald A. Pierce. 1994. "A note on Gauss-Hermite quadrature." *Biometrika* 81 (3):624–629. URL https://doi.org/10.1093/biomet/81.3.624.

Liu, Zhi and Nikhil Garg. 2021. "Test-optional Policies: Overcoming Strategic Behavior and Informational Gaps." URL https://arxiv.org/abs/2107.08922.

Long, Mark C. 2004. "Race and College Admissions: An Alternative to Affirmative Action?" *The Review of Economics and Statistics* 86 (4):1020–1033. URL http://www.jstor.org/stable/40042986.

Long, Mark C. and Marta Tienda. 2008. "Winners and Losers: Changes in Texas University Admissions Post-Hopwood." *Educational Evaluation and Policy Analysis* 30 (3):255–280. URL https://doi.org/10.3102/0162373708321384. PMID: 23136455.

McManus, Brian, Jessica Howell, and Michael Hurwitz. 2023. "Strategic Disclosure of Test Scores: Evidence from US College Admissions." Working Paper 23-843, Annenberg Institute at Brown University. URL http://www.edworkingpapers.com/ai23-843.

Nie, Lei. 2007. "Convergence rate of MLE in generalized linear and nonlinear mixed-effects models: Theory and applications." *Journal of Statistical Planning and Inference* 137 (6):1787–1804. URL https://www.sciencedirect.com/science/article/pii/S0378375806001972.

Office of the Under Secretary, U.S. Department of Education. 2023. "Strategies for Increasing Diversity and Opportunity in Higher Education." Tech. rep. URL https://sites.ed.gov/ous/files/2023/09/Diversity-and-Opportunity-in-Higher-Education.pdf.

Reber, Sarah, Gabriela Goodman, and Rina Nagashima. 2023. "Admissions at most colleges will be unaffected by Supreme Court ruling on affirmative action." Tech. rep., The Brookings Institution. URL https://www.brookings.edu/articles/admissions-at-most-colleges-will-be-unaffected-by-supreme-court-ruling-on-affirmative-action/.

Reeves, Richard V and Dimitrios Halikias. 2017. "Race gaps in SAT scores highlight inequality and hinder upward mobility." *The Brookings Institution* URL https://www.brookings.edu/articles/race-gaps-in-sat-scores-highlight-inequality-and-hinder-upward-mobility/.

Sacerdote, Bruce, Douglas O Staiger, and Michele Tine. 2025. "How Test Optional Policies in College Admissions Disproportionately Harm High Achieving Applicants from Disadvantaged Backgrounds." Working Paper 33389, National Bureau of Economic Research. URL http://www.nber.org/papers/w33389.

Skrondal, Anders and Sophia Rabe-Hesketh. 2009. "Prediction in Multilevel Generalized Linear Models." *Journal of the Royal Statistical Society Series A: Statistics in Society* 172 (3):659–687. URL https://doi.org/10.1111/j.1467-985X.2009.00587.x.

Smith, Ember and Richard V Reeves. 2020. "SAT math scores mirror and maintain racial inequity." *The Brookings Institution* URL https://www.brookings.edu/articles/sat-math-scores-mirror-and-maintain-racial-inequity/.

Tierney, Luke and Joseph B. Kadane. 1986. "Accurate Approximations for Posterior Moments and Marginal Densities." *Journal of the American Statistical Association* 81 (393):82–86. URL http://www.jstor.org/stable/2287970.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2nd ed.

Appendices

The Appendix is organized as follows.

– Appendix A contains additional descriptive charts and figures referenced throughout the paper.

– Appendix B reports estimated coefficients for our main model in tabular format; for year-by-year models including the 2021 and 2022 admission cycles; and for alternative model specifications which nest the main model. We also report a summary table of model fit statistics including pseudo-R^2 and AUROC. Moreover, we conduct additional robustness checks which (i) remove colleges which were already affected by state-level affirmative action bans from the sample; and (ii) restrict the sample to applicants who reported all admission outcomes.

– Appendix C contains an analysis of unobserved reporting bias.

– Appendix D contains several extensions to our main analysis, namely (i) an investigation of the impact of our simulated policies on tier-level *enrollment* rather than on unique admits; (ii) an alternative assumption regarding our test-optional simulation in which an applicant's decision to report her SAT score is based on her position in the distribution of past enrollees' SAT scores; (iii) simulations corresponding to universal *test-blind* and universal *race-blind, test-blind* policies; (iv) a discussion of results presented in Borghesan (2023), which studies similar questions under a general equilibrium approach; and (v) policy simulations which additionally assume that URM applicants in T1 and T2 send extra applications. We also provide details on the estimation of our model and the equilibrium assumptions underlying our test-optional policy simulation.

Appendix A Additional Tables and Figures

A.1 Descriptive Statistics and Charts

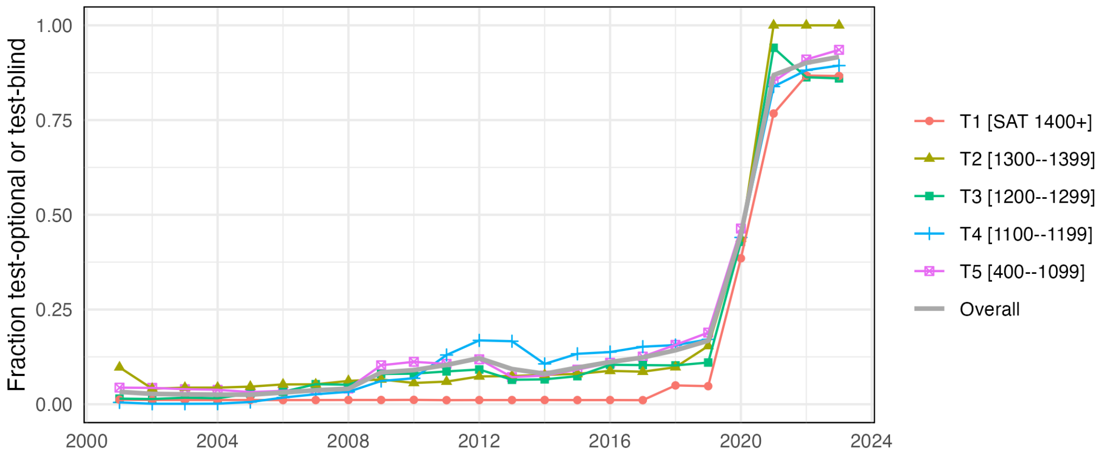

Figure 11: The enrollment-weighted fraction of 4-year colleges (Title IV participating; U.S. only; degree-granting, primarily baccalaureate or above) not requiring submission of standardized test scores, by selectivity tier. Tiers are defined according to the average SAT score of enrollees in 2015. Colleges which did not report average SAT scores are not included. Source: Integrated Postsecondary Education Data System (IPEDS).

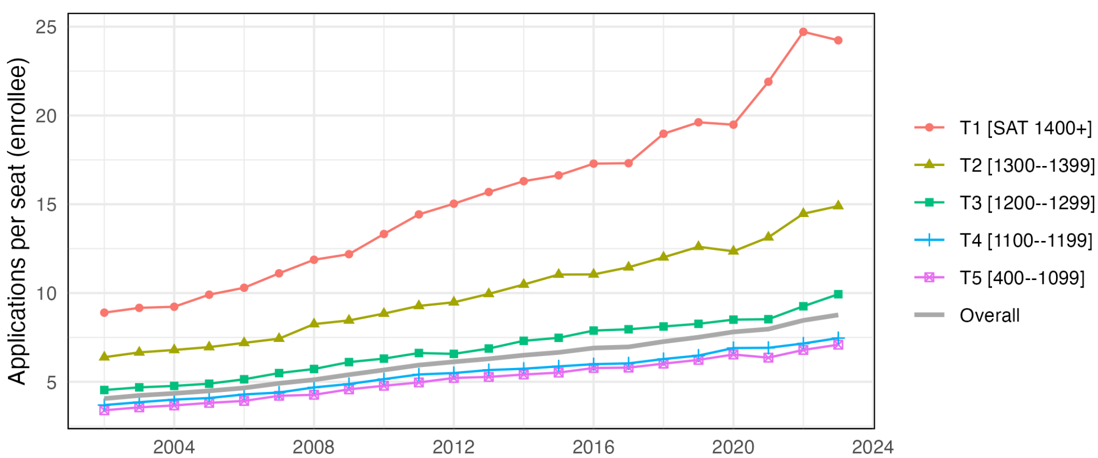

Figure 12: Number of applications per filled seat by tier and year, 2002–2023. Source: Integrated Postsecondary Education Data System (IPEDS).

	Number of:		
Filter	Applications	Applicants	Colleges
<No filter>	17,845,507	4,702,348	1,535
Applications with year applying between 2012 and 2020	12,075,438	2,784,602	1,533
Applicants with gender, ethnicity, home state	7,171,185	1,167,979	1,522
4-year colleges in IPEDS	6,724,553	1,167,979	1,521
Colleges with tier assignment	6,519,152	1,151,966	1,322
Applications with admission outcome	1,631,844	535,099	1,318
Colleges with enrollee SAT distribution	1,602,280	530,408	1,173
Colleges not excluded by enrollee SAT range	1,506,032	513,795	807
Colleges with both acceptances & rejections by year	1,483,297	508,459	794

Table 4: Number of applications, applicants, and colleges after each filter is applied to the data.

	mean	5%	25%	50%	75%	95%	99%	max
All applications	5.65	1	2	4	7	16	25	345
Applications with outcome	3.06	1	1	2	4	9	14	129
Applicants with all outcomes	3.58	1	1	3	5	10	16	128

Table 5: Distribution of number of applications per applicant (mean and selected quantiles). Data samples are the same as those defined in Table 1.

State of Residence	Applicant Share (The Data)	Enrollee Share (IPEDS)	Ratio
Michigan	19.37%	3.10%	6.26
Kentucky	7.20%	1.40%	5.14
Indiana	9.79%	2.41%	4.06
Wisconsin	5.46%	2.11%	2.58
South Carolina	2.99%	1.31%	2.28
Delaware	0.57%	0.30%	1.91
Minnesota	3.34%	1.89%	1.77
West Virginia	0.95%	0.54%	1.76
Virginia	4.71%	2.78%	1.69
Illinois	7.52%	4.44%	1.69
Arizona	2.16%	1.41%	1.53
Kansas	1.38%	0.92%	1.50
New Mexico	0.59%	0.47%	1.24
California	11.10%	10.01%	1.11
Alabama	1.37%	1.34%	1.02
Washington	1.53%	1.61%	0.95
North Dakota	0.18%	0.21%	0.89
Georgia	2.81%	3.62%	0.77
Texas	4.36%	7.60%	0.57
Florida	1.98%	3.87%	0.51
Nevada	0.29%	0.59%	0.50
Utah	0.48%	0.99%	0.49
Alaska	0.13%	0.27%	0.47
Idaho	0.21%	0.48%	0.44
Pennsylvania	1.87%	4.39%	0.42
Montana	0.10%	0.30%	0.34
Ohio	1.18%	3.77%	0.31
Tennessee	0.60%	1.94%	0.31
Mississippi	0.19%	0.61%	0.31
Missouri	0.51%	1.83%	0.28
Oregon	0.21%	0.87%	0.24
South Dakota	0.07%	0.30%	0.23
Other/Foreign	0.84%	3.78%	0.22
North Carolina	0.66%	3.02%	0.22
Nebraska	0.15%	0.71%	0.21
Maryland	0.33%	1.76%	0.19
Oklahoma	0.19%	1.06%	0.18
Colorado	0.28%	1.84%	0.15
New Jersey	0.53%	3.78%	0.14
Vermont	0.03%	0.20%	0.14
New Hampshire	0.06%	0.48%	0.13
District of Columbia	0.02%	0.17%	0.13
Massachusetts	0.34%	2.78%	0.12
New York	0.82%	6.68%	0.12
Rhode Island	0.04%	0.36%	0.12
Hawaii	0.04%	0.35%	0.11
Connecticut	0.19%	1.64%	0.11
Maine	0.05%	0.42%	0.11
Arkansas	0.08%	0.79%	0.10
Wyoming	0.01%	0.11%	0.09
Iowa	0.07%	0.83%	0.09
Louisiana	0.08%	1.53%	0.05

Table 6: Share of applicants (enrollees) by state of residence, aggregated across all years 2012–2020.

SAT / GPA	<2.0	2.0–2.99	3.0–3.19	3.2–3.39	3.4–3.59	3.6–3.79	3.8–4.0
1500–1600	0.0% / 1.5	0.1% / 3.1	0.1% / 3.6	0.2% / 5.6	0.3% / 5.8	0.9% / 5.7	6.1% / 5.6
1400–1490	0.0% / 3.3	0.2% / 3.3	0.2% / 4	0.5% / 4.6	1.0% / 4.7	2.1% / 4.7	8.2% / 4.5
1300–1390	0.0% / 1.6	0.3% / 3.5	0.4% / 4	0.8% / 4.1	1.5% / 4.1	2.7% / 4.1	7.2% / 4
1200–1290	0.0% / 2.9	0.8% / 3.4	1.1% / 3.5	1.7% / 3.6	3.0% / 3.7	4.4% / 3.8	8.5% / 3.6
1100–1190	0.0% / 2.8	1.7% / 3.2	1.8% / 3.3	2.5% / 3.4	3.7% / 3.4	4.3% / 3.4	6.2% / 3.3
1000–1090	0.0% / 2.7	2.1% / 3.1	1.8% / 3.2	2.2% / 3.2	2.7% / 3.3	2.5% / 3.3	2.6% / 3.3
800–990	0.1% / 2.5	3.5% / 2.9	2.1% / 3	2.0% / 3	2.0% / 2.9	1.5% / 3.1	1.2% / 3
<800	0.0% / 2.2	0.5% / 2.6	0.2% / 2.4	0.1% / 2.5	0.1% / 2.5	0.1% / 2.9	0.1% / 2.6

Table 7: Joint distribution of GPA and SAT in the data, and average number of applications per applicant in each cell.

Variable	Factor levels	Recoded levels
Gender	Male, Female	–
Ethnicity	Caucasian, East Asian, South Asian, Middle-Eastern, Native American, Hispanic, African-American, Other	Caucasian, East Asian, South Asian, Hispanic, African-American, Other
Class Rank	Valedictorian, Salutatorian, Top 1%, Top 2%, Top 3%, Top 4%, Top 5%, Top 10%, Top 15%, Top 20%, Top 25%, Top 30%, Top 35% Top 40%, Top 50%, Bottom 50%, Bottom 40%, Bottom 30%, Bottom 20%, Bottom 10%, Class Not Ranked	Top 1%, Top 5%, Top 10%, Top 25%, Top 50%, Bottom 50%, NR

Table 8: Original and recoded factor levels.

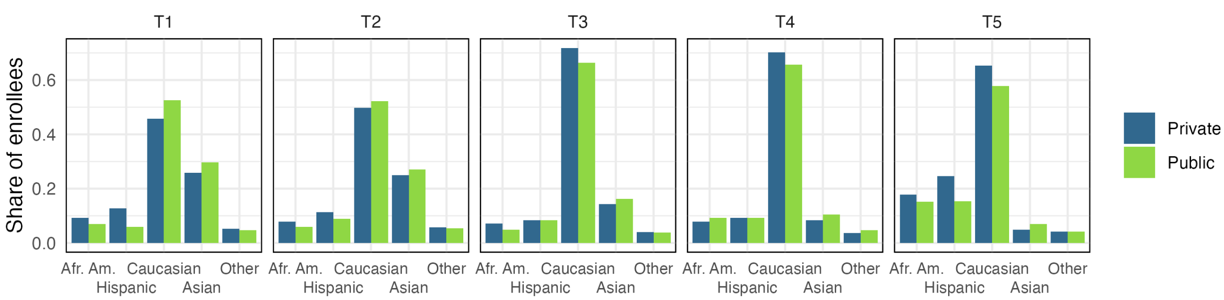

Figure 13: Share of enrollees in the data by race and tier, public vs. private colleges, aggregated across all years 2012–2020. The Native American category (accounting for less than 1% of all enrollees) is not included.

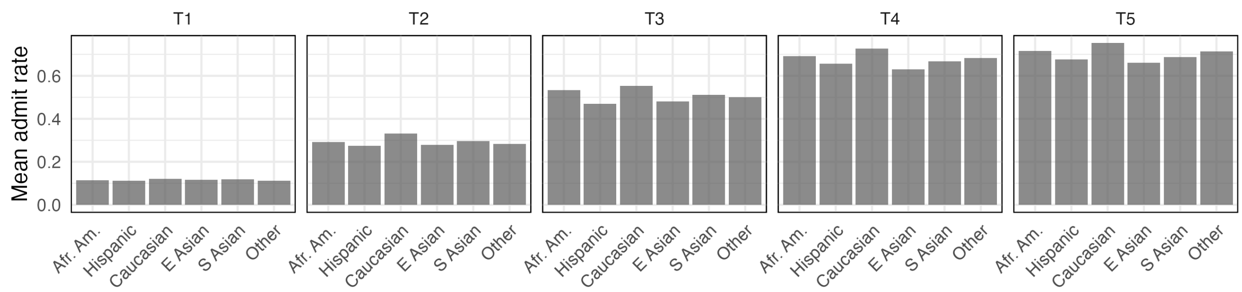

Figure 14: The application-weighted mean admit rate (IPEDS) of colleges applied to in the data, by race and tier.

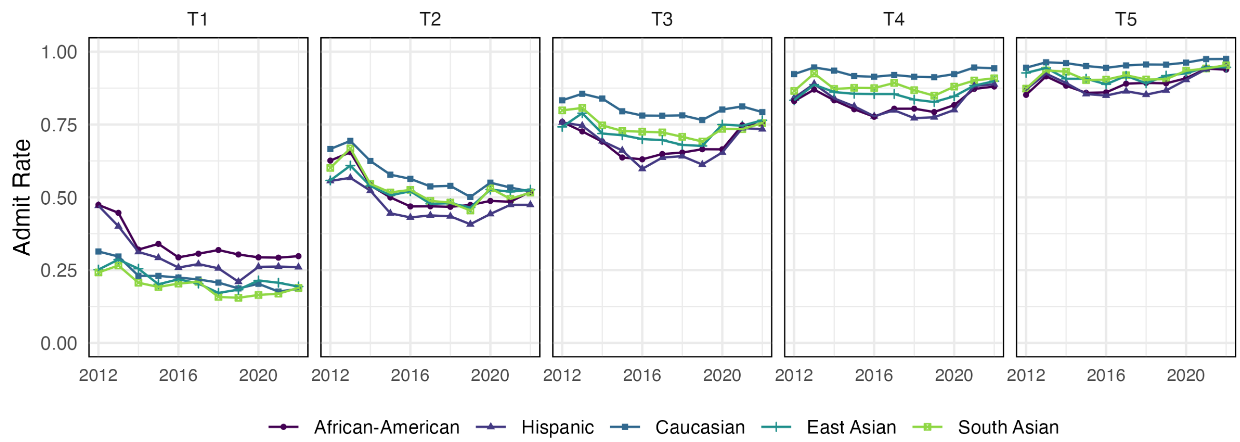

Figure 15: Admit rate by race, tier, and year, 2012–2022 (data). (Only applications with outcomes reported are counted in the denominator.)

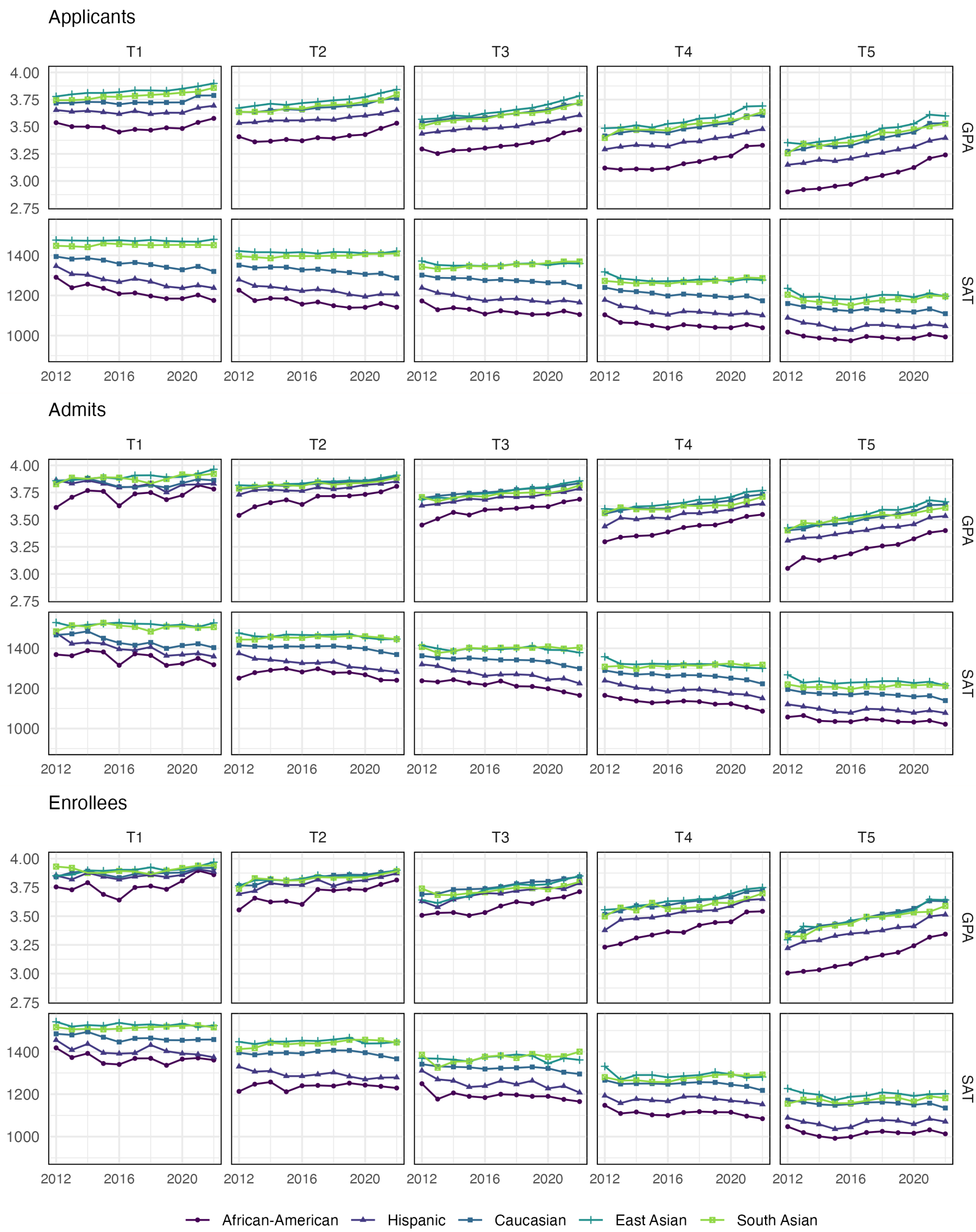

Figure 16: Mean GPA (unweighted) and SAT-equivalent score of applicants, admits, and enrollees, by race, tier, and year, 2012–2022 (data).

Median SAT	Colleges
1400–1600 (Tier 1)	Northeastern U., Stanford U., Cornell U., Harvard U., U. Pennsylvania, Columbia U. in the City of New York, Northwestern U., Vanderbilt U., Duke U., Brown U., Yale U., U. Chicago, Washington U. in St Louis, Princeton U., Georgia Institute of Technology, Johns Hopkins U., Carnegie Mellon U., Dartmouth College, Georgetown U., Tufts U., Massachusetts Institute of Technology, U. Notre Dame, Rice U.
1300–1399 (Tier 2)	U. California–Los Angeles, U. California–Berkeley, U. California–San Diego, New York U., Boston U., U. Southern California, U. Michigan–Ann Arbor, U. Minnesota–Twin Cities, U. Illinois Urbana–Champaign, U. North Carolina at Chapel Hill, U. Virginia, Binghamton U., Boston College, U. Maryland–College Park, Tulane U. Louisiana, Case Western Reserve U., Emory U., U. Rochester, Rensselaer Polytechnic Institute, William & Mary, Santa Clara U.
1200–1299 (Tier 3)	U. California–Santa Barbara, U. California–Davis, California Polytechnic State U–San Luis Obispo, Purdue U., Fordham U., U. Texas at Austin, Ohio State U., U. Massachusetts–Amherst, U. Washington–Seattle Campus, Rutgers U–New Brunswick, U. Connecticut, Stony Brook U., U. Miami, Baylor U., U. Pittsburgh–Pittsburgh Campus, U. Florida, Drexel U., U. Wisconsin–Madison, Miami U–Oxford, Clemson U., Virginia Polytechnic Institute and State U., U. Georgia, North Carolina State U., Raleigh, CUNY Bernard M Baruch College, George Washington U., Rochester Institute of Technology, American U., Villanova U., United States Naval Academy, U. Denver, United States Military Academy, U. San Diego, Saint Louis U., Brigham Young U–Provo, Loyola Marymount U.
1100–1199 (Tier 4)	U. California–Irvine, San Diego State U., Pennsylvania State U., U. California–Santa Cruz, U. Central Florida, U. Arizona, Michigan State U., Indiana U–Bloomington, Syracuse U., Texas A & M U–College Station, U. Colorado Boulder, U. South Florida, Florida State U., Temple U., CUNY Hunter College, Hofstra U., U. Iowa, U. South Carolina–Columbia, Arizona State U., U. Vermont, U. Delaware, U., Buffalo, Texas Tech U., U. Oregon, U. Missouri–Columbia, Loyola U. Chicago, James Madison U., U. Arkansas, Marquette U., CUNY City College, CUNY Brooklyn College, DePaul U., Auburn U., U. New Hampshire, Iowa State U., George Mason U., U. Kentucky, Colorado State U–Fort Collins, Texas Christian U., CUNY Queens College, U. Houston, Louisiana State U., U. Cincinnati, U. Tennessee–Knoxville, U. Dayton, Ithaca College, U. Illinois Chicago, U. San Francisco, Howard U., U. Kansas, Montana State U., State U. New York at New Paltz, U. the Pacific, Seton Hall U., Oregon State U., Loyola U. Maryland, Chapman U.
400–1099 (Tier 5)	California State U–Long Beach, Liberty U., California State U–Fullerton, U. California–Riverside, U. Alabama, St. John's U–New York, San Francisco State U., Grand Canyon U., California State U–Northridge, California State Polytechnic U–Pomona, California State U–Los Angeles, San Jose State U., Northern Arizona U., California State U–Sacramento, California State U–Chico, Quinnipiac U., SUNY at Albany, Ball State U., Florida International U., Ohio U., Texas State U., U. Rhode Island, California State U–Fresno, Washington State U., U. California–Merced, U. Tampa, Central Michigan U., U. Mississippi, Western Carolina U., Coastal Carolina U., Northern Illinois U., Pace U., Grand Valley State U., CUNY New York City College of Technology, East Carolina U., U. Memphis, U. North Carolina at Charlotte, U. North Texas, Utah State U., California State U–Dominguez Hills, Florida Atlantic U., Kent State U., U. Texas at San Antonio, California State U–Monterey Bay, West Virginia U., Sonoma State U., Western Michigan U., U. Akron, U. Hartford, Virginia Commonwealth U., California State U–East Bay, Bowling Green State U., Houston Baptist U., Eastern Michigan U., Kennesaw State U., CUNY Lehman College, Georgia State U., CUNY York College, California State U–San Bernardino, California State U–San Marcos, SUNY Buffalo State, Florida Gulf Coast U., Indiana U–Purdue University–Indianapolis, CUNY John Jay College of Criminal Justice

Table 9: Subset of colleges in each tier with the largest application volume in 2015 (IPEDS).

A.2 Comparison with External Data Sources

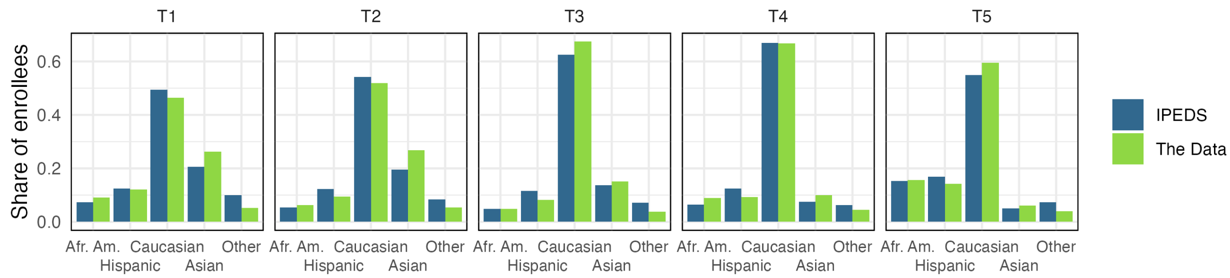

Figure 17: Share of enrollees by race and tier, IPEDS vs. Data, aggregated across all years 2012–2020. We fold the "Pacific Islander" IPEDS category into Asian, and the categories "Two or more races" and "Unknown" into Other. The Native American category (accounting for less than 1% of all enrollees) and the IPEDS category "Nonresident Alien" are not included.

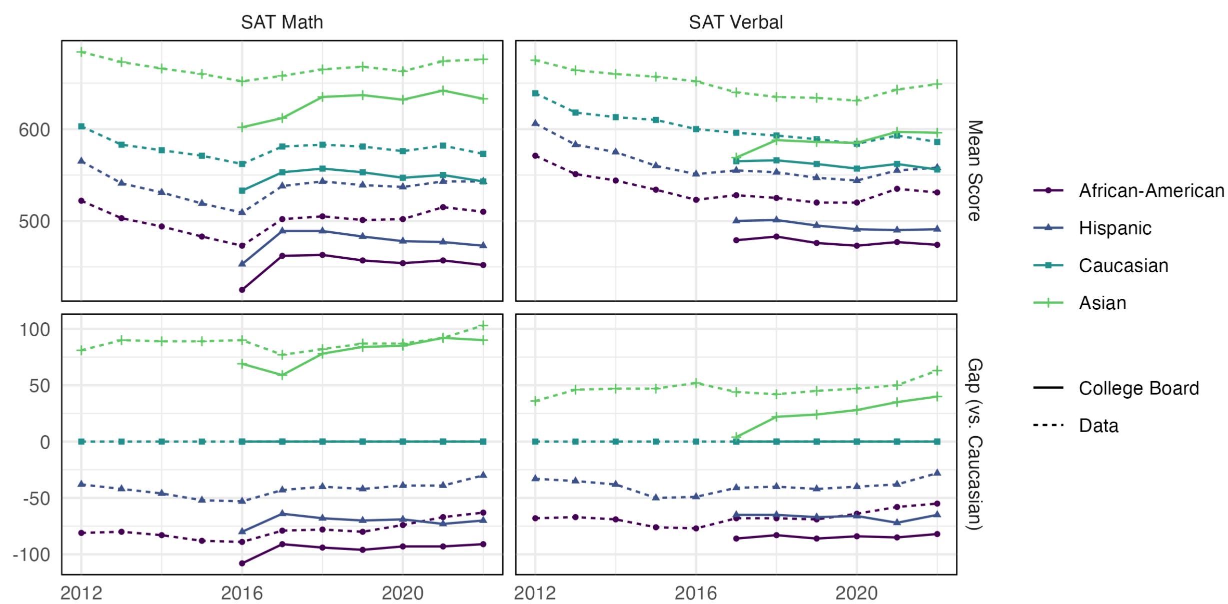

Figure 18: Mean SAT Math and Verbal scores (and the "gap" compared to Caucasians) by race and year, data versus College Board. The College Board numbers are obtained from the Suite of Assessments reports, which provide average scores by race for all test takers from 2016 to 2022 (see https://reports.collegeboard.org/sat-suite-program-results/data-archive, accessed August 28, 2024). We note that the higher scores observed in our data may reflect, in part, that users of the platform have greater intent to apply to college compared to the general test-taking population, which includes students in states where the SAT is a graduation requirement. Also note that we exclude the 2016 Verbal score from the College Board, as it is split into Critical Reading and Writing subsections preceding the redesign of the exam.

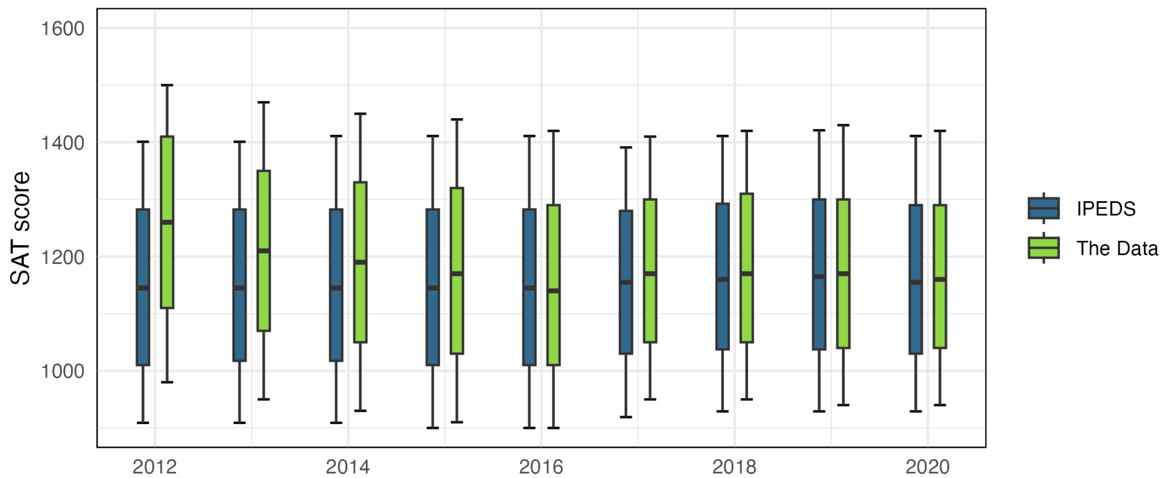

Figure 19: Distributional comparison of SAT scores among newly enrolled students in four-year colleges (IPEDS; estimated) and students in the data who reporting applying to a four-year college. Each "box" marks the 25th, 50th, and 75th percentiles; the "whiskers" mark the 10th and 90th percentiles.

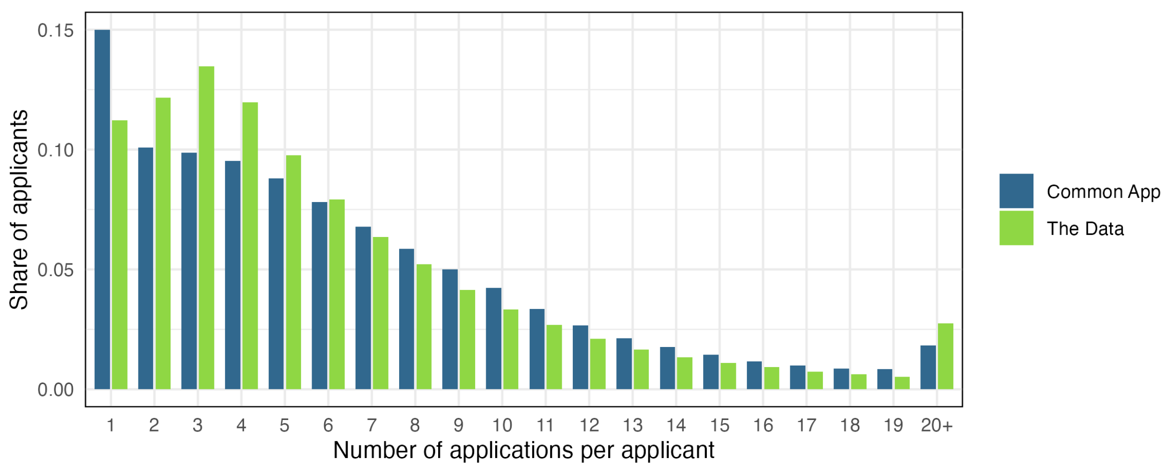

Figure 20: Distribution of the number of 4-year college applications per applicant. The Common App numbers are from Figure 1 in Kim et al. (2022). On the higher end, the Common App limits each user to a maximum of 20 applications (see `https://membersupport.commonapp.org/membersupport/s/article/What-is-the-maximum-number-of-colleges-to-which-a-First-year-applicant-can-submit-a-Common-App`, accessed August 28, 2024). On the lower end, students who apply to less selective colleges submit fewer applications on average (see Figure 12), and such colleges are less likely to use the Common App to manage their admissions.

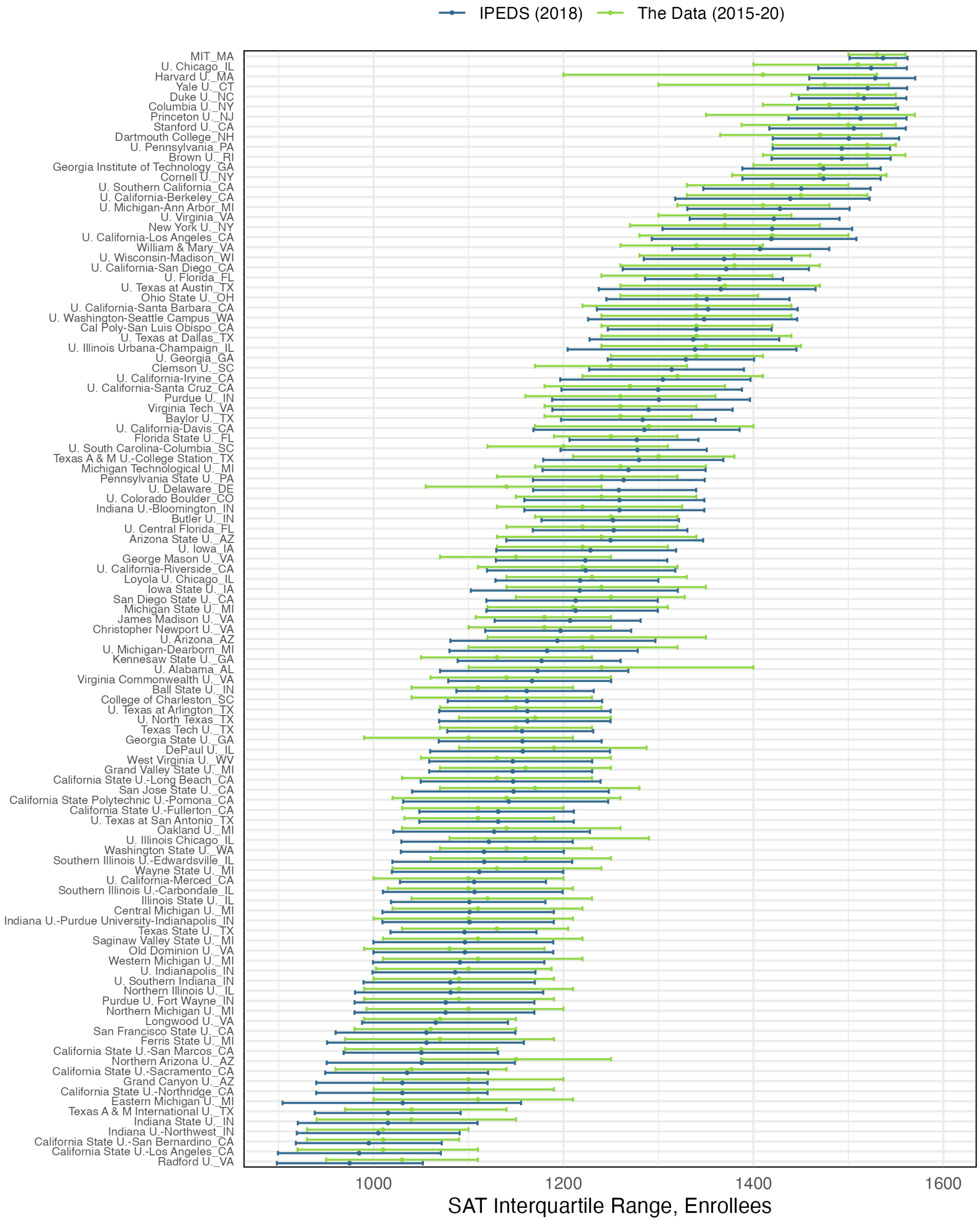

Figure 21: Comparison of SAT interquartile range of enrollees as estimated from IPEDS and as reported in the data (applicants who reported "accepted and attending"), for a selection of colleges with the most enrollees in the data. The estimation procedure is as described in Appendix D.2.

Appendix B Model Evaluation and Robustness Checks

Section B.1 presents additional tables and figures pertaining to the model described in Sections 3 and 4. Section B.2 provides the coefficients obtained from a year-by-year fit, including the 2021 and 2022 admission cycles. Section B.3 presents results given several alternative model specifications which nest the main model. Section B.4 considers a restricted sample which excludes applications to colleges in states in which affirmative action was already banned. Finally, Section B.5 restricts the sample to applicants who reported admissions outcomes to *all* 4-year colleges in their respective portfolios.

B.1 Main Model

We first provide additional details on the main model described in Section 3, from which the results in Section 4 are derived. Table 10 provides the model coefficients and standard errors, mirroring the results displayed in Figure 3. Figure 22 displays the estimated college-year cutoffs for a representative college from each tier. Figure 23 compares total unique admits and enrollees under each policy simulation to that under the status quo. Figures 24 and 25 show the degree to which the model's predicted admission probabilities are calibrated to the observed admission outcomes, as a function of SAT and GPA.

		T1		T2		T3		T4		T5	
Class Rank	05	–0.592	***	–0.489	***	–0.058		0.283	**	0.309	***
		(0.036)		(0.032)		(0.055)		(0.089)		(0.068)	
	10	–0.907	***	–0.841	***	–0.473	***	0.038		0.408	***
		(0.043)		(0.03)		(0.051)		(0.083)		(0.052)	
	25	–1.085	***	–1.221	***	–0.953	***	–0.214	**	0.535	***
		(0.053)		(0.031)		(0.049)		(0.081)		(0.043)	
	50	–0.523	***	–1.305	***	–1.344	***	–0.473	***	0.493	***
		(0.089)		(0.041)		(0.052)		(0.081)		(0.039)	
	>50	–0.415		–1.497	***	–1.389	***	–0.967	***	–0.216	***
		(0.231)		(0.119)		(0.087)		(0.091)		(0.051)	
	NR	–0.659	***	–0.878	***	–0.925	***	–0.442	***	0.312	***
		(0.043)		(0.025)		(0.047)		(0.079)		(0.036)	
Academics	GPA	–0.123	***	0.89	***	1.256	***	1.477	***	1.341	***
		(0.018)		(0.008)		(0.017)		(0.019)		(0.015)	
	SAT	0.2	**	0.894	***	1.076	***	1.647	***	1.347	***
		(0.062)		(0.013)		(0.015)		(0.017)		(0.013)	
	GPA:SAT	1.026	***	0.736	***	0.474	***	0.191	***	–0.037	**
		(0.014)		(0.014)		(0.015)		(0.015)		(0.011)	
Race	Afr. Am.	1.63	***	1.192	***	0.447	***	0.136	***	–0.17	***
		(0.066)		(0.033)		(0.032)		(0.027)		(0.028)	
	Hispanic	1.147	***	0.699	***	0.405	***	0.173	***	0.004	
		(0.054)		(0.027)		(0.027)		(0.028)		(0.03)	
	E Asian	–0.06		–0.051		–0.032		–0.133	***	–0.026	
		(0.046)		(0.027)		(0.037)		(0.039)		(0.051)	
	S Asian	–0.294	***	–0.064	*	–0.083	*	–0.093	*	0.019	
		(0.056)		(0.03)		(0.033)		(0.041)		(0.054)	
	Other	0.391	***	0.166	***	0.057		–0.032		–0.122	**
		(0.065)		(0.035)		(0.037)		(0.038)		(0.042)	
Gender	M	0.072		–0.088	***	–0.212	***	–0.008		–0.023	
		(0.048)		(0.019)		(0.018)		(0.019)		(0.021)	
In State	TRUE	0.291	***	0.046		–0.174	***	–0.022		0.042	
		(0.042)		(0.039)		(0.041)		(0.04)		(0.056)	
In State x Public	TRUE	0.889	***	0.868	***	0.873	***	–0.22	***	0.122	
		(0.098)		(0.044)		(0.045)		(0.048)		(0.069)	
δ		2.272	***	1.38	***	1.296	***	1.724	***	2.014	***
		(0.012)		(0.01)		(0.007)		(0.033)		(0.016)	

Table 10: Estimated logit coefficients for the pooled-year model (year entering college between 2012 and 2020) with college-year specific cutoffs. Standard errors in parentheses; stars correspond to significance levels .001 (***), .01 (**), .05 (*). Class Rank dummies correspond to "Top 5%", "Top 10%", etc. and are defined relative to the "Top 1%" reference class. Race dummies are defined relative to the Caucasian reference class.

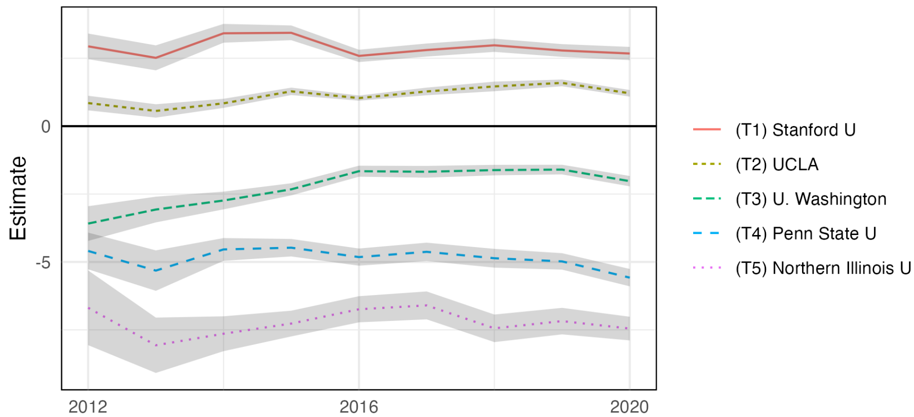

Figure 22: Estimated cutoffs $\{c_j\}$ with 95% confidence intervals, for a representative college from each tier. Estimates are relative to a reference college-year (University of Michigan–Ann Arbor, 2016).

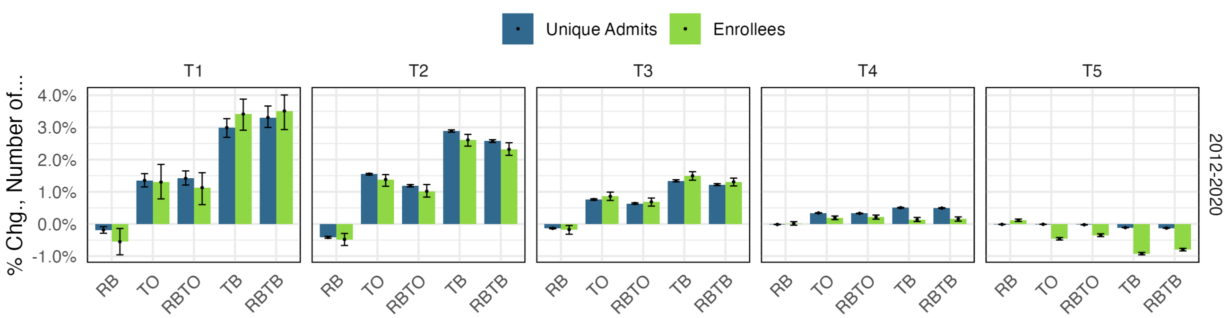

Figure 23: The relative difference in the expected number of unique admits and enrollees, by tier, under each policy studied in Section 4.2 in comparison to the *status quo*. By assumption, each college adjusts its cutoff such that the expected number of admits remains the same under each policy. The expected number of enrollees is estimated using the model described in Appendix D.

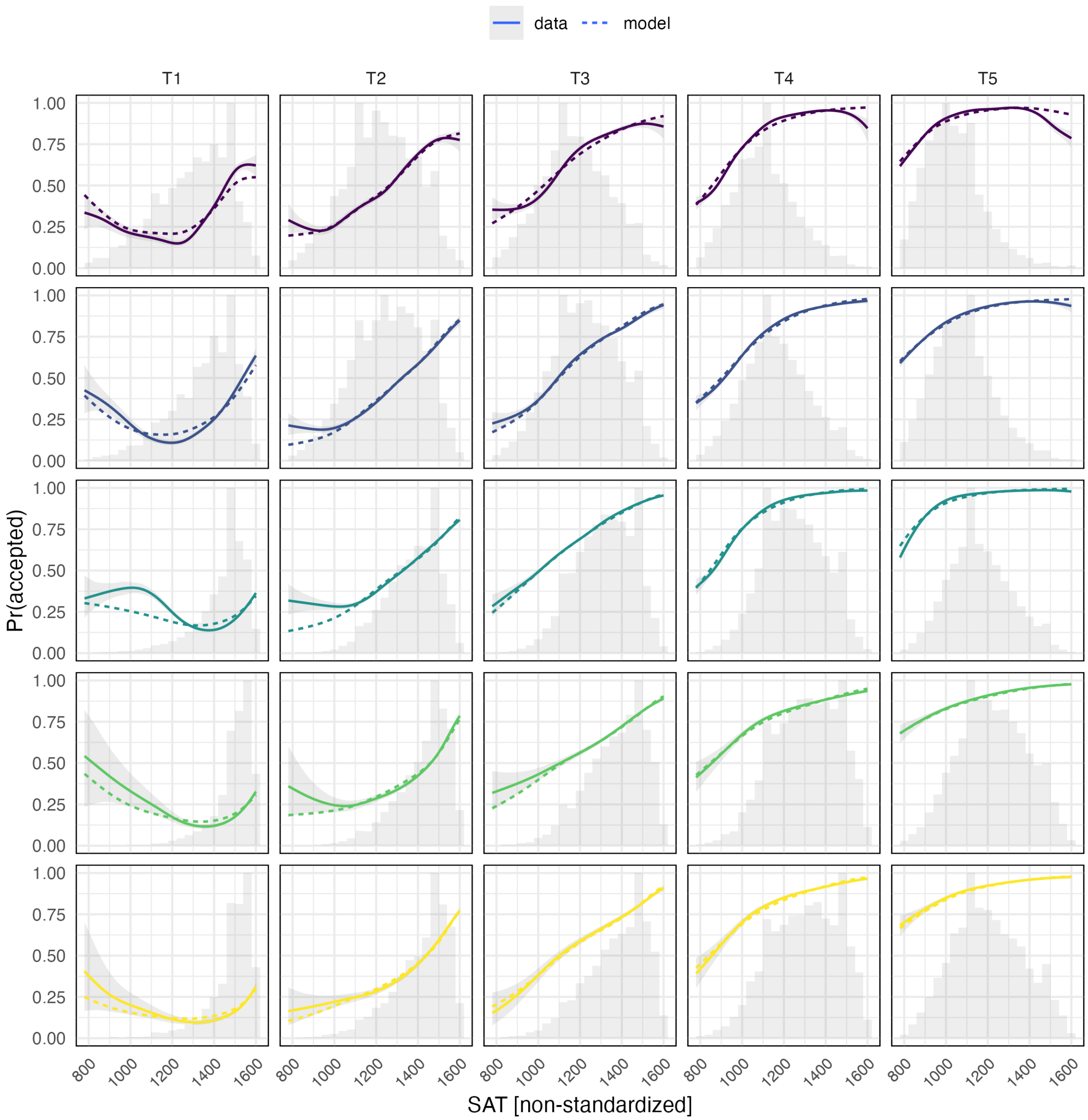

Figure 24: Calibration plot showing the level of agreement between observed outcomes and predicted admission probabilities, as a function of applicant SAT score on the original 1600-point scale. Each curve is a nonparametric smooth of actual admission outcomes y_{ij} (solid) or the predicted probabilities under our model $\hat{p}_{ij}$ (dashed), for all applications of the specified race and tier. In each case, the smoother is a logistic generalized additive model (GAM) with cubic spline in SAT as the linear predictor. The curves reflect the empirical distribution of non-SAT covariates ($X_{-\mathrm{SAT}}$) at different values of SAT; i.e., the curves estimate $\mathbb{E}[Y \mid \mathrm{SAT}] = \mathbb{E}_{X_{-\mathrm{SAT}}}[\mathbb{E}[Y \mid \mathrm{SAT}, X_{-\mathrm{SAT}}]]$. Note that $X_{-\mathrm{SAT}}$ includes the college applied to, such that higher admission probability for lower SAT scores reflects the fact that lower scorers tend to apply to less-selective colleges within each tier. Gray histograms show the distribution of SAT score by race and tier, with the modal bin frequency normalized to "1" on the vertical axis scale. See Figure 25 for the analogous calibration plot for GPA.

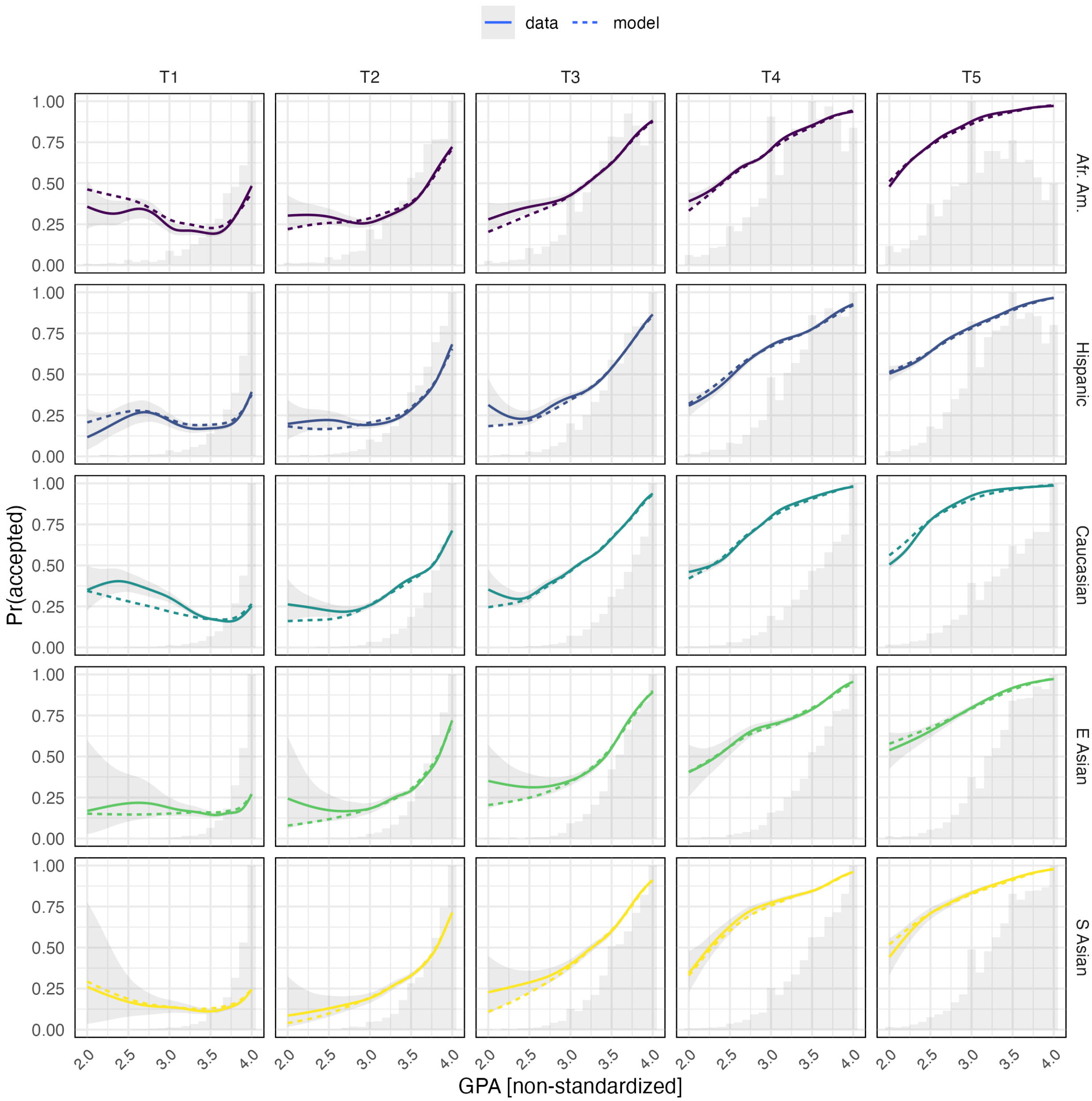

Figure 25: Calibration plot showing the level of agreement between observed outcomes and predicted admission probabilities as a function of applicant GPA (on the original 4.0 scale). Each curve is a nonparametric smooth of actual admission outcomes y_{ij} (solid) or the predicted probabilities under our model $\hat{p}_{ij}$ (dashed), for all applications of the specified race and tier. In each case, the smoother is a logistic generalized additive model (GAM) with cubic spline in SAT as the linear predictor. The curves reflect the empirical distribution of non-GPA covariates ($X_{-\mathrm{GPA}}$) at different values of GPA; i.e., the curves estimate $\mathbb{E}[Y \mid \mathrm{GPA}] = \mathbb{E}_{X_{-\mathrm{GPA}}}[\mathbb{E}[Y \mid \mathrm{GPA}, X_{-\mathrm{GPA}}]]$. Note that $X_{-\mathrm{GPA}}$ includes the college applied to, such that higher admission probability for lower GPA reflects the fact that lower-GPA applicants tend to apply to less-selective colleges within each tier. Gray histograms show the distribution of GPA by race and tier, with the modal bin frequency normalized to "1" on the vertical axis scale. See Figure 24 for the analogous calibration plot for SAT.

B.2 Year-by-Year Models

In this section, we present the coefficient estimates for year-by-year fits of our main model specification in Section 3. For each year from 2012 to 2022, we simply fit a separate model using the subset of data corresponding to that specific admission cycle.

We highlight in red the 2021 and 2022 cycles, which we excluded from our main (pooled-year) model due to changes in admission practices in response to the COVID-19 pandemic. Figure 28 shows that the magnitudes of the SAT coefficients (outside of T1) in 2021-22 are approximately half of those observed pre-2021, reflecting the widespread adoption of test-optional policies that de-emphasized the SAT as a determinant of admission. Note that the underlying applicant sample used for our estimation still only includes applicants who reported an SAT or ACT score to the platform, but we do not observe whether or not the applicants reported their test scores to the colleges in their portfolios.

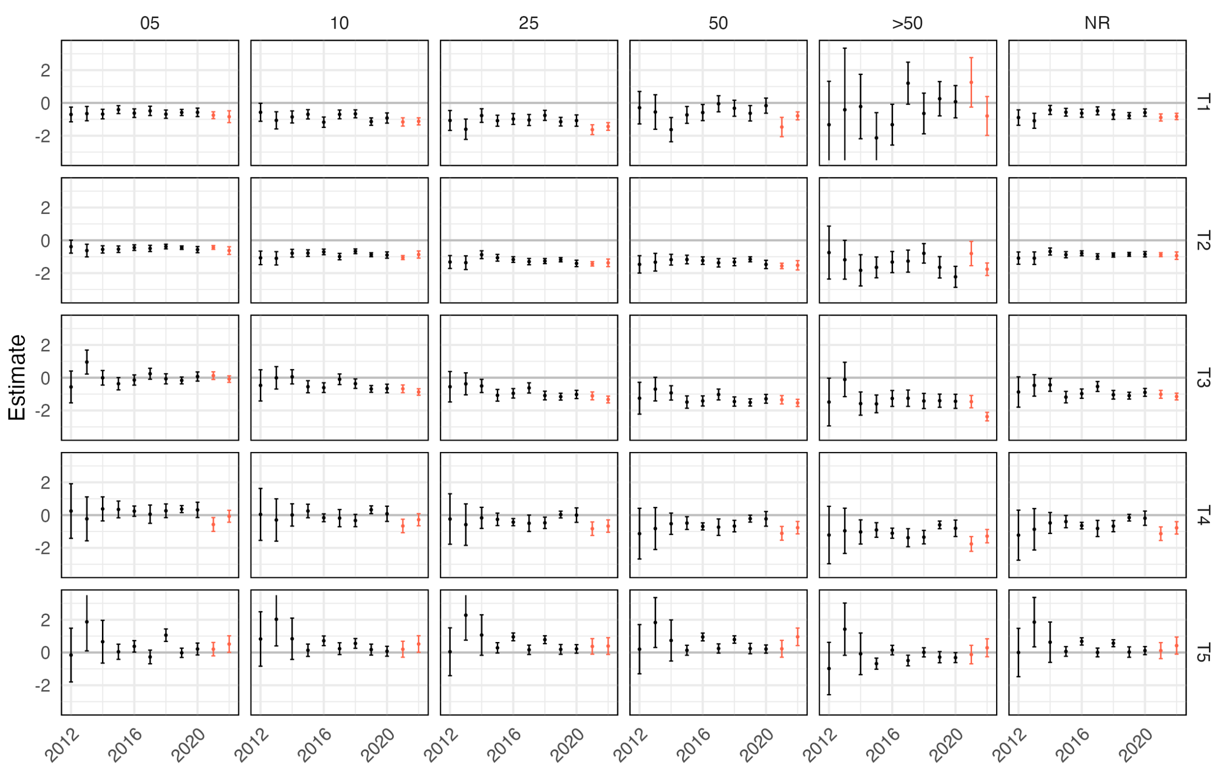

Figure 26: Estimated coefficients on CLASS_RANK terms with 95% confidence intervals. Estimates are relative to the "Top 1%" (`01`) reference class.

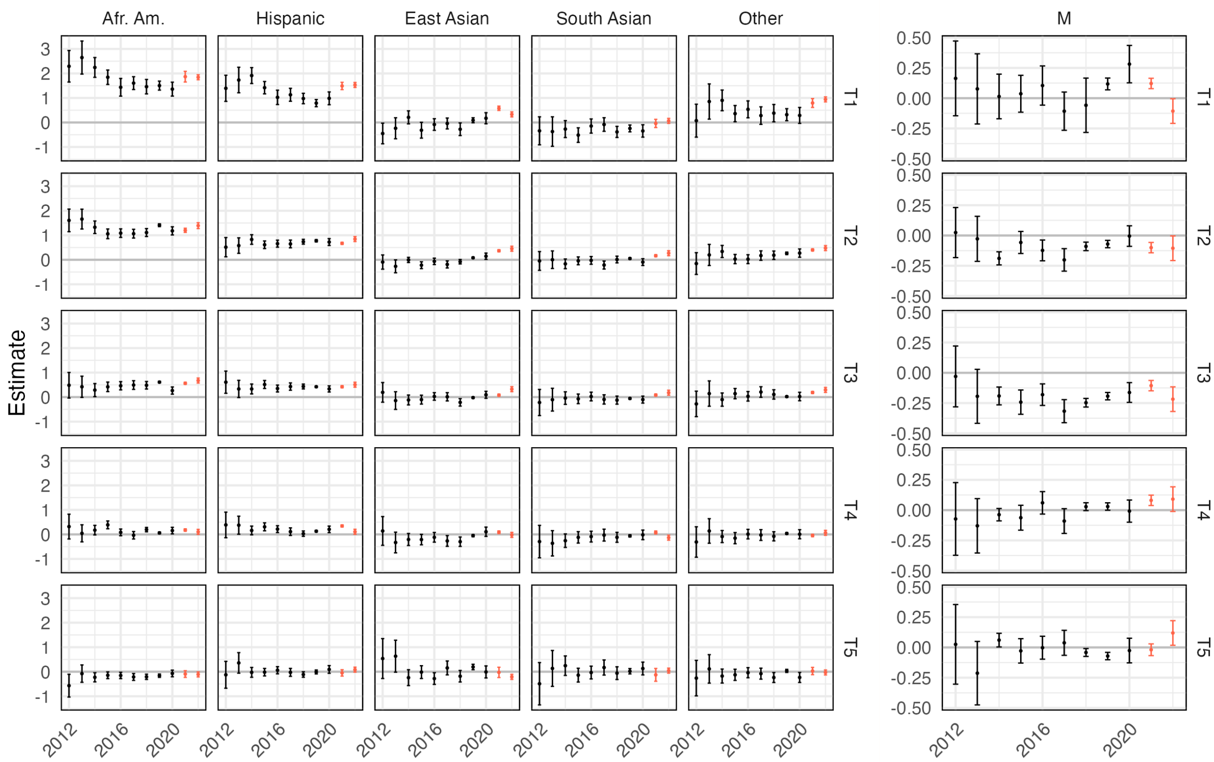

Figure 27: Estimated coefficients on RACE and GENDER terms with 95% confidence intervals. Race (gender) estimates are relative to the Caucasian (female) reference class.

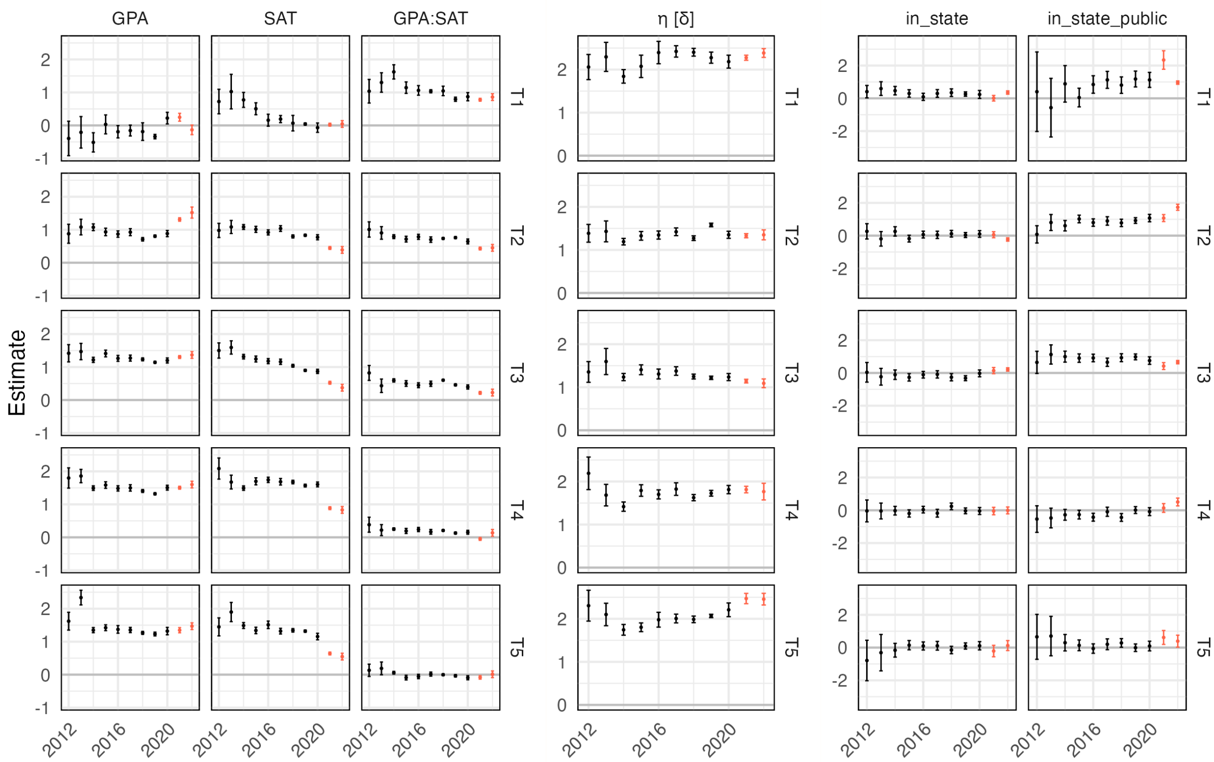

Figure 28: Left: Estimated coefficients on GPA, SAT, and their interaction, with 95% confidence intervals. GPA and SAT are standardized to have mean 0 and variance 1. Middle: Estimated coefficients $\delta_{T(j)}$ for the applicant-level random effect η_i. Right: Estimated coefficients on IN_STATE and IN_STATE:PUBLIC indicators.

B.3 Alternative Specifications

To allow more flexibility and non-linearity in the admission decisions, we consider in this section alternative specifications which nest the main model of Section 3. The first includes squared terms GPA^2 and SAT^2 in addition to the covariates in the main model; the second includes Race:SAT interactions; the third includes all second-order terms in GPA and SAT, as well as their interactions with race.

We first report model fit statistics for each model (Appendix B.3.1), followed by the estimated coefficients (Appendix B.3.2), and finally the results of our policy simulations (Appendix B.3.3). We find that the inclusion of the higher-order and interaction terms does not yield a material improvement in model fit. Moreover, the simulation results (i.e. the estimated impact of different admission policies on the racial composition of unique admits) are not particularly sensitive to the model specification, among those considered.

B.3.1 Model Fit

For the fit statistics presented below, we also include the null (intercept-only) model and a "fixed-effects only" model, the main model excluding the the applicant-level random effect η_i.

model	log.lik.	deviance	pseudo R^2	parameters	p-value (vs Main)
<null model>	-766578	1533156	0.000	1	-
Main excl. η_i	-423366	846733	0.441	5,527	-
Main	-399327	798655	0.479	5,538	-
Main + GPA^2 + SAT^2	-397257	794514	0.482	5,548	<1e-3
Main + Race:SAT	-398619	797238	0.480	5,588	<1e-3
Main + Race:GPA + Race:SAT (with 2nd-order terms)	-396014	792027	0.483	5,673	<1e-3

Table 11: Log-likelihood, deviance, pseudo R^2, parameter count, and p-value (of likelihood-ratio test against Main) for each model specification considered. The pseudo R^2 is defined as $1 - D_M/D_0$, where D_M and D_0 are the model deviance and null deviance, respectively.

model	T1	T2	T3	T4	T5
Main excl. η_i	0.732	0.834	0.864	0.900	0.870
Main	0.963	0.931	0.940	0.972	0.977
Main + GPA^2 + SAT^2	0.964	0.931	0.940	0.972	0.977
Main + Race:SAT	0.962	0.932	0.940	0.972	0.977
Main + Race:GPA + Race:SAT (with 2nd-order terms)	0.962	0.932	0.940	0.972	0.977

Table 12: In-sample AUROC for each model specification considered, by tier. The AUROC is computed using expected admission probabilities—i.e., $\hat{p}_{ij} = \mathbb{E}_{g_i}[p_{ij}(\eta_i)]$, where $\mathbb{E}_{g_i}[\cdot]$ denotes expectation over the posterior distribution of η_i—and varying the threshold t above which admission is predicted, i.e., $\hat{y}_{ij}(t) = \mathbb{1}\{\hat{p}_{ij} \geq t\}$.

B.3.2 Estimated Model Coefficients

Main + GPA^2 and SAT^2 Terms

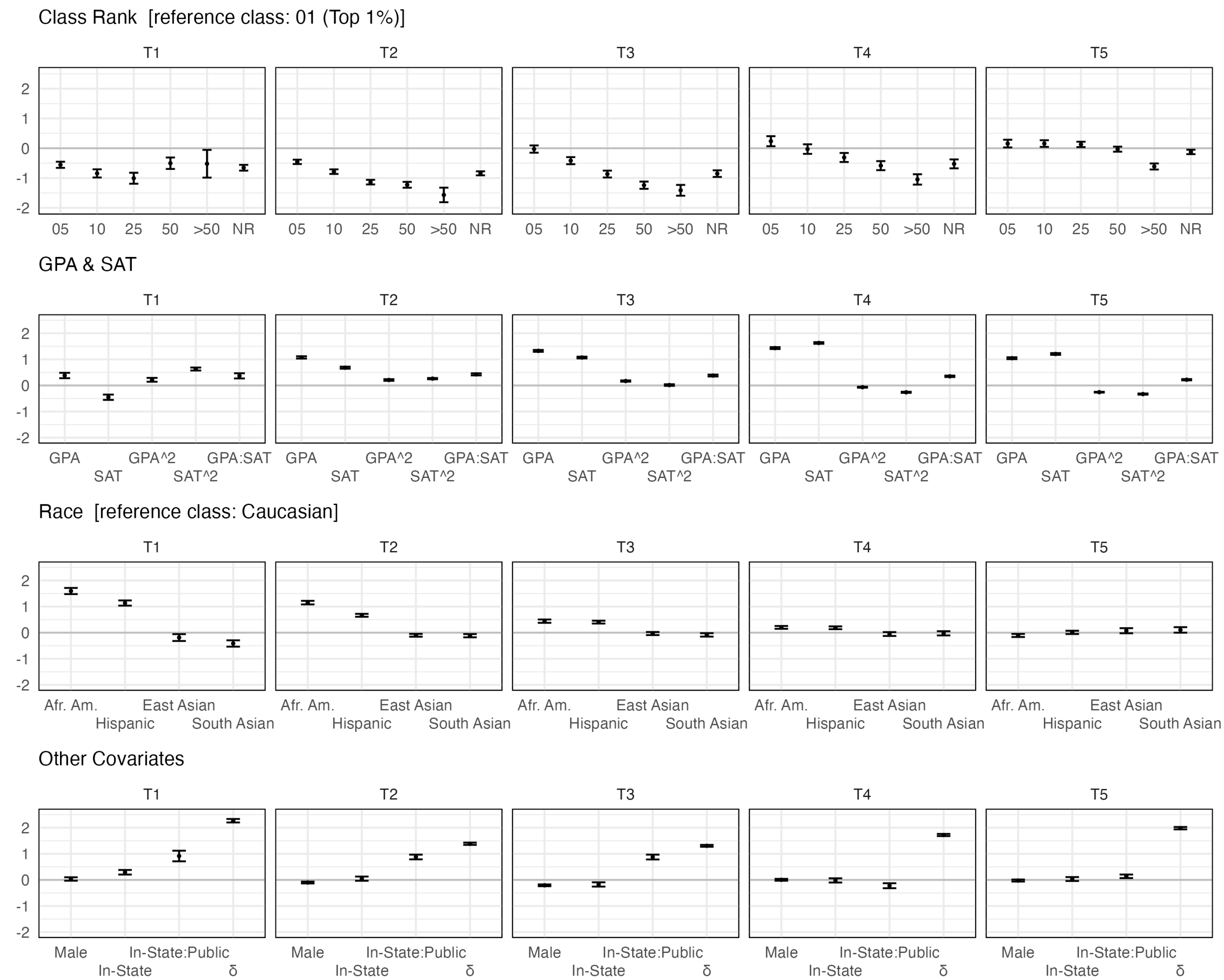

Figure 29: Estimated coefficients with 95% confidence intervals for model with all second-order terms in GPA, SAT. (Pooled model using years 2012–2020 with college-year cutoffs.)

Main + Race:SAT Interaction

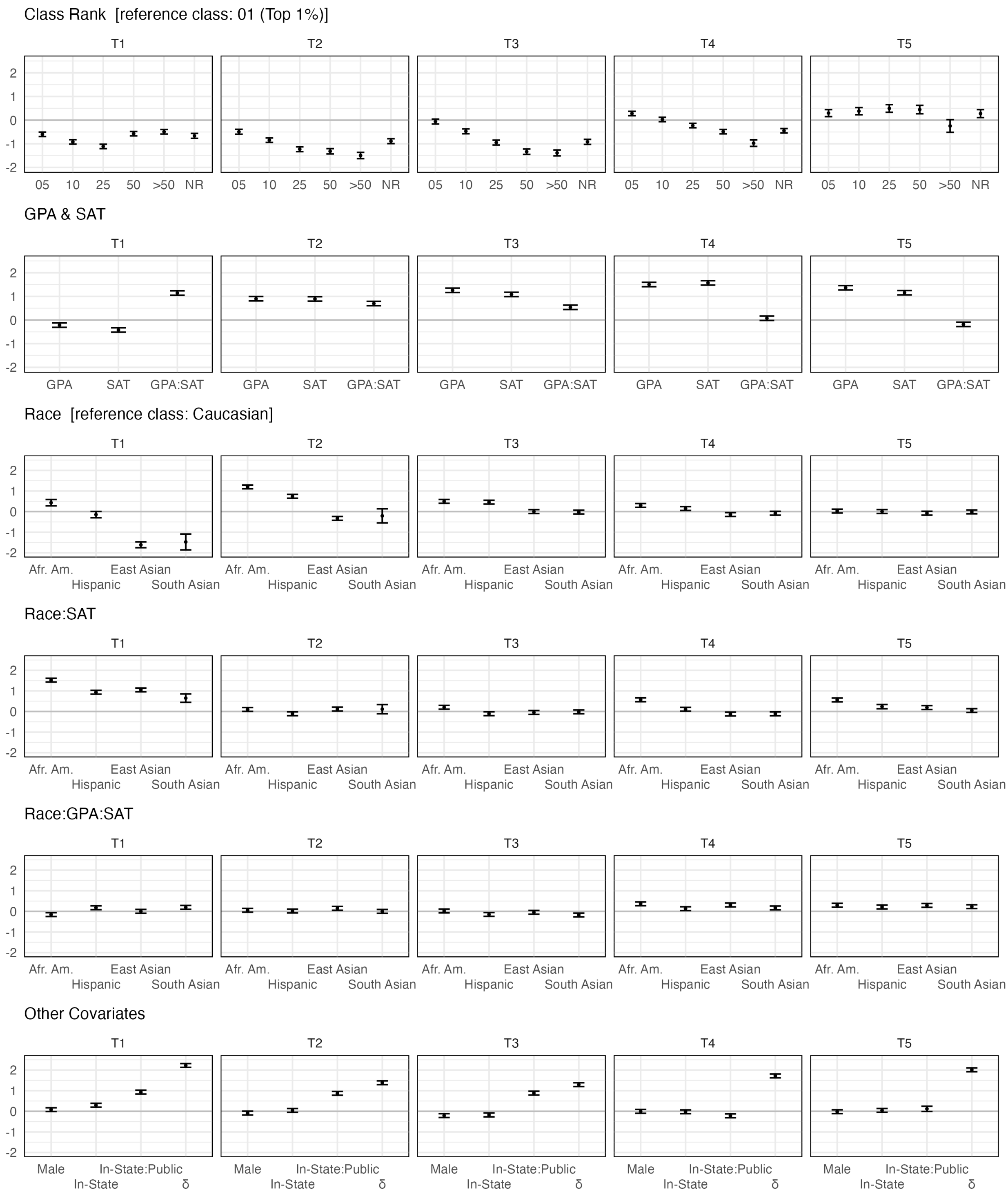

Figure 30: Estimated coefficients with 95% confidence intervals for model with Race:SAT and Race:GPA:SAT interactions. (Pooled model using years 2012–2020 with college-year cutoffs.)

Main + Race:GPA + Race:SAT (with 2nd-order terms)

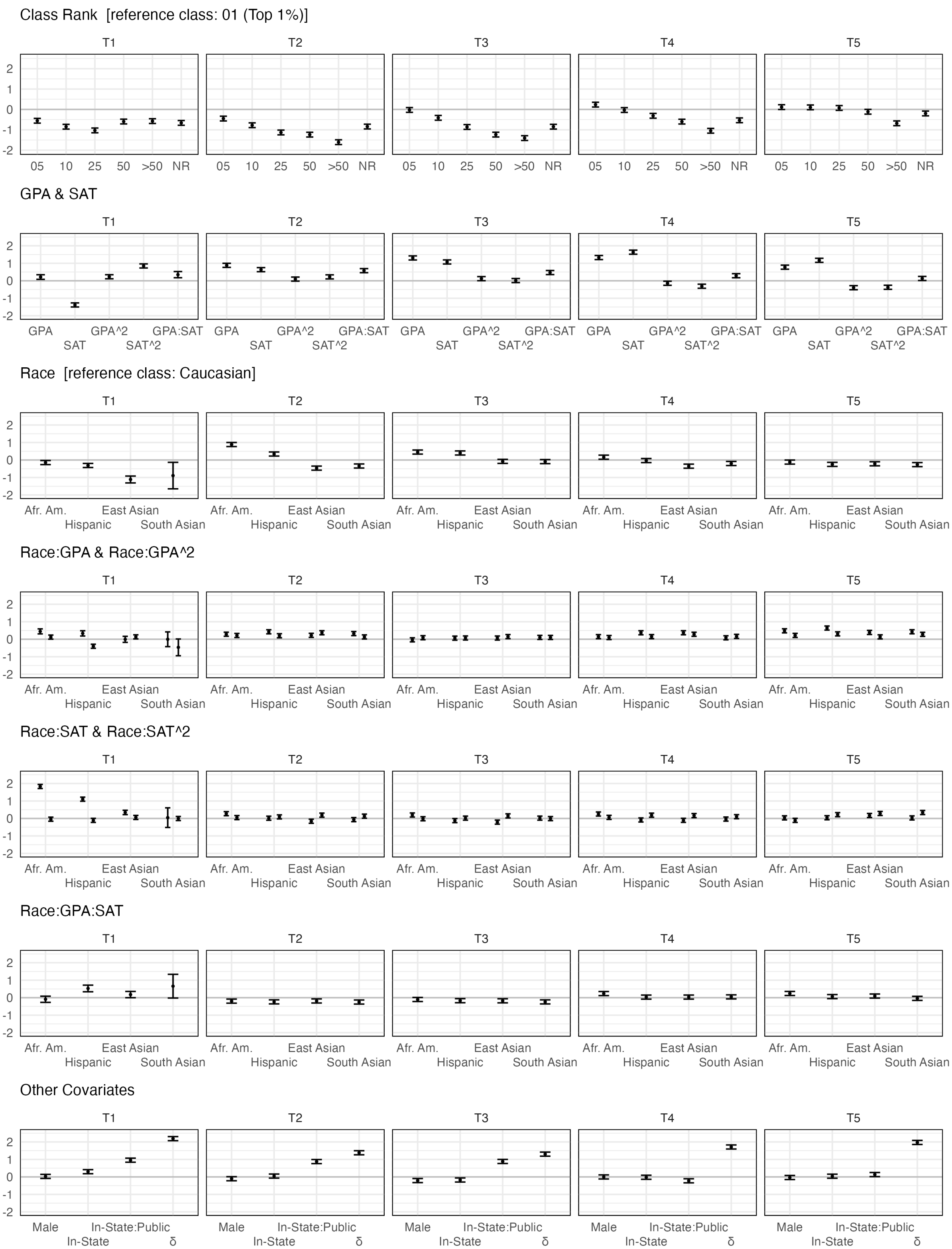

Figure 31: Estimated coefficients with 95% confidence intervals for model with all second-order terms in GPA, SAT, interacted with Race. (Pooled model using years 2012–2020 with college-year cutoffs.)

B.3.3 Policy Simulation Estimates

For each of the alternative model specifications considered in this section, we conduct the same set of policy simulations are as described in Section 4.2. The results are presented in Figure 32, which also includes the results under the main model for comparison.

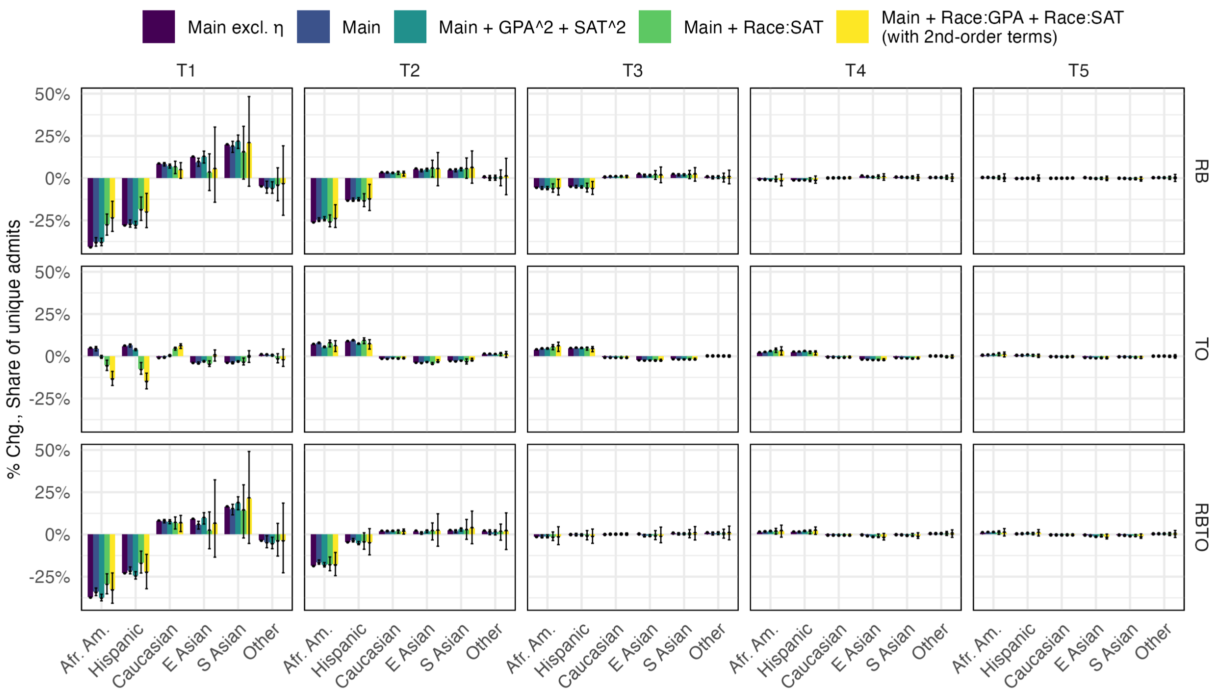

Figure 32: The simulated impact on race shares of unique admitted students in each tier under the "race-blind" (RB), "test-optional" (TO), and "race-blind test-optional" (RBTO) policies for each model specification.

Note that for the models with interactions between race and SAT, under the test-optional simulation, we eliminate the contribution of these interaction terms to the applicant's score (W) for the non-reporters. The reason is that these terms may be viewed as "contextualizing" an applicant's SAT score given his or her race. The positive coefficient on Race:SAT for URMs in Figure 30 and Figure 31 indicates that the higher a URM applicant's SAT score, the greater the "boost" in their systemic score (W) relative to a Caucasian applicant with the same SAT score. We contend that such contextualization is not valid when the applicant's score is imputed when the applicant chose to not submit her score, especially when the imputation itself does not take race into account, as we already assume. As a consequence, we see a decrease instead of increase in URM admissions to T1 under test-optional admissions.[49] Also note that this consideration does not affect the race-blind, test-optional (RBTO) simulation, as all coefficients on race are set to zero.

[49] A few colleges (including MIT and Dartmouth) state that reinstating standard test requirements *improves* the diversity of their campuses since the SAT scores are evaluated in context and allows students from disadvantaged background to demonstrate readiness for college. See https://mitadmissions.org/blogs/entry/we-are-reinstating-our-sat-act-requirement-for-future-admissions-cycles/ and https://president.dartmouth.edu/news/2024/02/reactivating-satact-requirement-dartmouth-undergraduate-admissions, accessed February 5, 2024.

B.4 Excluding Colleges Under State-Level Affirmative Action Bans

In this section, we exclude colleges from our analysis that are legally prohibited from considering race in their admission processes due to state-level affirmative action bans. This exclusion applies to public colleges in California, Washington, Florida, Michigan, Nebraska, Arizona, New Hampshire, and Oklahoma (since 2013), and to all colleges in Texas. The filtered sample, compared to the numbers in the middle column of Table 1, includes 1,173,156 applications to 4-year colleges (-28%); 430,020 applicants (-20%); and 1,237 unique colleges. The average number of applications per applicant is 5.15, and the average number of reported outcomes 2.73 per applicant.[50]

Pooling data from admission cycles from 2012 to 2020, Figure 33 displays the estimated coefficients for the main model specification. The coefficients are consistent with those estimated using data from all colleges as presented in Figure 3. Similarly, the outcomes of the policy simulations are also aligned, as illustrated in Figure 34.

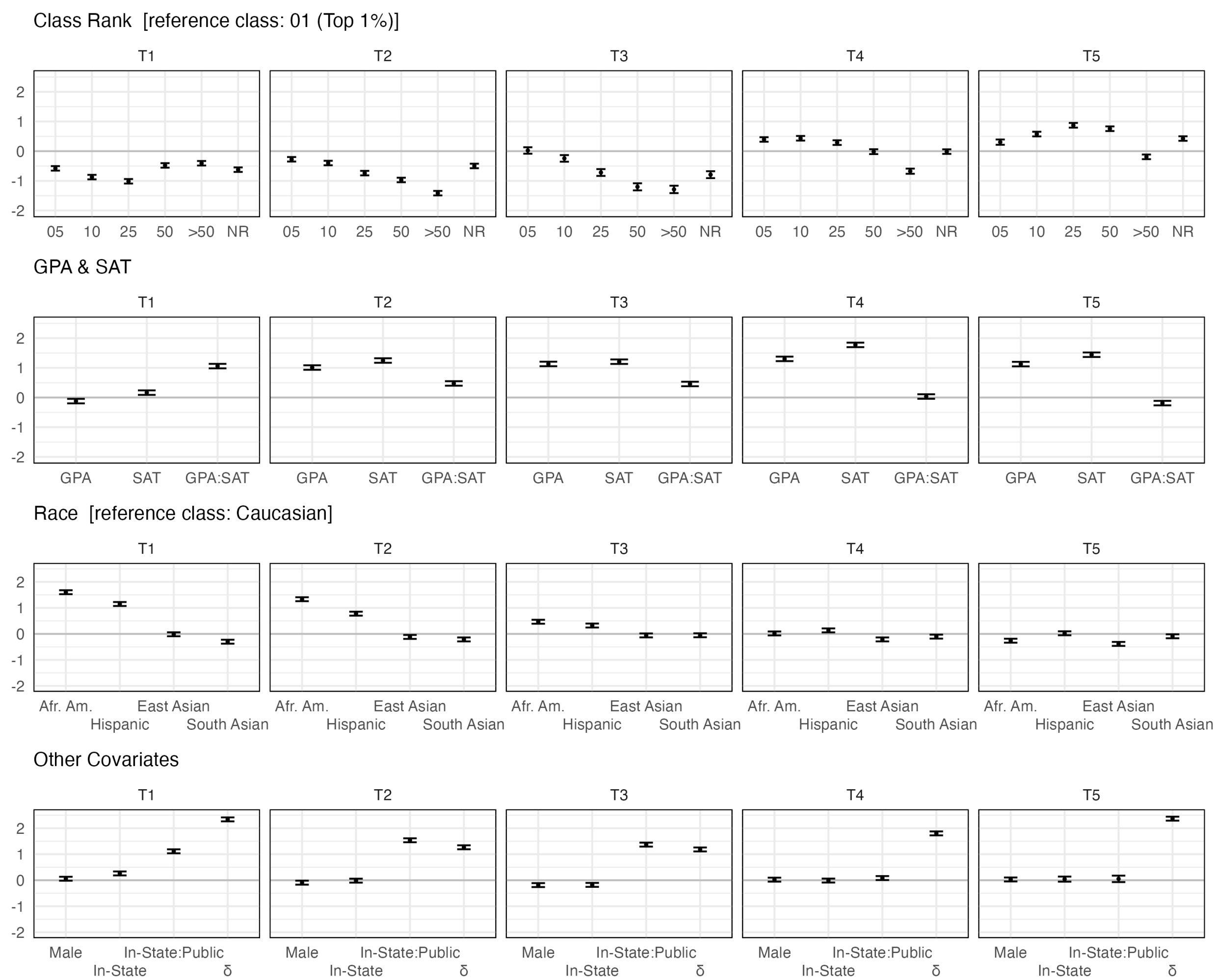

Figure 33: Estimated coefficients with 95% confidence intervals for the main model specification (pooled model using years 2012–2020 with college-year cutoffs), excluding colleges already affected by state-level affirmative action bans.

[50] At the tier level, the number of applications with reported outcomes decreases by 0% in T1; 41% in T2; 24% in T3; 29% in T4; and 31% in T5.

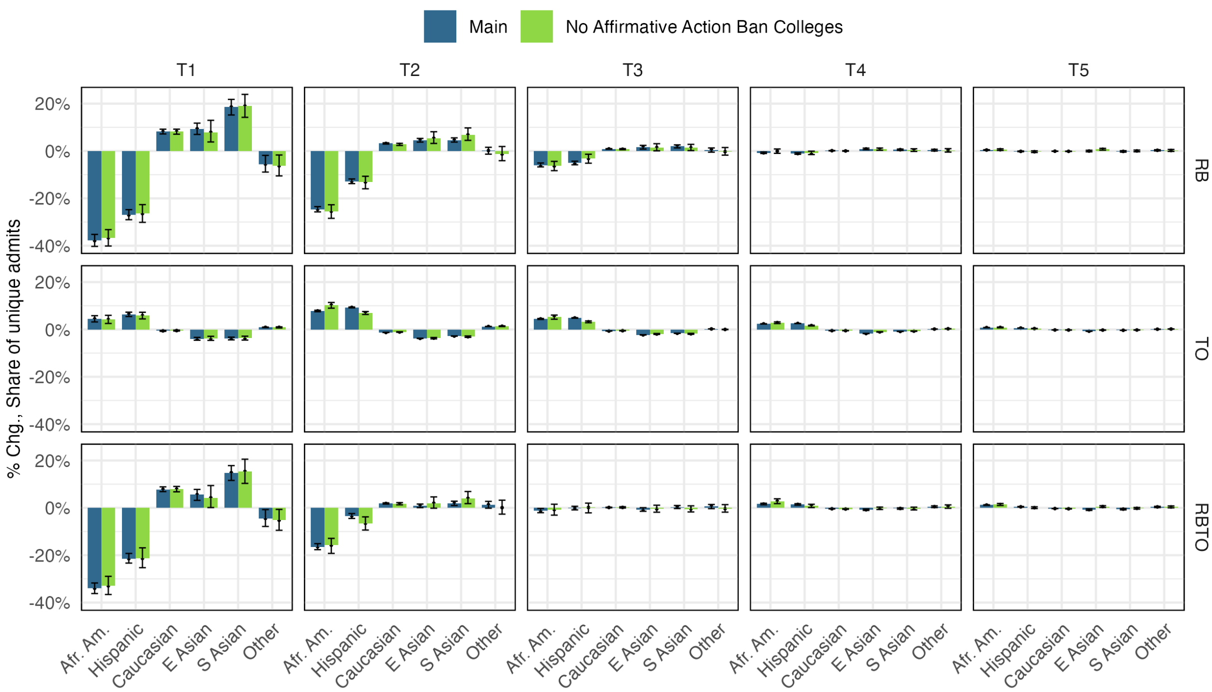

Figure 34: The simulated impact on race shares of unique admitted students in each tier under the "race-blind" (RB), "test-optional" (TO), and "race-blind test-optional" (RBTO) policies, for main model specification, excluding colleges already affected by state-level affirmative action bans.

B.5 Restricting to Applicants Who Reported All Admission Outcomes

In this section, we restrict our analysis to the sample of applicants who reported *all* admission outcomes to 4-year colleges (with tier assignment) in their respective portfolios. The filtered sample includes 683,785 applications to 4-year colleges; 192,425 applicants; and 1,308 unique colleges (see the rightmost column of Table 1). In this sample, applicants have 3.55 applications/reported outcomes on average. We do not observe qualitative differences in the estimated model coefficients (Figure 35) or policy simulation outcomes (Figure 36), in comparison to the results presented in Section 4 where all applicants with least one reported outcome are considered.

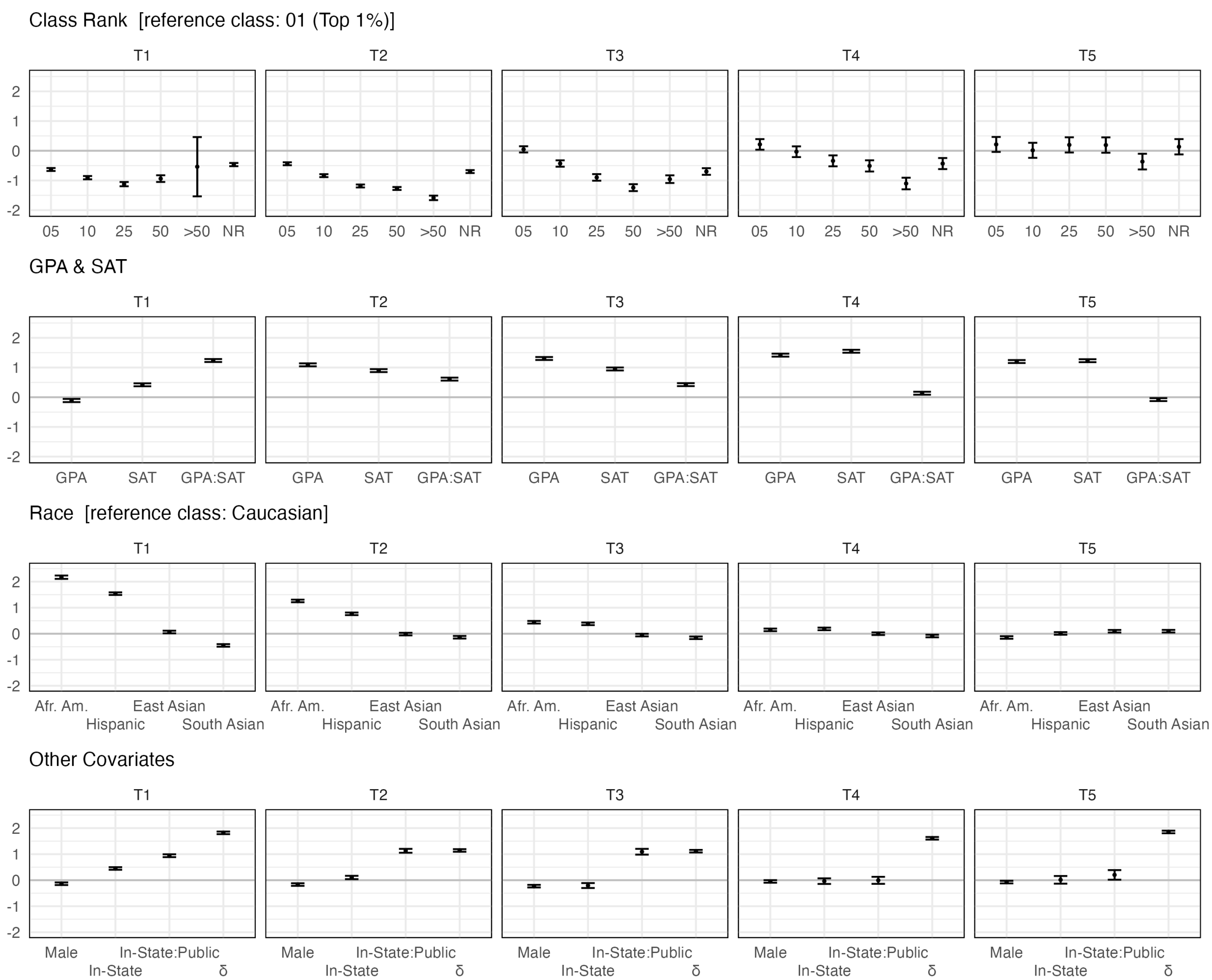

Figure 35: Estimated coefficients with 95% confidence intervals for main model specification, restricting sample to applicants who reported all admission outcomes. (Pooled model using years 2012–2020 with college-year cutoffs.)

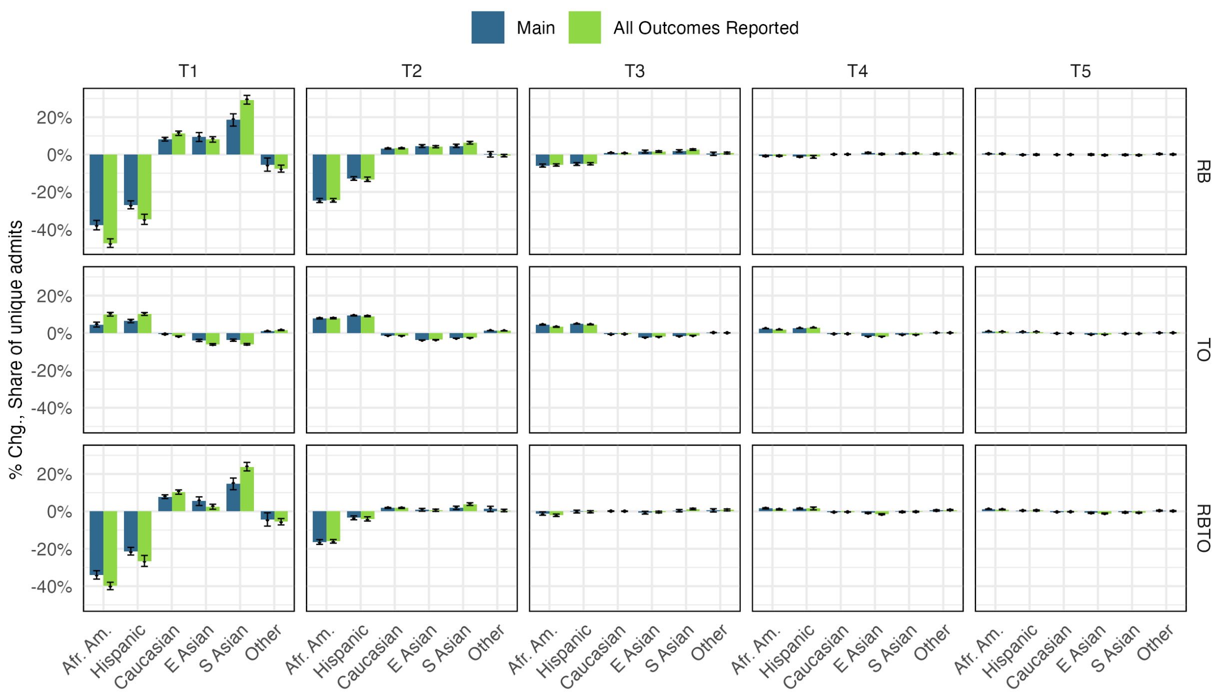

Figure 36: The simulated impact on race shares of unique admitted students in each tier under the "race-blind" (RB), "test-optional" (TO), and "race-blind test-optional" (RBTO) policies, for the restricted sample of applicants who reported *all* admission outcomes.

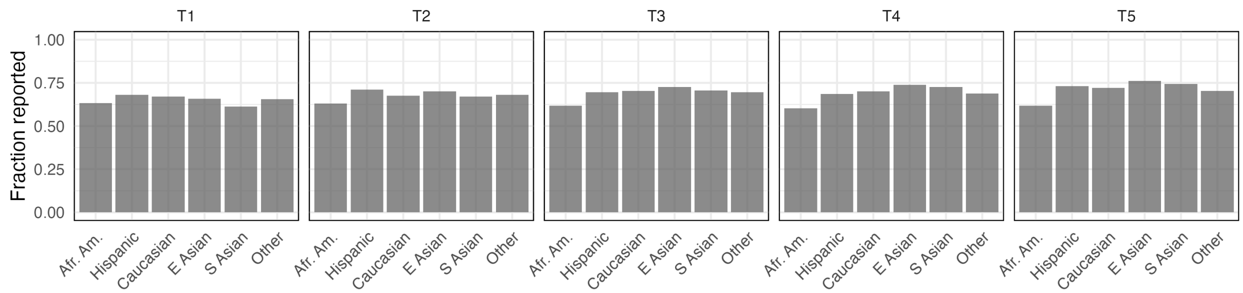

Figure 37: Fraction of applications which have an admission outcome reported—among those with either "Applied" status or with admission outcome reported—by tier and applicant's race.

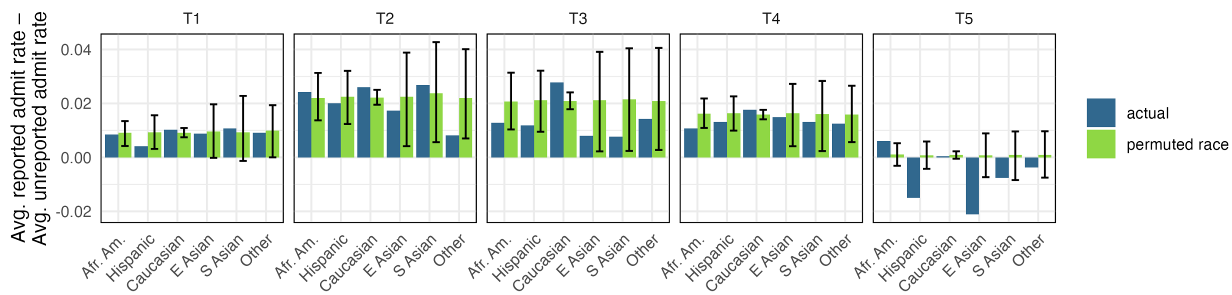

Figure 38: Difference between average IPEDS admit rate of applications with and without reported outcomes, by race and tier. Blue bars are based on observed applicant race, while green bars are based on permuted race.

Appendix C Reporting Bias

For our sample of nearly 550 thousand applicants who have reported at least one admissions outcome, roughly 30% of all applications do not have a reported outcome (in our dataset these colleges are listed as "Applied"). See Figure 37. One potential concern is that applicants might be more likely to report admissions than rejections. If this "reporting bias" is stronger or weaker for URM applicants, we may under- or over-estimate of the degree of racial preference in admissions. However, as discussed below, we find little evidence for differences in reporting behavior that would bias our results in this manner.

Since unreported outcomes are unobserved, we cannot directly quantify reporting bias as the difference between the likelihood of reporting an acceptance and that of reporting a rejection. Instead, we take as our assumption that unreported outcomes are more likely to be rejections at colleges with *lower overall admit rates*, irrespective of race. To this end, we examine the difference in the average IPEDS admit rate of applications with outcomes reported and of those without outcomes reported; these values are shown by the blue bars in Figure 38. In tiers 1–4, the difference is positive, suggesting that applicants are more likely to report outcomes for colleges with higher overall acceptance rates.

This differential, however, does not vary with race, which we conclude from a permutation test that randomly assigns race to applicants, with differences corresponding to the green bars in Figure 38. Using a sum-of-squares test statistic $Y_k = \sum_r (o_{rk} - \bar{o}_k)^2$, where o_{rk} is the observed difference for race r in tier k, and $\bar{o}_k$ is the overall difference in tier k, we find that differential reporting bias by race is significant only in T5 (one-sided p-values 0.913, 0.497, 0.117, 0.787, 0.000 in tiers 1–5).

Additionally, we perform a "rejection padding" exercise by which we sample applications without a reported outcome and assume them to be rejections, such that the reporting rate is equalized across racial groups within each tier (we select the highest reporting rate across racial groups to be the target). After refitting the model to this augmented data set and repeating our analysis, we do not find meaningful differences in our results, which is expected given that the difference in reporting rates among different groups is not large, as seen in Figure 37.

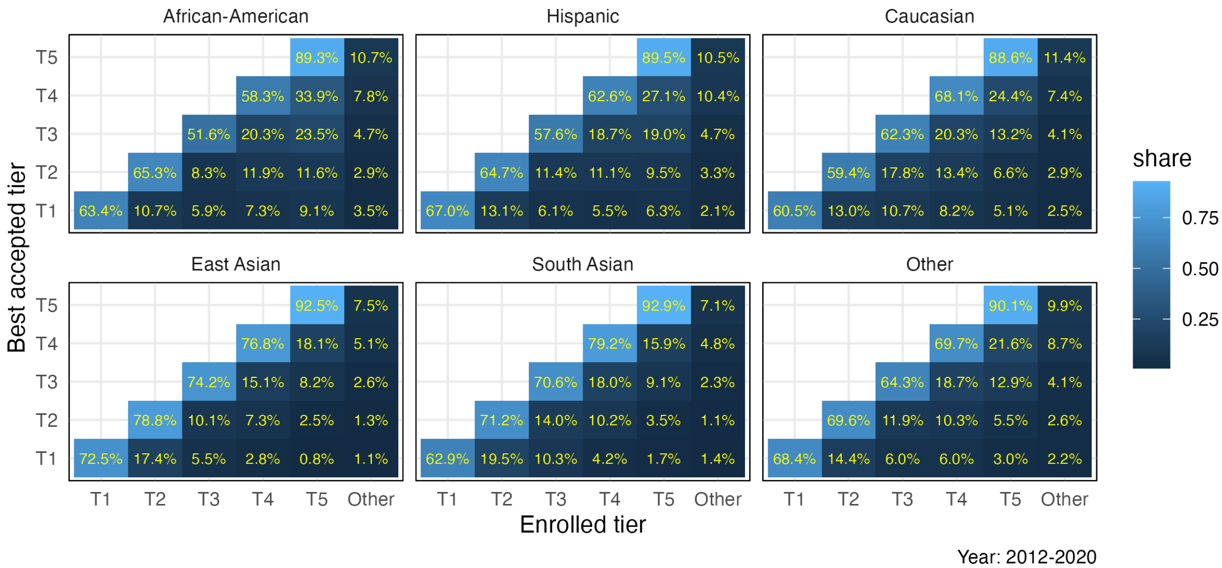

Figure 39: The distribution of enrolled tier, by race and best (i.e., most selective) accepted tier, for applicants who reported the college at which they ultimately enrolled, aggregated across all years 2012–2020. The "Other" enrollment category includes 4-year colleges which did not report SAT data to IPEDS (and are therefore not assigned a tier), as well as 2-year technical and community colleges.

Appendix D Additional Findings and Technical Details

D.1 Estimated Impact on Enrollment

In this section, we attempt to account for the effect of enrollment spillovers by modeling tier-level enrollment choices. Figure 39 shows, for the approximately 50 percent of applicants who reported their enrollment decision, the distribution of their choices, conditional on race and the most selective tier to which they reported an acceptance. (Applicants who reported their enrollment choice reported admission outcomes for 61% of their applications.) One notices that, while a majority of admits enroll in the most selective tier to which they are admitted, a considerable fraction of them enroll in a less selective tier. There are also noticeable differences between applicants of different races: for example, 79% of East Asian applicants whose highest acceptance was to a T2 college enrolled in a T2 college, while fewer than 60% of Caucasians did. This variability—even more pronounced in Figure 40—may reflect cross-racial differences in the importance placed on selectivity, cost, or other factors which determine enrollment choice. With these observations in mind, we model an applicant's probability of enrolling in tier k as

$$\Pr(\text{enroll in tier } k \mid n_1, \ldots, n_K \text{ offers, Race} = r) = \frac{\theta_{rk} n_k}{\sum_{k'=1}^{K} \theta_{rk'} n_{k'} + 1}$$

where n_k is the applicant's number of offers in tier k, and the "plus one" corresponds to a generic outside option available to all. We estimate the parameters $\theta_{r1}, \ldots, \theta_{rK}$ via maximum likelihood, using the observed enrollment choices of the approximately 50 percent of applicants in the data who reported them; these are shown in Table 13.

We use the model to estimate, for each tier, the racial composition of enrollees under the status quo, as shown in the top row of Figure 41; it is broadly similar to the racial composition of unique admits. Repeating our main exercise, we then calculate, for each of our simulated admission

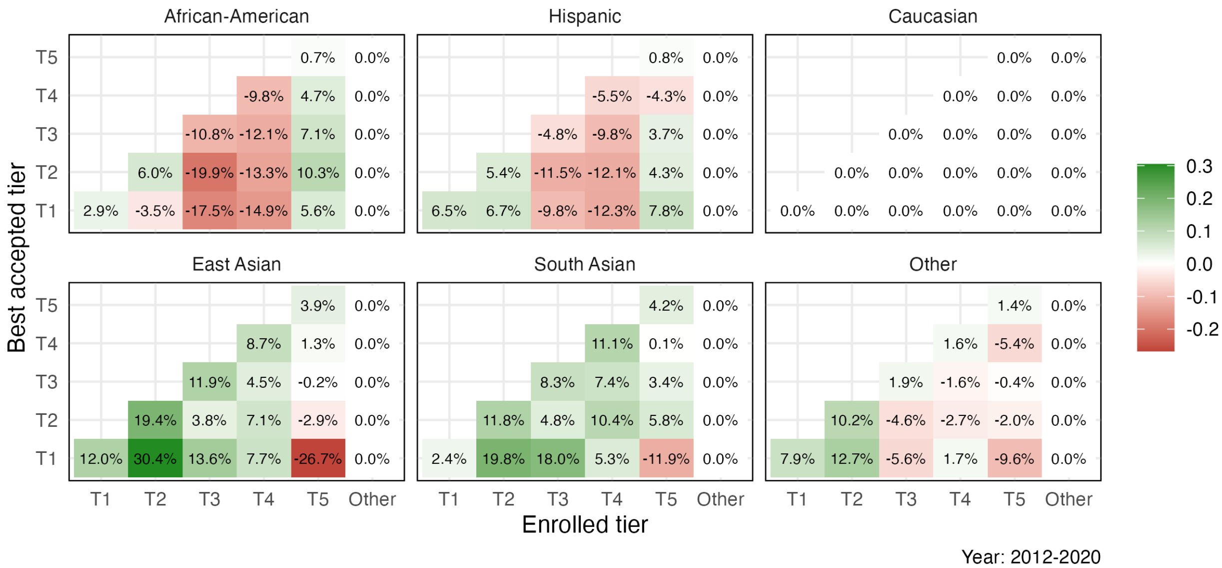

Figure 40: The percentage-point (pp) difference in the conditional frequency of enrolling in a tier (conditioned on not enrolling in a more selective tier), compared to Caucasian enrollees, by race and most selective accepted tier. For example: Conditional on not enrolling in a T3 college, East Asians whose highest acceptance was to a T3 college are 4.5pp more likely to enroll in a T4 college, compared to Caucasians. (Note that sample sizes are somewhat small in the bottom right cells.)

Race (r)	$\hat{\theta}_{r1}$	$\hat{\theta}_{r2}$	$\hat{\theta}_{r3}$	$\hat{\theta}_{r4}$	$\hat{\theta}_{r5}$
African-American	37.2	24.7	12.0	7.9	5.9
Hispanic	42.5	20.9	11.8	7.3	6.4
Caucasian	31.3	19.7	14.3	8.1	6.1
East Asian	115.8	59.4	27.8	14.3	10.2
South Asian	89.9	56.9	30.9	16.2	10.6
Other	47.1	28.4	15.5	8.5	6.6

Table 13: Estimated enrollment model parameters.

policies, the percent change in each race's expected share of enrollees, under the assumption that enrollment model is policy-invariant, and keeping with our assumption that colleges adjust their cutoffs so as make the same number of offers in expectation as under the status quo.[51] These results are displayed in Figure 41 and are markedly similar to those for unique admits. The reason is that—despite the fact that the set of admits shifts considerably—a typical Tier k admit (of any race) under the status quo is predicted to have a similar distribution of offers to that of a typical Tier k admit (of the same race) under each alternative policy, and hence is predicted to have similar enrollment probabilities. This is illustrated in Figures 42 and 43. In the case of T3, the average URM T3 admit under the race-blind policy has fewer T1 and T2 offers compared to the average URM T3 admit under the status quo, translating into a slightly higher probability of enrolling in T3, while the average non-URM T3 admit has a marginally lower probability of enrolling in T3

[51] An applicant's tier k enrollment probability is $\mathbb{E}_{N_1,\ldots,N_K}[\theta_{rk}N_k/(\sum_{k'=1}^{K}\theta_{rk'}N_{k'}+1)]$, which we estimate by Monte Carlo sampling admissions outcomes Y_{ij} from the fitted admission probabilities $\hat{p}_{ij}$—specifically, $Y_{ij}\sim$ Bernoulli$(\hat{p}_{ij})$ with $N_k := \sum_{T(j)=k} Y_{ij}$. We do not account for the uncertainty in the estimates of the enrollment model parameters.

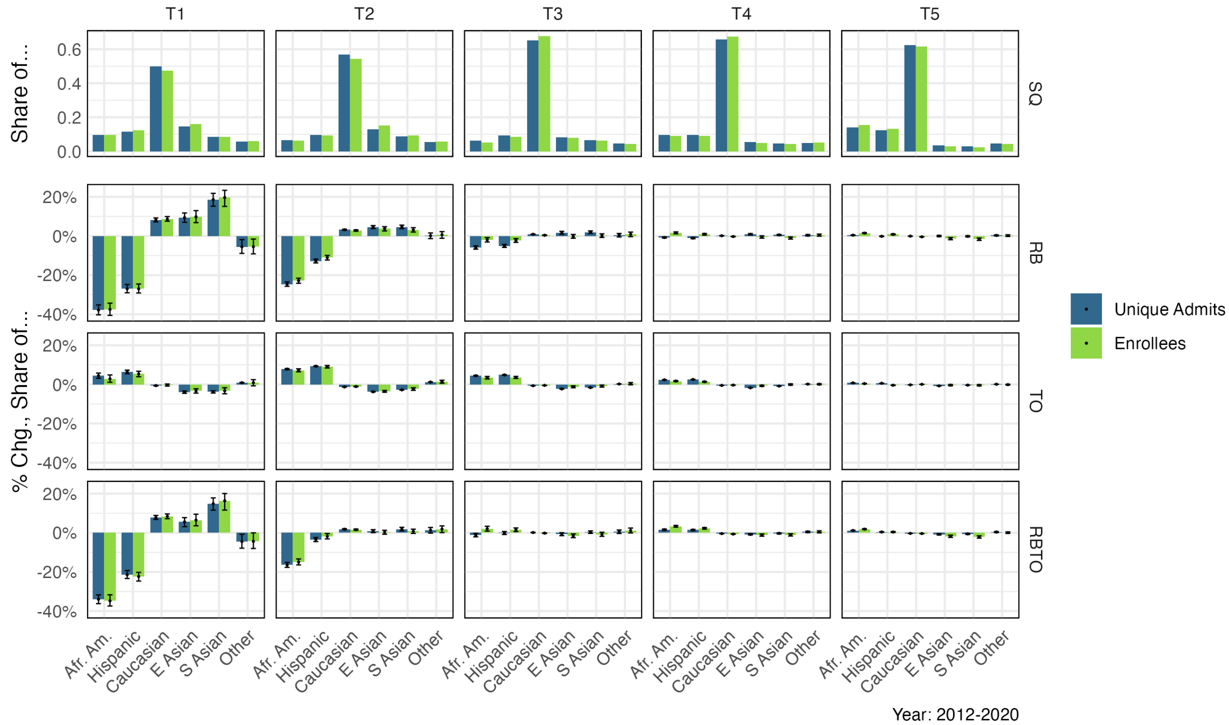

Figure 41: Top row: Race shares of (i) unique admitted students and (ii) enrollees, under the predictions of the baseline admissions model and the enrollment model. Next three rows: The simulated impact on race shares of unique admitted students and enrollment in each tier under a "race-blind" (RB) policy which sets race coefficients equal to zero; a "test-optional" (TO) policy in which colleges impute the mean SAT score conditioned on GPA and applicants withhold scores below this imputed value; and a "race-blind test-optional" policy (RBTO) which combines these elements.

(compared to the status quo). This explains the dampened impact on enrollment (compared to unique admits) under race-blind in T3.

Finally, we note that the results are virtually identical when using a coarser enrollment model in which the weights (θ) are not conditioned on race. Thus, while racial differences in enrollment choice affect the predicted racial composition of enrollees under each policy alternative, they appear to not affect its relative change when comparing these alternatives to the status quo.

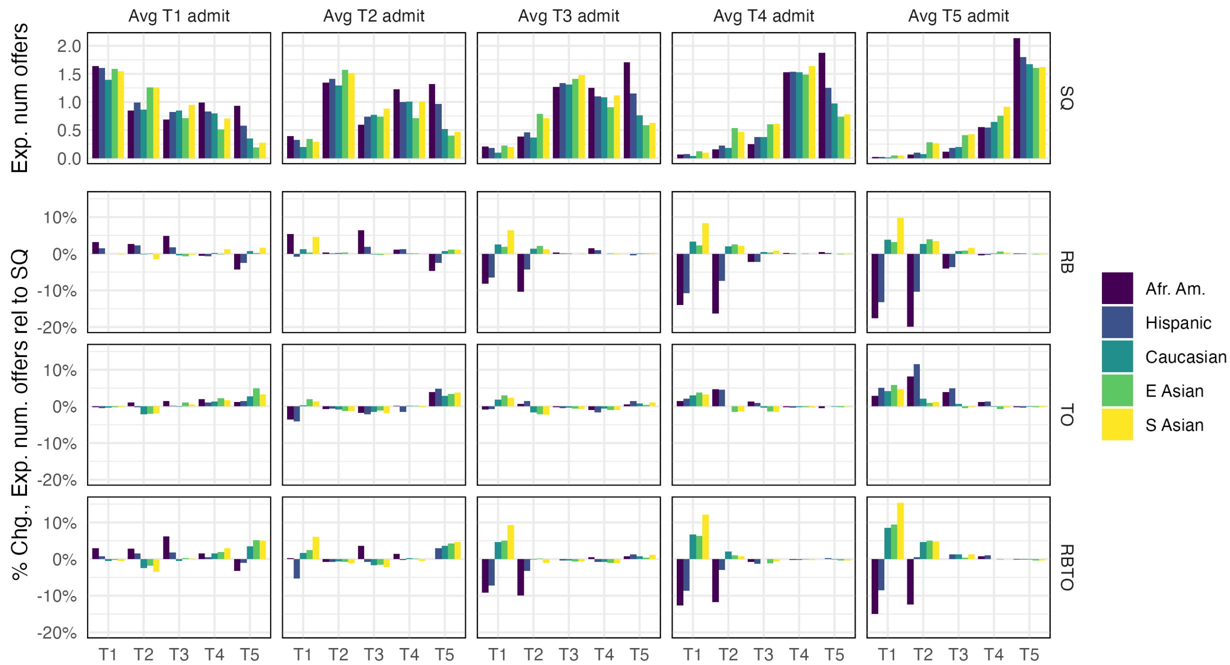

Figure 42: Top row: The expected number of offers to each tier for an average Tier k admit of each race, under the status quo (considering only applications with reported outcomes). Next three rows: The percent change in the expected number of offers for an average Tier k admit of each race, under the labeled policy (RB: race-blind; TO: test-optional), relative to the corresponding number in the top row.

Figure 43: Top row: The enrollment probability in each tier for an average Tier k admit of each race, under the status quo (considering only applications with reported outcomes). Next three rows: The percent change in enrollment probability for an average Tier k admit of each race, under the labeled policy (RB: race-blind; TO: test-optional), relative to the corresponding number in the top row.

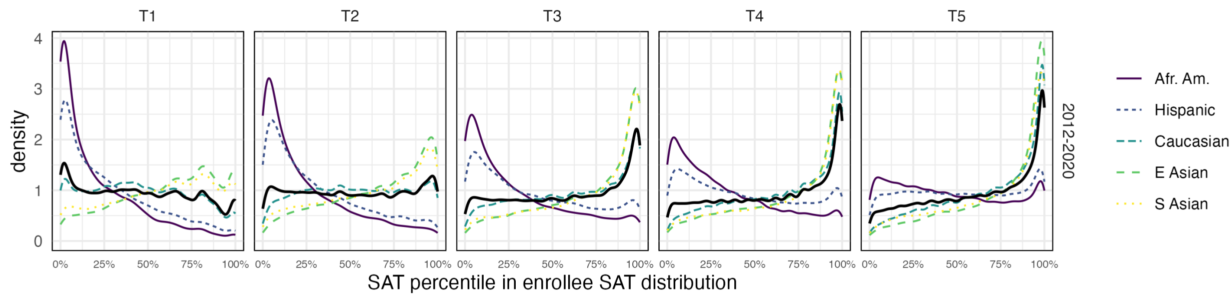

Figure 44: The distribution of the percentiles of applicants' SAT scores within each college's enrollee SAT distribution, by tier and race. Overall distribution in black.

D.2 Test-Optional Admissions: Alternative Score-Reporting Assumptions

In accordance with McManus, Howell, and Hurwitz (2023), we alternatively consider that an applicant's decision to disclose or withhold her SAT score is primarily influenced by comparison with the score distribution for a college's enrollees (i.e., from the previous admissions cycle). This information is discernible from various sources including colleges' own admissions websites, where colleges typically report the 25th and 75th percentiles of their enrolled students' SAT math and verbal scores. We consider a mechanistic rule in which an applicant discloses her score if and only if it is above a certain percentile of the enrollee score distribution.[52]

To estimate the entire distribution of enrollees' scores, we take a parametric approach which assumes that the SAT math and verbal score components, rescaled to the range $[0, 1]$, are correlated beta distributions. We estimate their correlation from the data; we find this to be approximately 0.65. Using the estimated beta distribution parameters, we calculate the mean and variance of the SAT composite score, and again estimate the parameters of a beta distribution which matches these moments.

Using the estimated enrollee SAT distribution for each college, we calculate each applicant's percentile within this distribution. Figure 44 shows the distribution of these percentiles by tier and race. In more selective tiers, URM applicants' scores are more concentrated among lower percentiles of the enrollee distribution.

The top row of Figure 45 shows the impact on race shares at different reporting thresholds. At a 0% threshold, all applicants report their scores and the admissions model remains unchanged from the status quo. The bottom row shows the impact of the corresponding race-blind, test-optional policy. Figure 46 shows the fraction of unique admitted students submitting scores at the different reporting thresholds.

D.3 Test-Blind and Race-Blind, Test-Blind Admissions

Test-Blind Admissions An admissions policy which doesn't consider applicants' SAT scores is called *test-blind*.[53] We conceive of test-blind admissions as a special case of the imputation rule that we define in Section 4 for the test-optional policy, in which each college j sets $\text{SAT}_{ij} = \widetilde{\text{SAT}}_{ij}$ for all applicants. The estimated impact on race shares of unique admitted students in shown in

[52] McManus, Howell, and Hurwitz (2023) finds that the disclosure probability is approximately 0.5 at the 25th percentile of the enrollee SAT distribution (see their Figure 1). However, while this percentile strongly predicts disclosure, there is still considerable variation in reporting.

[53] As of 2023, there are 75 test-blind colleges including the University of California system. (https://blog.collegevine.com/test-blind-colleges)

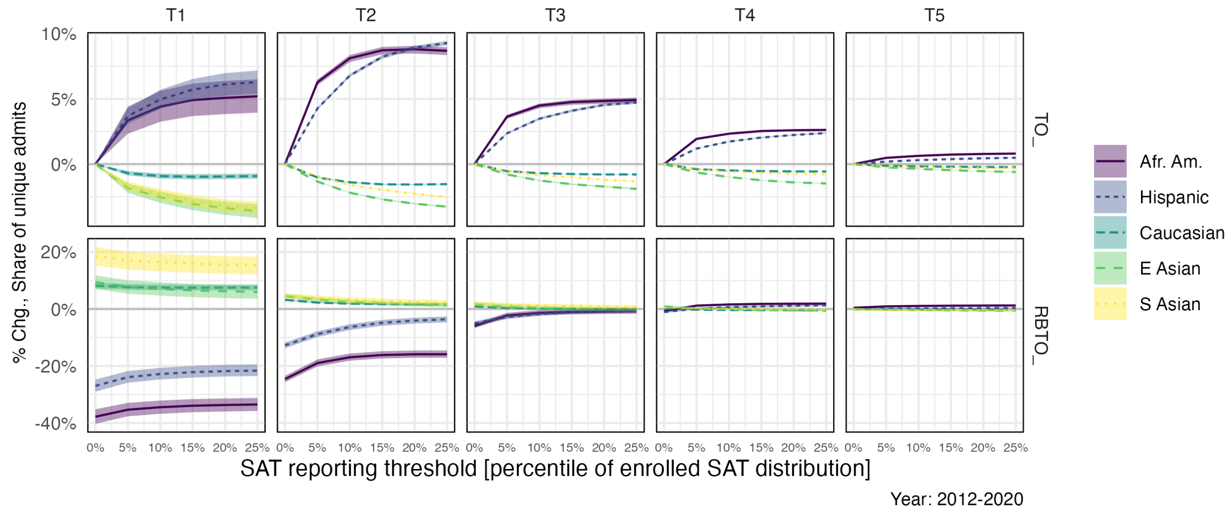

Figure 45: The simulated impact on race shares of admitted students under a "test-optional" (TO) and "race-blind, test-optional" (RBTO) policies, for different values of the assumed reporting threshold.

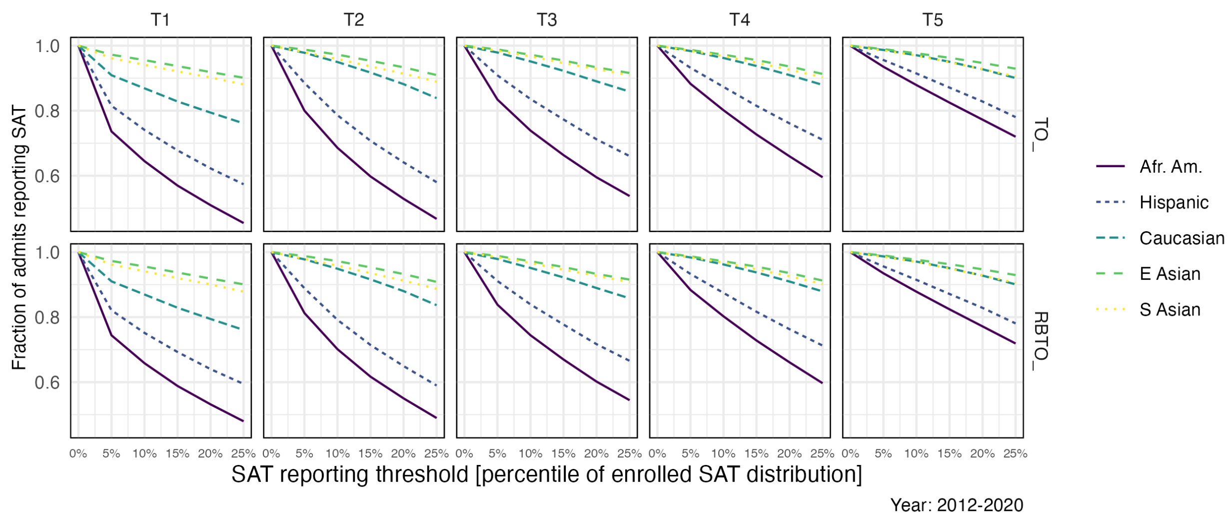

Figure 46: The expected fraction of admitted students submitting test scores under test-optional (TO) and race-blind, test-optional (RBTO) policies, for different values of the assumed reporting threshold.

Figure 47.

Race-Blind, Test-Blind Admissions We find that the negative impact on URMs of colleges going race-blind is partially mitigated by also going test-blind. However, in more selective tiers the estimated impact is more similar to race-blind than test-blind. In T1, we estimate the African-American and Hispanic shares of unique admits to decline by 30% and 18% relative to the status quo; see Figure 47.

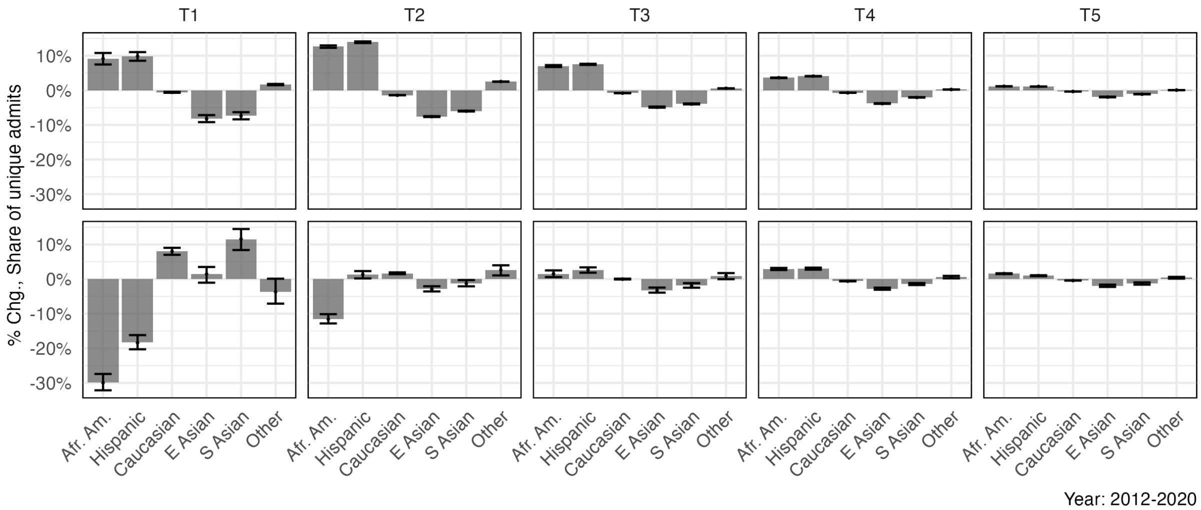

Figure 47: The simulated impact on race shares of admitted students under a "test-blind" (TB) policy in which colleges assign the mean applicant SAT score conditioned on GPA to all applicants; and a "race-blind test-blind" policy (RBTB) which additionally sets race coefficients equal to zero.

D.4 The Effect of Additional URM Applications

We consider how the estimated impact of different policies would change if URM applicants were to change their application behavior, whether by self-interest or at the encouragement of colleges via targeted recruitment strategies or other means. Specifically, we simulate additional applications for URM applicants to T1 and T2, holding both colleges' admissions scoring parameters and the applications of non-URM applicants fixed.

For each URM applicant who applied to either T1 or T2, we draw additional applications from the empirical distribution of applications among applicants of any race in the same GPA ventile, to the most selective tier to which the URM applicant applied, excluding colleges already in the applicant's portfolio. Figure 48 shows the simulated impact for zero, one, and two additional applications sent.

We find that, under the race-blind simulation, two additional URM applications substantially mitigates or reverses the decline in the URM share of (unique) admitted students in the baseline race-blind simulation. In T1, the African-American share of admitted students declines by only 11%, compared to 38%, and the Hispanic share increases by 3%, compared to a 27% decline. In T2, the African-American share declines by only 5%, compared to 25%, and the Hispanic share increases by 9%, compared to a 13% decline.

Under the test-optional simulation, two additional applications results in the African-American and Hispanic shares being boosted by 42% and 45% in T1, compared to 4% and 6% under the baseline test-optional simulation. In T2, the African-American and Hispanic shares increase by 32% and 35%, compared to 8% and 9%.

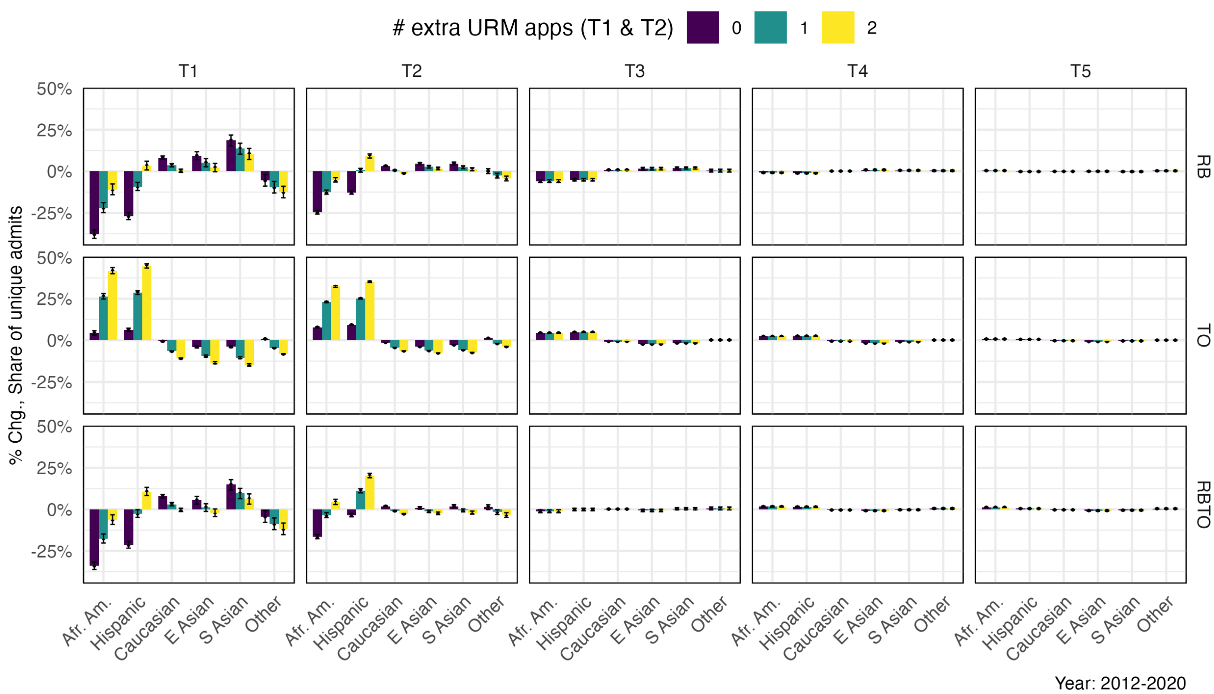

Figure 48: The simulated impact of zero, one, and two additional URM applications, under each policy considered. The extra applications are drawn from the empirical distribution of applications among applicants of any race in the same GPA ventile, to the most selective tier to which the URM applicant applied, excluding colleges already in the applicant's portfolio.

policy	Afr. Am.	Hispanic	Caucasian	E Asian	S Asian	Other
Tier 1						
SQ	9.5%	11.6%	50.0%	14.5%	8.6%	5.8%
RB	5.9% (-37.8%)	8.5% (-27.0%)	54.0% (8.2%)	15.9% (9.4%)	10.2% (18.6%)	5.5% (-5.5%)
TB	10.4% (9.1%)	12.7% (9.8%)	49.7% (-0.6%)	13.4% (-8.2%)	8.0% (-7.3%)	5.9% (1.7%)
RBTB	6.6% (-29.9%)	9.5% (-18.3%)	53.9% (8.0%)	14.7% (1.4%)	9.6% (11.5%)	5.6% (-3.6%)
TO	9.9% (4.5%)	12.3% (6.4%)	49.7% (-0.6%)	14.0% (-4.0%)	8.3% (-3.8%)	5.9% (0.9%)
RBTO	6.3% (-34.0%)	9.1% (-21.5%)	53.9% (7.8%)	15.3% (5.6%)	9.9% (14.9%)	5.6% (-4.5%)
Tier 2						
SQ	6.4%	9.6%	57.0%	12.9%	8.7%	5.4%
RB	4.8% (-24.6%)	8.4% (-12.8%)	58.8% (3.2%)	13.5% (4.5%)	9.1% (4.6%)	5.4% (0.2%)
TB	7.2% (12.7%)	11.0% (13.9%)	56.2% (-1.4%)	11.9% (-7.6%)	8.2% (-6.0%)	5.5% (2.5%)
RBTB	5.7% (-11.5%)	9.8% (1.3%)	57.9% (1.7%)	12.5% (-2.9%)	8.6% (-1.2%)	5.5% (2.6%)
TO	6.9% (7.8%)	10.5% (9.3%)	56.2% (-1.3%)	12.4% (-3.8%)	8.4% (-2.8%)	5.5% (1.2%)
RBTO	5.4% (-16.5%)	9.3% (-3.4%)	58.0% (1.8%)	13.0% (0.8%)	8.9% (1.9%)	5.5% (1.3%)
Tier 3						
SQ	6.2%	9.4%	65.3%	8.2%	6.4%	4.5%
RB	5.9% (-5.9%)	8.9% (-5.1%)	65.8% (0.9%)	8.3% (1.6%)	6.6% (1.9%)	4.5% (0.5%)
TB	6.7% (7.1%)	10.1% (7.5%)	64.7% (-0.8%)	7.8% (-4.9%)	6.2% (-3.9%)	4.5% (0.6%)
RBTB	6.3% (1.5%)	9.6% (2.6%)	65.3% (0.0%)	8.0% (-3.2%)	6.3% (-1.9%)	4.5% (0.9%)
TO	6.5% (4.4%)	9.8% (4.8%)	64.8% (-0.7%)	8.0% (-2.3%)	6.3% (-1.6%)	4.5% (0.2%)
RBTO	6.2% (-1.2%)	9.4% (-0.1%)	65.4% (0.1%)	8.2% (-0.7%)	6.5% (0.3%)	4.5% (0.6%)
Tier 4						
SQ	9.7%	9.6%	65.7%	5.4%	4.7%	5.0%
RB	9.6% (-0.8%)	9.5% (-1.1%)	65.8% (0.1%)	5.4% (0.9%)	4.7% (0.6%)	5.0% (0.4%)
TB	10.0% (3.6%)	10.0% (4.1%)	65.2% (-0.7%)	5.2% (-3.8%)	4.6% (-2.0%)	5.0% (0.3%)
RBTB	10.0% (2.9%)	9.9% (3.0%)	65.3% (-0.6%)	5.2% (-2.8%)	4.6% (-1.4%)	5.0% (0.6%)
TO	9.9% (2.3%)	9.9% (2.5%)	65.3% (-0.5%)	5.3% (-1.7%)	4.6% (-0.9%)	5.0% (0.1%)
RBTO	9.8% (1.6%)	9.8% (1.4%)	65.4% (-0.4%)	5.3% (-0.8%)	4.6% (-0.3%)	5.0% (0.5%)
Tier 5						
SQ	14.1%	12.3%	62.4%	3.6%	3.0%	4.6%
RB	14.2% (0.4%)	12.3% (-0.2%)	62.4% (-0.1%)	3.6% (0.0%)	3.0% (-0.2%)	4.6% (0.3%)
TB	14.3% (1.2%)	12.5% (1.1%)	62.2% (-0.3%)	3.5% (-1.9%)	2.9% (-1.1%)	4.6% (0.1%)
RBTB	14.4% (1.6%)	12.5% (1.0%)	62.2% (-0.4%)	3.5% (-1.9%)	2.9% (-1.3%)	4.6% (0.4%)
TO	14.2% (0.8%)	12.4% (0.6%)	62.3% (-0.2%)	3.6% (-0.8%)	2.9% (-0.4%)	4.6% (0.1%)
RBTO	14.3% (1.2%)	12.4% (0.4%)	62.2% (-0.3%)	3.6% (-0.8%)	2.9% (-0.5%)	4.6% (0.4%)

Table 14: Point estimates of the share of unique admits by race and (in parentheses) the relative change compared to the baseline model. The policies are abbreviated as SQ (status quo), RB (race blind), TB (test blind), and TO (test optional).

D.5 Comparison to Results in Borghesan (2023)

As noted in Section 1.2, Borghesan (2023) is similar to this paper in terms of the research questions explored, but adopts a structural modeling approach. To estimate the effects of an "SAT ban" and an affirmative action ban, the author builds an equilibrium framework which accounts for college cutoff adjustments (which we also allow for), as well as changes in application decisions and high school study effort (which we do not model). Before comparing the empirical results, we first briefly summarize Borghesan's model and the ways in which it differs from ours.

Model Borghesan (2023) classifies colleges as $\{\text{public}, \text{private}\} \times \{\text{elite}, \text{selective}, \text{non-selective}\}$ based on the Barron's selectivity rankings, corresponding to six tiers. Unlike in our model, all colleges within a tier are assumed to have the same cutoff. Colleges are assumed to (i) make admissions decisions on the basis of noisy signals (GPA, SAT) of an applicant's latent knowledge, with posterior inference depending on the applicant's observable characteristics (including race), and (ii) have explicit preferences for racial diversity which manifest through demographic-specific admission cutoffs. Our reduced-form model is similar in that we allow colleges to infer applicants' SAT scores on the basis of GPA, and our estimate of racial preferences may be viewed as race-specific admission cutoffs. However, (i) means that Borghesan's model, unlike ours, allows the effective weights on other characteristics—e.g., GPA—to change once the SAT is banned, in accordance with the college objective function.

On the applicant side, high school students decide (i) which college(s) to apply to (up to two colleges per tier), (ii) whether to take the SAT, absent a ban, and (iii) how much study effort to exert. A student's college-specific enrollment utility depends on net tuition, distance to home, and the probability of graduating, which itself depends on the college tier, the student's knowledge, and other characteristics. The cost of each application depends on household income, parental education, and other factors; unlike in our model, applicants are restricted to send at most two applications per tier for tractability. The decision to take the SAT depends on income and a measure of access based on the frequency of testing dates at the student's high school. Finally, the initial distribution of latent knowledge, and the returns to study effort, are allowed to vary by demographics.

Results The primary data source used to estimate the structural model in Borghesan (2023) is the Educational Longitudinal Study of 2002 (ELS 2002), which surveys "a nationally representative cohort of students who were in the tenth grade in 2002". This data yields 9,910 applicants. The survey contains detailed questions about demographics, high school GPA and test scores, time spent studying, college applications, admissions, matriculation, completion, and financial aid (financial aid is assumed to be exogenous). Figure 49 presents a comparison of our results with those presented in Borghesan (2023). To make the results directly comparable, we have repeated our analysis using the same 6-tier classification under Borghesan (2023), and also accounted for students' enrollment choices using the model described in Section D.1. Note that Borghesan's numbers are not directly reported, as the author focuses on overall URM enrollment levels rather than within-tier enrollment shares. However, these shares are calculable from numbers reported in the author's Table 13: Model Mechanisms; Table 2: Summary Statistics, II; and Table I-4: Access to College without Affirmative Action.

Race-Blind. Under an affirmative action ban with status quo testing policy (the first row in Figure 49), Borghesan (2023) reports an across-the-board decline in the URM enrollment share, leading to a 19% decline in the overall URM share among enrollees at any college.[54] The magnitudes are considerably larger than under our model: we find that the URM share of enrollment only changes substantially in elite private (-25%) and elite public (-10%) colleges, and is negligible in all other tiers.[55] Further, the decline of over or close to 50% at elite colleges seems substantially larger than the realized URM enrollment impact for the set of colleges that have so far reported their data

[54] Preliminary data for the Class of 2028 admission cycle shows that URM enrollment may in fact have increased by about 5% year-over-year, while the overall public and private (non-profit) 4-year college enrollment have increased by 3-4%. https://nscresearchcenter.org/current-term-enrollment-estimates/, accessed February 20, 2025.

[55] Note that our results may seem dampened compared to those in Figure 41. This is due to differences in college tier definitions: for instance, the 66 colleges which belong to T1 under Borghesan's classification contain colleges in both T1 and T2 under our original classification.

Figure 49: A comparison of the results on the relative change in URM enrollment in this paper and Borghesan (2023), the latter calculated from numbers reported in Table 13, Table 2, and Table I-4. (Confidence intervals are not available, and percentages may be imprecise due to rounding.) The classification of colleges into six tiers is based on the Barron's selectivity ranking, with private colleges in T1–3 and public colleges in T4–6. The author's corresponding descriptions are "elite", "highly selective", "less selective" for T1–3, and "elite", "most state flagships", and "satellite campus, some flagships" for T4–6 (see Table 3). The respective numbers of colleges in each tier (among those represented in our data) are 66, 163, 234, 33, 79, and 216; the overall enrollment shares are 5.3%, 15.6%, 13.0%, 10.3%, 15.3%, and 40.7%.

for the Class of 2028 (see Figure 10). Note that the race-blind results are only reported in the appendices in Borghesan (2023), and the author did not comment on the mechanism leading to this large effect. The author did note that URM enrollment is insensitive to testing policy under this regime (see the third row in Figure 49).

Test-Blind. Under a universal ban on SAT consideration (the second row in Figure 49, the main counterfactual explored in Borghesan (2023) and the analogue of our "test-blind" (TB) simulation presented in Appendix D.3), the author estimates that URM enrollment changes by +14% in elite privates; -8% in selective privates; +20% in non-selective privates; -20% in elite publics; -18% in selective publics; and +5% in non-selective publics, relative to the status quo outcome reflected in the ELS data. The author comments that these results are driven by an increase in the number of applications (31% for URMs and 18% for non-URMs), including by students who otherwise would not have applied; a compensating increase in college cutoffs; and a decrease in the level of sorting by knowledge among enrollees, reinforced by an increase (decrease) in high school study effort by previously low-knowledge (high-knowledge) students. However, as Borghesan (2023) does not present results in terms of the URM enrollment share, the author does not comment on the non-monotonicity with respect to selectivity among private colleges, or on the opposite directional effects among private elite and public elite colleges. In contrast, we find an immaterial change in URM representation in Tiers 2, 3, and 6, and an *increase* in representation in Tiers 3 and 4. The latter may be explained by our assumption that colleges do *not* use an applicant's race to impute SAT scores, which helps URM applicants who tend to have lower-than-average SAT scores (see Section 4.2.2). This is absent in the model in Borghesan (2023), which allows the inference of latent knowledge to depend on race.

Race-Blind, Test-Blind. Finally, the results for race-blind, test-blind (RBTB) are shown in the third row of Figure 49. Results from Borghesan (2023) are mostly similar to those under the affirmative action ban counterfactual (first row), with the exception of private less-selective colleges in which the effect size halves when the SAT is also banned. In contrast, our results for race-blind and test-blind are approximately additive (as we note in Section 4.2.3), such that a test-blind policy partially offsets the decline in URM enrollment shares at elite colleges when affirmative action is banned. Discussing the intuition for the apparent ineffectiveness of banning the SAT (in both race-aware and race-blind scenarios), the author argues that "there are a small number of non-SAT takers who would gain admission without SAT requirements, but these marginal applicants are largely shut out by the rise in admissions thresholds as more students apply."

D.6 Model Estimation

We use maximum likelihood estimation to fit the model parameters, namely the cutoffs for J colleges $c := (c_1, \ldots, c_J)$; the fixed-effect coefficients $\beta := (\beta_1, \ldots, \beta_K)$, where each β_k is the vector of tier-k coefficients; and the random-effect multipliers $\delta := (\delta_1, \ldots, \delta_K)$. For conciseness in what follows, we let $\alpha := (-c, \beta, \delta)$ denote the full parameter vector.

Let $\mathcal{J}_i \subseteq [J]$ denote applicant i's application set, and y_i the vector of observed outcomes ($y_{ij} = 1$ if applicant i received an acceptance from college $j \in \mathcal{J}_i$, and $y_{ij} = 0$ otherwise). We provide the following additional notation which allows us to write the linear predictor in (3a) as a linear function of α:

- Let $e_j \in \{0, 1\}^J$ denote the unit vector with a 1 in position j, and 0 otherwise.
- Let $\bar{X}_{ij}$ denote an expanded vector $(X_{ij}^{(1)}, \ldots, X_{ij}^{(K)})$ where $X_{ij}^{(k)} = X_{ij}$ is the vector of covariates corresponding to applicant i and college j if $T(j) = k$, and a vector of zeros otherwise.
- Let $H_{ij} \in \mathbb{R}^K$ denote the vector with η_i in position $T(j)$, and 0 otherwise.

Denoting $Z_{ij} := (e_j, \bar{X}_{ij}, H_{ij})$, we have $\alpha^\top Z_{ij} = -c_j + \beta_{T(j)}^\top X_{ij} + \delta_{T(j)}\eta_i$. For simplicity of notation below, we emphasize the dependence of p_{ij} on (α, η_i), and of Z_{ij} on η_i. Finally, we let $\varphi(\cdot)$ and $\phi(\cdot)$ denote the standard multivariate and univariate normal densities, respectively.

For each applicant i, the independence of the idiosyncratic factors ξ_{ij} across $j \in \mathcal{J}_i$ implies that the conditional probability (conditional on the value of the random effect η_i) of the observed outcomes y_i given parameters α is of the form

$$f_i(y_i \mid \eta_i; \alpha) = \prod_{j \in \mathcal{J}_i} p_{ij}(\alpha, \eta_i)^{y_{ij}} (1 - p_{ij}(\alpha, \eta_i))^{1 - y_{ij}}.$$

This can be written in exponential family form:

$$f_i(y_i \mid \eta_i; \alpha) = \exp\left\{ \sum_{j \in \mathcal{J}_i} y_{ij}\alpha^\top Z_{ij}(\eta_i) - b(\alpha^\top Z_{ij}(\eta_i)) \right\},$$

where $b(\cdot) = \log(1 + \exp(\cdot))$. The *marginal likelihood* of α for applicant i is obtained by integrating over the distribution of η_i:

$$L_i(\alpha; y_i) = \int f_i(y_i \mid \eta_i; \alpha) \times \phi(\eta_i)\, d\eta_i \tag{4}$$

$$= \frac{1}{\sqrt{2\pi}} \int \exp\Big\{ \underbrace{\sum_{j \in \mathcal{J}_i} \Big[y_{ij}\alpha^\top Z_{ij}(\eta_i) - b(\alpha^\top Z_{ij}(\eta_i)) \Big] - \frac{\eta_i^2}{2}}_{h_i(\eta_i)} \Big\}\, d\eta_i. \tag{5}$$

The marginal likelihood of α given observations from all applicants $\{y_i\}_{i=1}^n$ may be factorized using the independence of η_i across applicants:

$$L(\alpha; y) = \int \left[\prod_i f_i(y_i \mid \eta_i; \alpha) \right] \varphi(\eta) \, d\eta = \left(\frac{1}{\sqrt{2\pi}} \right)^n \prod_i \left\{ \int \exp\{h_i(\eta_i)\} \, d\eta_i \right\}. \tag{6}$$

We maximize the log-likelihood $\ell(\alpha; y) := \log L(\alpha, y)$ using an adaptive gradient ascent method, taking advantage of both (i) sparsity, namely that each $\mathcal{J}_i$ is a small subset of all colleges; and (ii) parallel computation, made possible by the factorization in (6). Since exact evaluation of the integrals is very costly, we approximate them by *adaptive Gauss-Hermite quadrature* (AGHQ), a standard approach in the literature.[56] The MLE is consistent and asymptotically normal as the number of applicants $n \to \infty$, a result in Nie (2007) and Bradley and Gart (1962). Thus, we use the Hessian of the log-likelihood to obtain standard errors under an asymptotic normality assumption.

Finally, we note that under a Bayesian interpretation of the model, the integrand in (4) is the conditional probability of observing y_i given η_i, multiplied by the prior on η_i. As such, it is proportional to the posterior distribution of η_i, which we denote as $g_i(\eta_i \mid y_i; \alpha)$. For subsequent inference (e.g., predicting the mix of applicants admitted to each tier under different admission policies), we take expectations with respect to an approximation of this posterior distribution. Specifically, we use Laplace's approximation of $g_i(\eta_i \mid y_i; \hat{\alpha})$ by a Gaussian density with mean $\hat{\eta}_i := \arg\max g_i(\eta_i \mid y_i; \hat{\alpha})$ and standard deviation $(-g_i(\hat{\eta}_i)/g_i''(\hat{\eta}_i))^{1/2} = \sqrt{-1/h_i''(\hat{\eta}_i)}$, a common approach in the GLMM literature.[57]

Moment Conditions Let X_{ij} denote the vector of observable covariates for applicant i and college j, and let $x_{ij} \in X_{ij}$ denote a specific covariate, for example SAT score or a dummy variable indicating whether i has a certain race (for the purposes of this section, X_{ij} may also include a dummy variable representing college j). The first-order condition for log-likelihood maximization, $\nabla_\alpha \ell(\hat{\alpha}; y) = 0$, is equivalent to the moment conditions

$$\sum_i \sum_{j \in \mathcal{J}_i : y_{ij} = 1} x_{ij} = \sum_i \sum_{j \in \mathcal{J}_i} \mathbb{E}_{g_i} \left[p_{ij}(\hat{\alpha}, \eta_i) \right] x_{ij},$$

where $\mathbb{E}_{g_i}[\cdot]$ denotes expectation under $g_i(\eta_i \mid y_i; \hat{\alpha})$, the posterior distribution of η_i. For a categorical variable such as race or college, the above condition implies that estimation will equalize each category's share of acceptances to its predicted share.

D.7 Test-Optional Admissions: Equilibrium Notion

As mentioned briefly in Section 4.2.2, colleges face an adverse selection problem if applicants rationally respond to a known test-optional imputation rule. Specifically, only applicants whose SAT scores are higher than their would-be imputed scores will voluntarily disclose them. We describe here an elementary unraveling result which follows; our discussion is informal but proofs of the result may be found in, e.g., Kelly (2022), Dessein, Frankel, and Kartik (2023), and Liu and Garg (2021).[58]

[56] AGHQ approximates integrals of the form $\int e^{-u^2} q(u)\, du$ as above by calculating $\sum_{\ell=1}^m w_\ell q(u_\ell)$ for appropriately chosen $\{(w_\ell, u_\ell)\}_{\ell=1}^m$. The u_ℓ's are the roots of Hermite polynomials, shifted and scaled based on the shape of the integrand. See Liu and Pierce (1994).

[57] See Section 4.1 of Skrondal and Rabe-Hesketh (2009) on the use of the empirical posterior distribution, and Tierney and Kadane (1986) for details on Laplace's approximation.

[58] For example, see Theorem 1 in Kelly (2022) and Lemma 4.1 in Liu and Garg (2021).

As suggested above, in a simple disclosure model the only equilibrium is one which entails all applicants reporting their scores. To see this, consider an applicant i to one college and suppose that:

1. The applicant's admission probability is non-decreasing in her SAT score, conditional on her non-SAT characteristics X_i;

2. The applicant must decide whether to report ($S_i = 1$) or withhold ($S_i = 0$) her test score, and that, if she chooses $S_i = 0$, the college imputes her score according to the conditional expectation
$$\widetilde{\mathrm{SAT}}_i = \mathbb{E}[\mathrm{SAT} \mid S_i = 0, X_i];$$

3. The college admits score reporters and non-reporters according to the same admission policy, the only difference being whether the true SAT score or the imputed score is used;

4. There is full information, in that the college knows the applicant's reporting strategy and the applicant knows the college's imputation rule;

5. The applicant behaves rationally by seeking to maximize her admission probability.

Under these assumptions, the applicant will withhold her score if and only if $\mathrm{SAT}_i < \widetilde{\mathrm{SAT}}_i$; however, if all applicants behave in this manner, it must be that $\widetilde{\mathrm{SAT}}_i > \mathbb{E}[\mathrm{SAT} \mid S_i = 0, X_i]$, i.e., the college's imputation rule is too optimistic. The same is true of any alternative imputation rule $t_0(X_i)$, as long as applicants rationally choose $S_i = 0$ if and only if $\mathrm{SAT}_i < t_0(X_i)$. The result is an "unraveling" whereby the imputation is so low that no applicants withhold scores in equilibrium—in clear contradiction of the observation that there are, in fact, some applicants who withhold their scores. This provides the motivation for our assumption that colleges do not account for adverse selection in score reporting (and which is also supported by the language that some colleges use to describe their admissions policy, namely that applicants who choose not to report will not suffer a "disadvantage" compared to those who do report).