Background and research question. Universities identify prospective students by purchasing "student lists" from College Board, ACT, and other vendors. A student list contains the contact information of prospective students who meet the search filter criteria (e.g., test score range, GPA, zip codes) specified by the university, who can then be recruited via mail, email, and social media. Research suggests student lists are important for the college access outcomes of millions of students each year (Howell, Hurwitz, Mabel, & Smith, 2021). After controlling for covariates, students who opt into College Board Student Search Service – allowing accredited institutions to "licence" their contact information – were 25% more likely to enroll in a 4-year college than students who opted out (see these results reproduced in Figure 1). Furthermore, the results were stronger for students who identified as Black, Latinx, and first-generation.

However, a series of recent reports by Author (XXXXa, XXXXb, XXXXc) argue student list products systematically exclude underrepresented student populations in two ways. The first source of exclusion is which prospective students are in the underlying database. Historically, student list products sold by College Board exclude non test-takers, but rates of test-taking differ by race and class. Second, several "search filters" (e.g., zip code, AP score) used to control which prospect profiles are purchased facilitate the exclusion of students from communities of color and low-income communities.

This paper examines the exclusionary effects of student list products. We ask, what is the relationship between student list search filters and the characteristics of students who are included vs. excluded in student lists purchased from College Board? We reconstruct College Board student list products using nationally representative data from the High School Longitudinal Study of 2009 (HSLS:09). We create measures for each search filter and measures that identify whether students are included in the underlying student list database. Empirical analyses investigate the relationship between particular filters and combinations of filters – our "X" variables of interest – and the racial and socioeconomic composition of students who are included vs. excluded from purchases that utilize these filters.

Literature Review. Sociology of education literature examines the relationship between recruiting and college access. For example, Holland (2019) analyzes recruiting from the perspective of high school students. Stevens (2007) analyzes recruiting from the perspective of a selective liberal arts college. Cottom (2017) analyzes the recruiting behavior of for-profit universities. Authors (XXXXD, XXXXE) investigated off-campus recruiting visits by public research universities. All of these studies assume that recruiting is primarily a practice of individual colleges and universities.

By contrast, Authors (XXXXA) argue that university recruiting behaviors are increasingly structured by products purchased from third-party software and consulting firms. Historically, the market for student list data has been dominated by College Board and ACT. However, recent market dynamics have created opportunities for edtech software and consulting firms to become suppliers of student list data.

Authors (XXXXB) issued public records requests to public universities to collect data about student list purchases. Empirically, Authors (XXXB) showed that student list purchases using particular combinations of filters were associated with profound racial and socioeconomic exclusion. Because of data limitations, Authors (XXXB) cannot determine which particular filters and filter thresholds are driving the exclusion. This paper overcomes this limitation, by recreating College Board student list products using HSLS:09 and then simulating who is included/excluded as filters and filter thresholds are changed.

Conceptual framework. We develop a conceptual framework from scholarship at the nexus of sociology of race and digital platforms (Burrell & Fourcade, 2021; Cottom, 2020; Hirschman & Bosk, 2020; Hirschman & Garbes, 2019; Norris, 2021; Ray, 2019). Student list products are algorithmic selection devices that – similar to Google Ads or Facebook – allow advertisers to control the prospective customers through the use of search filters. Structural racism is "systematic racial bias embedded in the 'normal' functions of laws and social relations" (Tiako, South, & Ray, 2021, p. 1143), whereby processes viewed as neutral or common-sense systematically advantage dominant groups. Structurally racist inputs are determinants of a selection device that are correlated with race because non-white people have been historically excluded from the input (Norris, 2021). We conceptualize several "geographic" and "academic" search filters as structurally racist inputs. For example, prospects may be filtered by zip code, but zip codes are correlated with race because of residential segregation (Benjamin, 2019). Next, we develop testable propositions about the relationship between specific filters and racial exclusion from purchased lists.

Data. Study data includes (1) HSLS:09 and (2) student list order summaries collected from public records requests to public universities. HSLS provides the extensive student-level demographic, geographic, and academic variables needed to recreate College Board products. Figure 2, based on data we collected from public records requests, shows that search filters can be categorized into the four buckets of geographic, academic, demographic, and student preferences.

Analyses and progress to date. Analyses for included/excluded simulations will consist of simple descriptive statistics – with appropriate tests of significance – and interactive maps. First, we will analyze the racial and socioeconomic composition of purchases that filter only on particular geographic filters discussed in the conceptual framework. Second, we will analyze purchases that filter on particular academic filters. However, most student list purchases utilize multiple search filters. Drawing from data on actual purchases made by public universities, the third stage of analyses will analyze purchases that filter on both geographic and academic filters. We will then simulate marginal changes to filters to identify the drivers of exclusion.

To date, we have created an analysis dataset from public-use HSLS:09 data. Our restricted HSLS:09 data application has been approved and we expect to receive data by 11/15/2022. Additionally, we have created scripts that produce desired descriptive statistics and interactive maps. These can be seen in Figures 3 and 4, respectively. The interactive map of (**ref?**)(fig:metromaps) can be accessed here.

Scholarly significance. Recent scholarship calls for increased attention to links between technology inputs and material mechanisms that drive racial bias (Benjamin, 2019). This study makes an educational contribution by analyzing structural racist inputs in student list products.

1 References

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new jim code*. Medford, MA: Polity.

Burrell, J., & Fourcade, M. (2021). The society of algorithms. *Annual Review of Sociology*, *47*, 213–237. Journal Article. https://doi.org/10.1146/annurev-soc-090820-020800

Cottom, T. M. (2017). *Lower ed: The troubling rise of for-profit colleges in the new economy*. New Press, The.

Cottom, T. M. (2020). Where platform capitalism and racial capitalism meet: The sociology of race and racism in the digital society. *Sociology of Race and Ethnicity*, *6*(4), 441–449. https://doi.org/10.1177/2332649220949473

Hirschman, D., & Bosk, E. A. (2020). Standardizing biases: Selection devices and the quantification of race. *Sociology of Race and Ethnicity*, *6*(3), 348–364. https://doi.org/10.1177/2332649219844797

Hirschman, D., & Garbes, L. (2019). Toward an economic sociology of race. *Socio-Economic Review*, *19*(3), 1171–1199. https://doi.org/10.1093/ser/mwz054

Holland, M. M. (2019). *Divergent paths to college: Race, class, and inequality in high schools*. Rutgers University Press. https://doi.org/10.36019/9780813590288

Howell, J., Hurwitz, M. H., Mabel, Z., & Smith, J. (2021). *Participation in student search service is associated with higher college enrollment and completion*. College Board. Retrieved from https://cbsearch.collegeboard.org/pdf/college-outreach-and-student-outcomes.pdf

Norris, D. (2021). Embedding racism: City government credit ratings and the institutionalization of race in markets. *Social Problems*. https://doi.org/10.1093/socpro/spab066

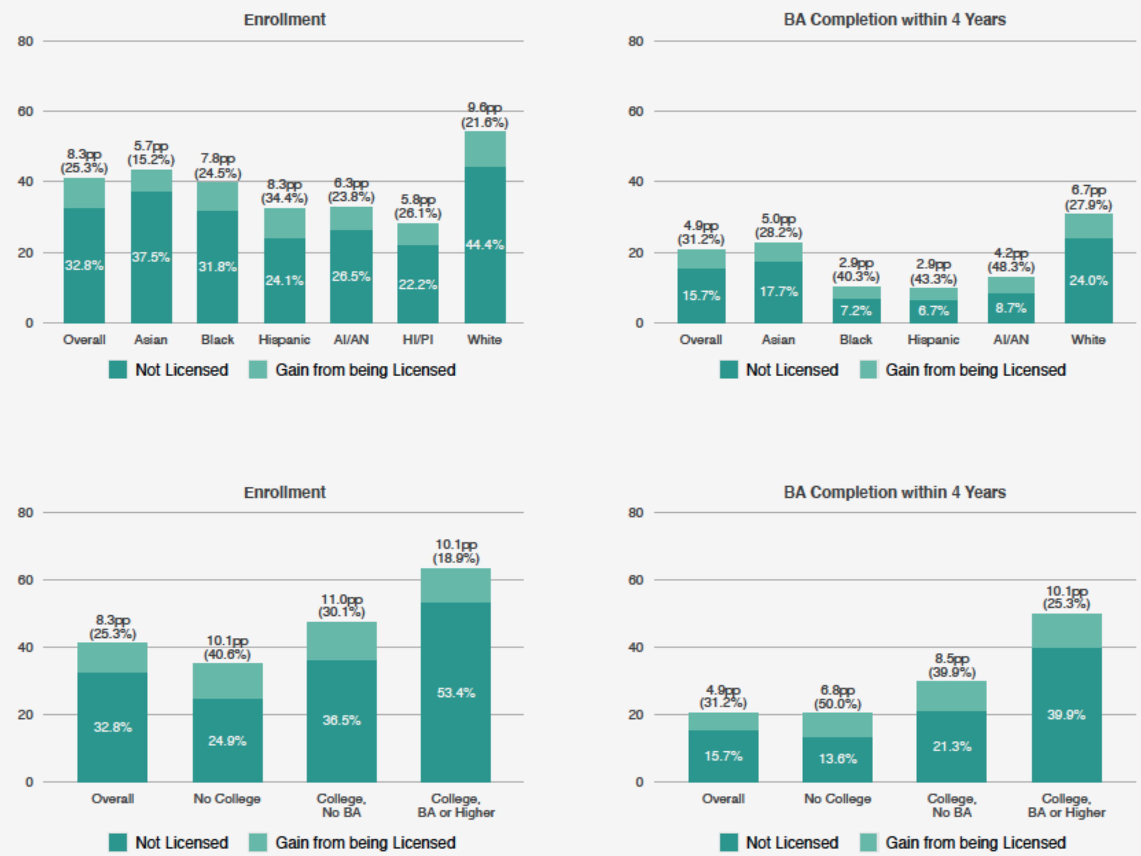
Ray, V. (2019). A theory of racialized organizations. *American Sociological Review*, *84*(1), 26–53. https://doi.org/10.1177/0003122418822335

Stevens, M. L. (2007). *Creating a class: College admissions and the education of elites* (p. 308). Cambridge, MA: Harvard University Press.

Tiako, M. J. N., South, E., & Ray, V. (2021). Medical schools as racialized organizations: A primer. *Annals of Internal Medicine*, *174*(8), 1143–1144. https://doi.org/10.7326/m21-0369 %m 34058105

2 Figures

Figure 1: Student Search Service: College Enrollment and Degree Completion



Note: AI/AN = American Indian or Alaska Native. HI/PI = Hawaiian or Pacific Islander. The sample for enrollment outcomes includes all SAT takers in the 2015–2018 high school graduation cohorts. The sample for completion outcomes is restricted to students in the 2015–2016 cohorts. Completion results are not reported for HI/PI students due to very small sample size (N=2,749), which returns imprecise estimates. Results are estimated from regressions that include student-level controls for: sex, race/ethnicity, SAT score, parental education level, last Student Search Service opt-in status, and graduation cohort and high school fixed effects. All differences between students whose names were licensed and those whose names were not licensed are statistically significant at the 1% level.

Figure 2: Filters Used in Order Purchases by Research vs. MA/Doctoral Universities

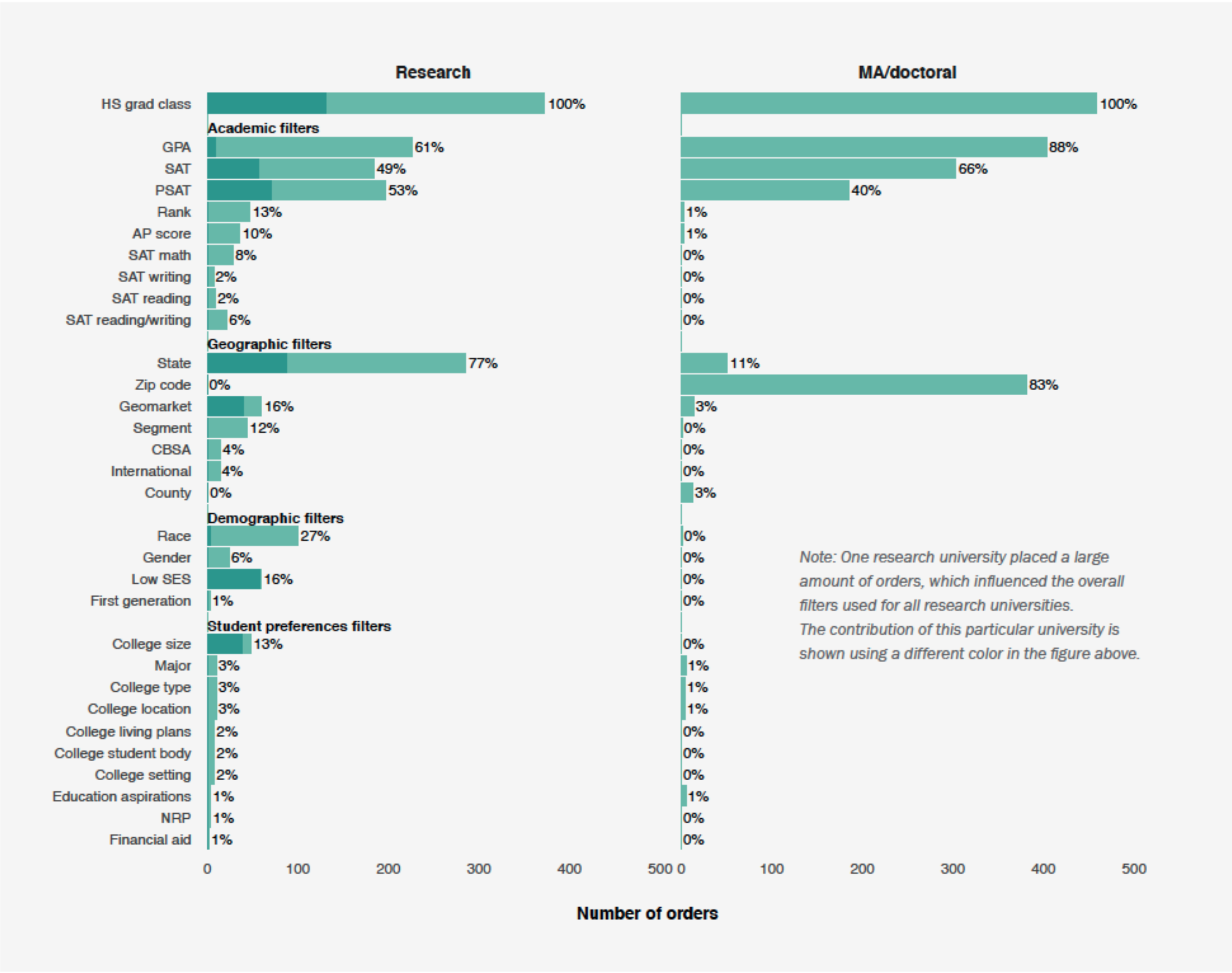
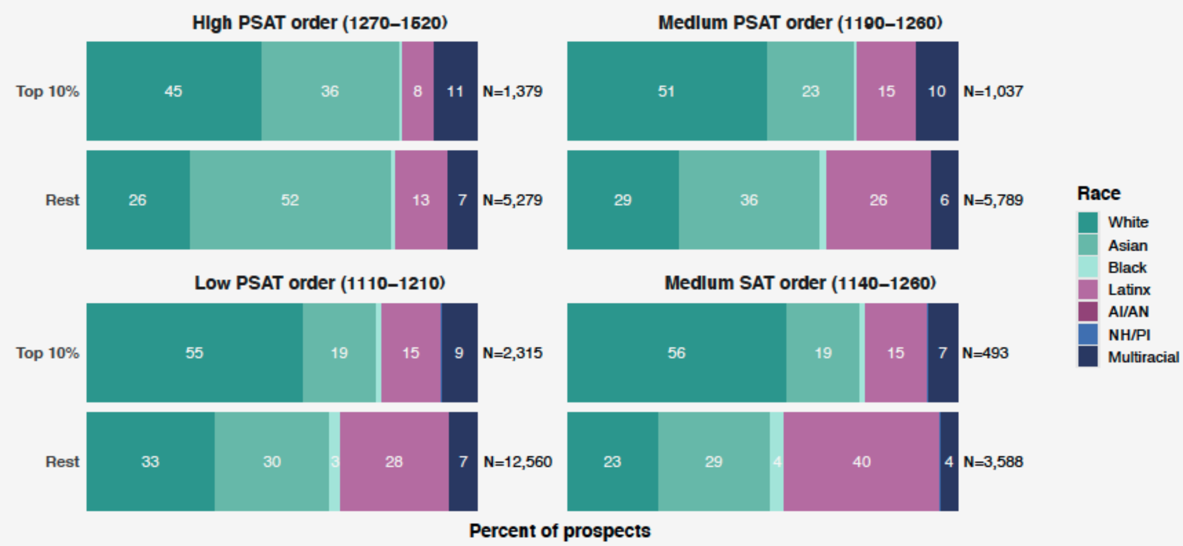


Figure 3: Los Angeles Prospects from Top Income Decile Zip Codes by Racial Composition



Note: Prospects whose race were unknown (1% top decile and 0.9% rest for high PSAT orders; 0.9% top decile and 0.9% rest for medium PSAT orders; 0.6% top decile and 0.6% rest for low PSAT orders; 0% top decile and 0.3% rest for medium SAT orders) or did not report their race (5% top decile and 4% rest for high PSAT orders; 4% top decile and 4% rest for medium PSAT orders; 4% top decile and 3% rest for low PSAT orders; 4% top decile and 3% rest for medium SAT orders) were excluded.

Figure 4: Segment Filter Prospects by Metro Maps (Average Income and Racial Composition)