A Sociological Analysis of Structural Racism in Student List Products

ABSTRACT

Universities identify prospective students by purchasing "student lists." Student list products are algorithmic selection devices that use search filters (e.g., test score, zip code) to select prospective students. We ask, what is the relationship between search filters and the racial composition of included versus excluded students? Drawing from the sociology of race, we conceptualize certain search filters as structurally racist inputs. Structurally racist inputs are determinants of selection devices that are correlated with race because some groups have been historically excluded from the input. We test propositions using a nationally representative sample of high school students. Several academic and geographic filters systematically exclude Black and Latinx students. We motivate critical policy research on third-party products and vendors in education.

1 Introduction

Racial inequality in college access remains an enduring barrier to social mobility [KARINA - ADD STATS/INFO AS REQUESTED BY REVIEWER2].

In economics the market for college access is "modeled as a two-sided matching problem in which the efficient outcome allocates students to colleges based on students' ability to benefit from the type and magnitude of the human capital investment that the college offers" (Hoxby, 2009, p. 106). Hoxby (2009) argues that information costs were the primary barrier to efficient matches. Students want to attend the best possible college but they don't know where they will be admitted and how much it will cost. Colleges want to enroll the best possible students, but they don't know who or where the "good" students are, or how to contact them. "In 1955," Hoxby (2009, p. 103) states, "there was *no* early national college aptitude test. Students and colleges simply did not know where students stood in the national distribution of high school graduates' achievement or aptitude." Colleges could not make an apples-to-apples comparison between students from different schools because the information on a high school transcript "is relative to a standard that a college will not understand unless it draws very often from the high school" (Hoxby, 2009, p. 103).

Hoxby (2009) credits the standardized college entrance exam for transforming U.S. higher education from a system of local autarkies to an efficient, national market by causing a "dramatic fall" (p. 102) in the cost of "colleges' information about students." From 1955 to 1990 the number of colleges requiring the SAT/ACT increased dramatically from 143 to to 1,839, while the number of SAT/ACT test-takers per freshman seat increased from 0.23 in 1955 to 0.69 in 1990 to 0.87 in 2005 (Hoxby, 2009). Test-takers can send scores to colleges they are interested in, allowing colleges to compare prospects from disparate places. Following the creation of Title IV federal financial aid programs, US higher education finance can be conceived as a national voucher system, whereby tuition revenue – including household savings and financial aid – follow students to whichever institution they enroll in. Most colleges cannot survive solely from inquiries by prospects who reach out on their own [CITE]. They must find desirable prospects who can be convinced to apply and enroll.

In 1972 the College Board began selling lists of prospective students to colleges (Belkin, 2019), enabling colleges to identify and target desirable students across the country. Student lists are a match-making intermediary connecting universities to prospective students. Lists sold by College Board and ACT became the primary means colleges used to identify college-going high school students [CITE RNL?]. A student list contains the contact information of prospective students who meet the search filter criteria (e.g., test score range, high school GPA, zip codes) specified by the university. Purchased lists are a fundamental input for undergraduate recruiting campaigns (EAB, 2018), which target prospects by mail, email, and on social media. Recent research suggests that student lists are important for college access and college choice decisions, particularly for first-generation and underrepresented minority students (Howell, Hurwitz, Mabel, & Smith, 2021; Moore, 2017; Smith, Howell, & Hurwitz, 2022).

However, we argue that student list products are a market mechanism that may exacerbate racial inequality in college access. Drawing from the sociology of race, we suggest that the prospect profiles of students of color from communities of color are less likely to be purchased for three reasons. First, the College Board and ACT student list products utilized by public and private non-profit colleges exclude non test-takers, but rates of test-taking differ by race (Blake & Langenkamp, 2022; Hyman, 2017). Second, colleges control which prospect profiles they purchase by filtering on "search filters" (e.g., SAT score, AP score, zip code). We conceptualize several student list search filters as "racialized inputs" (Norris, 2021) that systematically disadvantage underrepresented students of color because they have been historically excluded from this input. Third, because of admissions standards or because they do not know how search filters interact with local patterns of segregation, college administrators may select combinations of search filters that result in racial exclusion.

This manuscript focuses on search filters from the College Board Student Search Service that we conceptualize as racialized inputs (Norris, 2021). We address three research questions. First, What is relationship between individual search filters and racial composition of included vs excluded students? Second, in what ways do public universities racialized input search filters in concert with other search filters when purchasing student lists? Third, what is racial composition of student list purchases that utilize racialized input search filters in concert with other search filters?

We address RQ1 by reconstructing the College Board Student Search Service product using a nationally representative sample of 9th graders in 2009 from the High School Longitudinal Survey (HSLS:09). Analyses simulate and compare the racial composition of included versus excluded prospects when individual search filters are utilized, focusing on test score (e.g., SAT, AP) and geography (e.g., zip code) certain search filters are utilized. We address RQ2 analyzing 830 student lists purchased by 14 public universities, which were collected via public records requests. We address RQ3 by showing the racial composition of selected student list purchases – simulated purchases based on HSLS data and actual purchases based on public records requests – that filter on multiple search filters simultaneously.

The manuscript is organized as follows. First, we provide background on student list products, situating them vis-a-vis the process of recruiting students and summarizing recent dynamics in the market for student list data. Second, we review empirical scholarship on recruiting, focusing on scholarship from sociology. Third, we develop a a conceptual framework and propositions. Next, we describe methods and present results. Finally, we discuss implications for scholarship and policy. Student lists have been largely ignored by researchers and policymakers for more than 50 years. Thus, the primary goal of this research is to motivate future scholarship and policy debate about student list products.

[MOVE MOST OF NEXT TWO PARAGRAPHS TO DISCUSSION] We observe striking parallels between the origins and functions of student list products, which remain unregulated, and consumer credit report products, which are federally regulated. In the market for college access, SAT/ACT scores were viewed by colleges as a measure of the achievement or aptitude of potential students. Similarly, credit scores were designed to predict the probability of repayment for businesses that depend on the ability of customers to pay back debt (Leyshon & Thrift, 1999; Poon, 2007). Like SAT/ACT scores, credit scores were "designed to overcome the chronic problems of information asymmetries" (Leyshon & Thrift, 1999, p. 434) by providing firms information about potential customers (Hoxby, 1997; Leyshon & Thrift, 1999). SAT/ACT scores helped create a national higher education market by enabling colleges to compare prospects from disparate places. Similarly, credit scores allow firms “to distinguish ‘good’ from ‘bad’ customers ‘at-a-distance’ ” (Leyshon & Thrift, 1999, p. 434). Consumer reporting companies like Equifax wrap credit scores and other consumer information into products designed to filter and target customers (Equifax, 2023), enabling firms to transition from the model of approving/rejecting applicants to the more aggressive model of pre-approving desirable customers (Leyshon & Thrift, 1999). Similarly, student list products enabled colleges to transition from the model of accepting/rejecting applications to the model of identifying and targeting desirable prospects.

Consumer credit products are regulated under the Fair Credit Reporting Act – enforced by the FTC – and the Consumer Finance Protection Act – enforced by the Consumer Finance Protection Bureau – because these credit scores lead to the exension of credit. By contrast, student list products are not federally regulated. We suggest that student list products lead to the extension of credit through student loans because colleges identify prospects by buying lists at the top of the “enrollment funnel” and offer financial aid to convert admits to enrolled students in the last stage of the enrollment funnel. We argue that student list products should be regulated because they incorporate racialized inputs that disadvantage communities of color and because universities are likely to utilize student list products in ways that increase racial inequality. College Board argues that student list products can be used to increase racial diversity [CITE]. The question becomes, should policymakers continue to tolerate a product that is likely to do harm on the grounds that the product is also capable of doing good.

2 Background and Literature Review

2.1 Enrollment Management and the Enrollment Funnel

Enrollment management is the organizational behavior side of college access and student success. The term enrollment management can refer to a profession, an administrative structure, or an industry. As a profession, enrollment management integrates techniques from marketing and economics in order to “influence the characteristics and the size of enrolled student bodies” (Hossler & Bean, 1990, p. xiv). Beyond the basic goal of survival, universities pursue some combination of broad enrollment goals (e.g., tuition revenue, academic profile, racial diversity) (Hoxby, 2009); Winston (1999); Cheslock & Kroc (2012)], while also tending to specific needs of campus constituencies (e.g., College of Engineering needs majors, athletic teams need players) (Stevens, 2007). As an administrative structure, the office of enrollment management typically controls the activities of admissions, financial aid, and marketing and recruiting (Kraatz, Ventresca, & Deng, 2010). The enrollment management industry consists of university professionals (e.g., admissions counselors, VP for enrollment management, admissions counselors), professional associations (e.g., National Association for College Admission Counseling), and third-party vendors/consultancies that interact with universities and students (e.g., College Board, EAB, PowerSchool).

Universities cannot realize their enrollment goals solely from prospects who find the university on their own. They must incite demand and discover desirable prospects who can be convinced to enroll. The "enrollment funnel" – depicted in Figure 2 – is a conceptual model used by enrollment management industry to depict broad stages in the process of recruiting students. The funnel begins with a large pool of "prospects" (i.e., prospective students) that the university would like to convert into enrolled students. "Leads" are prospects whose contact information has been obtained. "Inquiries" are prospects that contact your institution and consist of two types: first, inquiries who respond to an initial solicitation (e.g., email) from the university; and second, "student-as-first-contact" inquiries who reach out to the university on their own (e.g., sending ACT scores). Applicants consist of inquiries who apply plus "stealth applicants" who do not contact the university before applying. The funnel narrows at each successive stage in order to convey the assumption of "melt" at each stage (e.g., a subset of "inquiries" will apply).

Practically, the purpose of the enrollment funnel is to inform recruiting interventions that target one or more stage. These interventions seek to increase the probability of "conversion" across stages (Campbell, 2017). At the top of the enrollment funnel, purchasing student lists is the primary means of converting prospects to leads. Purchased leads are served emails, brochures, and targeted social media designed to solicit inquiries and applications (Ruffalo Noel-Levitz, 2022b). Digital (e.g., Google display ads, YouTube) and traditional (e.g., TV, billboards) advertising are means of raising brand awareness and soliciting student-as-first-contact inquiries and (Cellini & Chaudhary, forthcoming; Ruffalo Noel-Levitz, 2022b). At the bottom of the funnel, colleges offer financial aid packages to convert admits to enrolled students (e.g., Hurwitz, 2012).

2.2 Scholarship on Recruiting

Scholarship on enrollment management behaviors targeting college access tend to focus on the latter stages of the enrollment funnel, particularly the process of deciding which applicants to admit [e.g.] (Bastedo, Howard, & Flaster, 2016; Killgore, 2009; Posselt, 2016) and the use of financial aid to convert admits to enrolled students (Doyle, 2010; Leeds & DesJardins, 2015; McPherson & Schapiro, 1998). Fewer studies investigate the earlier "recruiting" stages of identifying prospects, acquiring leads, and soliciting inquiries and applications.

Scholarship on recruiting from economics tends to analyze the causal effect of specific interventions on college access outcomes. Hoxby & Avery (2013) evaluate a nation-wide experiment that delivered customized information about admissions and financial aid to high-achieving (at least 90th percentile SAT/ACT score, 3.7 GPA), low-income students. The intervention positively affected applications, admission, and enrollment to selective colleges. These results catalyzed scholarship on information and advising interventions [e.g.] (Castleman & Goodman, 2018; Cunha, Miller, & Weisburst, 2018; Gurantz et al., 2021), but results have been mixed. Another set of studies evaluate interventions by flagship universities that combine outreach and financial aid (Andrews, Imberman, & Lovenheim, 2020; Dynarski, Libassi, Michelmore, & Owen, 2021).

A small number of studies from economics evaluate third-party products. Mulhern (2021) finds that customized information provided by Naviance college planning software causes students to apply to colleges where they have a high probability of admission and where their SAT scores are similar to previous admits, with affect sizes largest for Black, Latinx, and low-income students. In a report published by College Board, Howell et al. (2021) compared SAT test-takers who opted into the College Board Student Search Service – allowing accredited institutions to purchase their contact information – to those who opted out. After controlling for covariates (e.g., SAT score, parental education, high school), 41.1% of students who participated in Search attended a 4-year college compared to 32.8% of students who opted out, an 8.3 percentage point difference and a 25.3 (`(41.1-32.8)/32.8`) percent change in the relative probability. Participating in Search was associated with a larger change in the relative probability of attending a 4-year college for Black students (24.5%) and Hispanic student (34.4%) than White students (21.6%), and a larger change for students whose parents did not attend college (40.6%) than those whose parents had a BA (18.9%).[1] Leveraging a natural experiment in College Board student list purchases, Smith et al. (2022) find that the purchase of a prospect profile by a college increases the probability that the student will apply to and enroll at the college, with larger effects sizes Black, Latinx, low-income, and first-generation students.

Scholarship from sociology tends to document recruiting behavior "in the wild," often as part of broader analyses of college access or enrollment management [e.g.] (Cottom, 2017; Posecznick, 2017). Holland (2019) probes the structural holes between high school counseling and college recruiting efforts from the perspective of high school students. Underrepresented minority students reported "feeling like their school counselors had low expectations for them and were too quick to suggest that they attend community college" (p. 97). In turn, these students were drawn to colleges that made them feel wanted, often attending institutions with lower graduation rates and requiring larger loans than other options. An audit experiment by Thornhill (2019) sent inquiry email inquiries by fictitious Black high school students to white college admissions counselors. Counselors were less likely to respond to black students who expressed strong commitment to racial justice.

Several studies analyze connections between colleges and high schools from an organizational perspective (Khan, 2011; Salazar, 2022; Salazar, Jaquette, & Han, 2021; Stevens, 2007). Off-campus recruiting visits are often conceptualized as an indicator of enrollment priorities and/or a network tie indicating the existence of a substantive relationship (Author, XXXX). Stevens (2007) provides an ethnography of enrollment management at a selective liberal arts college. The college valued recruiting visits to high schools as a means of maintaining

[1] After controlling for covariates, Howell et al. (2021) find that 20.6% of students who participated in Search obtained a BA in four years compared to 15.7% of students who opted out, representing a 31.2% increase in the relative probability of graduation and this increase was larger for Black (40.3%) and Latinx (43.3%) students (48.3%) than it was for white students and larger for students whose parents did not attend college (50.0%) than it was for students whose parents had a BA (25.3%). The logical mechanism is that where students start college affects the probability of graduation (Long & Kurlaender, 2009; Melguizo, 2008).

relationships with guidance counselors at feeder schools, and tended to visit the same set of – disproportionately affluent, private, and white – high schools year after year. Salazar et al. (2021) analyzed off-campus recruiting visits by 15 public research universities. Most universities made more out-of-state than in-state visits. These out-of-state visits focused on affluent, predominantly white public and private schools.

Cottom (2017) shows that for-profit colleges found a niche in Black and Latinx communities precisely because traditional colleges ignored these communities (see also Dache-Gerbino, Kiyama, & Sapp (2018)). For-profits identified prospects by compiling and purchasing lists (Cottom, 2017). However, they did not rely on lists from College Board and ACT because their target audience of Black and Latinx adult women were not recent test-takers. Ironically, Black and Latinx adult women were vulnerable to marketing from for-profits because they were excluded from the College Board and ACT lists used by traditional colleges. The for-profit business model of encouraging students to take on federal and private loans exemplifies "predatory inclusion," the logic of "including marginalized consumer-citizens into ostensibly democratizing mobility schemes on extractive terms" (Cottom, 2020, p. 443). Another example of predatory inclusion is "payday loans," which target people with low credit scores based on data purchased from credit bureaus (Fourcade & Healy, 2013).

Reflecting on the recruiting literature, studies from both economics and sociology find that underrepresented student populations are particularly sensitive to recruiting interventions. Economics often evaluates the effect of outreach designed to increase underrepresented student enrollment at selective colleges, but sociology finds that the recruiting efforts of selective colleges prioritize students from privileged schools and communities. Scholarship from sociology assumes that recruiting is something done by individual colleges. While a few studies from economics analyze how students respond to interventions delivered by third-party products [e.g.] (Mulhern, 2021; Smith et al., 2022), no studies analyze third-party recruiting products as the fundamental object of analysis. We argue that third-party products and vendors structure college recruiting behavior and, in turn, college access. Prior research has not examined whether third-party recruiting products incorporate structural inequality in ways that systematically disadvantage underrepresented students and prior research has not examined how colleges utilize these products. This study deconstructs the College Board Student Search Service product. We simulate the racial composition of purchases that utilize search filters conceptualized as "racialized inputs" (Norris, 2021) and we examine the usage of these inputs in actual student lists purchased by public universities.

2.3 Background: Student List Products

List-based vs. behavioral-based leads. "Lead generation" is the process of connecting merchants who sell products to "leads" – consumers potentially interested in these products (Federal Trade Commission, 2016b). Student lists are an example of "list-based" lead generation. List-based lead generation is based on the direct mail business model (Singer, 1988) but has evolved into "database marketing," in which information about prospects is stored in a database and prospects are selected using search filters (e.g., Equifax, 2023). Behavioral-based targeting emerged from advances in digital technology and includes most advertising on websites and social media. Whereas list-based marketing proceeds in two steps – obtain customer contact information and then serve marketing material — behavioral-based targeting identifies targets based on their user profile and simultaneously serves advertisements to users while they are on the platform. An article on digital advertising by EAB (2018, p. 9) provides insight about usage of list-based and behavioral-based leads in higher education:

> For industries outside of higher education and for non-freshman recruitment, a primary aim of digital marketing is often that of identifying a pool of potentially interested customers . . . [By contrast] Where the recruitment of college-bound high school students is concerned, digital channels are less important from a lead-generation perspective, because the vast majority of likely candidates are already readily identifiable via testing and survey services (ACT, College Board, etc.). Digital marketing is, instead, of greatest value in further stages of the recruitment funnel, including inquiry generation and application generation.

When recruiting college-going high school students, EAB (2018) suggests that behavioral-based leads are less effective than purchasing names from College Board/ACT and then targeting these prospects on digital platforms (e.g., Meta allows colleges to serve ads to purchased names on Facebook/Instagram). Behavioral-based targeting (e.g., Google Ads, Google Display Network, Twitter) is the primary sources of leads when high-quality databases of prospective customers are not available. Therefore, postsecondary programs targeting working adults often rely on behavioral-based targeting (Carey, 2019; Ruffalo Noel-Levitz, 2022c).

Sources of student list data. Student list data are extracted from the user-data of students laboring on platforms (e.g., taking a test, searching for college). Historically, the student list business has been dominated by College Board and ACT, which derive student list data from their database of test-takers. In the 21st Century, advances in technology yielded new sources of student list data, particularly free online college search engines (e.g., Cappex) and college planning software sold to high schools used by high school students and guidance counselors (e.g., Naviance, Scoir) (Jaquette, Salazar, & Martin, 2022).

Who buys student lists. Extant knowledge about how colleges use student lists depends on market research by Ruffalo Noel Leviz, which publishes regular reports about recruiting practices based on survey responses from their clients (mostly public and private non-profit universities of mid-level size and mid-level selectivity). In an analysis of 120 4-year colleges, Ruffalo Noel-Levitz (2022b) reported that 87% of private and 86% of public institutions "purchase high school student names to generate inquiries and applicants" (p. 5). For public institutions, 20% purchased fewer than 50,000 names annually, 29% purchased 50-100K, 31% purchased 100-150K, and 20% purchased more than 150K names annually. Anecdotal evidence suggests that larger and more selective institutions purchase more names than smaller and less selective ones (Belkin, 2019; Jaquette et al., 2022). Ruffalo Noel-Levitz (2022a) reports that purchasing names was the top ranked expenditure item in the undergraduate marketing and recruiting budget for both private and public institutions. In 2022 the the average public institution allocated 15% of its budget to purchasing names (up from 12% in 2020), compared to 2% of its budget on behavioral-based leads.[2]

Buying College Board and ACT student lists. Each student list purchase is a subset

[2]For private colleges, 16% purchased fewer than 50,000 names annually, 35% purchased 50-100K, 28% purchased 100-150K, and 21% purchased more than 150K names annually (Ruffalo Noel-Levitz, 2022b). In 2022 the average private institution allocated 16% of its budget to purchasing names (up from 14% in 2020), compared to 7% of its budget on behavioral-based leads (Ruffalo Noel-Levitz, 2022a).

of prospects from a larger, underlying database. College Board, ACT, and other student list products (e.g., *Intersect*) incorporate search filters that allow customers to control which prospect profiles they select. Salazar, Jaquette, & Han (2022) categorizes search filters available in the College Board Student Search Service product into the four buckets of academic, geographic, demographic, and student preferences (e.g., desired campus size, intended major). Academic filters include high school graduating class, SAT score, PSAT score, AP score by subject, high school GPA, and high school class rank. Individual filters are specified as score ranges (e.g., XXX - XXX) and can be combined with other filters as `AND` or `OR` conditions. Geographic search filters include state, CBSA, county, zip code, and "geomarket" and "geodemographic" filters (described below). Demographic filters include race, ethnicity, gender, and first-generation status. Student preference filters include intended major, college size, and college type. Analyzing data about 830 student lists purchased by 14 public universities, Salazar et al. (2022) found that the average list purchase specified 4.44 criteria and 98.8% of purchases specified at least one academic and one geographic filter.

A purchased list (College Board template, ACT template) is essentially a spreadsheet with one row per prospect and columns for contact information and student characteristics from the pre-test questionnaire (e.g., graduation year, high school code, gender, ethnicity, race, first-generation status, intended major). Information about academic achievement is extremely limited, but can be inferred from search filters.

How lists are used. Much like the role of voter files in political campaigns (Culliford, 2020), purchased lists are the basic building block for data-informed undergraduate "recruiting campaigns." Enrollment managers use predictive models to inform recruiting interventions (Ruffalo Noel Levitz, 2021; Salazar et al., 2022). However, both the algorithms and the interventions must be fed data about prospects (e.g., cannot send brochures and emails without addresses). Therefore, at the top of the enrollment funnel (Figure 2), colleges buy the contact information of prospects they want to recruit. Once purchased, student lists are combined with behavioral based leads and inquiries who reach out on their own (e.g., took a virtual tour) and then layered with additional data sources, such as consumer data from credit bureaus, historical application/enrollment data about students who attended the same high school, etc. These layered data are the input to predictive models that inform decisions about recruiting interventions designed to push them to subsequent stages of the funnel (e.g., who gets a $0.50 postcard and who gets a $7 brochure). Ruffalo Noel-Levitz (2022b, p. 6) reports that email, targeted digital advertising (e.g., Instagram), and direct mail are the top three "preferred methods for first... contact with high school purchased names" and that the average number of times purchased names are contacted before giving up is 11 for private institutions and 8 for public institutions.

With respect to efficacy, Ruffalo Noel Levitz (2018) asked clients to rate different "first contact" interventions (e.g., off-campus recruiting visit) as sources of inquiries and enrolled students. For the median public college, purchased lists accounted for 26% of inquiries, which ranked #1, and accounted for 14% of enrolled students, which ranked fourth after "application as first contact" (19%), campus visit (17%), and off-campus visit (16%).[3]

3 Conceptual Framework

ADD INTRO PARAGRAPH

3.1 Selection Devices

The sociology of race is concerned with processes that allocate individuals to categories based on some set of input factors. Examples include college admissions, hiring, applications for credit, and prison sentencing. Selection devices are procedures or routines for making selection decisions (Hirschman & Bosk, 2020). One dimension selection devices differ is individual discretion versus standardization. Discretionary selection processes rely on the judgment of individual evaluators. In professional domains (e.g., psychiatric treatment, holistic admissions), evaluators exercise judgment about cases based on professional norms about how to evaluate inputs. By contrast, "standardized selection devices" make decisions based on a mathematical function in which the value of input variables determines the value of the outcome (Duncan, Ohlin, Reiss, & Stanton, 1953), for example a public university that admits applicants based on a function of ACT score and GPA.

[3] For the median private college, student list purchases were the highest source of inquiries, accounting for 32% of inquiries and were tied with off-campus recruiting visits as the highest source of enrolled students, accounting for 18% of enrolled students (Ruffalo Noel Levitz, 2018).

Student list products are selection devices that enable university administrators to select prospective students from a larger pool based on a set of input factors. Student list products are discretionary rather than standardized selection devices. For each purchase, the administrator may choose which inputs to filter on and which thresholds to apply to each filter. Although most College Board search filters measure descriptive characteristics of individual prospects (e.g., SAT score range, state), several search filters utilize predictive analytics (e.g., geomarket, geodemographic cluster).

Scholarship examines whether discretionary and standardized selection devices produce/reduce racial inequality [e.g.,](Benjamin, 2019; Korver-Glenn, 2018; Norris, 2021). Burrell & Fourcade (2021, p. 22) observe that, following 1970s anti-discrimination legislation, many industries adopted actuarial selection devices because "evidence had accumulated that both private and public decision-makers were routinely giving into vague intuitions, personal prejudices, and arbitrary opinions. Reviewing the literature, Hirschman & Bosk (2020) state that standardized selection devices can reduce racial inequality *if* the primary source of inequality is explicit or implicit racial bias from individual decision-makers.

However, standardized selection devices do not eliminate racial inequality stemming from structural racism. Bonilla-Silva (1997) criticizes social science disciplines (e.g., psychology, economics) for defining racism as an ideology held by individuals. These definitions cast attention to the attitudes and behaviors of individuals, ignoring the possibility that broader institutions can be racist. Bonilla-Silva (1997, p. 474) argues that racialized social systems allocate "economic, political, social, and even psychological rewards to groups along racial lines". In turn, structural racism is "a form of 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 and disadvantage marginalized groups. Amidst the growth of "colorblind" selection devices that do not include race as an input, scholarship from sociology finds that selection devices may produce racial inequality by utilizing seemingly neutral or objective determinants that are systematically correlated with race (e.g., Benjamin, 2019; Norris, 2021).

Racialized inputs. Norris (2021) reconstructs Moody's city government credit rating algorithm, which assigns credits scores to cities based on determinants thought to predict loan default. Norris (2021, p. 5) defines racialized inputs "those that are theoretically and empirically correlated with historical racial disadvantage," subjugation, and exclusion. By contrast, non-racialized inputs are "theoretically and emirically orthogonal or distant from racial disadvantage." Median family income is conceptualized as a racialized input; because of historical wage discrimination, income is correlated with race and cities with a greater share of Black residents have lower median income. Once median income is included in Moody's model, percent Black is no longer a significant predictor of city credit rating. Through the inclusion of seemingly neutral racialized inputs, "prior disadvantage and racism against Black individuals becomes institutionalized" [p. 2] and selection devices yield racially disparate outcomes "in ways that escape legibility/cognition as racially unequal" [p. 5].

Geography inputs. Geographic borders are the most commonly studied racialized inputs (Benjamin, 2019; Korver-Glenn, 2022; O'Neil, 2016) [ADD CITES]. These studies build on the fact that American communities and schools are racially segregated as a consequence of historic and contemporary laws, policies, and practices promoting racial segregation (e.g., Harris, 1993; Korver-Glenn, 2018; Rothstein, 2017). Algorithmic selection devices that categorize people based on geographic location without considering structures that produce segregation are likely to reproduce historical race-based inequality in opportunity. For example, O'Neil (2016) analyzes an algorithm uses zip code as an input to predict the probability of recidivism for previously incarcerated people. Because zip codes are correlated with race, using zip code to predict recidivism generates racial inequity in predicted risk.

Methods for using geography as a predictive input were pioneed by geodemography, a branch of market research that estimates the behavior of consumers based on where they live (Burrows & Gane, 2006). Geodemography emerged in the 1980s alongside efforts to fuse marketing and credit scoring (Leyshon & Thrift, 1999), as businesses that depended on customer credit transitioned from the model of approving/rejecting applicants to the more aggressive model of pre-approving desirable customers. Richard Webber (1988, p. 36), director of Experian Marketing Services and creator of Acorn and Mosiac, states:

> Geographical information can. . . be very useful at the recruitment stage. Addresses

> in postcodes with high levels of bad debt can be eliminated as can those where credit referencing activity is particularly low. Area classification systems, such as Mosaic and Acorn, yield further discriminators which can be used to reduce the recruitment of poor credit risks... The combination of all this information into a recruitment scorecard allows the credit operator to select the best possible addresses from rented lists, electoral rolls or the company's own customer file... and enables the recruitment of accounts to be redirected away from areas of high bad debt.

Initially, geodemographic systems scored individual localities (e.g., zip code) based on consumer behavior. Subsequent geodemographic classification systems (e.g., Mosaic by Experian) classify localities and individuals into like audience segments for marketers [CITE].

Predictive analytics inputs. MOVE PARAGRAPH ON PREDICTIVE ANALYTICS FROM START OF CF HERE

Scholarship on algorithmic bias has focused on the use of predictive analytics in selection devices [e.g.,](Burrell & Fourcade, 2021; Noble, 2018; Norris, 2021; O'Neil, 2016). Federal Trade Commission (2016a, p. 4) distinguishes selection devices that rely on descriptive analytics, which seek to "uncover and summarize patterns or features that exist in data sets" (e.g., high school GPA), versus those that rely on predictive analytics, which "refers to the use of statistical models to generate new data" (e.g., predicted probability of enrollment). Products that utilize predictive analytics apply statistical models to previous cases and apply the results of these analyses to predict outcomes for future cases.

In child welfare, for example, the "Structured Decision Making Model" (SDM) is a selection device that yields recommendations about whether children should be placed in protective care based on a standardized survey instrument designed to discern the likelihood of future abuse or neglect (Hirschman & Bosk, 2020).

Predictive analytics are another class of racialized inputs in selection devices. Predictive analytics are often used to predict outcomes for new cases based on analysis of previous cases. Federal Trade Commission (2016a) warns that using historical data to predict the behavior of a new group of people "creates new justifications for exclusion" (p. 10) and Burrell & Fourcade (2021, p. 224) state that "predicting the future on the basis of the past threatens to reify and reproduce existing inequalities of treatment by institutions." For example, predictive models of recidivism identify the determinants of recidivism for past cases and apply these results to predict the probability of recidivism for new cases (O'Neil, 2016). Ex-offenders with high predicted probabilities of recidivism receive greater police attention, which increases the probability of arrest. Harcourt (2007) refers to this phenomenon as the "ratchet effect," whereby disproportionately targeted populations are predicted to have higher risk of an outcome, which amplifies disproportionate targeting. Predictive analytics can also be used as an input. For example, geodemographic methods create geographic customer segments based on past consumer data and use these segments as one input to select target customers (College Board, 2011; Federal Trade Commission, 2014).

3.2 Student List Products

Test-taking and test score filters. RQ1 examines the relationship between individual student list product attributes and racial inequality, independent of how universities utilize student list products. We argue that the underlying architecture of student list products produce structural inequality in two broad ways.

The first source of structural inequality is which prospective students are included in the underlying database. For any selection device or statistical model, when individuals are excluded from an analysis sample because they have missing values for some or all variables, the possibility of selection bias emerges. If missingness is correlated with race, then the selection device yields systematic racial exclusion. Guidance from the Federal Trade Commission states that "If a data set is missing information from particular populations, using that data to build a... model may yield results that are unfair or inequitable to legally protected groups" (Jillson, 2021).

Historically, College Board and ACT student list products exclude students who do not take at least one of their assessments (e.g,. SAT, AP, PSAT).[4] Prior research shows that rates

[4]Recently, College Board and ACT, respectively, began allowing non test-takers to opt into student list products by participating in the College Board Big Future or the ACT Encourage college search engines, but it is unclear how many non test-takers opt in using these resources.

of SAT/ACT test-taking differ by race and class (e.g., Bastedo & Jaquette, 2011; Blake & Langenkamp, 2022; Hyman, 2017). Similarly, the percentage of students who take AP exams vary across race, particularly for STEM exams (Kolluri, 2018), and Black students are more likely than white students to attend a high school with few AP course offerings (Rodriguez & McGuire, 2019). These findings motivate the following proposition, which we analyze separately by assessment (SAT, PSAT, AP) and for taking any assessment.

P1: The condition of taking standardized assessments is associated with racial disparities in who is included versus excluded in student list products.

The second source of structural inequality in student list products is the use of racialized inputs as search filters. We argue that test score filters (e.g., SAT, PSAT, AP) meet Norris (2021) criteria – "theoretically and empirically correlated with historical racial disadvantage" (p.5) – of racialized inputs. Race-based differences in standardized test scores are a function of historical and contemporary segregation of U.S. communities in schools (Reardon, Kalogrides, & Shores, 2019), which drive race-based differences in school funding (Green III, Baker, & Oluwole, 2021) and race-based differences in access to college preparatory curriculum, including SAT/ACT test preparation (Park & Becks, 2015) and access to AP courses (Kolluri, 2018; Rodriguez & Hernandez-Hamed, 2020). Therefore, filtering prospects based on test scores without simultaneously considering the historical and contemporary structural inequalities that drive race-based differences in test scores is likely to reproduce racial inequality in educational opportunity.[5]

P2: As test score threshold increases, the proportion of underrepresented minority students included in student lists declines relative to the proportion who are excluded.

Geographic filters. [KARINA - MODIFY THIS SECTION] Geographic search filters enable universities to target prospects based on where they live. College Board geographic search filters include state, CBSA, county, zip code, geomarket, and geodemographic filters.

We conceptualize geographic search filters as racialized inputs because these filters are built

[5]By comparison, high school GPA is a strong predictor of postsecondary student success (Allensworth & Clark, 2020; Niu & Tienda, 2010) and is more theoretically and empirically distant from historical racial disadvantage (Alon & Tienda, 2007; Posselt, Jaquette, Bielby, & Bastedo, 2012). Therefore, we conceptualize the high school GPA search filter as a non-racialized input in student list products.

on top of historic and contemporary policies and practices promoting racial segregation. Targeting prospective students based on geographic location without consideration to macro and local structures that produce racial segregation is likely to reinforce historical race-based inequality in educational opportunity.

Prior research on recruiting consistently finds that selective private and public research universities disproportionately target affluent schools and communities (Author, XXXX; Salazar, 2022; Salazar et al., 2021; Stevens, 2007). These findings suggest that universities may filter on affluent zip codes when purchasing student lists. We expect that filtering for affluent neighborhoods is positively associated with racial exclusion because structures of racial segregation often prohibit people of color from living in affluent neighborhoods.

P3. As purchases filter on higher levels of zip-code affluence, the proportion of underrepresented minority students included in student lists declines relative to the proportion who are excluded.

University recruiting behavior often targets prospects in particular metropolitan areas (Salazar et al., 2021, 2022). When targeting metropolitan areas, we expect that utilizing finer geographic filters (e.g., zip code rather than county) is associated with greater racial dispairities in student list purchases because American residential segregation occurs at fine-grained geographic levels (Korver-Glenn, 2022).

P4. Filtering on smaller geographic localities is associated with greater racial disparities in included vs. excluded than filtering on larger geographic localities.

3.3 Utilizing Student List Products

Whereas the conceptual discussion thus far has focused on individual search filters, student list products are designed to filter on multiple search filters simultaneously. In contrast to standardized selection devices (e.g., XXXX), student list products are discretionary selection devices that allow administrators to choose which search filters and filter thresholds to select and how many purchases to execute. A university may utilize student list products in ways that reduce or amplify racial inequality in college access. This section draws from scholarship on product utilization to motivate analyses about how public universities utilize racialized input search filters in concert with other search filters when purchasing student lists (RQ2) and to motivate analyses about the racial composition of student list purchases that utilize multiple search filters (RQ3).

We highlight key findings from sociological scholarship on product utilization. On balance, scholarship tends to find that administrative discretion over selection devices causes structural inequality to increase [CITE; e.g.,] (Castilla, 2008; Cotter, Medeiros, Pak, & Thorson, 2021; Norris, 2022). Compared to standardized selection devices, discretionary selection devices allow the explicit or implicit racial bias of individual decision-makers to affect selection decisions (Burrell & Fourcade, 2021; Korver-Glenn, 2018). In the context of occupations/professions (e.g., medicine, social work), discretionary selection criteria reflect occupational/professional norms, which may be more or less cognizant of structural racism (Hirschman & Bosk, 2020; Tiako et al., 2021). For example, Korver-Glenn (2018) shows that homes in white neighborhoods received higher appraisal values than those in non-white neighborhoods because of appraiser discretion in selecting comparison homes for the valuation. Two necessary conditions to prevent ascriptive bias in discretionary selection processes are transparency – decisions about selection criteria are clear to all stakeholders – and accountability – decision-makers face consequences for utilizing racially biased selection criteria (Castilla, 2008; Norris, 2022). Finally, Americans dramatically underestimate the magnitude of racial income inequality (Kraus, Onyeador, Daumeyer, Rucker, & Richeson, 2019). Discretionary selection devices that incorporate racialized inputs are likely to amplify racial inequality because decision-makers may be ignorant about how selection inputs interact with on-the-ground patterns of racial inequality (Korver-Glenn, 2018) and/or because decision-makers may view racialized inputs as objective, colorblind measures of merit (Krippner & Hirschman, 2022; O'Brien & Kiviat, 2018).

Drawing from the utilization literature, we motivate analyses about administrative discretion and racial inequality in student list purchases. First, universities may select academic achievement filters that are consistent with their admissions standards, which are a function of university stakeholders and the admissions profession [CITE]. Although race-based differences in standardized testing may be driven by structural inequality, most admissions offices viewed test scores objective measures of achievement or aptitude (Hoxby & Avery, 2013), until quite recently. Therefore, we expect that selective institutions are more likely to filter on standardized test scores compared to less selective institutions and are likely to filter on higher score thresholds.

Second, filtering on multiple search filters facilitates micro-targeting of desired prospects. Indeed, micro-targeting has become a branding strategy for student list products. College Board Student Search promises to "create a real pipeline of best-fit prospects" (College Board, n.d.) while ACT Encoura uses the tag-line "find and engage your best-fit students" (Encoura, n.d.). Consulting firms encourage universities to execute multiple student list purchases, each targeting different market segments (e.g., Waxman, 2019).[6] Thus, student list products are powerful, complicated tools for precisely identifying prospects. However, the flip-side of micro-targeting is exclusion. Purchased lists do not show how the characteristics of targeted prospects compare to the demographics of their surrounding community. Thus, specifying multiple filters can yield unintended racial inequality because administrators have incomplete knowledge about how the intersections of these filters interact with local patterns of racial segregation. Considering a non-racialized filter (high school GPA), adding a racialized input filter (e.g., SAT score, AP score) may increase racial inequality. Additionally, filtering on multiple structurally racist inputs (e.g., SAT score and zip-code) may have a compounding effect on racial inequality.

Third, amidst the data science revolution, College Board and ACT have added search filters based on predictive analytics (e.g., College Board "Geomarket", ACT "probability of enrollment"). The College Board Geomarket filter sub-divides states/metropolitan areas into distinct markets based on historical data about college enrollment. The College Board Geodemographic Segment filter allocates individual census tracts and individual high schools into distinct clusters based on past college enrollment. Drawing from the sociology of race (@ Burrell & Fourcade, 2021; Harcourt, 2007), creating new geographic borders based on historical patterns for the purpose of including/excluding future students serves to amplify the

[6] For example, in data collected by (Salazar et al., 2022), colleges assigned names to individual searches, such as, "Women in STEM," "International 2022 PSAT 1200," "CA 2021 SAT URM to 1290," "Performing arts," etc.

effect of historic race-based inquality in educational opportunity. Furthermore, administrators utilizing these filters likely have incomplete knowledge about how these borders interact with local patterns of segregation. We expect that using Geomarket or Geodemographic filters, in concert with other search filters, is associated with racial inequality in targeted versus excluded prospects.

Finally, some colleges may utilize student list products to increase enrollment by underrepresented populations. College Board and ACT student list products are designed to facilitate this goal by incorporating filters for race, ethnicity, and first-generation status. For example, colleges may purchase separate lists for particular racial/ethnic groups, specifying different test score thresholds for different groups. Purchases that explicitly target underrepresented populations may be more or less inclusive of students from communities of color. Given the complexity of student list products and incomplete knowledge about race-based income inequality (Kraus et al., 2019), policymakers should be concerned about potential unintended harm from these efforts. Student list purchases designed to overcome one inequality in college access may unintentionally amplify other social inequalities. For example, purchases designed to target "women in STEM" may yield racial or socioeconomic inequality. Additionally, purchases that explicitly target underrepresented minority students with high test scores may systematically exclude students from predominantly non-white communities.

4 Methods

4.1 Data

Our analyses utilize two data sources. First, the primary data source is the High School Longitudinal Study of 2009 (HSLS09) conducted by the National Center for Education Statistics (NCES). HSLS09 is a nationally representative survey that follows a cohort of more than 23,000 students from 944 schools entering the ninth grade in Fall 2009. Follow-up surveys were administered to students in Spring 2012 (when most were in 11th grade), in 2013, in 2016, and NCES collected high school transcripts in 2013-14. HSLS provides the extensive student-level demographic, geographic, and academic variables needed to create academic and geographic filters used within student list purchases.

Our analysis sample includes students who meet all of the following conditions: completed Spring 2012 first follow-up survey; completed 2013 update survey; and obtained high school transcript data. Of the 23,503 respondents included in HSLS09, our unweighted analysis sample consists of the 16,530 students who meet all conditions.[7] The survey weight variable W3W2STUTR is designed for respondents who meet these conditions. After weighting, these 16,530 students represent the population of approximately 4.187 million U.S. 9th graders in 2009.

The second data source consists of "order summaries" for student lists that public universities purchased from College Board. These data are used to inform hypothetical student list purchases in the final set of analyses. As described in Salazar et al. (2022), we collected these data by issuing public records requests to all public universities in five states (CA, IL, TX, MN, and one anonymous state) about student lists purchased from 2016-2020. Salazar et al. (2022) analyzed 830 College Board orders, which yielded more than 3.6 million prospect profiles. These orders were placed by 14 public universities. Figure 3 shows the filters utilized in these orders.

4.2 Variables

Our research question is, what is the relationship between student list search filters and the racial composition of students who are included versus excluded from College Board student list purchases? In turn, our dependent variable measures student demographic characteristics and our independent variables are measures of student list filters. Descriptive statistics for analysis variables are shown in Table 2.

Dependent variable. Our primary dependent variable is the student race/ethnicity composite variable X2RACE, which includes the following seven categories: American Indian/Alaska Native, non-Hispanic; Asian, non-Hispanic; Black/African-American, non-Hispanic; Hispanic; More than one race, non-Hispanic; Native Hawaiian/Pacific Islander, non-Hispanic; and White, non-Hispanic.[8]

Independent variables. Independent variables are measures of student list filters. Choices

[7] Unwighted sample was rounded to 10 to meet restricted data regulations by NCES

[8] We collapse "Hispanic, no race specified" and "Hispanic, race specified" into a single category.

about independent variables were based on our conceptual framework and the set of student list filters observed in our public records request data collection, shown in Figure 3. Our conceptual framework restricts analytic focus to academic filters and geographic filters.

Propositions **P1** and **P2** focus on academic filters. **P1** is concerned about which students take standardized assessments, which determines inclusion in the underlying College Board student list database. **P2** is concerned with test score thresholds utilized to filter prospects. For **P1**, we create dichotomous measures for each of the following assessments (input variables in parentheses) based on test score variables from the high school transcript file: PSAT/PreACT (X3TXPSATCOM); SAT/ACT (X3TXSATCOMP); any AP exam (variables with names that start with X3TXAP); and any STEM AP exam. For **P2**, we use these same input variables to create test score measures for PSAT/PreACT; SAT/ACT; highest AP exam score; and highest AP STEM exam score. We also create a measure of high school GPA in academic courses (X3TGPAACAD), which is a question asked in the pre-test questionnaire of College Board assessments. Consistent with how College Board filters work, **P2** variables are analyzed as categorical rather than continuous variables. To select thresholds for **P2** variables – for example, an SAT score thresholds of less than 1000, 1000+, 1200+, 1300+, etc. – we considered what the product allows, what we observed in orders collected via public records requests, and the goal of parsimony.

Propositions **P3** and **P4** focus on geographic filters. Drawing from Figure 3, we create measures for student county (X2GCNTY), zip code (X2GZIPCD), and CBSA (based on crosswalk with home zip code). Next, we attach income data to localities by merging in data from the American Community Survey (ACS) 2012 5-year estimates. We do not create independent variables for geomarket filter or geodemograhic segment filter because these filters utilize geographic borders based on proprietary College Board data.

4.3 Analyses

Analyses utilize simple descriptive statistics, with appropriate statistical tests. All analyses compare the racial composition of included versus excluded prospects when particular filters and/or filter thresholds are utilized to purchase prospect profiles.

Consider a hypothetical purchase that all prospects took an AP STEM exam. We compare the racial composition of the included group to the racial composition of the excluded group. For example, Black students comprise 5.05% (`=91/1803`) of AP STEM test-takers and Black students comprise 10.6% (`=1564/14722`) of students who do not take an AP STEM exam. The test for difference in proportions compares whether the proportion of included prospects who identify as Black differs from the proportion of excluded prospects who identify as Black, and this test is run separately for each race/ethnicity group. This comparison focuses on the racial composition of prospects targeted from the university perspective; that is, what is the racial composition of prospects who are targeted by a particular set of filters versus the racial composition of prospects who are excluded by these filters?

Analyses for propositions P1 through P4 examine purchases that utilize individual filters in isolation. The final set of analyses examine purchases that utilize academic and geographic filters in combination with one another, with choice of filters informed by commonly observed combinations from the public request data and also by theoretical considerations.

4.4 Limitations

This manuscript uses HSLS09 to recreate the College Board Student Search Service. One limitation is that HSLS variables for SAT test-taking and test scores also include ACT test-takers, with ACT scores converted to the SAT scale. The same is true for the PSAT and PreACT. The Student Search service includes students who take at least one College Board assessment, but we cannot differentiate between College Board and ACT test-takers, so our analyses incorrectly treat ACT test-takers as College Board test-takers. We considered restricting the analysis sample to states where the majority of students take the SAT rather than the ACT. We chose not to take this approach because the ACT "Educational Opportunity Service" student list product – now, named Encoura – includes academic and geographic filters that are nearly identical to the College Board filters that are the focus of this manuscript (Schmidt, 2022). Thus, analyses can be interpreted as who would be included/excluded by both College Board and ACT student list products.

Second, test-takers have the opportunity to opt-out of the College Board Student Search Service and the ACT Educational Opportunity Service but HSLS09 has no reasonable proxy for whether students opt-in or opt-out. Moore (2017) finds that 86% of ACT test-takers opt-in, but does not investigate the student characteristics associated with opting in. Third, the HSL09 cohort pre-dates the increase in test-optional admissions policies and decline in test-takers which occurred since the onset of COVID19. This undermines the external validity of our findings with respect to current cohorts of high school students. That said, several for-profit vendors have developed student list products (e.g., Intersect by PowerSchool) poised to acquire market share ceded by College Board and ACT, and these products use filters that are similar to College Board and ACT products (Feathers, 2022). Our analysis of structurally racist inputs and exclusion yields insights across student list products. Fourth, we could not make measures for high school class rank, an academic filter, or for geomarket and geodemographic filters, which utilize proprietary College Board data.

5 Findings

5.1 Student List Products

Individual Filters. We address RQ1 by first describing the racial characteristics of HSLS09 students who completed standardized assessments in comparison to those who did not, which would determine inclusion in the underlying College Board student list database. Figure 4 presents the racial/ethnic composition of prospects included (i.e., completed assessment) and excluded (i.e., did not complete assessment) across SAT, PSAT, and AP exams. For example, the top left graph shows that more than 1.8 million prospects completed the SAT and would have presumably been included in the College Board student list database. In comparison, more than 2.3 million prospects did not complete the SAT and would be excluded from the database. White students make up 53% of included students who completed the SAT and 51% of excluded students who did not. Table 3 reports statistical tests for proportions between included and excluded students by race/ethnicity. Differences in White student proportions across included and excluded prospects are statistically significant ($p<0.000$). While Asian and Multiracial students make up nearly equal proportions in both included and excluded prospects, Hispanic students make up only 17% of included prospects relative to 26% of excluded prospects ($p<0.000$). Black students similarly make up 12% of included prospects but 15% of excluded prospects ($p<0.000$).

Other standardized assessments resulted in similar included prospects that were on average made up of larger proportions of White and Asian students and smaller proportions of Hispanic, Black, and American Indian/Alaska Native students than excluded groups, lending support for Proposition P1. For example, Figure 4 also shows the composition of included versus excluded prospects by AP exam completion in any subject on the top right panel. Similar to SAT, White (54% versus 51% White) and Asian students (8% versus 3%) make up statistically significant ($p<0.000$) larger proportions of included prospects. While an equal proportion of included and excluded students are Hispanic (22%), Black students make up a smaller share ($p<0.000$) of included prospects (8%) than excluded prospects (15%). When inclusion versus exclusion is determined by completing an AP exam in a STEM subject, total included students declines to nearly half (383,669) of those included via completing any AP exam. Moreover, Black (7% versus 14%), Hispanic (19% versus 22%), Multiracial (7% versus 14%), and American Indian/Alaska Native (0.2% versus 0.7%) students make up smaller statistically significant proportions of the included prospects relative to excluded prospects based on completion of an AP STEM exam.

Proposition P2 suggests the proportion of underrepresented minority students included in student lists decline relative to the proportion who are excluded as assessment score thresholds increase. In order to test this proposition, we analyze the racial composition of included versus excluded students at minimum score thresholds commonly used across student list purchase orders for SAT, PSAT, and AP exams. For example, Figure 5 presents these results for SAT (left panel) and PSAT assessments (right panel). For the top left panel, each bar represents the racial composition of included prospects who completed the SAT exam and scored at the minimum threshold indicated. On the bottom left panel of Figure 5, each bar represents the racial composition of excluded prospects who did not complete the SAT exam in addition to students who did complete the exam but did not meet the minimum score threshold indicated. Statistical tests for differences in proportion for Figure 5 are reported in online appendices for space considerations.

As SAT score thresholds increase from less than 1000 to greater than 1400 in Figure 5, proportions of included White and Asian students increase while proportions of included Hispanic and Black students decrease. For example, White students make up a statistically significant ($p<0.000$) smaller share of included (47%) than excluded (53%) prospects scoring less than 1000 on the SAT, which results in an equal share of Hispanic students (22%) and a greater share of included Black students (19% versus 12%) relative to excluded prospects at this score threshold. However, Hispanic student proportions in included versus excluded prospects decrease to 12% versus 25% at scores greater than 1000, 9% versus 23% at scores greater than 1200, and down to 5% versus 22% at scores greater than 1400. Similarly, Black student proportions in included versus excluded prospects decrease to 6% versus 16% at scores greater than 1000, 4% versus 14% at scores greater than 1200, 2% versus 14% at scores greater than 1300, and down to making up 0% of included prospects at scores greater than 1400. These proportional differences across score thresholds are statistically significant ($p<0.05$) for both Hispanic and Black students (online appendix).

While making up relatively small proportions of the overall sample, declines in proportions of American Indian/Alaska Native students and Native Hawaiian/Pacific Islander students within included versus excluded prospect groups are statistically significant as score thresholds increase (online Appendix A). In order to more equitably capture these differences, we report the number of students rather than their overall representational proportion within included versus excluded groups. For instance, more than 7,600 American Indian/Alaska Native students and nearly 2,500 Native Hawaiian/Pacific Islander students are represented in the included prospects relative to the more than 20,800 and 16,300 represented in the excluded prospects at SAT scores less than 1000, respectively. However, American Indian/Alaska Native students decline to zero and Hawaiian/Pacific Islander students decline to 435 students in the included prospects group by the 1300 or greater SAT score threshold.

PSAT results are also shown in Figure 5 for composite scores that range from 60 to 240. [9] Similar to SAT, as PSAT composite score thresholds increase from less than 120 to greater than 220, proportions of included White and Asian students increase while proportions of included Hispanic and Black students decrease relative to excluded prospects. Online appendices show all comparisons between included and excluded students across PSAT score

[9] PSAT exams taken 2014 or before receive composite scores that range from from 60 to 240. PSAT exams taken 2015 or later are scored via a range from 320 to 1520. Our lower bound PSAT composite score thresholds of 120, 170, 200, and 220 for HSLS students who completed the exam prior to 2014 equate to minimum score thresholds of 890, 1220, 1410, and 1510 on the 2015 or later PSAT scale, respectively. CITE

thresholds are statistically significant at the p<0.000 level, with the exception of multiracial students at the 220 or greater minimum score threshold.

We find similar racial disparities in included versus excluded prospects across AP exam score thresholds, providing strong support for Proposition P2. Figure 6 shows similar results as Figure 5 for AP exams. As AP score thresholds for any subject exam (left panel) increase from one to five, proportions of included White and Asian students increase while proportions of included Hispanic, Black, Multiracial, American Indian/Alaska Native, and Native Hawaiian/Pacific Islander decrease relative to excluded prospects. For example, the 110,360 included prospects (relative to excluded prospects) who had a score of one on any subject AP exam were on average 38% White (52% excluded), 6% Asian (4% excluded), 27% Hispanic (22% excluded), and 21% Black (14% excluded). By an AP score threshold of four or greater, included prospect proportions shift (relative to excluded) to 60% White (51% excluded), 10% Asian (3% excluded), 20% Hispanic (22% excluded), and 4% Black (4% excluded).[10] Similar patterns are evident for AP STEM exam completion (right panel).

Proposition P3 suggests as purchases filter on higher levels of zip code affluence, the proportion of underrepresented minority students included in student lists will decline relative to the proportion who are excluded. In order to test this proposition, we analyze the racial composition of included versus excluded students when filtering by zip code median household income. In order to deal with median household incomes varying widely across the U.S., we categorized all zip codes into percentiles based on levels of median household income within their respective Core Based Statistical Areas (CBSA). For example, median household income percentiles based on the 378 zip codes within the Los Angeles metropolitan area are \$55,256 at the 20th percentile, \$70,804 at the 40th percentile, \$89,709 at the 60th percentile, and \$108,316 at the 80th percentile (in 2022 CPI). So the Los Angeles zip code 92649, which captures parts of the Huntington Beach area, with a median household income of \$109,159 (in 2022 CPI) would be categorized as zip code in the 80th percentile of affluence within CBSA. This approach also aligns with common ways in which student list orders purchase prospect's contact information by filtering on zip codes within specific CBSAs.

[10]Proportional differences for these specific racial/ethnic categories at reported score thresholds are statistically significant at the p<0.000 level and reported in online appendices

Figure 8 presents the racial composition of zip codes that included (top panel) versus excluded prospects (bottom panel) when filtering based on percentile of affluence within CBSA. The figure suggests that as zip code affluence increases, included prospects have larger proportions of White students and smaller proportions of Hispanic and Black students relative to excluded prospects. For example, Hispanic and Black students make up 30% and 27% of included prospects and 20% and 11% of excluded prospects at zip codes below the 20th percentile of affluence, respectively. The proportions of Hispanic and Black students within included prospects decline as zip code affluence increases up though the 89th percentile. For zip codes in 90th percentile or higher of affluence within CBSA, the proportions of Hispanic students within included prospects declines to 11% relative to making up 23% of excluded prospects ($p<0.000$). Similarly, Black students make up 9% of included prospects relative to making up 14% of excluded prospects ($p<0.000$) within the most affluent zip codes. [11]

These findings suggest that purchases filtering on higher levels of zip-code affluence lead to smaller proportions of underrepresented minority students included in student lists relative to the proportion who are excluded, providing support for Proposition P3. However, we acknowledge that categorizing zip codes within CBSA limits the number of rural zip codes captured within the included prospect groups. By categorizing zip code affluence within CBSA, only rural zip codes within micropolitan statistical areas (i.e., areas that have at least one urban cluster of at least 10,000 people with commuting ties to adjacent metropolitan areas that have higher degrees of social and economic activity) will be captured via CBSA.

5.2 Utilizing Student List Products

Filter Usage. Our second research question asks in what ways do public universities use racialized input search filters in concert with other search filters. We begin by analyzing how often filters were used for actual College Board students list purchases across research versus ma/doctoral universities in our project sample. While student list purchases typically filter on multiple criteria, Figure 3 illustrates the prevalence of each individual filter by institution type. Beyond all orders using a high school graduating class filter, we group filters across academic, geographic, demographic, and student preferences. Across both research

[11] Online appendices report statistical tests for proportions between included and excluded students by race/ethnicity for zip code affluence.

and ma/doctoral universities, commonly used academic filters include GPA , SAT , and PSAT. Compared to ma/doctoral universities, research universities were less likely to filter on GPA or SAT score and more likely to filter on PSAT score. Geographic filters used to purchase student lists differed substantially across institution type. Orders by research universities most commonly used a state filter whereas ma/doctoral universities used a zip code filter. Research universities also used filters utilizing geography as a predictive input for assessing the college going behavior of prospective students living in those areas (e.g., geomarket, segment). Although not as commonly used as academic and geographic filters, research universities also used demographic and student preference filters. Filters for race, low socioeconomic status, and gender were the most commonly used demographic filters, whereas college size, major, college type, and location were most commonly used student preferences.

Table 1 shows the top ten filter combinations used for student list purchases across institution type for our project sample. [12] For ma/doctoral universities, the top 10 filter combinations account for 95% of all orders. This is a function of nearly half of all orders using a combination of high school graduation class, zip code, SAT scores, and GPA as filters and another 32% of orders using the same filters with PSAT scores instead of SAT. For research universities, the top ten filter combinations account for 51% of all orders. The most common filter combination, making up 10% of all orders, included high school graduation class, state, race, SAT, PSAT, GPA, and high school rank. The second most common combination only used high school graduation class, state, and PSAT score filters, which were also used for the remaining top five combinations in addition to filters like race, gpa, low-socioeconomic status, and students' preference for nationally recognized programs.

Racial Composition of Resulting Lists. Our last research question focuses on assessing the racial composition of student list purchases that utilize racialized input search filters in concert with other search filters. We begin by simulating student list purchases via filtering the HSLS09 sample based on common combinations of filters used by public and ma/doctoral

[12] While the majority of orders specified multiple filters as "AND" conditions, some orders specified particular filters as "OR" conditions (e.g., SAT score in some range or PSAT score in some range). Additionally, filter combinations are skewed by universities that made large numbers of orders relative to other universities

universities in Figure 3. For example, Figure 11 is a simulated student list resulting from combining the two most common academic filters: GPA and SAT. Drawing on utilization literature, high school GPA is a strong predictor of postsecondary student success (CITE) and is less likely to be theoretically and empirically correlated to racial disadvantage than standardized test scores. On the other hand, universities may filter on test score thresholds that are consistent with their admissions standards (e.g., more prestigious universities recruit and enroll students with higher test scores) (CITE) but hold race-based differences driven by structural inequality. We therefore explore the racial composition of prospects included in student lists when filtering on GPA greater than or equal to 3.0 while simulating increases to minimum thresholds for standardized tests.

The top panel of Figure 11 presents the racial composition of prospects included when filtering on GPA greater than or equal to 3.0 across SAT thresholds at increments of 50 beginning at scores just above the HSLS09 sample median of 1010. For space considerations, we only present included prospect groups across all combinations. The figure suggests that even at the the lowest SAT score, White students make up much larger proportions of included prospects while Black and Hispanic students make up significantly smaller proportions when filtering for both GPA and SAT. For example, White students make up 72% of included prospects when filtering for GPAs greater than or equal to 3.0 in combination with SAT scores greater than 1050, whereas Hispanic and Black students make up 10% and 3%, respectively. Racial disparities only grow as SAT thresholds increase. Moreover, these racial disparities are greater than when filtering for similar thresholds for GPA (Figure 7) and SAT score (Figure 5) individually. The bottom panel of Figure 11 suggest similar results are evident when combining a GPA filter greater than or equal to 3.0 and a PSAT filter. While Hispanic and Black students make up larger proportions at lower thresholds of PSAT in comparison to SAT when combined with GPA, the racial disparity for Black students is still greater in the combination of filters than filtering for similar thresholds of PSAT score (Figure 5) individually.

We also simulate the racial composition of included prospects when filtering on both GPA and AP scores in Figure 13. The top panel presents the racial composition of prospects included when filtering on GPA greater than or equal to 3.0 while simulating increases in AP scores in any subject exam. Across all AP score thresholds, White students make up a larger proportion while Black and Hispanic students make up smaller proportions of included prospects when filtering for both GPA and AP relative to filtering for these individually. For example, White students make up 49% of included prospects when filtering for GPAs greater than or equal to 3.0 in combination with an AP score of 1 in any subject exam, whereas Hispanic and Black students make up 17% and 21%, respectively. Hispanic and Black student proportions decline as AP thresholds increase. This decline is most significant for Black students, which result in less than 1% of Black students making up included prospects when filtering for GPAs greater than or equal to 3.0 in combination with an AP score of 5. Moreover, these racial disparities are greater than when filtering for similar thresholds for GPA (Figure 7) and AP score (Figure 6) individually. The bottom panel of Figure 13 suggest similar results are evident when combining a GPA filter greater than or equal to 3.0 and an AP STEM exam filter.

[REVISE THIS BASED ON GEOGRAPHY DECISION] In order to assess the microtargeting effects of combining academic and geographic filters, Figure 12 adds a zip code filter to the GPA and SAT/PSAT order simulations presented above. We again deal with median household incomes varying widely across the U.S. by categorizing all zip codes into percentiles based on levels of median household income within CBSAs. The top panel of Figure 12 presents the racial composition of included students when filtering for GPAs greater than or equal to 3.0, SAT scores greater than or equal to 1050, and zip codes at various levels of affluence. In comparison to racial disparities in included versus excluded prospects driven by just zip code affluence in Figure 5, the combination of zip code with GPA and SAT filters leads much greater disparities even at lower levels of affluence. For example, Figure 12 shows White students make up 72% of included prospects when filtering for GPAs greater than or equal to 3.0 in combination with SAT scores greater than 1050 within the lowest income zip codes (<20th percentile), whereas Hispanic and Black students make up 9% and 7%, respectively. The proportions of Hispanic and Black included prospects resulting from the combination of filters are considerably lower than the 30% of Hispanic and 27% of Black included prospects resulting from only filtering by zip code (Figure 8). Greater racial disparities result from the the combination of filters across all levels of zip code affluence in comparison to only filtering by zip code, although proportional differences are modest at higher incomes. Similar patterns are evident when combining similar zip code and GPA filters with a PSAT filter for composite scores greater than or equal to 150.

By drawing on actual student lists purchased by universities in our project sample, we can explore the effects of utilizing filters beyond admissions standards and microtargeting. For example, while this study is unable to use HSLS09 to recreate predictive analytics filters that subdivide geographic areas into distinct markets based on past college enrollments, we can analyze the resulting racial composition of student lists using such filters via our project sample. Several universities in our project sample that used College Board's Segment filter. To build Segment, College Board integrates information about test-takers and their neighborhood and school – including historical college going behavior. These data are used used "to group the 33,000+ high schools and 44,000 neighborhoods into 29 unique high-school types and 33 unique neighborhood types" (College Board, 2011b, p. 4), resulting in high school (HS) clusters HS:51-HS:79 and educational neighborhood (EN) clusters EN:51-EN:83. When making student list purchases, a universities may purchase the profiles of prospects who are associated with particular combinations of neighborhood and high school cluster (e.g., neighborhood cluster EN:73 and high school categories HS:65 or HS:70).

Table 4 and Table 5, recreated from College Board (2011b), show the characteristics of Segment neighborhood clusters and school clusters, respectively. Drawing on our conceptual framework, Table 4 and Table 5 suggest that predictive analytics use to create clusters are likely drawing on historical and geographical correlations between racial and income demographics. For example, neighborhood cluster EN78 is 26% nonwhite and has median income of $134,400 while neighborhood cluster EN:71 is 97% nonwhite and has median income of $42,661.

We therefore analyze the racial composition of student lists from eight orders by a public research university that utilized Segment filters and specified very similar academic criteria across orders. These eight orders targeted 2019-2023 high school graduating classes, and resulted in 131,562 prospects whose profiles were purchased. All eight Segment orders filtered on GPAs ranging from a low of B- to a high of A+. The orders specified minimum PSAT/SAT scores ranging from 1220-1240 and maximum PSAT/SAT scores of 1450. Student lists were also geographically filtered across state/CBSAs and segments. All eight orders filtered on the same Segment high school and neighborhood clusters, which are highlighted in Table 4 and Table 5. [13]

Figure 14 compares racial and income the characteristics of prospects whose profiles were purchased via Segment to the characteristics of all high school students in the top four metropolitan areas where many prospects whose profiles were purchased reside: New York (27,932 prospects, rank #1), Los Angeles (12,307 prospects, rank #2), Philadelphia (9,126 prospects, rank #3), Washington, DC (5,728 prospects, rank #4). For each metropolitan area, we show two figures: on the left column, we show the racial composition of prospects whose profiles were purchased living in the metropolitan area compared to the racial composition of all public high school students in the metropolitan area; on the right column, we show the income of prospects whose profiles were purchased living in the metropolitan area – defined as the average median household income of prospects' home zip codes– compared to the overall median income of the metropolitan area.

For New York, Figure 14 shows White and Asian students comprised 58% and 27% of prospects whose profiles were purchased via Segment, respectively, compared to making up 30% and 9% of students in public high schools in the metropolitan area. By contrast, Black and Latinx students comprised just 1% and 8% of prospects, respectively, compared to 26% and 34% of students in public high schools. Furthermore, prospects whose profiles were purchased lived in zip codes that were much more affluent – an average of $153,000 – than the overall New York metropolitan area median income of $91,000. Figure 14 shows similar racial patterns in the other three metropolitan areas.

[13]All orders filtered for the following school and neighborhood clusters combinations: 1) Neighborhood cluster 51, with any high school cluster; 2) Neighborhood cluster 53, with high school cluster 70; 3) Neighborhood cluster 58, with any high school cluster; 3) Neighborhood cluster 60, with high school clusters 65, 70, or 79; 4) Neighborhood cluster 61, with high school cluster 65; 5) Neighborhood cluster 63, with high school clusters 68 or 70; 6) Neighborhood cluster 69, with high school clusters 65 or 79; 7) Neighborhood cluster 70, with high school clusters 65, 68, 70, or 75; 8) Neighborhood cluster 73, with any high school cluster; 9) Neighborhood cluster 78, with high school cluster 66; 10) High school cluster 79, with any neighborhood cluster.

We also draw on the project sample to explore how colleges may utilize student list products to increase enrollment by underrepresented populations. For instance, some universities in the study made orders targeting prospective students who are women interested in science, technology, engineering, and math (STEM) via two different filter combinations. The first combination used SAT scores, GPA (ranging from a low of B and high of A+), a state filter (in-state versus out-of-state), and prospects' self-reported intended major. SAT score filters for these orders ranged from 1200 to 1600 for in-state prospects and 1300 to 1600 for out-of-state prospects. The second pattern also used the same GPA and state filters, but STEM interest was proxied via AP test scores. Orders for in-state prospects filtered for scores ranging from 3 to 5 on AP STEM tests whereas orders for out-of-state prospects filtered for scores ranging from 4 to 5 on AP STEM tests.

We analyze the resulting student lists from Women in STEM orders – focusing on prospects from four metropolitan areas – from one public research university in order to analyze in-depth patterns in the racial and economic characteristics of prospects that result from the combination of achievement, geographic, and gender filters used to target women interested in STEM. Because nearly 85% of prospects from Women in STEM orders were out-of-state prospective students (N=10,668), we select four out-of-state metropolitan areas and compare prospective students to the characteristics of public high school women students in those metropolitan areas. [14]

Figure 15 compares racial and income characteristics of prospective students whose profiles were purchased to the characteristics of all female public high school students in each metropolitan area. For each metropolitan area, we show two figures. The figure on the left column provides the average median household income for the overall metropolitan area, for prospects whose profiles were purchased using AP scores, and prospects whose profiles were purchased using SAT scores. Prospect-level economic characteristics are measured by taking the average median household income of all prospects' home zip codes. The figure on the right column provides the racial/ethnic composition of all public high school women 12th

[14] The four out-of-state metropolitan areas were selected based on where the greatest number of prospects whose profiles were purchased and based on regional variation: New York (1,663 prospects, rank #1), Atlanta (548 prospects, rank #2), Chicago (485 prospects,rank #3), and Seattle (279 prospects, rank #11).

grade students in the metropolitan area, of prospects whose profiles were purchased using AP scores, and of prospects whose profiles were purchased using SAT scores.

For example, Figure 15 shows the overall median household income for the Atlanta metropolitan area is $76,000. Relative to this overall median household income, Women in STEM prospects whose profiles were purchased by the university across both types of orders are more affluent. Prospects that scored a 4 or 5 on an AP STEM exam live in Atlanta zip codes where the average median household income is $96,000. Whereas prospects that scored a 1200 to 1600 on the SAT and indicated an interest in STEM majors live in Atlanta zip codes where the average median household income is $105,000. Women in STEM prospects are also consistently more affluent than their overall metropolitan areas across Chicago, New York, and Seattle, with prospects whose profiles were purchased through SAT filters having the highest overall incomes.

Figure 15 also shows the racial/ethnic composition of prospects relative to public high school women students in their metropolitan areas. For example, public high school women students in New York are 29% White, 9% Asian, 26% Black, and 34% Latinx. However, Women in STEM prospects from New York that scored a 4 or 5 on an AP STEM exam whose profiles were purchased by the university are 38% White, 44% Asian, 3% Black, 11% Latinx, 5% multiracial. Only 2 women in STEM prospects in the New York metro whose profiles were purchased via AP scores identified as Native American. These racial disparities are most pronounced in orders using SAT scores. Of the 821 prospects from the New York metro that scored a 1200 to 1600 on the SAT and indicated an interest in STEM majors, 54% were White, 43% Asian, and 3% multiracial. Only one prospect identified as Latinx and zero prospects identified as Black or Native American. Similar patterns in racial/ethnic composition of Women in STEM prospects are also evident across Atlanta, Chicago, and Seattle.

6 Discussion

CUT/RANDOM TEXT THIS ARTICLE IS INTERESTING; COULD USE WHEN DISCUSING CREDIT REPORTS https://nationalmortgageprofessional.com/news/42294/using-credit-bureau-data-your-marketing

SEE IF YOU CAN WORK CELLINI ARTICLE ON ADVERTISING SPEND INTO THE MANUSCRIPT (although see Cellini & Chaudhary (forthcoming) analysis of advertising spend by colleges)

In turn, enrollment management consulting firms buy names on behalf of colleges, utilize predictive models help colleges "build the class you want more efficiently than ever" (Ruffalo Noel Levitz, 2023b) by identifying "right fit" prospects "who align with your enrollment goals" (Ruffalo Noel Levitz, 2023a).

Often, student list purchases are also informed by predictive models (Fire Engine RED, 2022).

The test-optional movement and growth in the number of high schools using college planning software threaten the competitive advantage of College Board and ACT in the domain of coverage.

Prior scholarship on recruiting (e.g., Salazar et al., 2021) assumes that recruiting is done by individual colleges and universities. Universities identify prospective students by purchasing student lists, but prior research has not investigated how student list products structure the connection between universities and prospective students. We ask, what is the relationship between student list search filters and the racial composition of students who are included versus excluded in student lists purchased from College Board? We develop a conceptual framework about structural racism in algorithmic products by drawing from sociology and critical data studies. Structurally racist inputs are determinants of a selection device that are correlated with race because historically dominated racial groups have been historically excluded from the input (Hirschman & Bosk, 2020). We propose that several academic filters and geographic filters are structurally racist inputs. We assess propositions about the relationship between search filters and racial exclusion using a nationally representative sample of 9th graders from 2009.

Results for Proposition P1 suggest conditioning on test-taking is associated with racial disparities in included versus excluded prospects across SAT, PSAT, and AP exams. Test-takers are on average made up of larger proportions of White and Asian students and smaller proportions of Hispanic, Black, and American Indian/Alaska Native students. While some proportional differences between test-takers were modest, these proportions determine inclusion in the underlying College Board student list database and are then exacerbated by filtering across score thresholds in results for Proposition P2. When analyzing geographic filters, results for Proposition P3 suggest that included prospects have larger proportions of White students and smaller proportions of Hispanic and Black students as student list purchases filter on higher levels of zip code affluence. Additionally, results for Proposition P4 find that filtering at smaller levels of geography (zipcode) is associated with larger racial disparities between included versus excluded Hispanic and Black prospects in comparison to filtering at larger levels of geography (county).

Lastly, results suggest filtering on multiple criteria compound the effect of racial disparities between included versus excluded prospects. For example, we find that combining multiple academic filters such GPA and SAT/PSAT scores or GPA and AP scores leads to larger proportions of White students and smaller proportions of Black and Hispanic students within included prospects. Moroever, racial disparities between included versus excluded prospects are generally larger when combining multiple academic filters than when filtering at similar thresholds for any one filter individually. Similarly, orders that combine academic filters (GPA, SAT, PSAT) with zip code lead to lower proportions of Hispanic and Black students within included prospects across all levels of affluence and at greater disparities (relative to excluded prospects) than when filtering for zip code alone.

These results have policy implications for federal regulatory agencies concerned with consumer protection and equality of opportunity. Consider zip code filters. Given the history of racial segregation, there is no equality of opportunity rationale for products that enable universities to target particular zip codes. Over the last decade, the Federal Trade Commission (FTC) has become concerned about algorithmic products that "categorize consumers in ways that can result in exclusion of certain populations" (Federal Trade Commission, 2016a, p. 9). The FTC enforces the FTC Act, which applies to all organizations engaged in interstate commerce. Section 5 of the FTC Act prohibits "unfair" practices, defined as practices that meet three criteria: (1) causes substantial harm to consumers; (2) harm cannot be reasonably avoided; and (3) harm not outweighed by benefits to other consumers and to competition (FDIC, 2018).

Zip code filters may cause substantial harm to consumers (criterion #1) because students who live in nearby non-targeted zip codes are excluded from college access opportunities. Consumers cannot reasonably avoid the injury (criterion #2) because they cannot easily move to a different zip code. The benefit to targeted consumers may not outweigh the harm to excluded consumers (criterion #3).

This manuscript is the first word on student list products, not the last word. Other filters may satisfy the FTC unfair practices criteria more unequivocally than zip code. Future research should examine filters based on predictive analytics, which model past cases to make predictions about future cases. One example is ACT's "Enrollment Predictor" filter, in which "every student in the Encoura®Data Cloud is scored on their likelihood to enroll at your institution" (Schmidt, 2022). College Board developed several geographic filters that create geographic borders based on historic, proprietary data on college enrollment. The "geomarket" filter carves metropolitan areas into distinct markets. Geodemographic segment filters utilize cluster analysis to allocate individual high schools and individual census tracts into distinct clusters based on historic college-going behavior. The analysis of Moody's city government credit rating algorithm by Norris (2021) suggests that these filters can be recreated – or closely approximated – using publicly available data sources.

Another topic for future research is demographic search filters, which allow universities to target prospects by race, ethnicity, gender, and first-generation status. The equity rationale is that these filters facilitate access for underrepresented populations, particularly in a post affirmative action landscape. However, analyses by Salazar et al. (2022) found that purchases that filtered for underrepresented racial/ethnic groups often disproportionately targeted students from affluent, predominantly white schools and communities. Additionally, "women in STEM" purchases yielded profound racial and socioeconomic inequality.

In addition to empirical analyses, legal scholarship informs how regulatory agencies interpret the law. For example, Lawler & Dold (2021) argues that the Consumer Financial Protection Bureau (CFPB) has regulatory authority over for-profit colleges because organizations that provide financial advisory services to consumers seeking loans – the activities of any financial aid office – are "covered persons" under the Consumer Financial Protection Act (CFPA).

Similarly, legal scholarship can inform how the FTC and the CFPB intepret regulatory authority over student list vendors and products. One issue is whether and how student list vendors can be regulated as "consumer reporting agencies," which are regulated by the Fair Credit Reporting Act and the CFPA. A consumer reporting agency is an entity that sells information about prospective consumers that leads to the extension of credit (i.e., loans) (15 U.S.C §1681a). Student list vendors may qualify as consumer reporting agencies because of the systematic link between student lists and student loans. That is, the first stage of the enrollment funnel is to identify "leads" by purchasing student lists and the last stage is convert admits to enrolled students by offering financial aid packages.

The broader contribution of this manuscript is to motivate critical education policy research that focuses on third-party products and vendors. The majority of policy research in education analyzes students, schools, or universities, often in relation to federal, state, or local policies. Scholarship from critical data studies and sociology shows that structural racism is "a feature, not a bug" of digital platforms (Benjamin, 2019; Hicks, 2017; Noble, 2018) because racial exploitation is the defining feature of capitalism (Robinson, 2000) and the defining feature of platform capitalism (Cottom, 2020). By contrast, Nichols & Garcia (2022) observes that scholarship on technology and education is dominated by technocratic analyses of instruction and student learning outcomes. The nascent "platform studies in education" literature urges education research to follow the example of critical data studies and "go beyond pedagogical and technical questions toward social, political, and economic critiques" (Napier & Orrick, 2022, p. 207). However, this literature has not yet investigated how platforms structure educational opportunity along racial, class, and geographic dimensions. We propose an empirical literature on third-party products and vendors in education that bridges scholarship on education policy and platform studies. This literature will incorporate structural theories of inequality and theories of organizational behavior from sociology and economics. Student list products represent a model topic for this literature.

One thread within this research agenda examines the business models of edtech platform capitalism (e.g., Komljenovic, 2022; Williamson, 2021). "Data rent" refers to "digital traces" created by users interacting with a platform (Sadowski, 2020), which often become the basis for new products. Drawing from Marx (1978), Sadowski (2019) develops the concept “data as capital” to describe how platforms monetize user-data. The formula $M - C - M'$ represents economic capital, whereby money M is invested to produce commodity C, which is sold for a larger amount of money M' (Marx, 1978). Student list products exemplify this process. Student list data are derived from the user-data of students laboring on a platform, whether that be taking a standardized assessment or searching for scholarships on a free college search engine. Processes that profit from student list data follow the formula for economic capital. College Board uses the cycle $M - C - M' - C - M''$, investing money (M) to create tests (C), which are sold to households for M' and also yield student list data (C), which are sold to universities (M'') looking for students. New entrants to the market for student list data (e.g., PowerSchool, EAB) add another link to the cycle. Instead of selling names at a price-per-prospect (e.g., \$0.50) like College Board, they wrap proprietary databases of prospects within software-as-service products that recruit these prospects (e.g., Intersect, Enroll360), which are then sold to universities for an annual subscription.

Future critical policy research should also examine how college access is structured by vendors and consultancies in the broader enrollment management industry. Many universities depend on enrollment management consulting firms to recruit students (Jaquette et al., 2022). In our public records data collection, roughly half the public universities outsourced student list buys to consulting firms (Salazar et al., 2022). These universities tended to be uninformed about who they were recruiting. Several enrollment management consultancies sell algorithmic products designed to make recommendations about list buys (Fire Engine RED, 2022; Ruffalo Noel Levitz, 2021). For example, Ruffalo Noel Levitz offers an algorithm that tells universities how many names to buy from each zip code (James Madison University, 2017). Beyond name buys, consultancies develop and implement strategy about digital advertising, direct mail, which high schools to visit, and tuition pricing and financial aid. To the extent that unviersities outsource enrollment management to consultancies, these consultancies substantially structure college access. However, extant scholarship on college access assumes that universities perform these functions in-house.

The enrollment management industry also structures “student success” in higher education and K-12. For example, EAB's Starfish student success software incorporates university administrative data, predicts student outcomes using past data, issues notifications when students get "off-track," and "make[s] student intervention easy and integrated" (EAB, 2022). A growing number of scholars are harnessing advances in machine learning to predict student success (e.g., Cardona, Cudney, Hoerl, & Snyder, 2022). Commercial student success products also utilize machine learning, but scholarship has not investigated the fairness of third-party predictive models sold to universities. Do these models use race/ethicity as an input? Do commercial student success platforms achieve higher graduation rates by pushing students out of certain majors?

Difficulty obtaining data is an obstacle to empirical scholarship on third-party vendors and products. Pasquale (2015) notes that "deconstructing the black boxes of Big Data isn't easy" because platform capitalism creates intentional barriers to inspection (p.6). Cottom (2020) argues that "administrative opacity is a deliberate strategy to manage regulatory environments. It shields organizations, both public and private, from democratic appeals for access and equity" (p. 443).

Student list products exemplify this opacity. College Board began selling student lists in 1972 (Belkin, 2019), but prior research never investigated student list products because few people know they exist and because of difficulty obtaining the data. We spent three years – years we would like back! – attempting to collect data about student list products by issuing public records requests to public universities (Salazar et al., 2022). We gained traction only after obtaining pro bono representation from four multinational law firms.

This manuscript suggests that investigating third-party products need not be so laborious. Salazar et al. (2022) assumed that quantitative analysis of College Board student list products required (1) order summary data (i.e., filter criteria) and (2) prospect-level student list data for each purchase. By contrast, this manuscript shows that analyses of student list products require (1) knowledge of product specifications and (2) student-level survey data containing variables necessary to recreate these product specifications. The analyses presented here additionally utilized order summaries collected via public records requests. However, obtaining order summaries is less daunting than obtaining both the order summary and the associated prospect-level data for each purchase.

Future research can investigate third-party products by combining rich NCES longitudinal survey data with methods from investigative journalism (e.g., Feathers, 2022). Consider commercial "student success" products like EAB's Starfish. First, researchers can learn about product specifications from internet searches by attending trade shows (e.g., NACAC) where vendors peddle their wares. State contract databases, for example the Illinois Procurement Bulletin, show which public universities purchased Starfish. Additional information about product specifications can be obtained by issuing public records requests to universities that purchased a Starfish contract (the best use of public records requests is obtaining contracts and product documentation). Second, use NCES survey data to recreate – or approximate – the input measures utilized by the product. Third, recreate the analytic approaches utilized by the product. One caution, algorithimic products often facilitate targeting by geographic locality, but NCES survey sample sizes are often too small for analyses of particular states or metropolitan areas.

The payoff for developing this critical policy literature is great, and so is the cost of inaction. Third-party providers do not want to be the object of research because scrutiny from scholars will lead to scrutiny from regulators, which may disrupt profitable practices (Cottom, 2020). Increasingly, third-party providers perform core functions of schools and universities (Komljenovic, 2022; Nichols & Garcia, 2022). If researchers continues to ignore these products, then policy research will have diminishing influence on core functions of schools and universities. As conservative courts challenge progressive education policies like affirmative action, policy research should go on the offensive by applying theory about structural mechanisms to investigate structural racism by third-party products and vendors.

In higher education, third-party providers now dwarf the for-profit college market. Nevertheless, the Higher Education Act (HEA) – enforced by the US Department of Education – regulates for-profit Title IV institutions, but remains agnostic about "third-party servicers" (aside from lenders and guaranty agencies). For example, responding to concerns about incentive-based compensation for online program management (OPM) companies, the Department of Education (2011) argued that tuition sharing with third-party vendors is not problematic because enrollment goals are determined by the institution, not the vendor.[15]

Therefore, developing a critical policy literature on third-party vendors/products demands that researchers be "price-makers" rather than "price-takers" when it comes to which issues demand policy attention. Instead of doing research that fits within the constraints of federal education policy, do research that shifts the focus of federal policy. Given the narrow focus of the HEA and the Department of Education, this research agenda should target the Federal Trade Commission, the Consumer Finance Protection Bureau and other agencies that enforce laws concerned with equality of opportunity for consumers. This shift in target audience requires researchers to learn more about these regulatory agencies, the laws they enforce, and to develop relationships with key staff. This new research focus represents a "paradigm shift" in education policy research, as opposed to "normal science" (Kuhn, 1962). This paradigm shift will be well remunerated with policy impact and scholarly productivity.

[15]US Department of Education (2011) states that, "the independence of the third party (both as a corporate matter and as a decision-maker) from the institution that provides the actual teaching and educational services is a significant safeguard against the abuses the Department has seen heretofore. When the institution determines the number of enrollments and hires an unaffiliated third party to provide bundled services that include recruitment, payment based on the amount of tuition generated does not incentivize the recruiting as it does when the recruiter is determining the enrollment numbers" (p.11)

7 References

Allensworth, E. M., & Clark, K. (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. https://doi.org/10.3102/0013189X20902110

Alon, S., & Tienda, M. (2007). Diversity, opportunity, and the shifting meritocracy in higher education. *American Sociological Review*, *72*(4), 487–511. Journal Article. Retrieved from <Go to ISI>://000248696500001

Andrews, R. J., Imberman, S. A., & Lovenheim, M. F. (2020). Recruiting and supporting low-income, high-achieving students at flagship universities. *Economics of Education Review*, *74*, 101923. https://doi.org/https://doi.org/10.1016/j.econedurev.2019.101923

Author. (XXXX).

Bastedo, M. N., Howard, J. E., & Flaster, A. (2016). Holistic admissions after affirmative action: Does "maximizing" the high school curriculum matter? *Educational Evaluation and Policy Analysis*. Journal Article. https://doi.org/10.3102/0162373716642407

Bastedo, M. N., & Jaquette, O. (2011). Running in place: Low income students and the dynamics of higher education stratification. *Educational Evaluation and Policy Analysis*, *33*(3), 318–339.

Belkin, D. (2019). For sale: SAT-Takers' names. Colleges buy student data and boost exclusivity. *The Wall Street Journal*. Retrieved from https://www.wsj.com/articles/for-sale-sat-takers-names-colleges-buy-student-data-and-boost-exclusivity-11572976621

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

Blake, M. K., & Langenkamp, A. G. (2022). Differences between latinx and white students in college-going behaviors. *Research in Higher Education*, *63*(2), 248–270. https://doi.org/10.1007/s11162-021-09648-6

Bonilla-Silva, E. (1997). Rethinking racism: Toward a structural interpretation. *American Sociological Review*, *62*(3), 465–480. https://doi.org/10.2307/2657316

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

Burrows, R., & Gane, N. (2006). Geodemographics, software and class. *Sociology-the Journal of the British Sociological Association*, *40*(5), 793–812. Journal Article. https://doi.org/10.1177/0038038506067507

Campbell, A. (2017). *Higher education marketing: How to master your admissions funnel*. Hop Online. Retrieved from https://hop-online.com/blog/higher-education-marketing-admissions-process/

Cardona, T., Cudney, E. A., Hoerl, R., & Snyder, J. (2022). Data mining and machine learning retention models in higher education. *Journal of College Student Retention: Research, Theory & Practice*, *0*(0), 1521025120964920. https://doi.org/10.1177/1521025120964920

Carey, K. (2019). The creeping capitalist takeover of higher education. *HuffPost*. Retrieved from https://www.huffpost.com/highline/article/capitalist-takeover-college/

Castilla, E. J. (2008). Gender, race, and meritocracy in organizational careers. *American Journal of Sociology*, *113*(6), 1479–1526. https://doi.org/10.1086/588738

Castleman, B., & Goodman, J. (2018). Intensive college counseling and the enrollment and persistence of low-income students. *Education Finance and Policy*, *13*(1), 19–41. https://doi.org/10.1162/edfp_a_00204

Cellini, S. R., & Chaudhary, L. (forthcoming). Where do colleges advertise? Demographic targeting by u.s. colleges. *Economics of Education Review*.

Cheslock, J. J., & Kroc, R. (2012). Managing college enrollments. In R. Howard, B. Knight, & G. McLaughlin (Eds.), *The handbook for institutional researchers* (pp. 221–236). San Francisco, CA: Jossey-Bass.

College Board. (n.d.). College board search solutions. The College Board. Retrieved from https://cbsearch.collegeboard.org/solutions

College Board. (2011). *Segment Analysis Service: An educationally relevant geodemographic tagging service*. College Board. Retrieved from https://secure-media.collegeboard.org/mSSS/media/pdf/segment-analysis-service-overview.pdf

Cotter, K., Medeiros, M., Pak, C., & Thorson, K. (2021). "Reach the right people": The politics of "interests" in facebook's classification system for ad targeting. *Big Data & Society*, *8*(1), 16. https://doi.org/10.1177/2053951721996046

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

Culliford, E. (2020). How political campaigns use your data: What campaigns know about U.S. Voters and how they use it to shape their strategies. *Reuters*. Retrieved from https://graphics.reuters.com/USA-ELECTION/DATA-VISUAL/yxmvjjgojvr/

Cunha, J. M., Miller, T., & Weisburst, E. (2018). Information and college decisions: Evidence from the texas GO center project. *Educational Evaluation and Policy Analysis*, *40*(1), 151–170. Journal Article. https://doi.org/10.3102/0162373717739349

Dache-Gerbino, A., Kiyama, J. M., & Sapp, V. T. (2018). The dangling carrot: Proprietary institutions and the mirage of college choice for Latina students. *The Review of Higher Education*, *42*(1), 29–60.

Doyle, W. R. (2010). Changes in institutional aid, 1992-2003: The evolving role of merit aid. *Research in Higher Education*, *51*(8), 789–810. Journal Article.

Duncan, O. D., Ohlin, L. E., Reiss, A. J., & Stanton, H. R. (1953). Formal devices for making selection decisions. *American Journal of Sociology*, *58*(6), 573–584. Retrieved from http://www.jstor.org/stable/2772558

Dynarski, S., Libassi, C., Michelmore, K., & Owen, S. (2021). Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students. *American Economic Review*, *111*(6), 1721–1756. https://doi.org/10.1257/aer.20200451

EAB. (2018). *Making your digital ads count: 15 lessons on new and emerging techniques in undergraduate recruitment marketing*. EAB.

EAB. (2022). Starfish: Scale your student success efforts. Retrieved from https://eab.com/products/starfish/

Encoura. (n.d.). Encoura. ACT. Retrieved from https://encoura.org/

Equifax. (2023). Drive customer growth: Best practices for financial services firms to acquire the right customers and grow share. Retrieved from https://assets.equifax.com/marketing/US/assets/drive-customer-growth-ebook.pdf

FDIC. (2018). VII. Unfair and deceptive practices. In *FDIC consumer compliance examination* *manual* (pp. 1–8).

Feathers, T. (2022). College prep software naviance is selling advertising access to millions of students. *The Markup*. Retrieved from https://graphics.reuters.com/USA-ELECTION/DATA-VISUAL/yxmvjjgojvr/https://themarkup.org/machine-learning/2022/01/13/college-prep-software-naviance-is-selling-advertising-access-to-millions-of-students

Federal Trade Commission. (2014). *Data brokers: A call for transparency and accountability*. Federal Trade Commission. Retrieved from https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf

Federal Trade Commission. (2016a). *Big data: A tool for inclusion or exclusion? Understanding the issues*. Federal Trade Commission. Retrieved from https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf

Federal Trade Commission. (2016b). *"Follow the lead" workshop*. Retrieved from http://www.ftc.gov/system/files/documents/reports/staff-perspective-follow-lead/staff_perspective_follow_the_lead_workshop.pdf

Fire Engine RED. (2022). *Data services: Search modeling*. Fire Engine RED. Retrieved from https://www.fire-engine-red.com/data-services/

Fourcade, M., & Healy, K. (2013). Classification situations: Life-chances in the neoliberal era. *Accounting Organizations and Society*, *38*(8), 559–572. Journal Article. https://doi.org/10.1016/j.aos.2013.11.002

Green III, P. C., Baker, B. D., & Oluwole, J. O. (2021). School finance, race, and reparations. *Wash. And Lee Journal of Civil Rights and Social Justice*, *27*(2), 484–558. Retrieved from https://scholarlycommons.law.wlu.edu/crsj/vol27/iss2/7/

Gurantz, O., Howell, J., Hurwitz, M., Larson, C., Pender, M., & White, B. (2021). A national-level informational experiment to promote enrollment in selective colleges. *Journal of Policy Analysis and Management*, *40*(2), 453–479. https://doi.org/https://doi.org/10.1002/pam.22262

Harcourt, B. E. (2007). *Against prediction: Profiling, policing, and punishing in an actuarial age* (pp. viii, 336 p.). Chicago: University of Chicago Press.

Harris, C. I. (1993). Whiteness as property. *Harvard Law Review*, *106*(8), 1707–1791. https://doi.org/10.2307/1341787

Hicks, M. (2017). A feature, not a bug. *Technology Stories*, *15*(4). Retrieved from https://www.technologystories.org/a-feature-not-a-bug/

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

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

Hossler, D., & Bean, J. P. (1990). *The strategic management of college enrollments* (pp. xxii, 330 p.). Jossey-Bass.

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

Hoxby, C. M. (1997). *How the changing market structure of u.s. Higher education explains college tuition* (No. Working Paper 6323). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/W6323.pdf

Hoxby, C. M. (2009). The changing selectivity of american colleges. *Journal of Economic Perspectives*, *23*(4), 95–118. Retrieved from <Go to ISI>://000272677100005

Hoxby, C. M., & Avery, C. (2013). The missing "one-offs": The hidden supply of high merit students for highly selective colleges. *Brookings Papers on Economic Activity*, *2013*(1), 1–65. https://doi.org/10.1353/eca.2013.0000

Hurwitz, M. (2012). The impact of institutional grant aid on college choice. *Educational Evaluation and Policy Analysis*, *34*(3), 344–363. https://doi.org/10.3102/0162373712448957

Hyman, J. (2017). ACT for all: The effect of mandatory college entrance exams on postsecondary attainment and choice. *Education Finance and Policy*, *12*(3), 281–311. https://doi.org/10.1162/EDFP_a_00206

James Madison University. (2017). Commonwealth of virginia standard contract, contract no. UCPJMU4648.

Jaquette, O., Salazar, K. G., & Martin, P. (2022). *The student list business: Primer and market dynamics.* washington, DC: TICAS. Retrieved from https://ticas.org/wp-content/uploads/2022/09/The-Student-List-Business_-Primer-and-Market-Dynamics.pdf

Jillson, E. (2021). *Aiming for truth, fairness, and equity in your company's use of AI.* Federal Trade Commission. Retrieved from https://www.ftc.gov/news-events/blogs/business-blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai

Khan, S. R. (2011). *Privilege: The making of an adolescent elite at st. Paul's school* (p. 232). Princeton, N.J.: Princeton University Press.

Killgore, L. (2009). Merit and competition in selective college admissions. *Review of Higher Education*, *32*(4), 469–488. Retrieved from <Go to ISI>://WOS:000266737500002

Kolluri, S. (2018). Advanced placement: The dual challenge of equal access and effectiveness. *Review of Educational Research*, *88*(5), 671–711. https://doi.org/10.3102/0034654318787268

Komljenovic, J. (2022). The future of value in digitalised higher education: Why data privacy should not be our biggest concern. *Higher Education*, *83*(1), 119–135. https://doi.org/10.1007/s10734-020-00639-7

Korver-Glenn, E. (2018). Compounding inequalities: How racial stereotypes and discrimination accumulate across the stages of housing exchange. *American Sociological Review*, *83*(4), 627–656. https://doi.org/10.1177/0003122418781774

Korver-Glenn, E. (2022). *Race brokers: Housing markets and racial segregation in 21st century urban america* (pp. pages cm). New York, NY: Oxford University Press.

Kraatz, M. S., Ventresca, M. J., & Deng, L. N. (2010). Precarious values and mundane innovations: Enrollment management in american liberal arts colleges. *Academy of Management Journal*, *53*(6), 1521–1545.

Kraus, M. W., Onyeador, I. N., Daumeyer, N. M., Rucker, J. M., & Richeson, J. A. (2019). The misperception of racial economic inequality. *Perspectives on Psychological Science*, *14*(6), 899–921. https://doi.org/10.1177/1745691619863049

Krippner, G. R., & Hirschman, D. (2022). The person of the category: The pricing of risk and the politics of classification in insurance and credit. *Theory and Society*, *51*(5), 685–727.

https://doi.org/10.1007/s11186-022-09500-5
Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
Lawler, M., & Dold, M. (2021). For-profit schools as covered persons under the CFPA. *Berkeley Law, Center for Consumer Law & Economic Justice*. Retrieved from https://protectborrowers.org/wp-content/uploads/2021/03/SBPC_UCB_For-Profits.pdf
Leeds, D. M., & DesJardins, S. L. (2015). The effect of merit aid on enrollment: A regression discontinuity analysis of iowa's national scholars award. *Research in Higher Education*, *56*(7), 471–495. https://doi.org/10.1007/s11162-014-9359-2
Leyshon, A., & Thrift, N. (1999). Lists come alive: Electronic systems of knowledge and the rise of credit-scoring in retail banking. *Economy and Society*, *28*(3), 434–466. https://doi.org/10.1080/03085149900000013
Long, B. T., & Kurlaender, M. (2009). Do community colleges provide a viable pathway to a baccalaureate degree? *Educational Evaluation and Policy Analysis*, *31*(1), 30–53. https://doi.org/10.3102/0162373708327756
Marx, K. (1978). Capital. In R. C. Tucker (Ed.), *The marx-engels reader* (second). New York: W.W. Norton; Company.
McPherson, M. S., & Schapiro, M. O. (1998). *The student aid game*. Princeton, NJ: Princeton University Press.
Melguizo, T. (2008). Quality matters: Assessing the impact of attending more selective institutions on college completion rates of minorities. *Research in Higher Education*, *49*(3), 214–236. https://doi.org/10.1007/s11162-007-9076-1
Moore, J. (2017). *Do students who opt into ACT's educational opportunity service (EOS) enroll in college at higher rates?* ACT, Inc. Retrieved from https://www.act.org/content/dam/act/unsecured/documents/R1652-benefits-of-act-eos-opt-in-2017-08.pdf
Mulhern, C. (2021). Changing college choices with personalized admissions information at scale: Evidence on naviance. *Journal of Labor Economics*, *39*(1), 219–262. https://doi.org/10.1086/708839
Napier, A., & Orrick, A. (2022). The economic, social, and political dimensions of platform studies in education. *Harvard Educational Review*, *92*(2), 206–208. https://doi.org/10.17763/1943-5045-92.2.206
Nichols, T. P., & Garcia, A. (2022). Platform studies in education. *Harvard Educational Review*, *92*(2), 209–230. https://doi.org/10.17763/1943-5045-92.2.209
Niu, S. X., & Tienda, M. (2010). Minority student academic performance under the uniform admission law: Evidence from the university of texas at austin. *Educational Evaluation and Policy Analysis*, *32*(1), 44–69. https://doi.org/10.3102/0162373709360063
Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: New York University Press.
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
Norris, D. (2022). The illusion of transparency: The political double standard in city credit ratings. *Socio-Economic Review*. Journal Article. https://doi.org/10.1093/ser/mwac016
O'Brien, R. L., & Kiviat, B. (2018). Disparate impact? Race, sex, and credit reports in hiring. *Socius*, *4*, 2378023118770069. https://doi.org/10.1177/2378023118770069
O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and*

threatens democracy (First edition., pp. x, 259 pages). New York: Crown.

Park, J. J., & Becks, A. H. (2015). Who benefits from SAT prep?: An examination of high school context and race/ethnicity. *Review of Higher Education*, *39*(1), 1–23. Retrieved from <Go to ISI>://WOS:000360777300001

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information* (p. 311). Cambridge: Harvard University Press.

Poon, M. (2007). Scorecards as devices for consumer credit: The case of fair, isaac & company incorporated. *Sociological Review*, *55*, 284–306. https://doi.org/10.1111/j.1467-954X.2007.00740.x

Posecznick, A. (2017). *Selling hope and college merit, markets, and recruitment in an unranked school*. Ithaca ; London: Cornell University Press. Retrieved from https://muse.jhu.edu/book/52616

Posselt, J. R. (2016). *Inside graduate admissions: Merit, diversity, and faculty gatekeeping* (pp. x, 250 pages). Cambridge, MA: Harvard University Press.

Posselt, J. R., Jaquette, O., Bielby, R., & Bastedo, M. N. (2012). Access without equity: Longitudinal analyses of institutional stratification by race and ethnicity, 1972-2004. *American Educational Research Journal*, *49*(6), 1074–1111.

Reardon, S. F., Kalogrides, D., & Shores, K. (2019). The geography of racial/ethnic test score gaps. *American Journal of Sociology*, *124*(4), 1164–1221. https://doi.org/10.1086/700678

Robinson, C. J. (2000). *Black marxism: The making of the black radical tradition* (pp. xxxiii, 436). Chapel Hill, N.C.: University of North Carolina Press. Retrieved from Table of Contents http://lcweb.loc.gov/catdir/toc/99030995.html

Rodriguez, A., & Hernandez-Hamed, E. (2020). Understanding unfulfilled AP potential across the participation pipeline by race and income. *Teachers College Record*, *122*(9), 38. Retrieved from <Go to ISI>://WOS:000595649100009

Rodriguez, A., & McGuire, K. M. (2019). More casses, more acess? Understanding the effects of course offerings on back-white gaps in advanced placement course-taking. *Review of Higher Education*, *42*(2), 641–679. Retrieved from <Go to ISI>://WOS:000454537100016

Rothstein, R. (2017). *The color of law: A forgotten history of how our government segregated America*. Liveright Publishing.

Ruffalo Noel Levitz. (2018). *2018 marketing and student recruitment report of effective practices*. Ruffalo Noel Levitz. Retrieved from http://learn.ruffalonl.com/rs/395-EOG-977/images/RNL_2018_Student_Recruitment_Marketing_Report_EM-19.pdf

Ruffalo Noel Levitz. (2021). RNL student search and engagement. Retrieved from https://www.ruffalonl.com/enrollment-management-solutions/building-demand/student-search-and-engagement/

Ruffalo Noel Levitz. (2023a). RNL enrollment marketing. Retrieved from https://www.ruffalonl.com/enrollment-management-solutions/building-demand/

Ruffalo Noel Levitz. (2023b). RNL student search and engagement. Retrieved from https://www.ruffalonl.com/enrollment-management-solutions/building-demand/student-search-and-engagement/

Ruffalo Noel-Levitz. (2022a). *2022 cost of recruiting an undergraduate student report*. Ruffalo Noel-Levitz. Retrieved from https://www.ruffalonl.com/thank-you/download-the-2022-cost-of-recruiting-an-undergraduate-student-report/

Ruffalo Noel-Levitz. (2022b). *2022 marketing and recruitment practices for undergraduate* *students report*. Ruffalo Noel-Levitz. Retrieved from https://www.ruffalonl.com/papers-research-higher-education-fundraising/marketing-and-recruitment-practices-for-undergraduate-students/

Ruffalo Noel-Levitz. (2022c). *2022 online program marketing and recruitment practices*. Ruffalo Noel-Levitz. Retrieved from https://www.ruffalonl.com/thank-you/download-the-2022-online-marketing-and-recruitment-practices-report/

Sadowski, J. (2019). When data is capital: Datafication, accumulation, and extraction. *Big Data & Society*, *6*(1), 12. https://doi.org/10.1177/2053951718820549

Sadowski, J. (2020). The internet of landlords: Digital platforms and new mechanisms of rentier capitalism. *Antipode*, *52*(2), 19. https://doi.org/10.1111/anti.12595

Salazar, K. G. (2022). Recruitment redlining by public research universities in the los angeles and dallas metropolitan areas. *The Journal of Higher Education*, 1–37. https://doi.org/10.1080/00221546.2021.2004811

Salazar, K. G., Jaquette, O., & Han, C. (2021). Coming soon to a neighborhood near you? Off-campus recruiting by public research universities. *American Educational Research Journal*, *58*(6), 1270–1314. https://doi.org/10.3102/00028312211001810

Salazar, K. G., Jaquette, O., & Han, C. (2022). *Geodemographics of student list purchases: A first look*. Washington, DC: TICAS. Retrieved from https://ticas.org/wp-content/uploads/2022/09/Geodemographics-of-Student-List-Purchases_A-First-Look.pdf

Schmidt, D. (2022). *Prospect search filters*. Encoura. Retrieved from https://helpcenter.encoura.org/hc/en-us/articles/360035260452-Prospect-Search-Filters-

Singer, P. (1988). Behind The Flood of Mail: Name Brokers. *The New York Times*. Retrieved from https://www.nytimes.com/1988/01/24/nyregion/behind-the-flood-of-mail-name-brokers.html

Smith, J., Howell, J., & Hurwitz, M. (2022). The impact of college outreach on high schoolers' college choices: Results from over one thousand natural experiments. *Education Finance and Policy*, *17*(1), 105–128. https://doi.org/10.1162/edfp_a_00334

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

Thornhill, T. (2019). We want black students, just not you: How white admissions counselors screen black prospective students. *Sociology of Race and Ethnicity*, *5*(4), 456–470. https://doi.org/10.1177/2332649218792579

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

US Department of Education. (2011). *Implementation of program integrity*. US Department of Education. Retrieved from https://fsapartners.ed.gov/sites/default/files/attachments/dpcletters/GEN1105.pdf

Waxman, B. (2019). The power of list segmentation part i: Are you already doing this? *intead, Global & Local Academic Branding*. Retrieved from https://services.intead.com/blog/the-power-of-list-segmentation-part-i-are-you-already-doing-this

Webber, R. (1988). Exploiting marketing opportunities. *Credit Management*, 36–37.

Williamson, B. (2021). Making markets through digital platforms: Pearson, edu-business, and the (e)valuation of higher education. *Critical Studies in Education*, *62*(1), 50–66. https://doi.org/10.1080/17508487.2020.1737556

Winston, G. C. (1999). Subsidies, hierarchy and peers: The awkward economics of higher education. *Journal of Economic Perspectives*, *13*(1), 13–36. Retrieved from <Go to ISI>://000078904300002

8 Tables

Table 1: Top filter combinations used in College Board orders purchased purchased by research vs. ma/doctoral

Research			MA/doctoral		
Filters	**Count**	**Percent**	**Filters**	**Count**	**Percent**
HS grad class, GPA, SAT, PSAT, Rank, State, Race	39	10%	HS grad class, GPA, SAT, Zip code	206	45%
HS grad class, PSAT, State	27	7%	HS grad class, GPA, PSAT, Zip code	145	32%
HS grad class, GPA, PSAT, State, Race	20	5%	HS grad class, SAT, State	31	7%
HS grad class, PSAT, State, Low SES	20	5%	HS grad class, GPA, SAT, PSAT, Zip code	28	6%
HS grad class, GPA, PSAT, State	17	5%	HS grad class, GPA, SAT, State	7	2%
HS grad class, GPA, SAT, State	16	4%	HS grad class, SAT, Geomarket	6	1%
HS grad class, GPA, AP score, Geomarket	15	4%	HS grad class, GPA, SAT, County	5	1%
HS grad class, GPA, SAT, PSAT, State, Segment, Gender	13	3%	HS grad class, GPA, SAT, PSAT, County	4	1%
HS grad class, PSAT, Geomarket	12	3%	HS grad class, GPA, PSAT, State	2	0%
HS grad class, SAT, State, Low SES, College size	11	3%	HS grad class, SAT, Geomarket, College type	2	0%

Table 2: HSLS09 Descriptive Statistics

	Unweighted	**N**	**SE**	**Pct**
Race/Ethnicity				
White	9,390	2,163,043	45,293	51.7
Asian	1,370	150,222	15,373	3.6
Hisp	2,520	920,384	41,451	22.0
Black	1,660	574,370	36,346	13.7
Multi	1,410	332,043	12,921	7.9
NH/PI	70	18,784	5,241	0.4
AI/AN	110	28,519	6,288	0.7
Academic Filters				
SAT Test-Taker	7,910	1,860,677	54,277	55.6
SAT Non-Test-Taker	8,610	2,326,689	54,249	44.4
PSAT Test-Taker	4,780	3,086,739	51,247	73.7
PSAT Non-Test-Taker	11,760	1,100,627	51,417	26.3
AP test-taker (any)	2,990	694,359	33,918	16.6
AP non-test-taker (any)	13,530	3,493,007	34,022	83.4
AP test-taker (STEM)	1,800	383,669	23,721	9.2
AP non-test-taker (STEM)	14,720	3,803,697	23,893	90.8
Academic GPA	16,480	4,177,402	6,863	99.8
Missing Academic GPA	40	9,964	6,562	0.2

* Unweighted sample sizes rounded to nearest 10 per NCES restricted data license regulations

Table 3: Test Taker Differences in Proportion

	Included	Excluded	Difference	Lower CI	Upper CI
SAT					
White	0.573	0.472	0.101***	0.100	0.102
Asian	0.045	0.029	0.016***	0.015	0.016
Hisp	0.173	0.257	-0.084***	-0.085	-0.083
Black	0.124	0.147	-0.023***	-0.024	-0.022
Multi	0.077	0.081	-0.004***	-0.005	-0.004
NH/PI	0.003	0.006	-0.003***	-0.003	-0.002
AI/AN	0.005	0.008	-0.003***	-0.003	-0.003
PSAT					
White	0.533	0.511	0.022***	0.021	0.023
Asian	0.058	0.028	0.030***	0.030	0.031
Hisp	0.199	0.227	-0.028***	-0.029	-0.027
Black	0.125	0.142	-0.017***	-0.018	-0.016
Multi	0.08	0.079	0.001***	0.001	0.002
NH/PI	0.003	0.005	-0.002***	-0.002	-0.002
AI/AN	0.002	0.009	-0.007***	-0.007	-0.007
AP					
White	0.542	0.512	0.030***	0.029	0.031
Asian	0.083	0.026	0.057***	0.056	0.058
Hisp	0.216	0.221	-0.005***	-0.006	-0.004
Black	0.081	0.148	-0.067***	-0.068	-0.067
Multi	0.072	0.081	-0.009***	-0.009	-0.008
NH/PI	0.005	0.004	0.001***	0.000	0.001
AI/AN	0.001	0.008	-0.007***	-0.007	-0.006
AP STEM					
White	0.557	0.513	0.044***	0.043	0.046
Asian	0.11	0.028	0.082***	0.081	0.083
Hisp	0.193	0.223	-0.030***	-0.031	-0.029
Black	0.074	0.144	-0.070***	-0.071	-0.069
Multi	0.06	0.081	-0.021***	-0.022	-0.020
NH/PI	0.004	0.005	-0.001	-0.000	0.000
AI/AN	0.002	0.007	-0.005***	-0.006	-0.005

Table 4: Filter by neighborhood segments

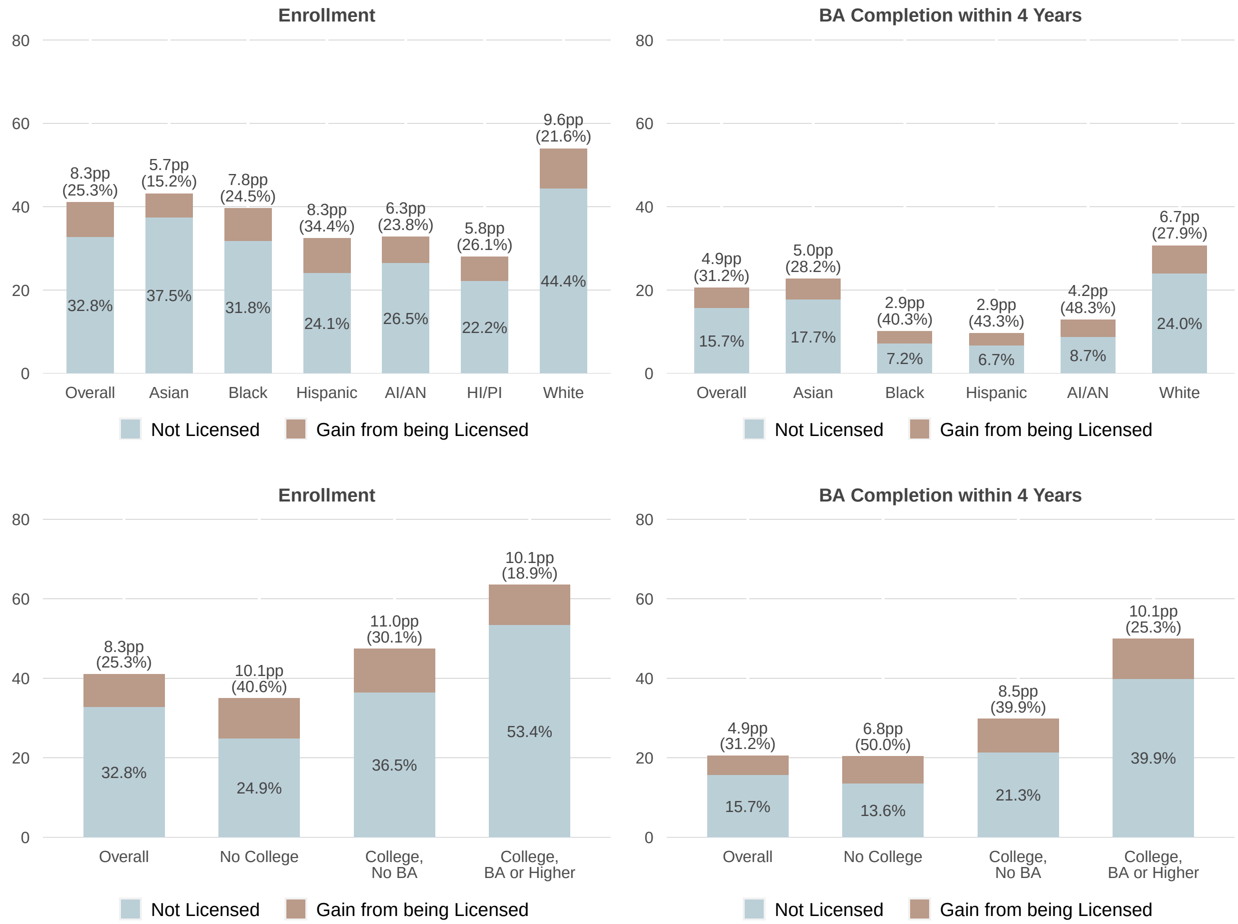
2011 D+ Cluster	SAT Math	SAT CR	Going Out of State	Percent NonWhite	Need Financial Aid	Med Income
51	546	533	32%	30%	57%	$95,432
52	480	470	30%	58%	71%	$63,578
53	561	544	32%	50%	55%	$92,581
54	458	443	25%	83%	76%	$38,977
55	566	565	52%	24%	63%	$71,576
56	420	411	29%	93%	66%	$35,308
57	541	519	52%	47%	43%	$67,394
58	533	489	28%	87%	69%	$68,213
59	561	562	52%	24%	74%	$54,750
60	589	590	63%	37%	36%	$104,174
61	585	567	51%	30%	40%	$123,858
62	596	595	67%	24%	72%	$59,824
63	548	541	39%	23%	65%	$69,347
64	466	466	48%	34%	29%	$49,829
65	440	433	23%	93%	78%	$45,081
66	499	492	20%	12%	76%	$50,453
67	519	501	27%	53%	59%	$60,960
68	552	558	52%	35%	65%	$57,902
69	534	521	37%	19%	65%	$88,100
70	613	598	65%	29%	61%	$86,381
71	405	408	39%	97%	68%	$42,661
72	399	397	31%	87%	47%	$32,708
73	528	514	29%	42%	62%	$90,849
74	433	435	29%	84%	79%	$44,065
75	459	457	28%	85%	72%	$50,421
76	514	509	27%	38%	64%	$61,332
77	502	492	26%	18%	75%	$62,372
78	594	578	56%	26%	39%	$134,400
79	550	551	57%	32%	74%	$40,909
80	534	527	39%	39%	65%	$49,877
81	491	483	27%	57%	72%	$63,030
82	496	491	29%	21%	75%	$53,465
83	500	490	19%	26%	71%	$49,335
Total	**512**	**502**	**32%**	**43%**	**65%**	**$70,231**

Table 5: Filter by high school segments

2011 D+ Cluster	SAT Math	SAT CR	Going Out of State	Percent NonWhite	Need Financial Aid	Med Income
51	462	457	14%	33%	68%	$40,918
52	489	496	81%	99%	77%	$64,730
53	471	484	28%	38%	62%	$60,833
54	376	371	33%	96%	38%	$38,146
55	489	481	39%	46%	44%	$71,845
56	536	508	73%	43%	49%	$63,967
57	434	435	29%	82%	79%	$48,301
58	592	577	51%	27%	32%	$104,509
59	499	489	19%	18%	74%	$47,685
60	523	549	23%	30%	33%	$70,175
61	485	370	33%	89%	9%	$61,385
62	474	473	34%	92%	67%	$55,515
63	440	427	28%	86%	72%	$49,238
64	606	542	37%	89%	57%	$81,911
65	515	503	28%	43%	65%	$72,692
66	498	515	37%	37%	73%	$60,272
67	526	546	48%	41%	69%	$71,279
68	541	540	41%	26%	62%	$79,260
69	390	395	36%	92%	74%	$43,391
70	595	581	56%	33%	48%	$105,721
71	400	412	57%	98%	80%	$43,137
72	528	544	35%	25%	64%	$70,018
73	451	438	24%	89%	76%	$48,406
74	654	579	76%	80%	46%	$59,089
75	514	502	31%	20%	71%	$72,850
76	600	584	72%	50%	28%	$90,265
77	595	508	64%	75%	39%	$39,490
78	473	468	48%	43%	22%	$56,703
79	594	585	61%	26%	71%	$65,180
Total	**514**	**502**	**32%**	**44%**	**65%**	**$70,223**

9 Figures

Figure 1: Student Search Service and four-year college enrollment/completion



Notes: AI/AN = American Indian or Alaska Native. HI/PI = Hawaiian or Pacific Islander. Sample for enrollment outcomes is all SAT takers in the 2015–2018 high school graduation cohorts. Sample for completion outcomes is students in the 2015–2016 cohorts. 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, graduation cohort, and high school fixed effects. All differences between licensed versus non-licensed students are statistically significant at the 1% level.

Figure 2: The enrollment funnel

Figure 3: Filters used in College Board orders purchased by 14 public universities

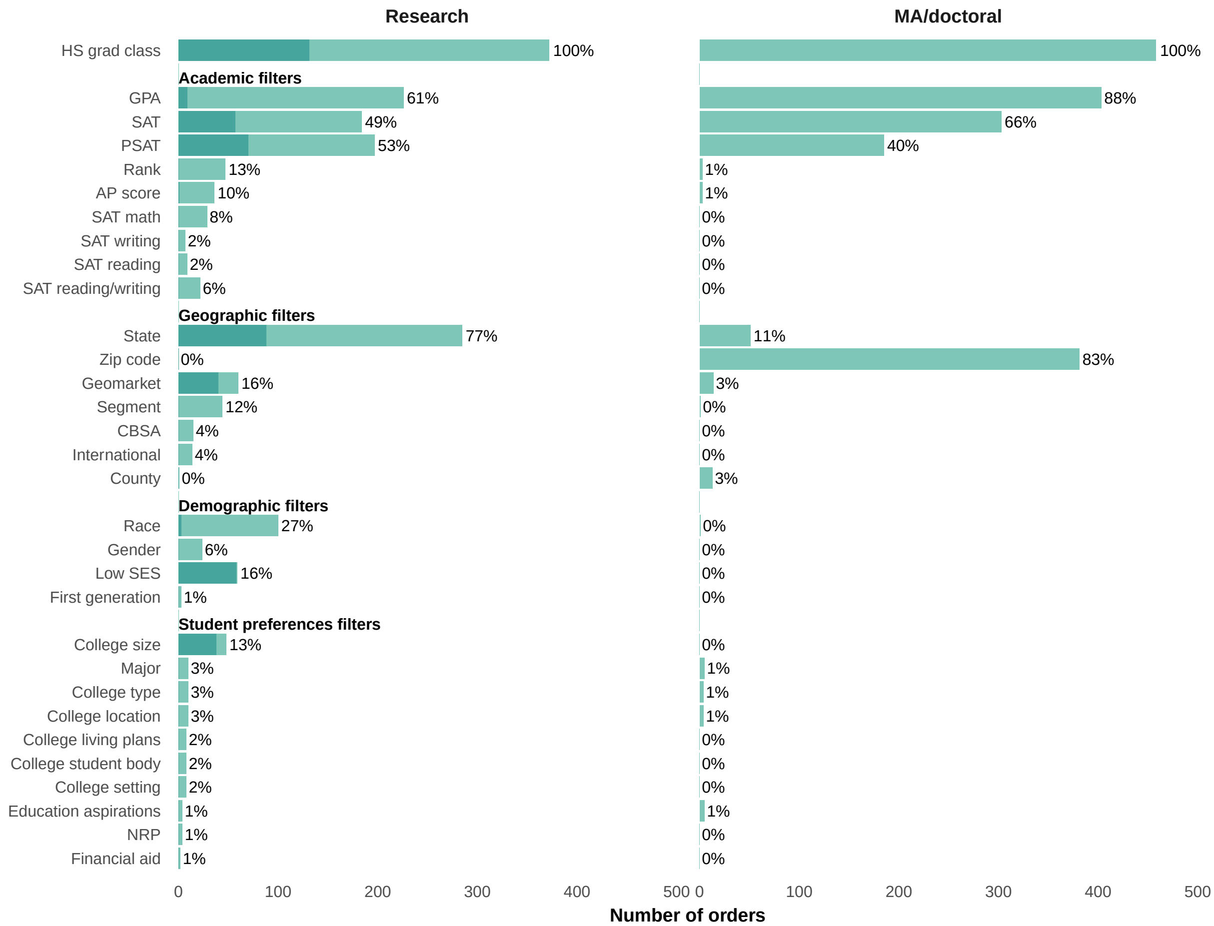


Figure 4: Test Takers Across SAT, PSAT, and AP Assessments

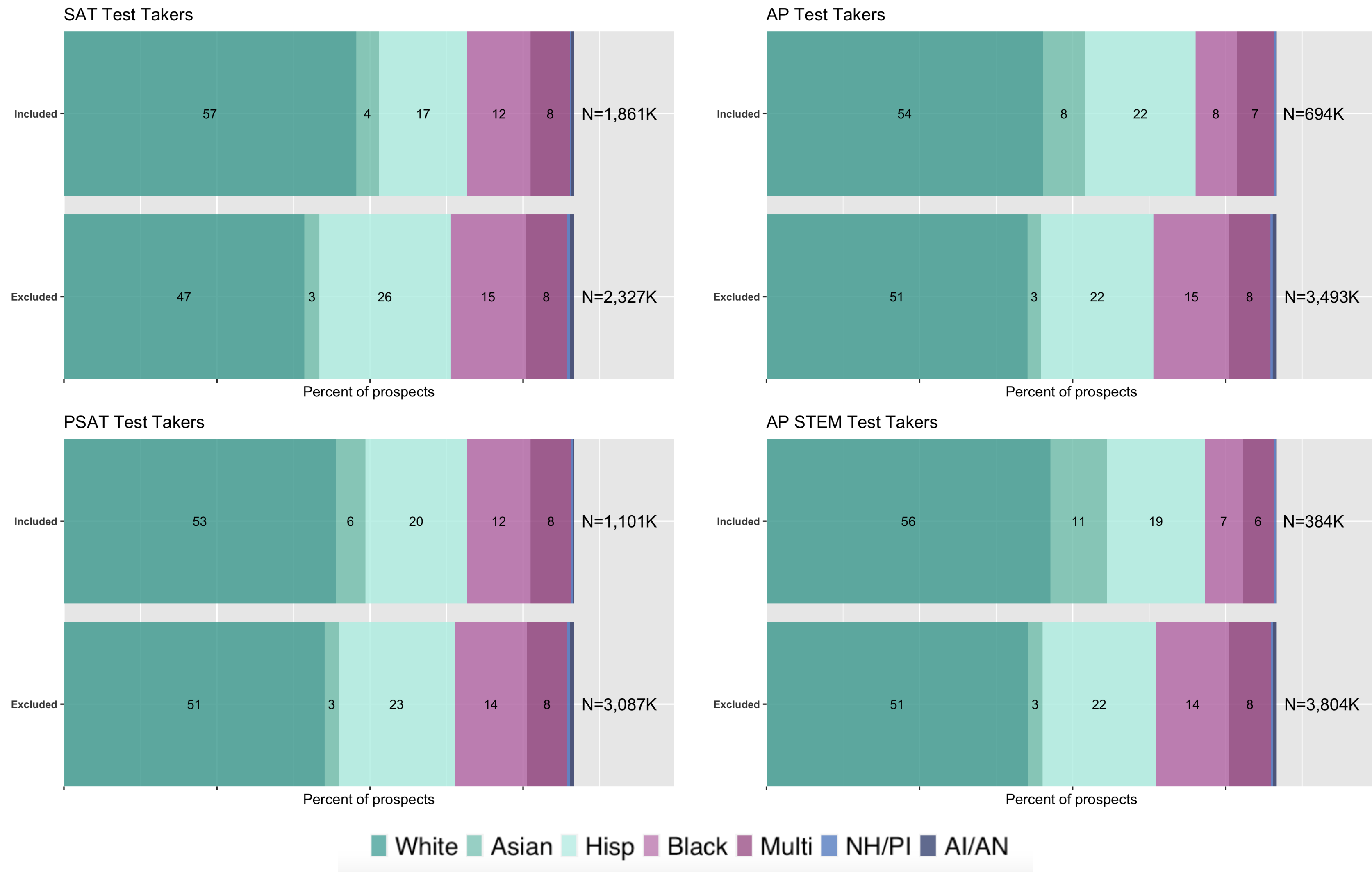


Figure 5: SAT and PSAT Filters Across Thresholds

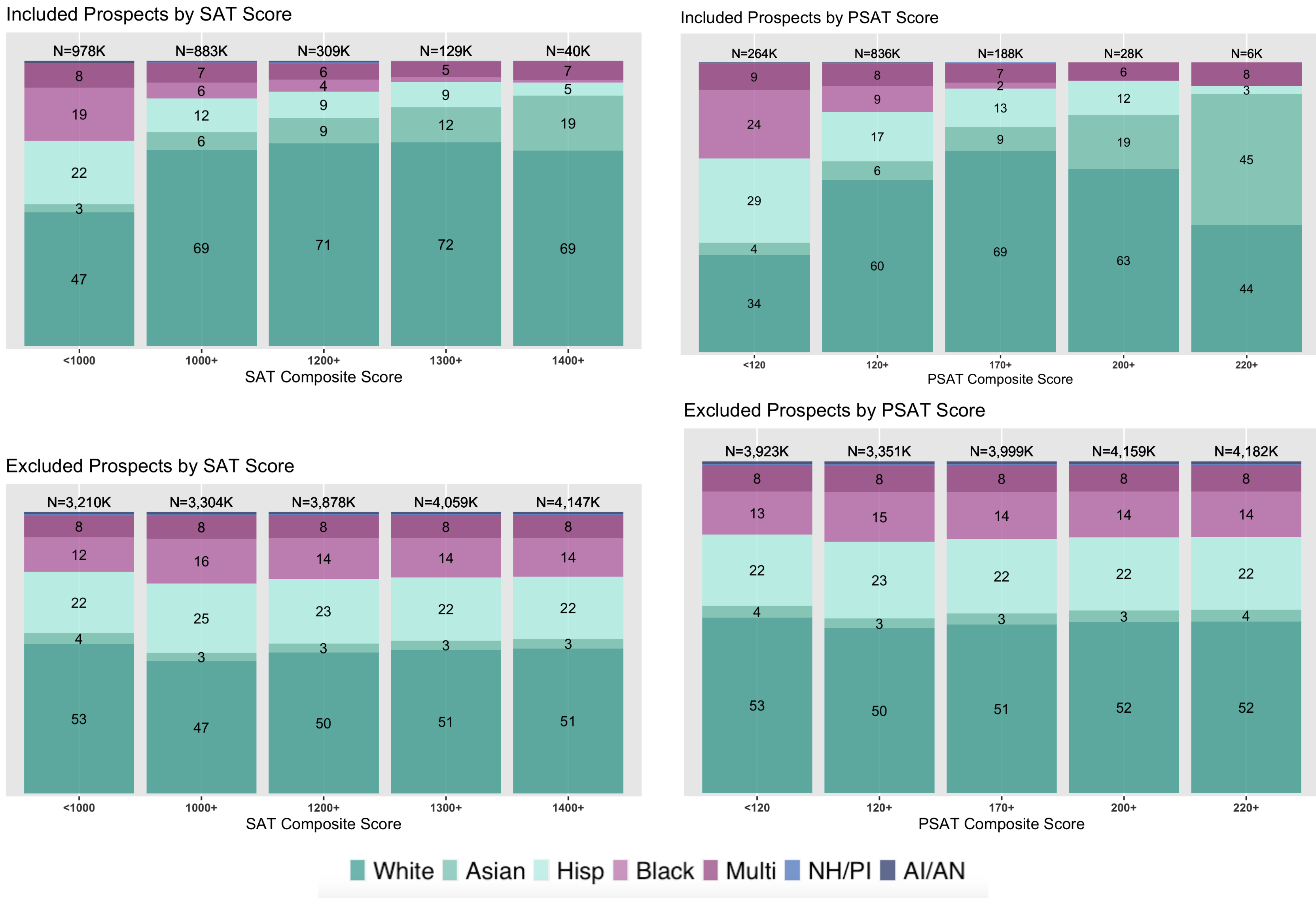


Figure 6: AP Filter Across Thresholds

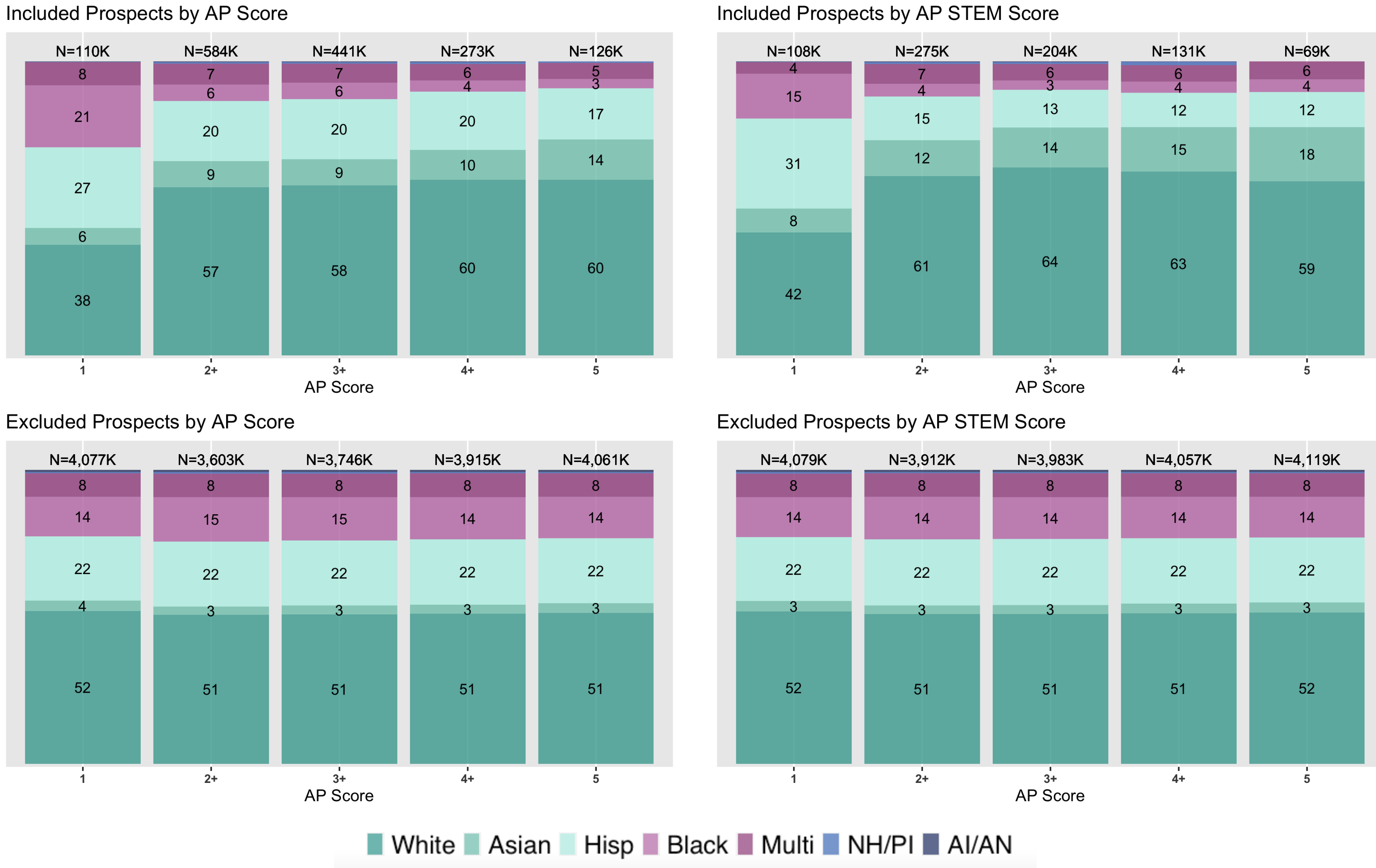


Figure 7: GPA Filter Across Thresholds

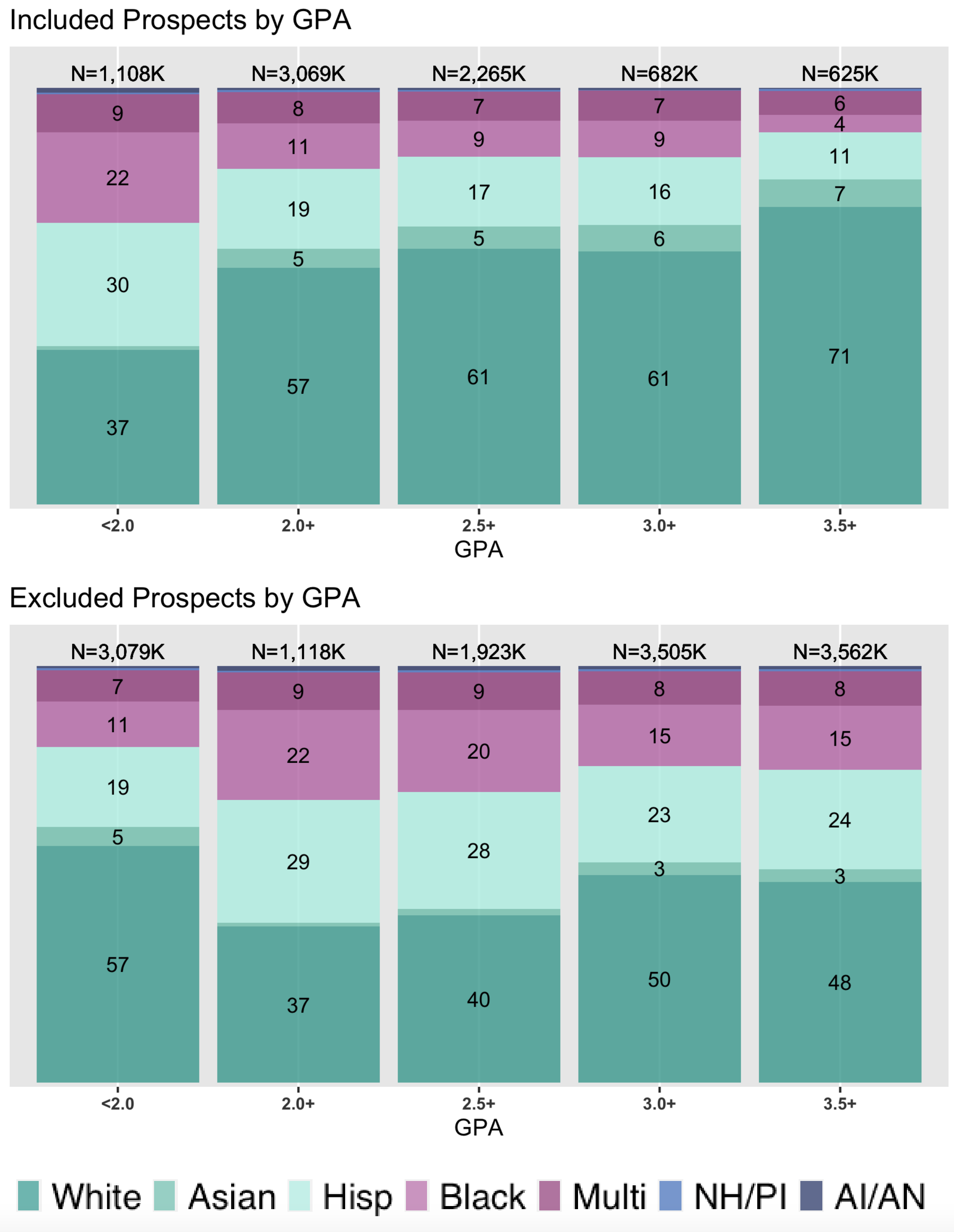


Figure 8: Zip Code Filter Across Affluence Percentiles

Figure 9: Zip Code Filter Across Affluence Percentiles for Los Angeles and New York

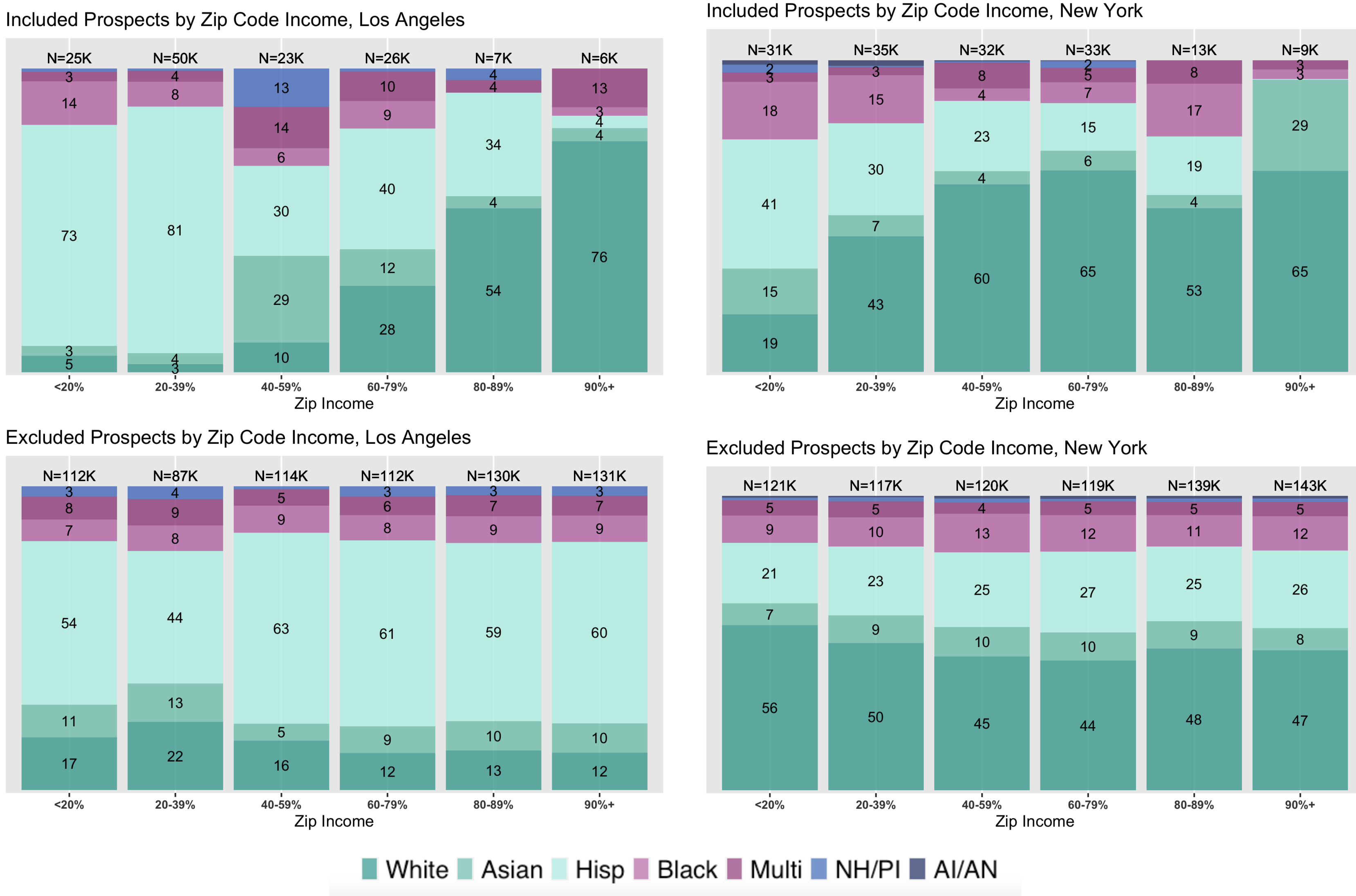


Figure 10: Zip Code and County Filters Across Affluence Levels

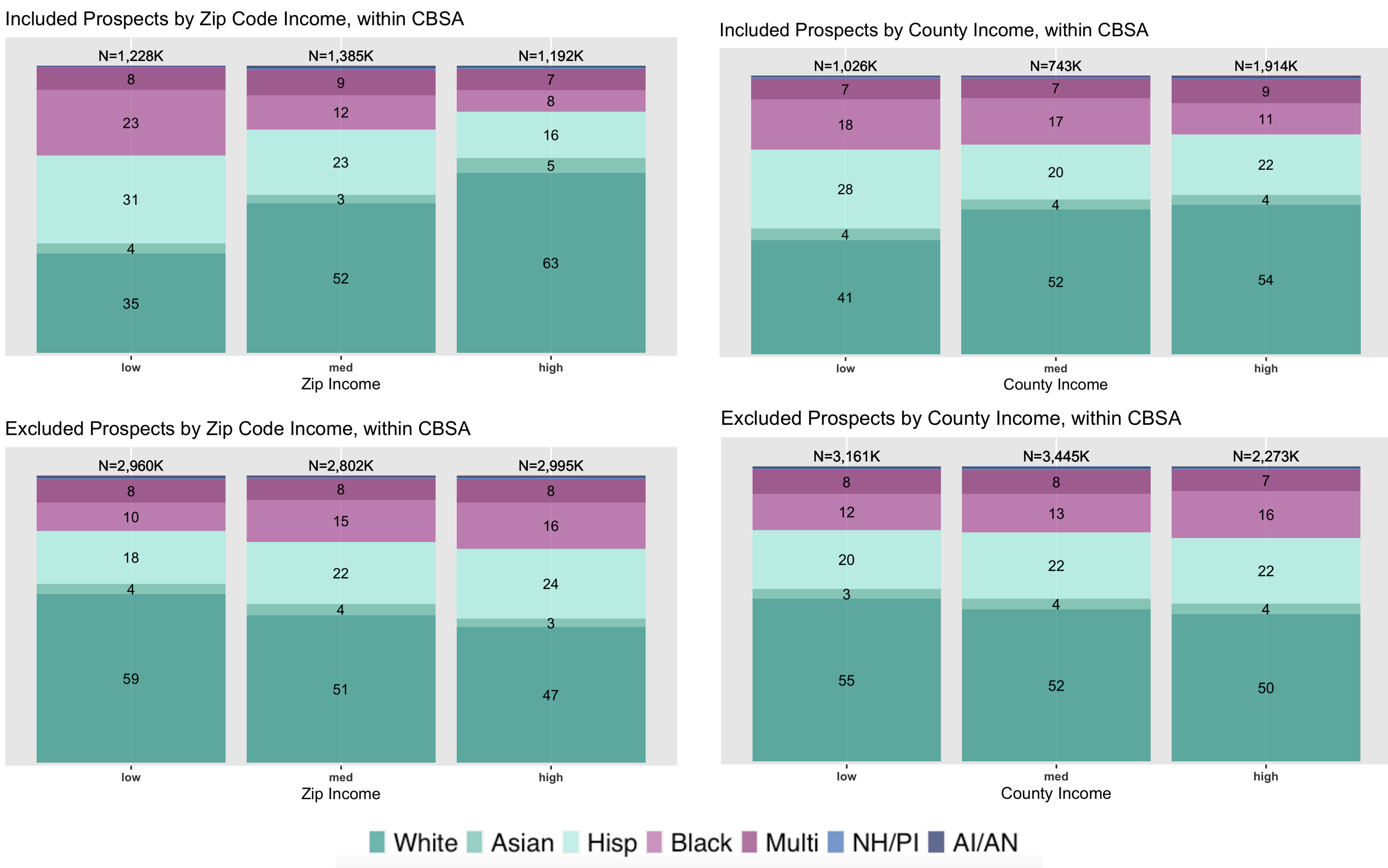


Figure 11: Academic and Geographic Combination: GPA (3.0+) and SAT or PSAT (across score thresholds)

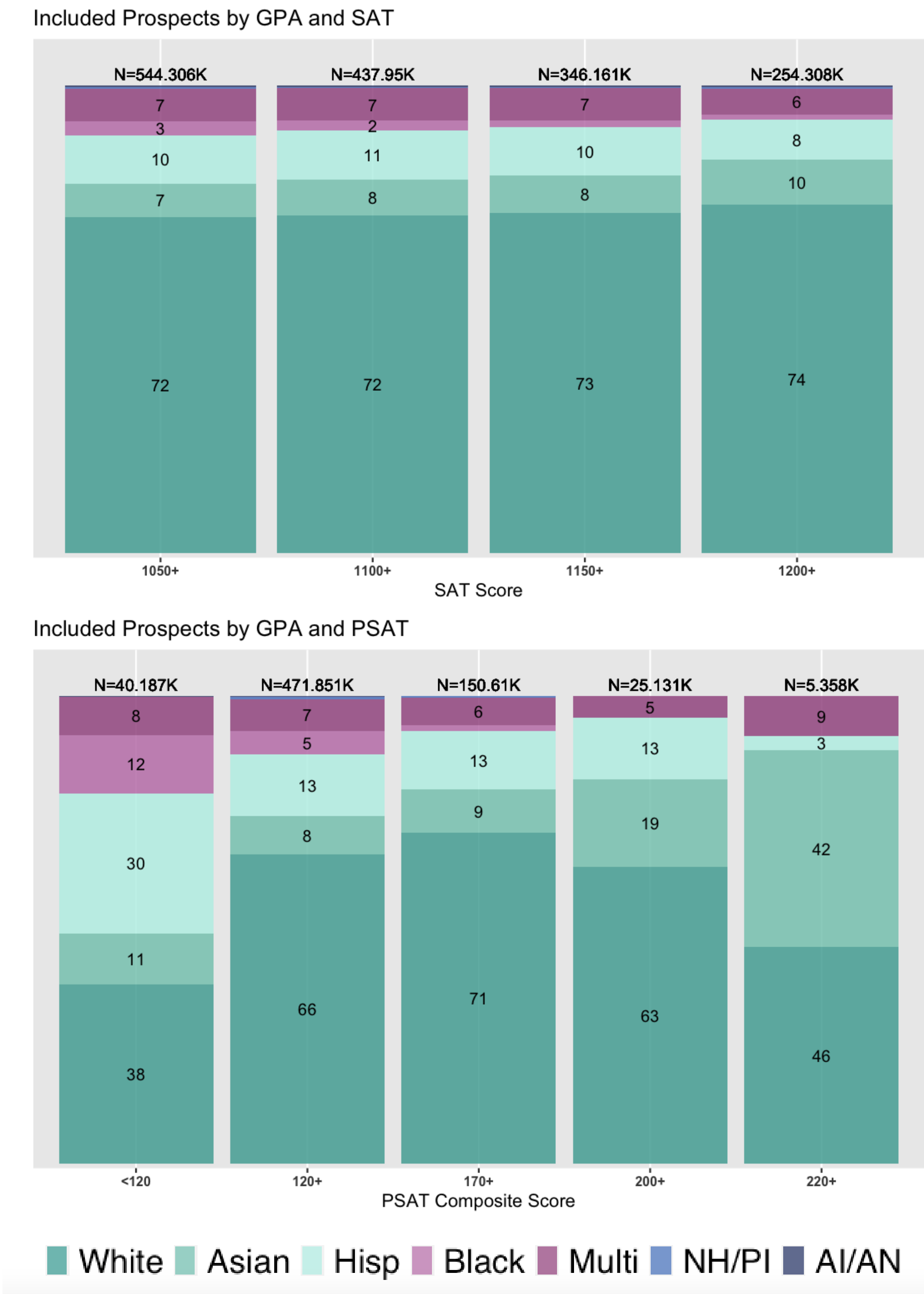


Figure 12: Academic and Geographic Combination: GPA (3.0+), PSAT (150+) or SAT (1050+), and Zip (across income thresholds)

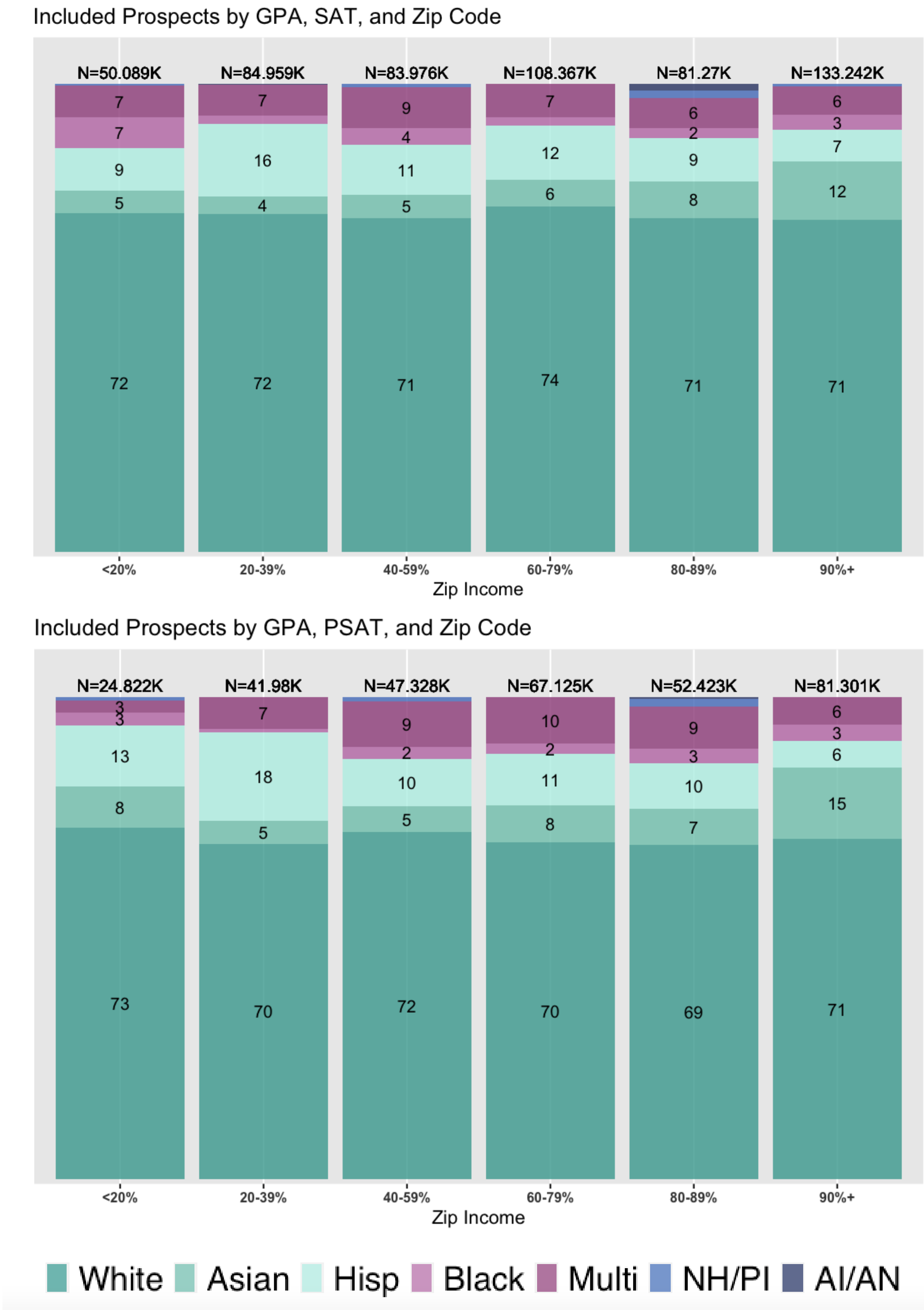


Figure 13: Academic and Geographic Combination: GPA (3.0+) and AP (acrossscore thresholds)

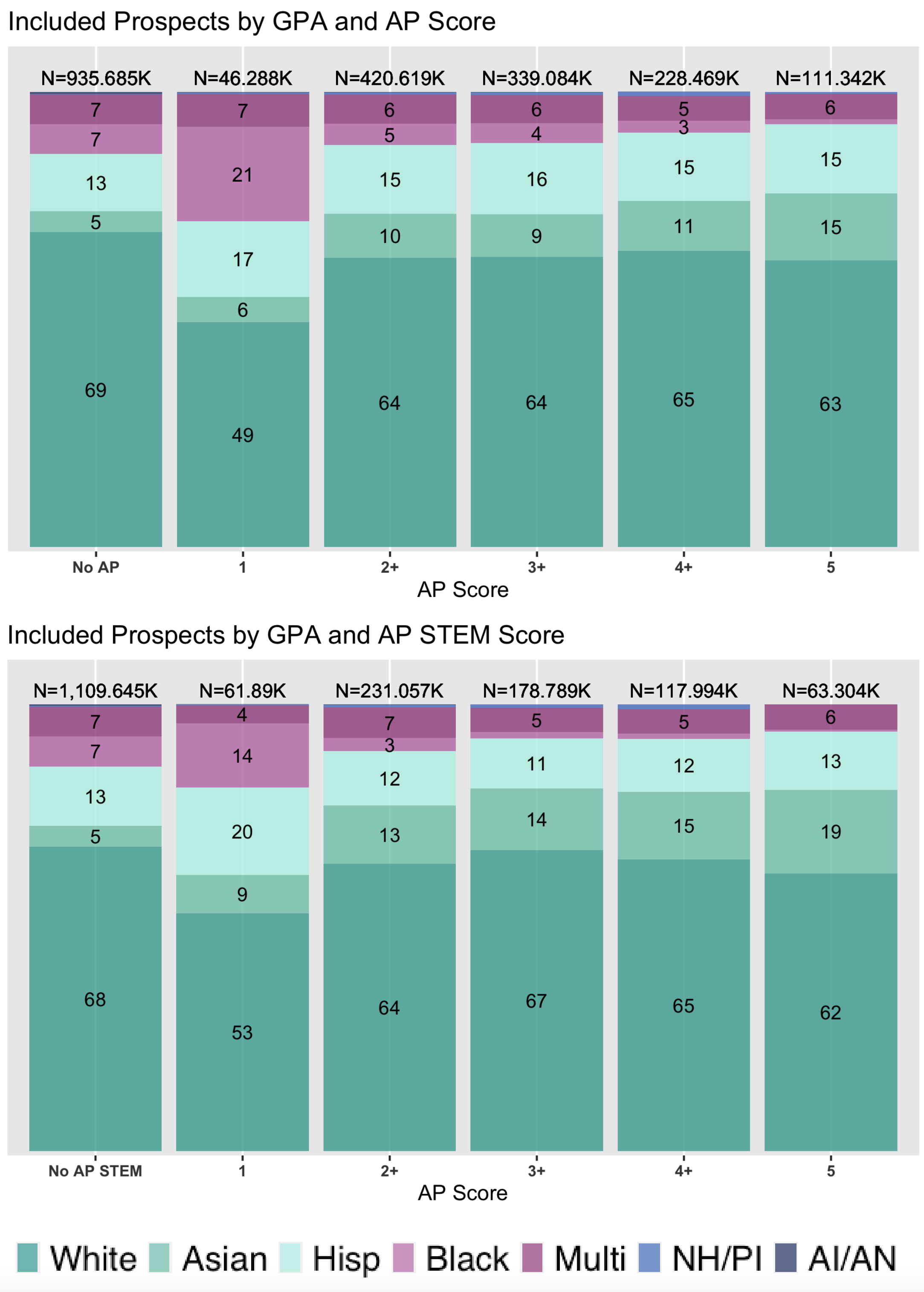


Figure 14: Segment filter prospects by metro (average income and racial composition)

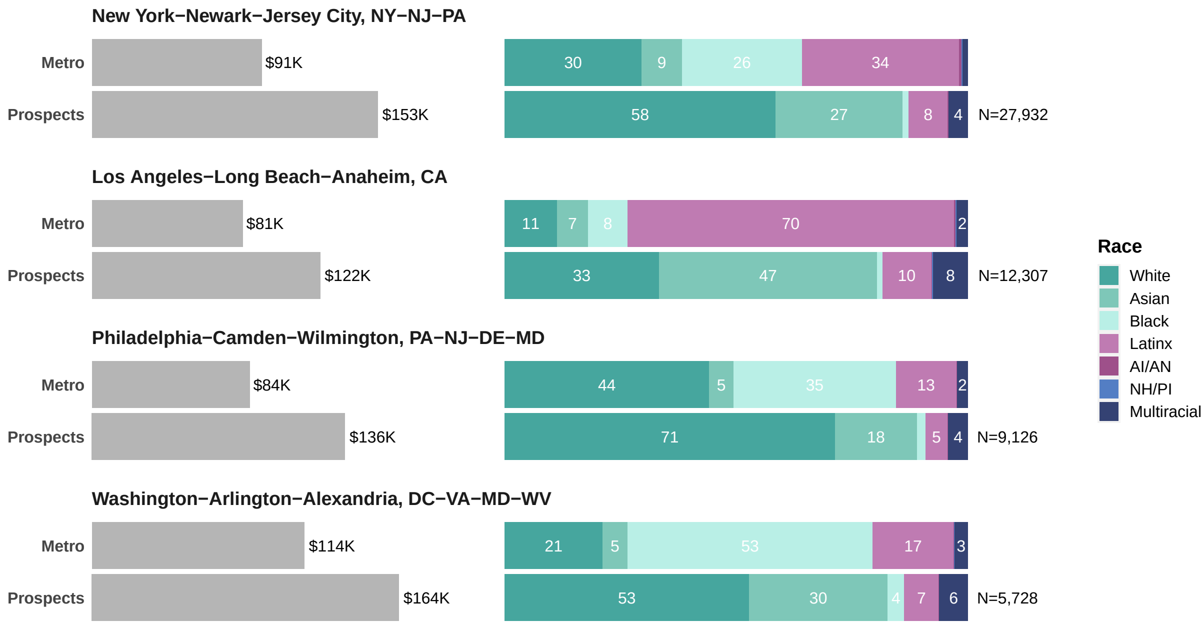
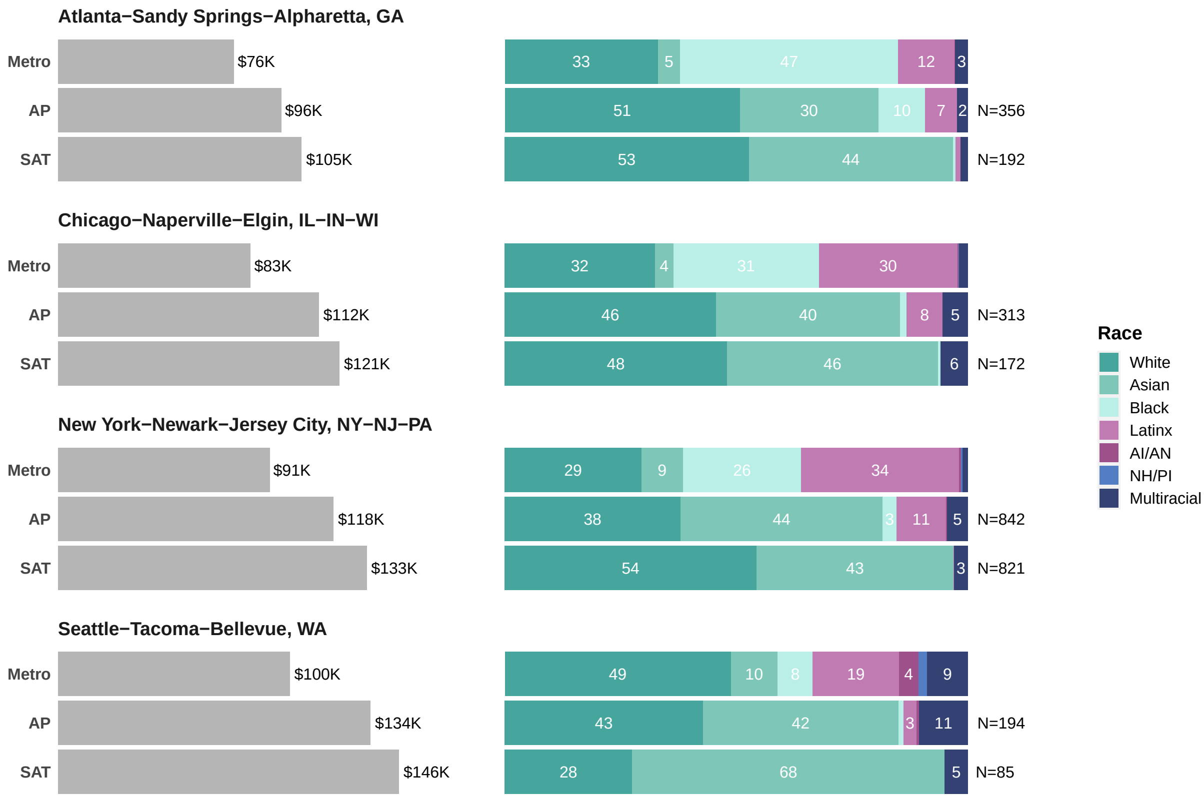


Figure 15: Women in STEM Prospects by Metro (Average Income and Racial Composition)



10 Online Appendix

Table 6: Score Threshold Proportion Differences in Included vs. Excluded across Race/Ethnicity

	White	Asian	Hisp	Black	Multi	NH/PI	AI/AN
SAT							
Less than 1000	-0.062***	-0.010***	0.004***	0.065***	0.005***	-0.003***	0.002***
1000+	0.217***	0.033***	-0.129***	-0.103***	-0.012***	-0.001***	-0.006***
1200+	0.209***	0.057***	-0.137***	-0.103***	-0.025***	0.001***	-0.003***
1300+	0.205***	0.090***	-0.137***	-0.122***	-0.027***	-0.002***	NA
1400+	0.170***	0.160***	-0.175***	-0.130***	-0.012***	NA	NA
PSAT							
Less than 120	-0.194***	0.007***	0.076***	0.105***	0.014***	-0.004***	-0.005***
120+	0.098***	0.034***	-0.062***	-0.060***	-0.003***	-0.001***	-0.006***
170+	0.185***	0.051***	-0.093***	-0.123***	-0.015***	0.001*	NA
200+	0.116***	0.151***	-0.102***	NA	-0.015***	NA	NA
220+	-0.078***	0.417***	-0.191***	NA	0.002	NA	NA
AP							
1	-0.143***	0.022***	0.056***	0.077***	-0.001	-0.003***	-0.006***
2+	0.066***	0.061***	-0.017***	-0.094***	-0.010***	0.001***	-0.007***
3+	0.070***	0.058***	-0.017***	-0.091***	-0.015***	0.001*	-0.005***
4+	0.086***	0.071***	-0.022***	-0.106***	-0.026***	0.003***	-0.005***
5+	0.084***	0.104***	-0.046***	-0.109***	-0.025***	0.001	NA
AP STEM							
1	-0.100***	0.046***	0.090***	0.015***	-0.040***	-0.004***	-0.007***
2+	0.101***	0.092***	-0.077***	-0.101***	-0.012***	0.002***	-0.005***
3+	0.130***	0.105***	-0.097***	-0.111***	-0.025***	0.003***	-0.005***
4+	0.113***	0.119***	-0.106***	-0.103***	-0.023***	0.007***	-0.006***
5+	0.078***	0.151***	-0.101***	-0.096***	-0.020***	NA	NA
GPA							
Less than 2.0	-0.198***	-0.036***	0.104***	0.109***	0.016***	-0.002***	0.006***
2.0+	0.193***	0.037***	-0.103***	-0.107***	-0.016***	0.002***	-0.006***
2.5+	0.212***	0.037***	-0.112***	-0.112***	-0.018***	-0.001***	-0.006***
3.0+	0.109***	0.032***	-0.068***	-0.060***	-0.008***	-0.004***	-0.002***
3.5+	0.233***	0.034***	-0.124***	-0.112***	-0.026***	0.000	-0.006***

Table 7: Zip Affluence Proportion Differences in Included vs. Excluded across Race/Ethnicity

	White	Asian	Hisp	Black	Multi	NH/PI	AI/AN
Affluence Percentile							
Less than 20%	-0.260***	0.008***	0.094***	0.163***	0.000	0.001***	-0.005***
20-39%	-0.089***	-0.018***	0.108***	0.015***	-0.010***	-0.002***	-0.003***
40-59%	0.021***	-0.002***	-0.012***	-0.028***	0.010***	0.001***	0.009***
60-79%	0.031***	0.010***	0.011***	-0.068***	0.012***	0.007***	-0.003***
80-89%	0.169***	0.011***	-0.091***	-0.065***	-0.015***	-0.003***	-0.006***
Greater than 90%	0.160***	0.020***	-0.120***	-0.056***	-0.006***	-0.003***	0.004***

Table 8: Zip and County Proportion Differences in Included vs. Excluded across Race/Ethnicity

	White	Asian	Hisp	Black	Multi	NH/PI	AI/AN
Zip							
Low Income	-0.242***	0.000	0.123***	0.129***	-0.005***	-0.001***	-0.005***
Moderate Income	0.007***	-0.010***	0.012***	-0.026***	0.011***	0.002***	0.003***
High Income	0.154***	0.023***	-0.083***	-0.085***	-0.008***	0.001	-0.001***
County							
Low Income	-0.141***	0.008***	0.084***	0.059***	-0.008***	0.003***	-0.004***
Moderate Income	0.004***	0.000	-0.027***	0.035***	-0.010***	0.001	-0.002***
High Income	0.037***	0.000	-0.005***	-0.047***	0.014***	-0.001***	0.002***

Table 9: GPA and PSAT/SAT Score Threshold Proportion Differences in Included vs. Excluded across Race/Ethnicity

	White	Asian	Hisp	Black	Multi	NH/PI	AI/AN
SAT							
1050+	0.233***	0.041***	-0.133***	-0.124***	-0.011***	0.001	-0.006***
1100+	0.229***	0.045***	-0.127***	-0.130***	-0.011***	-0.002***	-0.004***
1150+	0.229***	0.048***	-0.126***	-0.135***	-0.010***	-0.002***	-0.005***
PSAT							
1200+	0.243***	0.066***	-0.144***	-0.135***	-0.025***	-0.001***	-0.004***
<120	-0.135***	0.075***	0.081***	-0.012***	0.003*	NA	-0.006***
120+	0.164***	0.051***	-0.099***	-0.097***	-0.012***	-0.001***	-0.005***
170+	0.199***	0.059***	-0.098***	-0.130***	-0.021***	-0.002***	NA
200+	0.119***	0.154***	-0.089***	NA	-0.033***	NA	NA
220+	-0.053***	0.386***	-0.190***	NA	0.007	NA	NA

Table 10: GPA, PSAT/SAT, and Zip Code Proportion Differences in Included vs. Excluded across Race/Ethnicity

	White	Asian	Hisp	Black	Multi	NH/PI	AI/AN
SAT (1050+)							
Less than 20%	0.098***	0.028***	-0.077***	-0.056***	0.016***	-0.002***	NA
20-39%	0.166***	0.031***	-0.087***	-0.095***	-0.012***	0.003***	-0.006***
40-59%	0.252***	0.020***	-0.117***	-0.113***	-0.032***	NA	-0.005***
60-79%	0.129***	0.045***	-0.041***	-0.129***	-0.008***	0.007***	-0.004***
80-89%	0.269***	0.067***	-0.189***	-0.120***	-0.017***	NA	NA
Greater than 90%	0.230***	0.058***	-0.174***	-0.109***	-0.007***	0.002***	-0.002***
PSAT (150+)							
Less than 20%	0.172***	0.023***	-0.084***	-0.117***	0.013***	0.001	NA
20-39%	0.125***	0.050***	-0.057***	-0.108***	-0.007***	0.006***	NA
40-59%	0.250***	0.029***	-0.149***	-0.104***	-0.018***	-0.005***	-0.004***
60-79%	0.101***	0.049***	-0.051***	-0.121***	0.020***	0.009***	NA
80-89%	0.240***	0.091***	-0.179***	-0.119***	-0.023***	NA	NA
Greater than 90%	0.216***	0.082***	-0.157***	-0.106***	-0.024***	NA	NA

Table 11: GPA and AP Proportion Differences in Included vs. Excluded across Race/Ethnicity

	White	Asian	Hisp	Black	Multi	NH/PI	AI/AN
AP							
1	-0.023***	0.020***	-0.054***	0.072***	-0.007***	0.000	-0.006***
2+	0.132***	0.068***	-0.077***	-0.100***	-0.017***	0.001***	-0.006***
3+	0.131***	0.063***	-0.068***	-0.102***	-0.019***	0.002***	-0.006***
4+	0.141***	0.078***	-0.074***	-0.118***	-0.026***	0.004***	-0.006***
5+	0.116***	0.114***	-0.070***	-0.130***	-0.025***	0.001***	NA
AP STEM							
1	0.016***	0.050***	-0.024***	0.006***	-0.039***	-0.003***	-0.006***
2+	0.134***	0.100***	-0.104***	-0.113***	-0.012***	0.002***	-0.006***
3+	0.163***	0.108***	-0.114***	-0.129***	-0.025***	0.004***	NA
4+	0.139***	0.119***	-0.105***	-0.129***	-0.026***	0.008***	NA
5+	0.106***	0.153***	-0.091***	-0.134***	-0.024***	NA	NA