

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 1,839, while the number of SAT/ACT test-takers per freshman seat increased from 0.23 in 1955 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 – follows 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. A student list contains the contact information of prospective students who meet the search filter criteria (e.g., test score, GPA) specified by the university. Student lists are the fundamental input for undergraduate recruiting campaigns because purchased names – alongside “inquiries” who reach out on their own – constitute the set of prospects who receive subsequent recruiting interventions (e.g., mail, email) designed to push them towards the application and enrollment stages of the “enrollment funnel”. Ruffalo Noel-Levitz (2022b, p. 5) reports that 86% of public universities “purchase high school student names to generate inquiries and applicants.” Of these, 80% purchase more than 50,000 names annually. Howell, Hurwitz, Mabel, & Smith (2021) show that after controlling for covariates, 41.1% of students who participated in College Board Search – allowing accredited institutions to purchase their contact information — attended a 4-year college compared to 32.8% of students who opted out. Smith, Howell, & Hurwitz (2022) find that when a prospect’s contact information is purchased by a college, they are more likely to enroll at the college.

Prior research has not investigated how the architecture of student list products makes certain prospects more or less likely to be targeted by colleges. We argue that student list products incorporate societal structural inequality in ways that exacerbate racial inequality in college access. Drawing from the sociology of race, we argue that the architecture of student list products reproduce structural inequality in two ways. First, the College Board and ACT student list products have historically excluded non test-takers, but rates of test-taking differ by race (Blake & Langenkamp, 2022). Second, colleges control which prospect profiles they purchase by filtering on “search filters” (e.g., AP score, zip code). We conceptualize several student list search filters as “racialized inputs” (Norris, 2021) that disadvantage underrepresented students of color because they have been historically excluded from this input. In turn, administrators utilizing student list products may disproportionately exclude communities of color because of college admissions criteria or because they have incomplete knowledge about how search filters interact with local patterns of 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 information about recruiting and student lists vis-a-vis scholarship on recruiting. Second, drawing from the sociology of race, we develop a conceptual framework about the architecture of student list products and how these products are utilized. Next, we describe methods and present results. Simulating purchases using HSLs:09, we find that search filters conceptualized as racialized inputs – filtering on taking a college entrance exam, test-score thresholds, or zip code affluence – are associated with racial inequality in targeted prospects. Additional simulations show that filtering on multiple racialized inputs may have a compounding effect on racial inequality. Finally, targeted analyses of lists purchased by public universities reveal extreme structural inequality, even in purchases that explicitly target underrepresented groups.

The manuscript concludes by discussion implications for policy and scholarship. Although some colleges may use student list products to increase access for underrepresented groups, the potential for unintended inequities seems high. The question becomes, should policy-makers tolerate a product that is likely to do harm on the grounds that it is capable of

doing good? As The US Department of Education (2023a) considers new federal regulations for “third-party servicers”, we observe striking similarities between consumer report products, which are federally regulated, and student list products, which are not. Like consumer reports, student lists systematically lead to the extension of credit vis-a-vis student loans. With respect to state policy, we argue that state longitudinal data systems can create a more equitable alternative to private, third-party student lists. State-provided lists could serve an important function in direct admissions policies being adopted by several states (Odle & Delaney, 2022). We close by discussing future analyses of student list products and also broader scholarship about the role of third-party products and vendors in college access.

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) (Cheslock & Kroc, 2012; Hoxby, 2009; Winston, 1999), 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 enroll. “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 estimate the effect of specific interven-

tions on college access outcomes. Hoxby & Avery (2013) evaluate a nation-wide experiment that delivered customized information about admissions and financial aid to high-achieving, 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 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 for Black, Latinx, and low-income students.

¹After 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).

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 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. Whereas economics often evaluates the effect of outreach designed to increase underrepresented student enrollment at selective colleges, sociology finds that the recruiting efforts of selective colleges prioritize students from privileged schools and communities. 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), scholarship from sociology assumes that recruiting is something done by individual colleges. However, 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. 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 conceptually motivated search filters and we examine the usage of these filters 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, such as programs targeting working adults (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 Levitz, 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 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 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 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

²For 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).

and 98.8% of purchases specified at least one academic and one geographic filter.

A purchased list ([College Board template](#), [ACT template](#)) is 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, ethnicity, race, intended major). Information about academic achievement is very 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. Purchased lists are combined with student-as-first-contact inquiries 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, a \$7 brochure).

Ruffalo Noel-Levitz (2022b) reports that email, targeted digital advertising (e.g., Instagram), and direct mail are the top three methods for first contact with purchased high school names. Additionally, the average number of times purchased names are contacted before giving up is 8 and 11, respectively. for public and private 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

³For 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).

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.

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, administrators choose which inputs to filter on and which thresholds to apply to each filter.

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.

Neither standardized nor discretionary selection devices eliminate racial inequality stemming from structural racism. Bonilla-Silva (1997) criticizes social science disciplines (e.g., psychol-

ogy, 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. 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 credit 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 empirically 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.

Using geography as a predictive input was pioneered by geodemography, a branch of market research that estimates the behavior of consumers based on where they live (Burrows & Gane, 2006). Initial geodemographic systems scored individual localities based on consumer behavior. Subsequent systems (e.g., Mosaic by Experian) classify localities and individuals into like audience segments for marketers [CITE]. Geodemography emerged in the 1980s alongside efforts to fuse marketing and credit scoring, at a time when businesses dependent on customer credit transitioned from approving/rejecting applicants to the more aggressive model of pre-approving desirable customers (Leyshon & Thrift, 1999). Richard Webber – director of Experian Marketing Services UK and creator of Acorn and Mosaic – has been called the “founder of geodemographics” (McElhatton, 2004). Webber (1988, p. 36) described the integration of credit scores and geographic information to recruit customers, which is similar to the process of integrating SAT scores and geographic filters in student list products:

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.

Predictive analytics inputs. Another class of racialized inputs come from predictive analytics, which have been a focus of scholarship on algorithmic bias [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 based on “features that exist in data sets” (e.g., high school GPA, gender), versus those that rely on predictive

analytics, which “refers to the use of statistical models to generate new data” (e.g., predicted probability of recidivism). The creation of predictions proceeds in two steps: first, apply statistical models to previous cases to determine the predictors of an outcome; second, apply the results of these analyses to predict the outcome for future cases.

Predictive analytics are commonly utilized as the outcome variable in standardized selection devices. For example, algorithms assign credit scores to individuals (Poon, 2007) and to cities (Norris, 2021) based on analyses of predictors of default for past cases. Additionally, discretionary selection devices such as student list products utilize predictive analytics as another input to select on (Federal Trade Commission, 2014). For example, the ACT student list product offers the “Enrollment Predictor” search filter, which allows colleges to filter prospects based on their predicted probability of enrolling in your college (Schmidt, 2019). College Board uses predictive analytics to create new geographic filters (e.g., “Geomarket” filter, “Geodemographic Segment” filter), which essentially draw geographic borders for including/excluding future prospects based on analyses of the college-going behavior of past prospects (College Board, 2011). Whether predictive inputs are used as an outcome or an input, 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.” Harcourt (2007) refers to this phenomenon as the “ratchet effect,” whereby disproportionately targeted/excluded populations are predicted to have higher risk of an outcome, which amplifies disproportionate targeting/exclusion.

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. Sample selection bias is a concern whenever individuals are excluded from a selection device or statistical model because of missing values for some or all variables. Jillson (2021, para 4) warns that missingness correlated with race results in systematic racial

bias: “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.”

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 of SAT, ACT, and AP test-taking differ by race (Blake & Langenkamp, 2022; e.g., Hyman, 2017; Kolluri, 2018). Additionally, 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:

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 the Norris (2021, p. 5) racialized input criteria of being “theoretically and empirically correlated with historical racial disadvantage.” 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 drive 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. Geographic search filters enable universities to target prospects based

⁴Recently, 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.

⁵By 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).

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 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) while for-profit colleges systematically target poor, communities of color [Cottom (2017); Dache-Gerbino et al. (2018);]. These findings suggest that selective private colleges and public flagship universities may filter on affluent zip codes when purchasing student lists, while for-profits may filter on low-income, urban zip codes. 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 declines relative to the proportion excluded.

Many public universities filter on larger localities (e.g., county, state, CBSA) as means of targeting their local catchment area. We do not conduct analyses to this effect because this manuscript is primarily concerned with the potential for student list products to do harm. We would like to test propositions about the College Board Geomarket and Geodemographic Segment filters because scholarship raises concerns about filters that incorporate predictive analytics. We cannot recreate these filters using HSLS data because they rely on proprietary data, but we can examine actual student list purchases that utilized these filters.

3.3 Utilizing Student List Products

While the discussion so far has focused on individual filters, student list products are designed to filter on multiple search filters simultaneously and they grant administrators discretion over which filters to select and how many purchases to execute. Individual universities may

utilize student list products in ways that reduce or amplify racial inequality in college access. This section motivates analyses about how public universities utilize racialized search filters in concert with other search filters when purchasing student lists (RQ2) and about the racial composition of student list purchases that utilize multiple search filters (RQ3).

First, 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 (Castilla, 2008; Cotter, Medeiros, Pak, & Thorson, 2021; Norris, 2022). Discretionary selection devices allow the explicit or implicit individual bias to affect selection decisions (Burrell & Fourcade, 2021; Korver-Glenn, 2018). 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. Discretionary selection criteria often reflect occupational or professional norms, which may conceive of racialized inputs as objective, colorblind measures of merit (Hirschman & Bosk, 2020; Krippner & Hirschman, 2022; Tiako et al., 2021). Two necessary conditions to prevent bias in discretionary selection processes are transparency – selection criteria are clear to all stakeholders – and accountability – utilizing biased selection criteria yields consequences (Castilla, 2008; Norris, 2022). Finally, because 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 these inputs interact with local patterns of racial inequality (Korver-Glenn, 2018).

These findings motivate analyses about administrative discretion and racial inequality in student list purchases. First, universities may select academic achievement filters based on admissions standards, which are a function of university stakeholders and the admissions profession [CITE]. Until quite recently, most admissions offices viewed test scores as objective measures of achievement or aptitude (Hoxby & Avery, 2013). 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,

which has become a branding strategy for student list products. For example, College Board Student Search promising to “create a real pipeline of best-fit prospects” (College Board, n.d.) and consultancies encouraging universities to execute multiple student list purchases, each targeting different market segments (e.g., Waxman, 2019).⁶ 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 segregation. Considering a non-racialized filter (high school GPA), adding a racialized input filter (e.g., 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, College Board and ACT have added search filters based on predictive analytics (e.g., College Board “Geomarket”, ACT “Enrollment Predictor”). The Geomarket filter sub-divides states/metropolitan areas into distinct markets based on historical data about college enrollment. Geodemographic Segment filters allocate individual census tracts and individual high schools into distinct clusters based on past college enrollment. Creating new geographic borders based on historical patterns amplifies the effect of historic race-based inequality (Burrell & Fourcade, 2021). Furthermore, administrators utilizing these filters 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, 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. Considering the complexity of student list products and incomplete knowledge about race-based income inequality (Kraus et al., 2019), student list purchases designed to overcome one inequality may unintentionally amplify other so-

⁶In 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.”

cial 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 (HSL09) conducted by the National Center for Education Statistics (NCES). HSL09 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. HSL09 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 HSL09, 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

⁷Unweighted sample was rounded to 10 to meet restricted data regulations by NCES

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 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

⁸We collapse “Hispanic, no race specified” and “Hispanic, race specified” into a single category.

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 geodemographic 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 HSLS09 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 Academic Filters

We begin by describing the racial characteristics of prospects who completed standardized assessments in comparison to those who did not, which determines 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 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

⁹PSAT exams taken 2014 or before receive composite scores that range 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](#)

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).

Given differences in completion rates for standardized assessments by race, our conceptual framework outlines an interest in whether using GPA filters leads to greater racial parity between included versus excluded students relative to standardized assessments as thresholds increase. We therefore analyze the racial composition of included versus excluded students at minimum thresholds commonly used across student list purchase orders for high school GPA. Figure 7 shows the racial composition of included (top panel) and excluded (bottom panel) students across less than 2.0, 2.0 or greater, 3.0 or greater, and 3.5 or greater thresholds of GPA. Similar to standardized assessments, Figure 7 suggests proportions of included prospects increase for White students and Asian students (although modestly) while proportions of included prospects decrease for Hispanic and Black students as GPA thresholds increase. For instance, White student proportions increase from 37% (relative to 57% excluded, $p < 0.000$) at GPA less than 2.0 to 71% (relative to 57% excluded, $p < 0.000$) at GPA 3.5 or greater. On the other hand, Hispanic and Black included student proportions (relative to excluded proportions) decrease from 30% (19% excluded, $p < 0.000$) and 22% (11% excluded, $p < 0.000$) at GPA less than 2.0 to 11% (24% excluded, $p < 0.000$) and 4% (15% excluded, $p < 0.000$) at GPA 4.0 or greater, respectively. The 12,591 American Indian/Alaska Native students included at GPA less than 2.0 (15,928 excluded, $p < 0.000$) also decline to 1,304 (27,216 excluded, $p < 0.000$) at GPA 3.5 or greater score threshold.

However, Figure 7 shows that GPA filters at “middle” thresholds (2.0 to 3.0) lead to smaller declines in proportions of included Hispanic and Black students relative to middle thresholds of SAT and PSAT filters. In increasing GPA from 2.0 or greater to 3.0 or greater, the proportions of Hispanic and Black included students decrease by 3 percentage points or less. In comparison, increasing PSAT from scores 120 or greater to 200 or greater results

¹⁰Proportional 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

in an up to nine percentage point decrease in the number of included Black students. This pattern of lesser relative declines in the proportion of underrepresented minority students included at “middle” thresholds is also evident by AP filters (see Figure 6). However, given the disparities in AP course availability and exam completion rates, a considerable smaller number of overall included students are captured by AP filters than GPA.

5.2 Geographic Filters

Proposition P3 and Proposition P4 conceptualize how the use of geographic filters may result in greater racial disparities in proportions of included prospects relative to excluded prospects. For instance, 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.

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 through 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

To contextualize these findings, Figure 9 presents similar results as Figure 8 for two specific CBSAs: Los Angeles and New York. The left panel of Figure 6 shows the racial composition of included (top) and excluded (bottom) students across percentiles of zip code affluence for Los Angeles. Similar to results across all CBSAs in Figure 8, proportions of included prospects relative to excluded prospects increase for White students while proportions decline for Hispanic and Black students as zip codes become more affluent. For instance, White student proportions increase from 5% (relative to 17% excluded, $p < 0.000$) at the 20th percentile, to 28% (relative to 12% excluded, $p < 0.000$) at 50th-79th percentiles, and up to making up 76% of included prospects (relative to 12% excluded, $p < 0.000$) for zip codes at the 90th percentile or higher of affluence. On the other hand, Hispanic and Black included student proportions (relative to excluded proportions) decrease from 73% (54% excluded, $p < 0.000$) and 14% (7% excluded, $p < 0.000$) at zip codes in the lower 20th percentiles of affluence to 4% (60% excluded, $p < 0.000$) and 3% (9% excluded, $p < 0.000$) at the 90th percentile or higher of affluence, respectively. While New York provides a different racial composition of students than Los Angeles, similar and statistically significant patterns persist.

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

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

areas that have higher degrees of social and economic activity) will be captured via CBSA.

To analyze whether filtering on smaller geographic localities is associated with greater racial disparities in included prospects relative to excluded prospects (Proposition P4), we compare the racial characteristics of prospects based on zip code filters versus county filters. We categorize all zip codes and counties based on levels of median household income within their respective Core Based Statistical Areas (CBSA), given results for Proposition P3, to analyze whether relationships between racial composition and geographic level change across levels of affluence. Therefore, zip codes and counties are categorized as low, moderate, and high income based on their median household incomes falling below 30th percentile, within 30th-70th percentiles, or greater than 70th percentile of affluence within their respective CBSAs.

Figure 10 presents the racial composition of prospects when using zip code filters (left panel) in comparison to county filters (right panel). There are modest differences between included versus excluded groups when comparing zip code filters to county filters at low and moderate levels of affluence. For example, Hispanic students are slightly overrepresented within included relative to excluded prospects (23% versus 22%, $p < 0.000$) by zip code at moderate levels of affluence but underrepresented (20% versus 22%, $p < 0.000$) at moderate affluence levels for county. Similarly, Black students are slightly underrepresented within included relative to excluded prospects (12% versus 15%, $p < 0.000$) by zip code at moderate levels of affluence but overrepresented (20% versus 13%, $p < 0.000$) at moderate affluence levels for county.

Differences between the use of zip code filters and county filters are most evident in high levels of affluence. For instance, Hispanic and Black students make up 16% and 8% of included prospects and 24% and 16% of excluded prospects for the most affluent zip codes, respectively. However, in the most affluent counties, Hispanic students make up an equal share of included and excluded prospects (22%). Similarly, filtering on the most affluent counties leads to smaller differences in proportions of Black students within included and excluded prospects (11% versus 16%, $p < 0.000$). These results suggest filtering for smaller geographic localities (i.e., zip codes) is associated with greater racial disparities in included prospects

relative to excluded prospects in comparison to larger geographic localities (i.e., counties) at higher levels of affluence, which results for Proposition P3 suggest are the thresholds with the greatest lack of racial parity between included versus excluded students.

5.3 Combinations of Filters

Our last set of analyses focus on assessing whether filtering on multiple criteria compound the effect of racial disparities in which prospects are included versus excluded. We draw on Figure 3 to select common filters used across orders. We begin by combining the two most common academic filters: GPA and SAT. Figure 11 (top panel) presents the racial composition of prospects included when filtering on GPA greater than or equal to 3.0 while simulating increases to minimum SAT thresholds at increments of 50 beginning at scores just above the 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 while Black and Hispanic students make up significantly smaller proportions of included prospects 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 increases. 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 when filtering for similar thresholds of PSAT score (Figure 5) individually.

In order to assess the 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

GPA's 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 GPA's 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.

Lastly, we assess 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 GPA's 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 GPA's 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.

6 Discussion

6.1 Summary

Racial inequality in college access remains an enduring challenge for policymakers concerned about equality of opportunity. The market for college access depends on students knowing where they want to enroll and colleges knowing who they want to enroll. Student lists are a match-making intermediary connecting colleges to prospective students. Recent research suggests that participating in student list products positively affect college access, with stronger effects for underrepresented students (Howell et al., 2021; Moore, 2017; Smith et al., 2022). However, the underlying architecture of student list products may incorporate structural inequality in ways that systematically disadvantage underrepresented students.

This manuscript investigates the College Board student list search filters, how those filters are utilized, and the racial composition of students who are included versus excluded in student list purchases. We develop a conceptual framework by drawing from the sociology of race. We argue that several academic and geographic search filters are “racialized inputs” (Norris, 2021), which are correlated with race because of historical, race-based exclusion from the input. We recreate selected College Board search filters using a nationally representative sample of 9th graders from 2009. Results for proposition P1 show that conditioning on college entrance exam test-taking (SAT/ACT, PSAT/PreACT, and/or AP) results in racial disparities in students are included in the underlying student list database. Filtering on higher test-score thresholds (P2) is associated with larger proportions of white and Asian students and smaller proportions of Black and Latinx students being included in student list purchases. Filtering on higher levels of zip code affluence (P3) is associated with higher proportions of white students and lower proportions of Black and Latinx students.

Actual student list purchases select multiple search filters simultaneously. Simulations based on HSLs data suggest that filtering on multiple criteria (GPA and test score, test score and zip code affluence, KARINA FILL OUT) can have a compounding effect on racial disparities between included versus excluded prospects. Student list products are discretionary rather than standardized selection devices. Scholarship on product utilization suggests that

administrator discretion likely increases racial inequality in recruiting. In short, these are powerful, complicated products that incorporate structurally racist inputs, and any person affiliated with a Title IV institution can execute student list purchases. We find that public research universities tended to utilize higher test score thresholds than MA/doctoral universities, yielding racial disparities that can be construed as a function of admissions standards. However, specifying multiple filters can easily yield unintended racial inequality because administrators may have incomplete knowledge about how these filters intersect with local patterns of segregation, which differ substantially within and between metropolitan areas. Analyses of student lists purchased by public universities support these concerns, showing dramatic inequality in purchases that filter on predictive analytics (geodemography) and purchases that targeted underrepresented populations.

6.2 Policy Implications

Federal policy. We observe striking parallels between the functions of consumer credit reports and student list products. Credit scores are designed to predict the probability of repayment, thereby overcoming “the chronic problems of information asymmetries” (Leyshon & Thrift, 1999, p. 434) by enabling firms dependent on customer credit “to distinguish ‘good’ from ‘bad’ customers ‘at-a-distance’” (Leyshon & Thrift, 1999, p. 434). Similarly, in the market for college access, SAT/ACT scores were viewed as a measure of achievement or aptitude that helped create a national higher education market by enabling colleges to make apples-to-apples comparisons between applicants from different places (Hoxby, 1997, 2009). Consumer reporting companies like Equifax (2023) wrap credit scores, geographic information, and other consumer information into products that filter prospective customers, enabling firms to transition from approving/rejecting applicants to the more aggressive model of pre-approving desirable customers (Leyshon & Thrift, 1999). Similarly, student list products filter prospects by academic achievement, geographic location, and other characteristics. Consumer credit report 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 products lead to the extension of credit. We argue that student list products systematically lead to the extension of credit. At the top of

the enrollment funnel (Figure 2), colleges obtain the contact information of prospective students by purchasing student lists. These purchased leads are systematically targeted with recruiting interventions designed to maximize the probability of conversion in subsequent stages of the enrollment funnel. At the end of the funnel, admits are offered financial aid packages – including loan aid – designed to increase the probability of enrollment. Student loans are the second highest source of debt in the U.S [CITE]. Clearly, there is a systematic link between student list products and student debt and, therefore, a compelling rationale for federal regulation of student list products and student list vendors.

Historically, the U.S. Department of education has focused on regulating Title IV institutions (direct providers) and “third-party servicers” that administer federal financial aid. The US Department of Education (2023a) “Dear Colleague” letter broadened the definition of third-party servicers to include entities “interact with prospective students for the purposes of recruiting or securing enrollment.” Although the regulations clearly target the activities of online program managers (OPMs), it remains unclear whether they cover student list vendors and products. Given that these regulations seek “to identify and remedy the root causes of unaffordable debts” (US Department of Education, 2023b), we recommend that the definition of third-party servicers be revised to include student list vendors, which are the root intermediaries linking students to Title Iv institutions.

Any future federal regulatory efforts should consider recent dynamics in the market for student list data. Although XX states now require high school students to take the SAT/ACT [CITE], the diffusion of test-optional admissions policies may reduce the number of test-takers from other states who are included in College Board/ACT student list products. New sources of student list data – particularly online college search engines and college planning software purchased by high schools – and new student list vendors (e.g., PowerSchool) have entered the market. Second, The distinction between student list vendor and consultant is disappearing. Both College Board and ACT offer enrollment management consulting services. At the same time, large enrollment management consulting firms (e.g., EAB, Ruffalo Noel Levitz) have become student list vendors. Third, rather than selling names at a price per-prospect, the emerging trend is to wrap a proprietary database of prospects

within a software-as-service recruiting products (e.g., EAB’s Enroll360) that universities must purchase in order to obtain access to these prospects. Jaquette et al. (2022) describe these dynamics in greater detail.

State policy. State policymakers are positioned to develop “public option” alternatives to private, third-party student lists products. This idea is inspired by national voter databases used in political campaigns, which are based on voter files collected by local and state governments (Culliford, 2020). Similarly, we argue that states can create student lists from statewide longitudinal student data systems, which contain information about academic achievement and contact information. Students and their parents would have the opportunity to opt in or opt out. Importantly, colleges would receive the “names” of students who opt in for free. Therefore, the public option would have no need for search filters that help colleges micro-target the “right” students. Jaquette & Salazar (2022) provides more detail about essential product features and challenges to overcome.

Our proposed public option complements the adoption of direct admissions policies by state policymakers. Similar to the model of pre-approving customers in consumer markets, direct admissions policies make proactive admissions offers to high school students that meet admissions criteria. Odle & Delaney (2022) report that five states have adopted direct admissions or are considering adoption. Analyses by Odle & Delaney (2022) found that adoption by Idaho increased first-time undergraduate enrollment at the campus-level and the state-level. Brown & Burns (2023) state that identifying eligible students and their contact information is a significant barrier to equality of opportunity in direct admissions. To date, privately-led and state-led direct admissions policies to date rely on student lists purchased from third-party vendors. To make sure all eligible students in the state are included, Brown & Burns (2023) recommend that states adopting direct admissions should develop student lists based on state longitudinal data systems. At present, one state higher education system informed us that they are creating a student list product based on Jaquette & Salazar (2022) as part of their broader effort to implement a direct admissions system.

6.3 Implications for Scholarship

Scholarship on recruiting largely assumes that recruiting is done by individual colleges and universities. Although Smith et al. (2022) analyze the effect of a college purchasing a prospect profile on the prospect’s college choice, prior research has not investigated how the underlying architecture of student list products makes prospects more or likely to be targeted. As the first manuscript to focus on student list product architecture and utilization, we developed a broad conceptual framework to create scaffolding for more targeted research.

Future research should examine filters based on predictive analytics, which model past cases to make predictions about future cases. 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. Manuscripts that analyze a single filter can yield broader insights about the implications of predictive analytics. 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. In 2021, College Board (2021) released three new “Environmental Attributes” geodemographic search filters that allocate individual high schools to categories: (1) Travel Rates (out of state); (2) Travel Rates (distance from home); and (3) AP engagement rates.

Another topic for future research is demographic search filters, which allow universities to target prospects by race, ethnicity, sex, and first-generation status. The equity rationale is that these filters facilitate access for underrepresented populations, particularly in a post affirmative action landscape. However, purchases that target one inequity may reproduce others. For example, depending on the set of filters selected, purchases that filter for underrepresented racial/ethnic groups may disproportionately target students from affluent, predominantly white schools and communities. Demographic search filters may be the subject of future legal debate. The role of scholarship is to analyze demographic filters systematically

so that legal and policy debates are based on empirical evidence.

Future research should examine student lists in a test-optional world. On one hand, test-optional policies may increase application and enrollment from historically underrepresented groups. On the other hand, a long-term decline in test-takers may erode the coverage of College Board and ACT student list products, creating opportunities for other products (e.g., Cappex, Intersect, Scoir). Future research should examine these products, which prospects they include and how search filters facilitate inclusion or exclusion.

Finally, we hope that this manuscript motivates critical education policy research that focuses on third-party products and vendors more broadly. 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 “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 third-party products and vendors structure educational opportunity along racial, class, and geographic dimensions.

For example, future research should examine how college access is structured by enrollment management consulting firms. Anecdotally, in our public records data collection, roughly half the public universities outsourced student list buys to consulting firms (Salazar et al., 2022). Several enrollment management consultancies sell algorithmic products designed to make recommendations about list buys (Fire Engine RED, 2022; Ruffalo Noel Levitz, 2021). For example, one product from Ruffalo Noel Levitz recommends how many names a college should buy from each zip code (James Madison University, 2017). These firms also serve recruiting interventions to purchased names (Jaquette et al., 2022). Beyond recruiting, consultancies advise on tuition pricing, and financial aid. However, extant scholarship on college access assumes that universities perform these functions in-house. To the extent that universities outsource enrollment management to consultancies, these consultancies substantially structure which students are funneled to which institution, and the financial aid offers they receive. Scholarship can inform how enrollment management consulting firms are incorporated into The US Department of Education (2023b) third-party servicer regulations.

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>
- 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>
- Brown, L., & Burns, R. (2023). *The link between student lists and direct admissions: Pros, cons, and policy implications*. State Higher Education Executive Officers.
- 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/>
- 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>
- College Board. (2021). Introducing environmental attributes. *YouTube*. Retrieved from <https://www.youtube.com/watch?v=VmTU9sb4ZiY>
- 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., Micheltmore, 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.
- 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>
- 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>
- 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 post-secondary 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. (2022). *Student list policy: Problems, regulations, and a solution*. Washington, DC: TICAS. Retrieved from #
- 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>
- 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>
- 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/10.1080/0308514991000167688>

[//doi.org/10.1080/030851499000000013](https://doi.org/10.1080/030851499000000013)

- 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>
- McElhatton, N. (2004). Secrets of my success: Richard webber, founder of geodemographics. *Campaign Live UK*. Retrieved from <https://www.campaignlive.co.uk/article/secrets-success-richard-webber-founder-geodemographics/226747>
- 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'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy* (First edition., pp. x, 259 pages). New York: Crown.
- Odle, T. K., & Delaney, J. A. (2022). You are admitted! Early evidence on enrollment from idaho's direct admissions system. *Research in Higher Education*, 63(6), 899–932. Journal Article. <https://doi.org/10.1007/s11162-022-09675-x>
- 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
- 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>
- Posiecznik, A. (2017). *Selling hope and college merit, markets, and recruitment in an unranked school*. Ithaca ; London: Cornell University Press. Retrieved from <https://doi.org/10.1080/030851499000000013>

[//muse.jhu.edu/book/52616](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>
- 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 cases, more access? 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. (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/>
- 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. (2019). *Prospect search filters*. Encoura. Retrieved from <https://helpcenter.encoura.org/hc/en-us/articles/360035260452-Prospect-Search-Filters->

- 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. (2023a). *Requirements and responsibilities for third-party servicers and institutions*. US Department of Education. Retrieved from <https://fsapartners.ed.gov/knowledge-center/library/dear-colleague-letters/2023-02-15/requirements-and-responsibilities-third-party-servicers-and-institutions-updated-feb-28-2023>
- US Department of Education. (2023b). *Update on the department of education's third-party servicer guidance*. US Department of Education. Retrieved from <https://blog.ed.gov/2023/04/update-on-the-department-of-educations-third-party-servicer-guidance/>
- 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.
- 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](https://doi.org/10.1215/00141801-1999-002)

8 Tables

Table 1: Top filter combinations used in College Board orders purchased by 14 public universities

Filters	Count	Pct	Cum count	Cum pct
HS grad class, GPA, SAT, Zip code	206	24.8	206	24.8
HS grad class, GPA, PSAT, Zip code	145	17.5	351	42.3
HS grad class, GPA, SAT, PSAT, Rank, State, Race	39	4.7	390	47.0
HS grad class, SAT, State	38	4.6	428	51.6
HS grad class, GPA, SAT, PSAT, Zip code	28	3.4	456	54.9
HS grad class, PSAT, State	28	3.4	484	58.3
HS grad class, GPA, SAT, State	23	2.8	507	61.1
HS grad class, GPA, PSAT, State, Race	20	2.4	527	63.5
HS grad class, PSAT, State, Low SES	20	2.4	547	65.9
HS grad class, GPA, PSAT, State	19	2.3	566	68.2
HS grad class, GPA, AP score, Geomarket	15	1.8	581	70.0
HS grad class, PSAT, Geomarket	14	1.7	595	71.7
HS grad class, GPA, SAT, PSAT, State, Segment, Gender	13	1.6	608	73.3
HS grad class, SAT, State, College size	11	1.3	619	74.6
HS grad class, SAT, State, Low SES, College size	11	1.3	630	75.9
HS grad class, GPA, SAT math, SAT reading/writing, State, Segment	10	1.2	640	77.1
HS grad class, PSAT, State, Segment	10	1.2	650	78.3
HS grad class, PSAT, Geomarket, Low SES	9	1.1	659	79.4
HS grad class, SAT, Geomarket	9	1.1	668	80.5
HS grad class, GPA, AP score, State	8	1.0	676	81.4

Table 2: Descriptive Statistics

	Unweighted	N	SE	Pct
<i>Race/Ethnicity</i>				
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
<i>Academic Filters</i>				
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

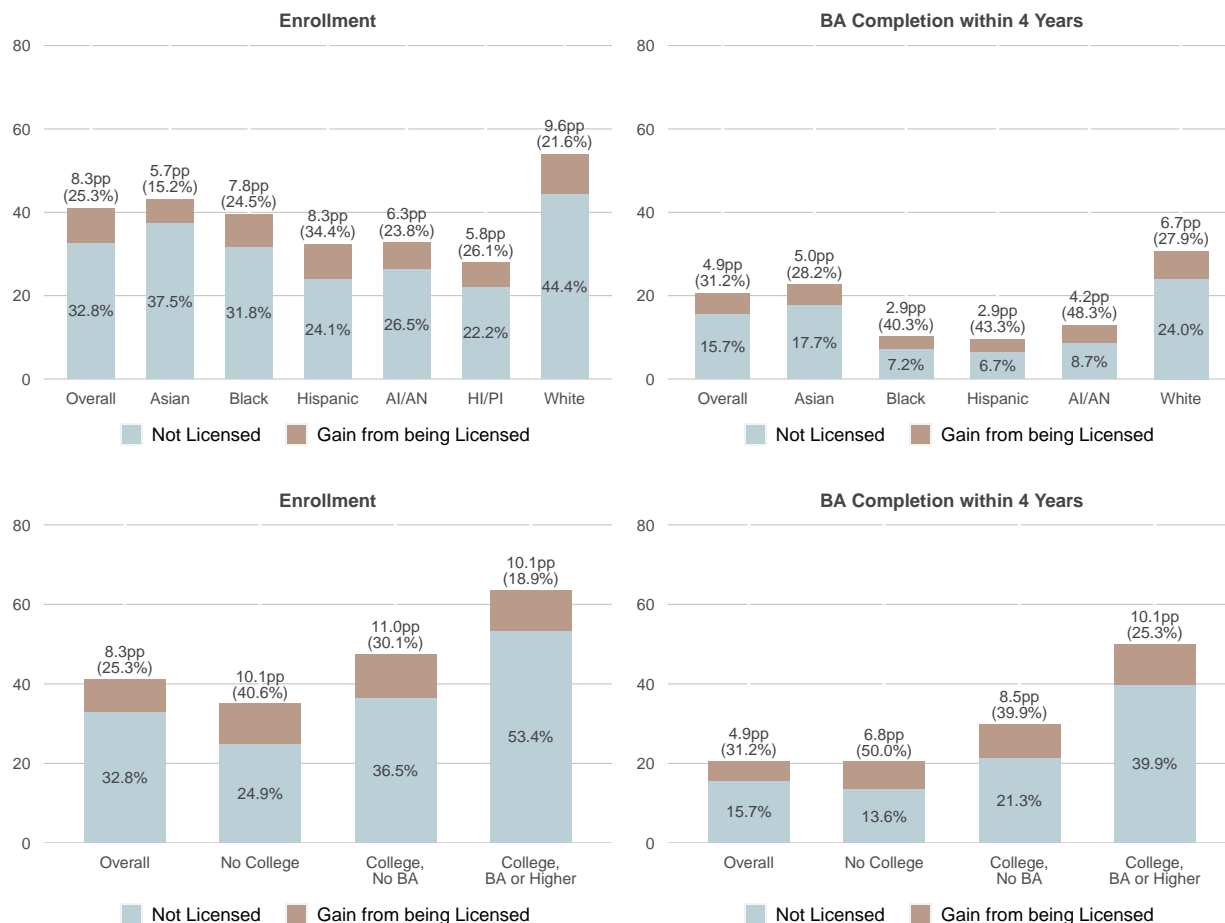
* Unweighted sample sizes rounded to nearest 10 per NCES restricted data license regulations

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

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

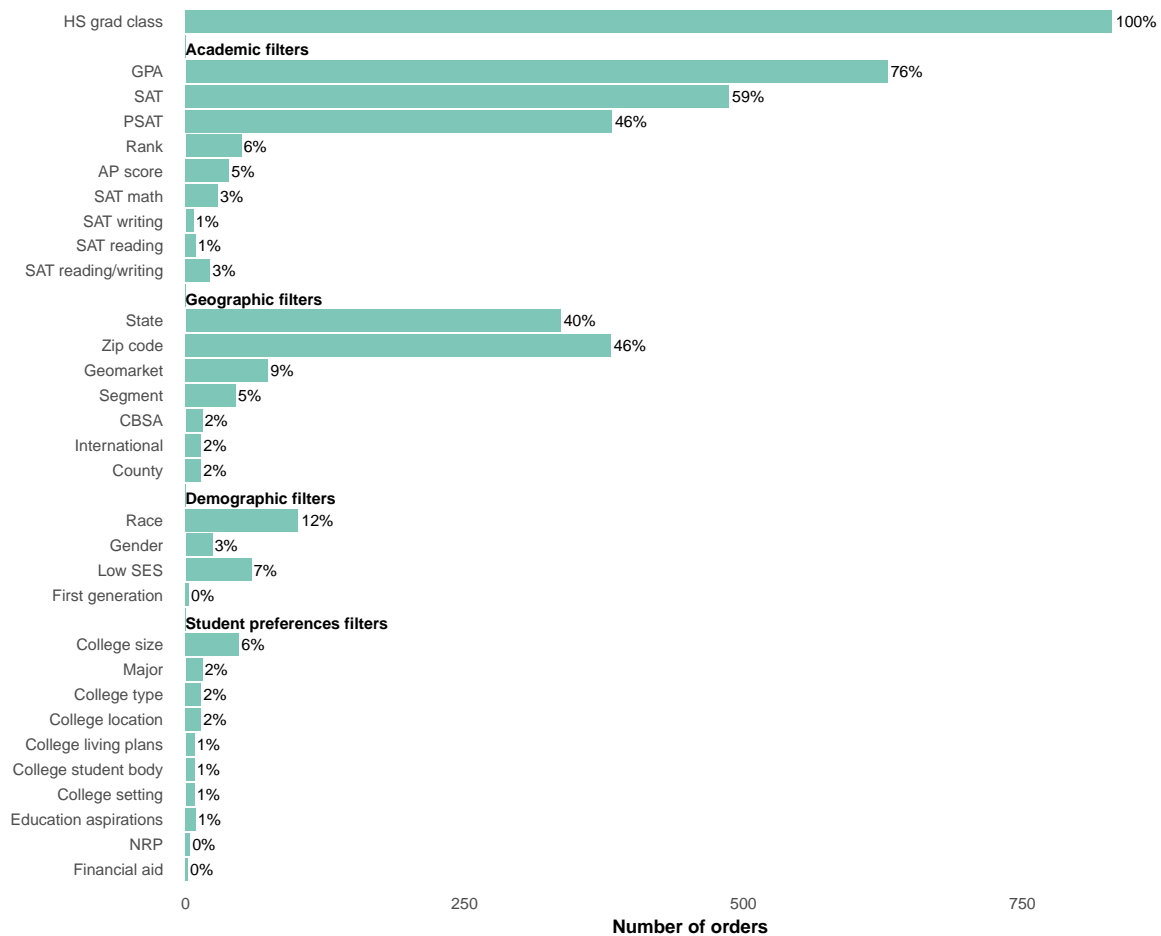


Figure 4: Test Takers Across SAT, PSAT, and AP Assessments

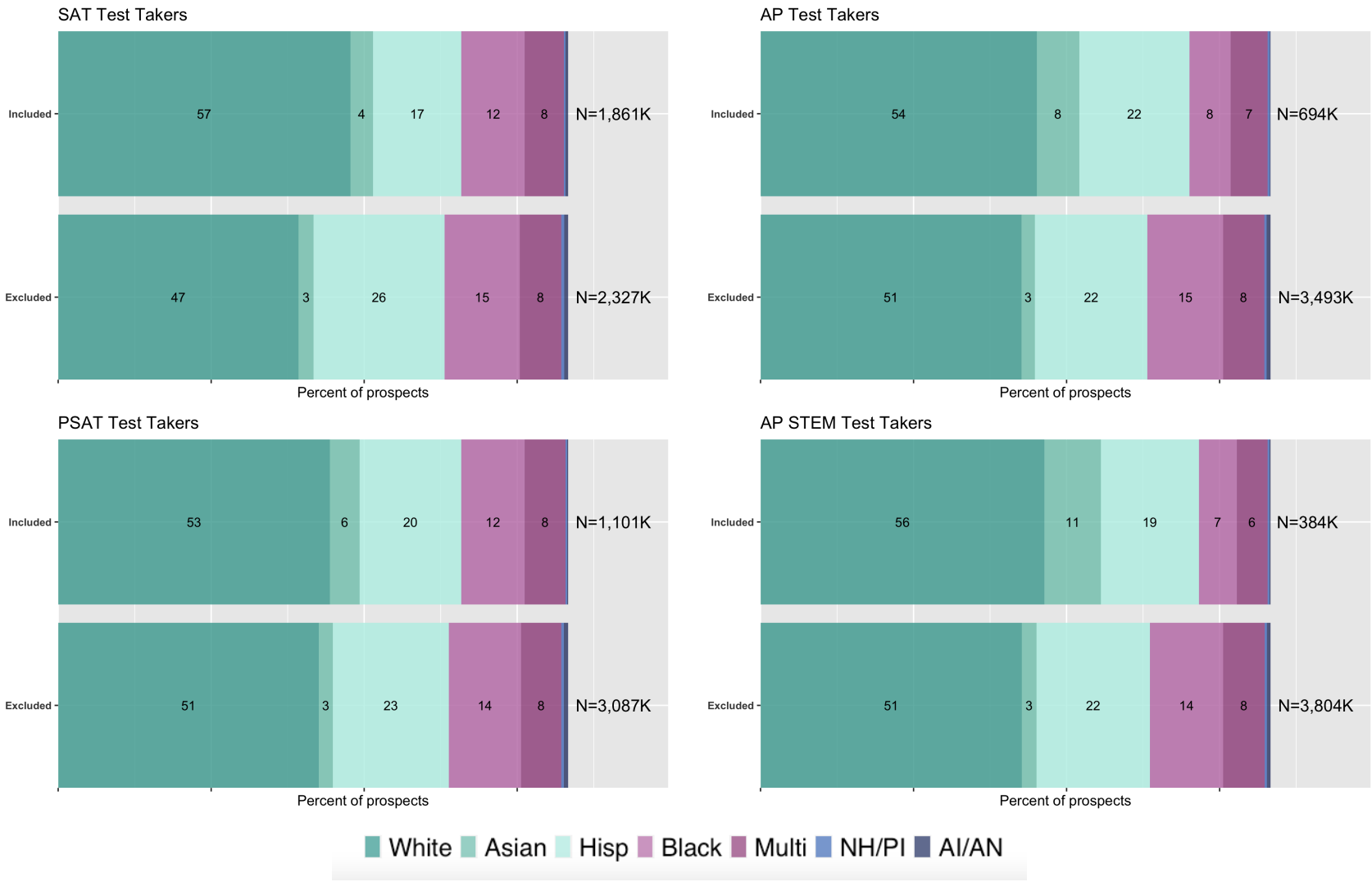
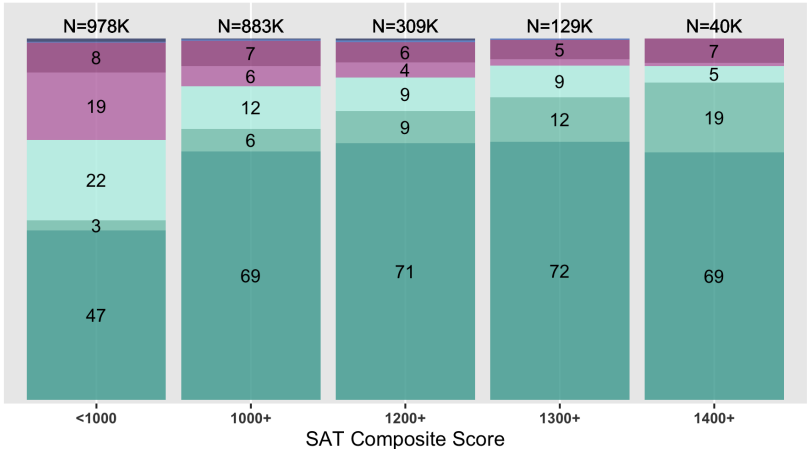
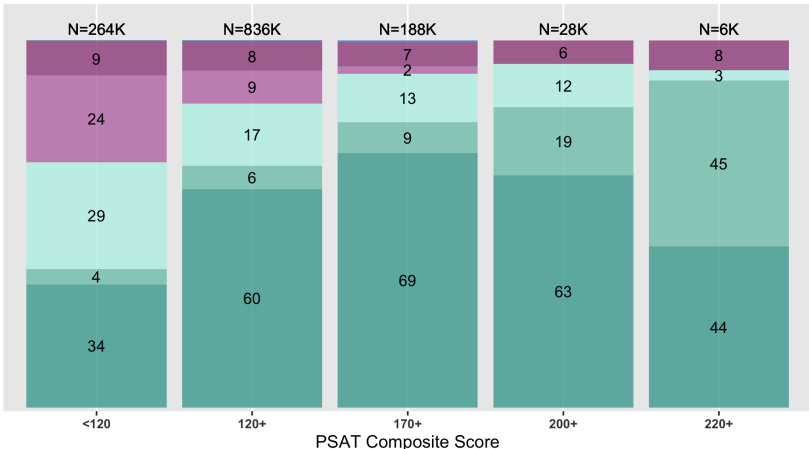


Figure 5: SAT and PSAT Filters Across Thresholds

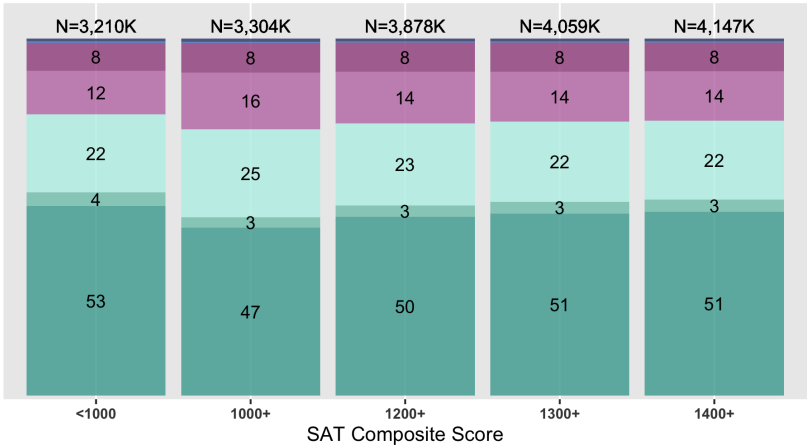
Included Prospects by SAT Score



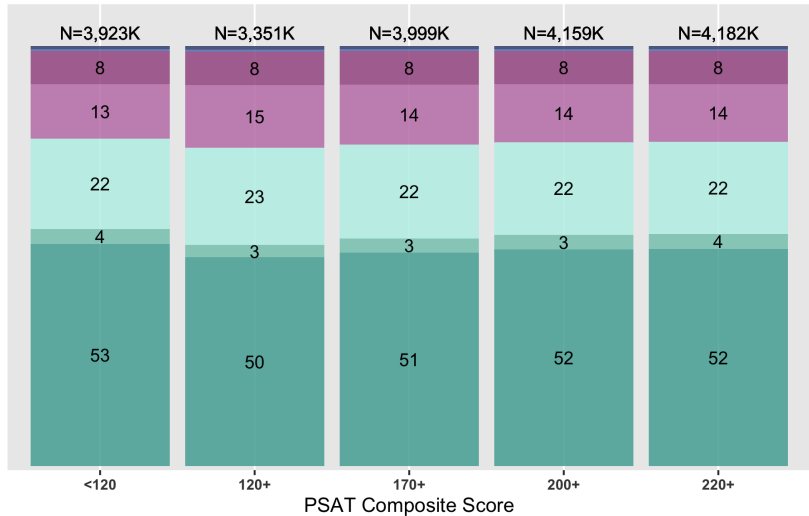
Included Prospects by PSAT Score



Excluded Prospects by SAT Score



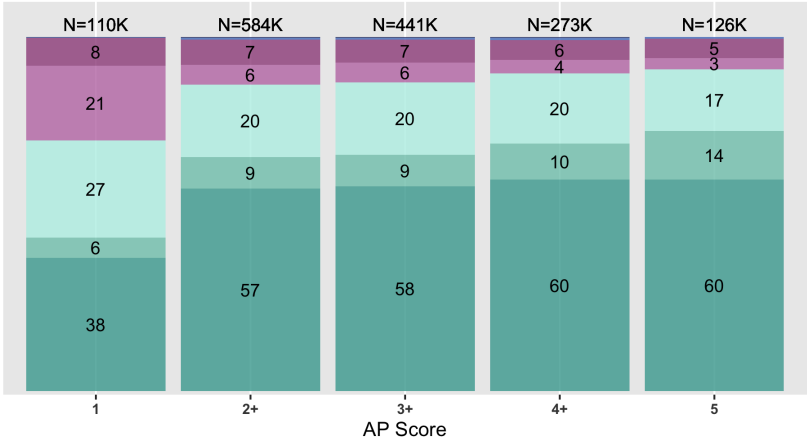
Excluded Prospects by PSAT Score



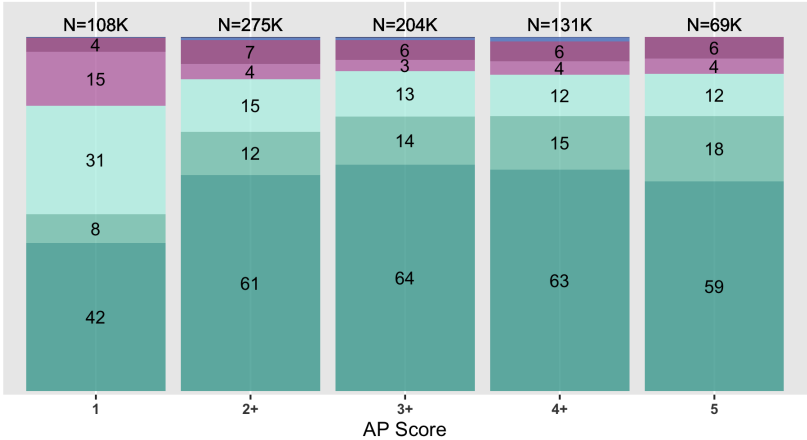
White Asian Hisp Black Multi NH/PI AI/AN

Figure 6: AP Filter Across Thresholds

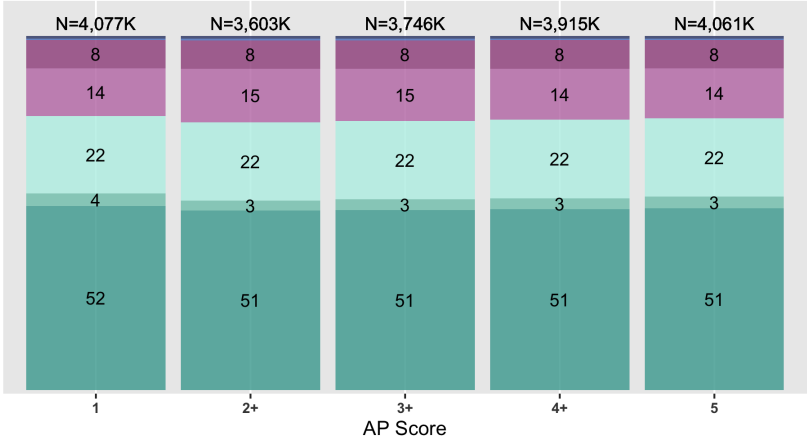
Included Prospects by AP Score



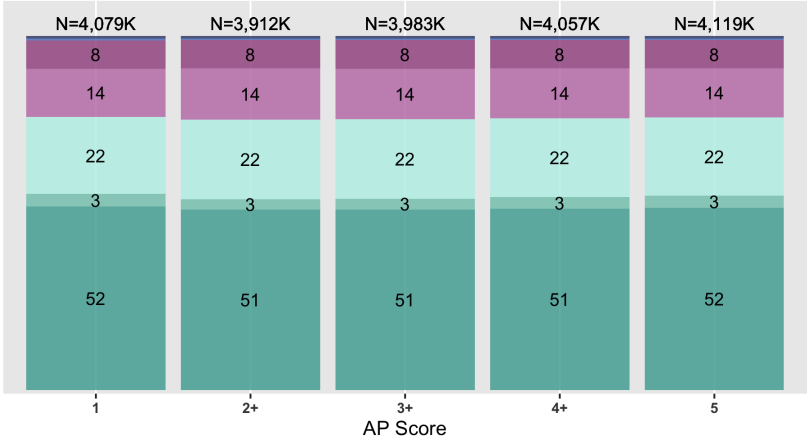
Included Prospects by AP STEM Score



Excluded Prospects by AP Score



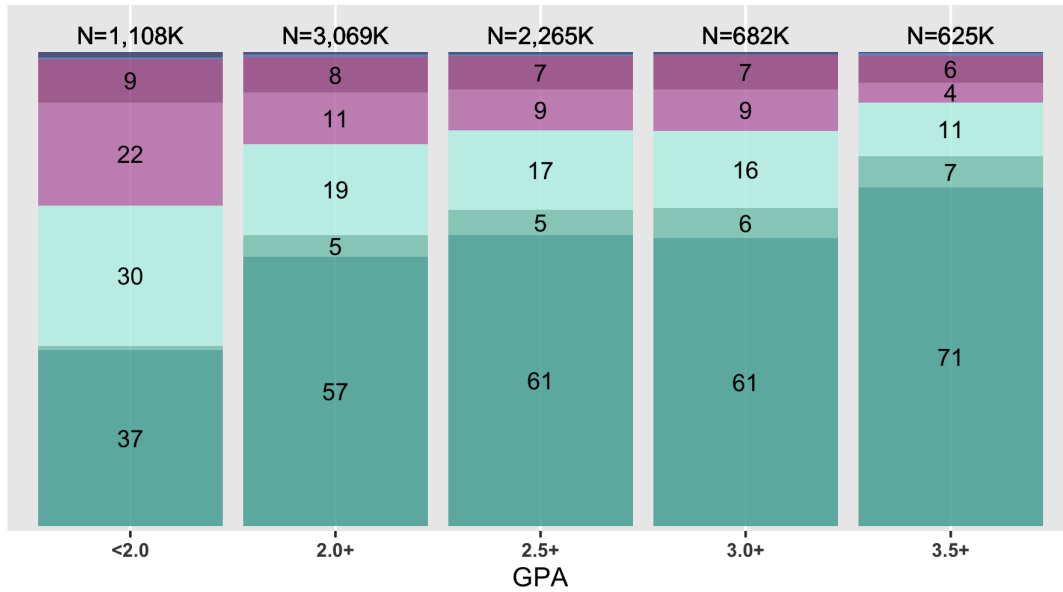
Excluded Prospects by AP STEM Score



White Asian Hisp Black Multi NH/PI AI/AN

Figure 7: GPA Filter Across Thresholds

Included Prospects by GPA



Excluded Prospects by GPA

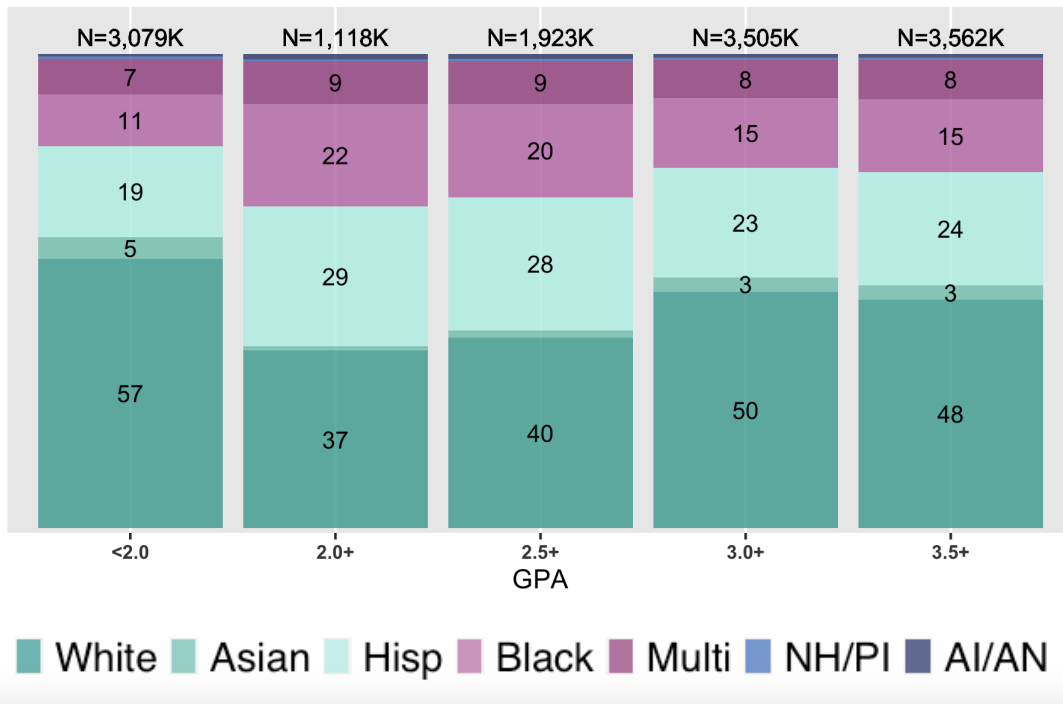
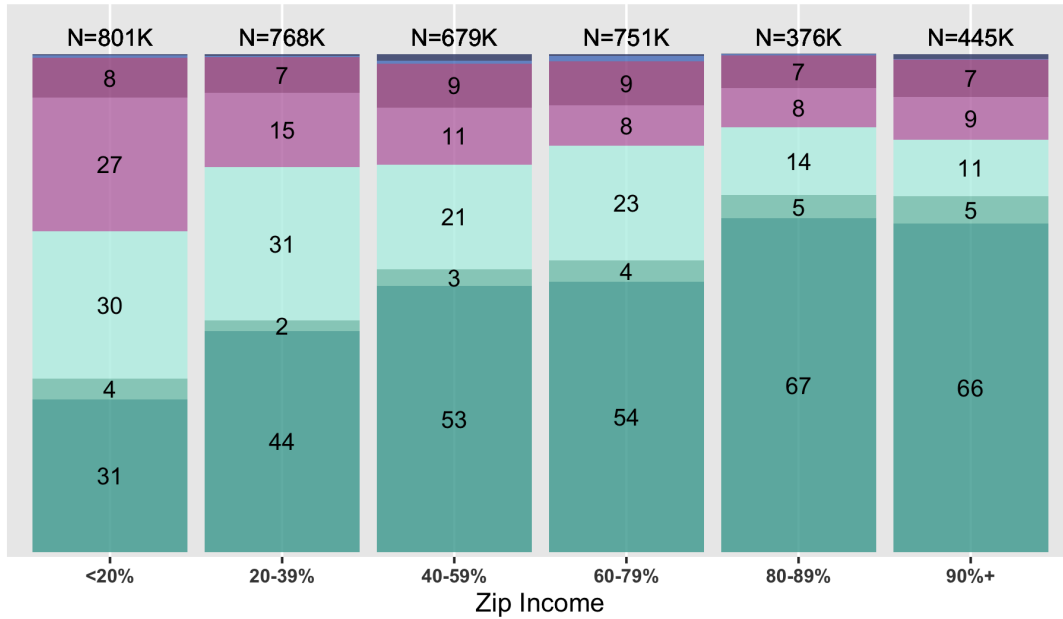
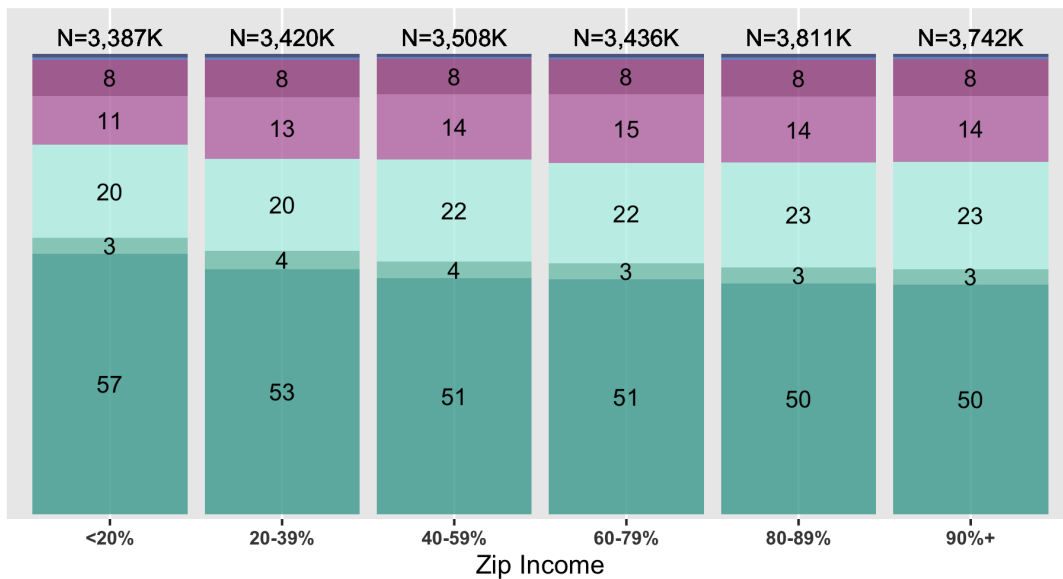


Figure 8: Zip Code Filter Across Affluence Percentiles

Included Prospects by Zip Code Income, within CBSA



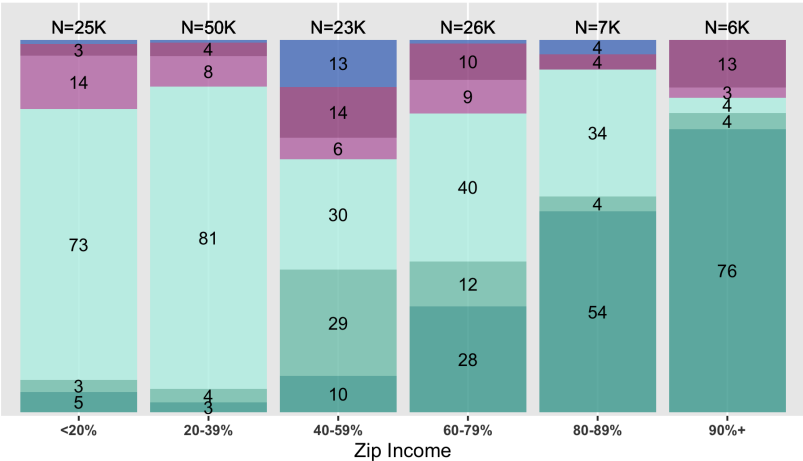
Excluded Prospects by Zip Code Income, within CBSA



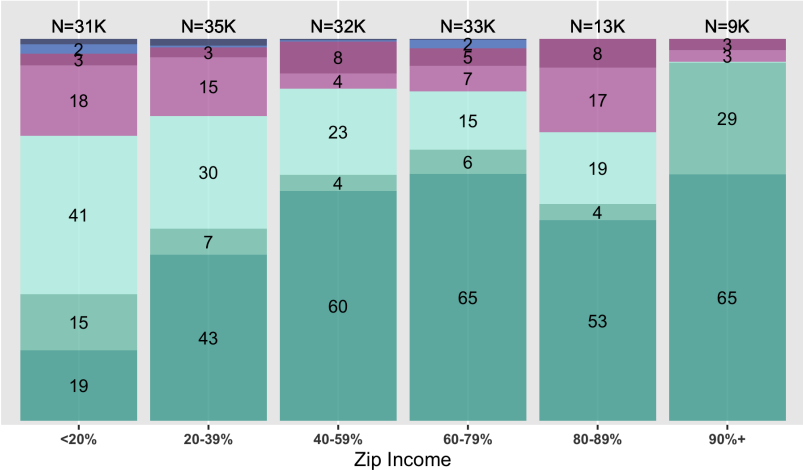
White Asian Hisp Black Multi NH/PI AI/AN

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

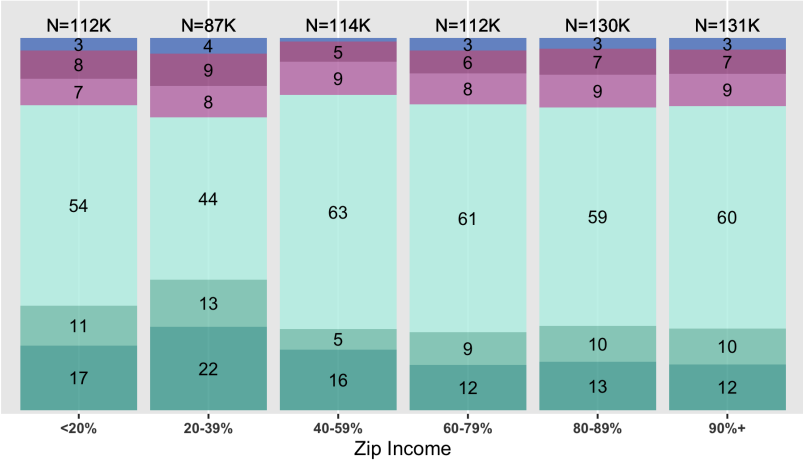
Included Prospects by Zip Code Income, Los Angeles



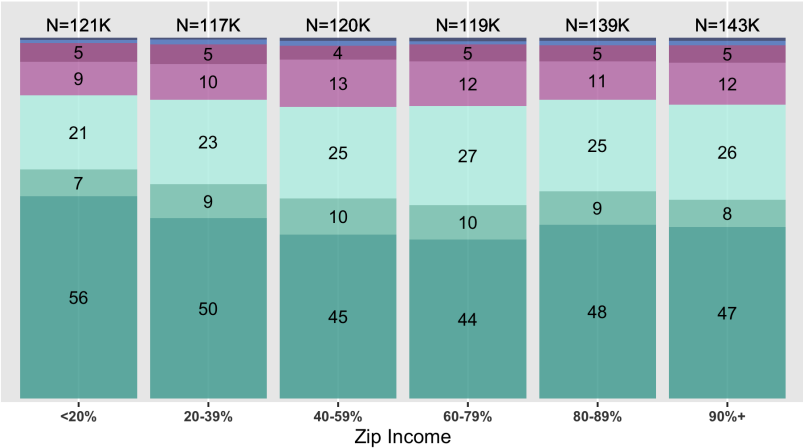
Included Prospects by Zip Code Income, New York



Excluded Prospects by Zip Code Income, Los Angeles



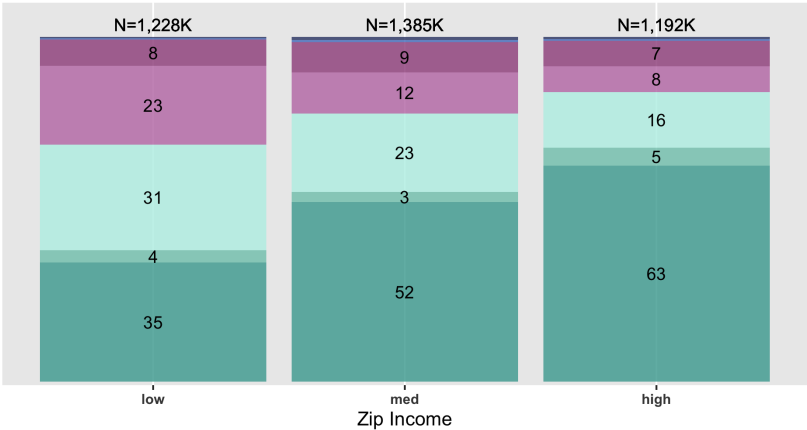
Excluded Prospects by Zip Code Income, New York



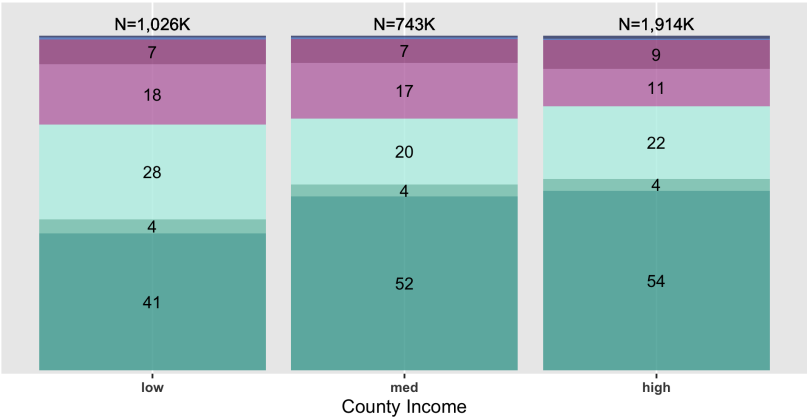
White Asian Hisp Black Multi NH/PI AI/AN

Figure 10: Zip Code and County Filters Across Affluence Levels

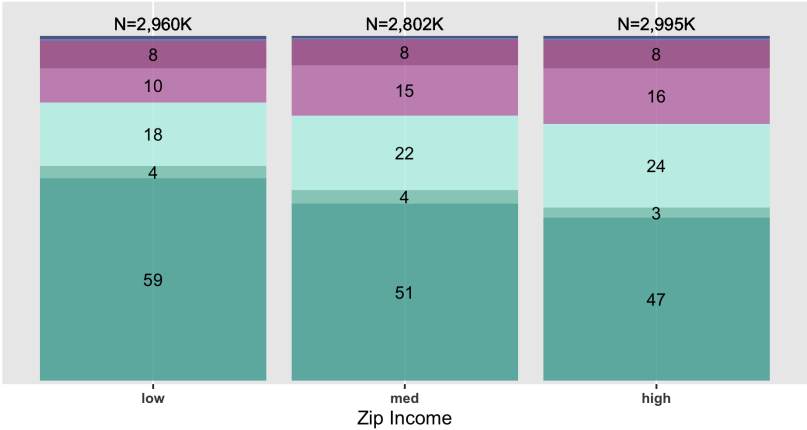
Included Prospects by Zip Code Income, within CBSA



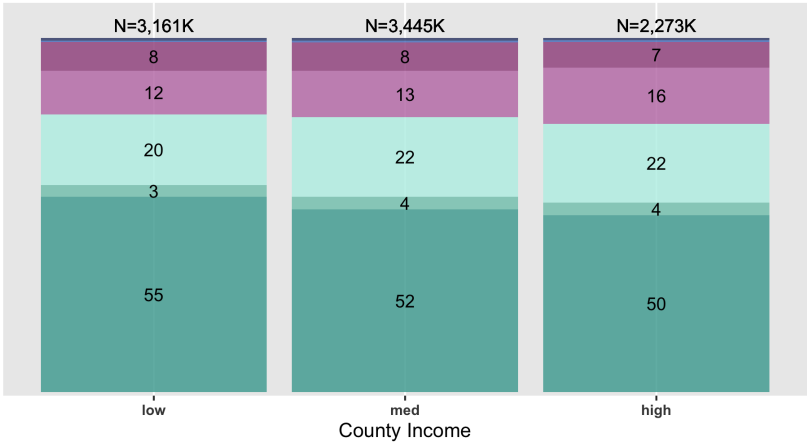
Included Prospects by County Income, within CBSA



Excluded Prospects by Zip Code Income, within CBSA



Excluded Prospects by County Income, within CBSA



White Asian Hisp Black Multi NH/PI AI/AN

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

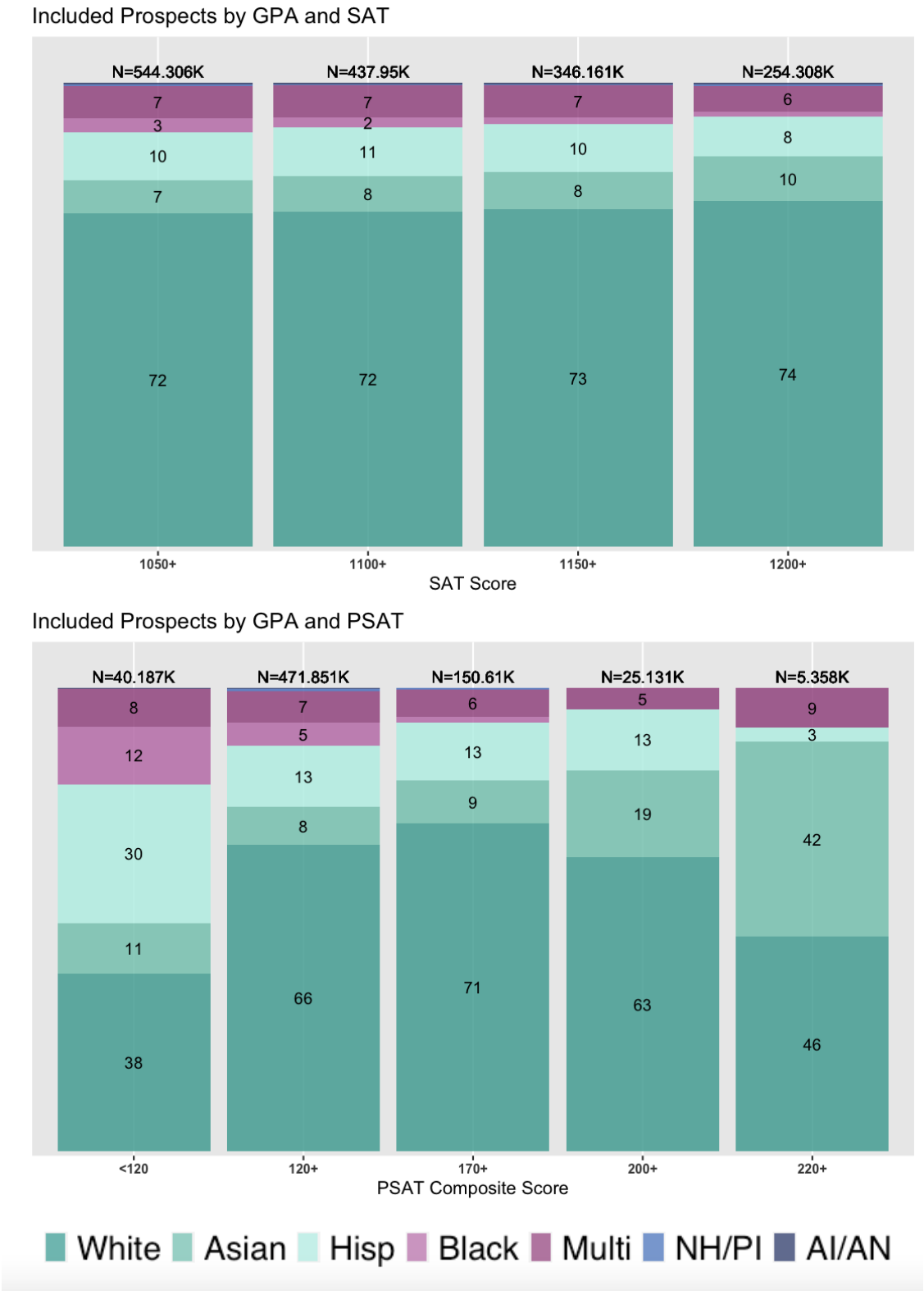
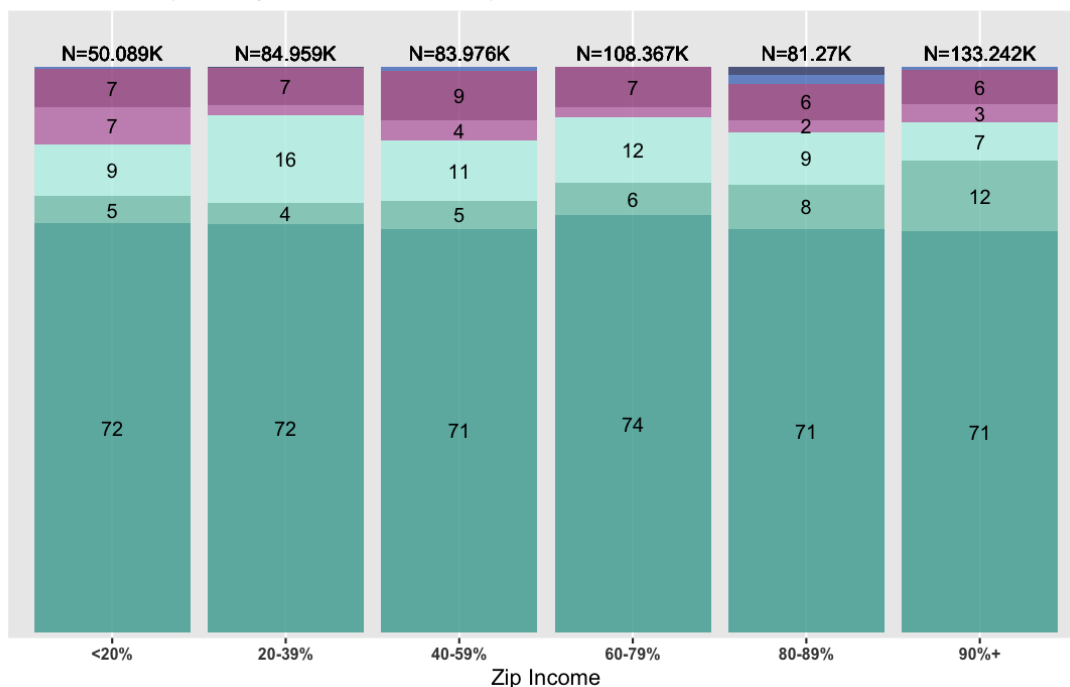
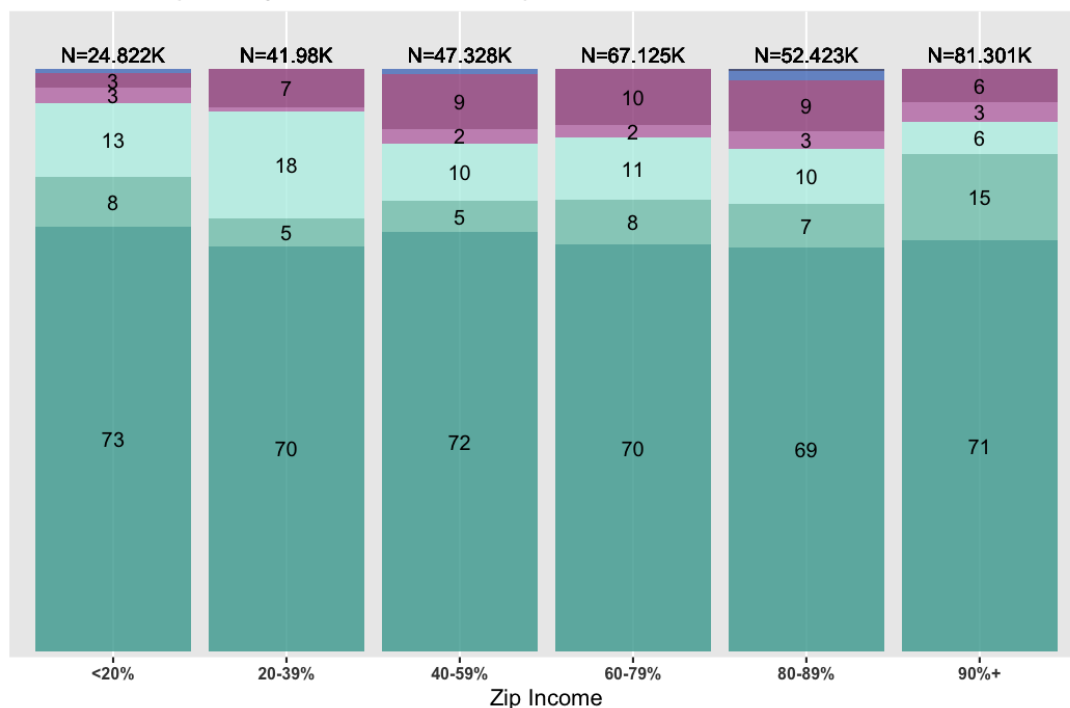


Figure 12: Academic and Geographic Combination: GPA (3.0+), PSAT (150+) or SAT (1050+), and Zip (across income thresholds)

Included Prospects by GPA, SAT, and Zip Code



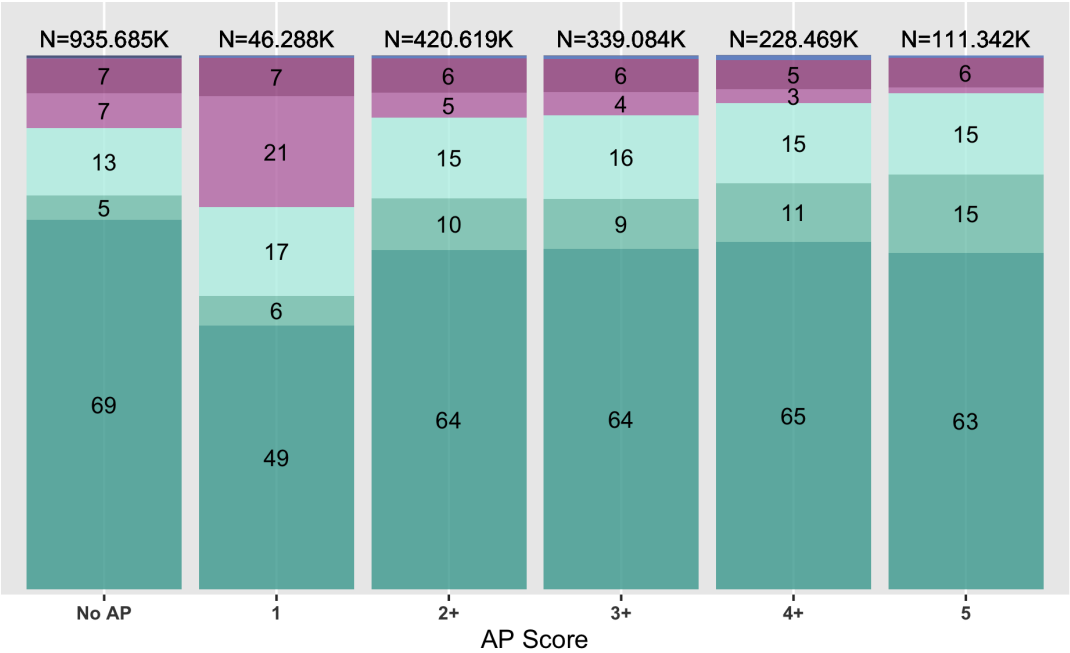
Included Prospects by GPA, PSAT, and Zip Code



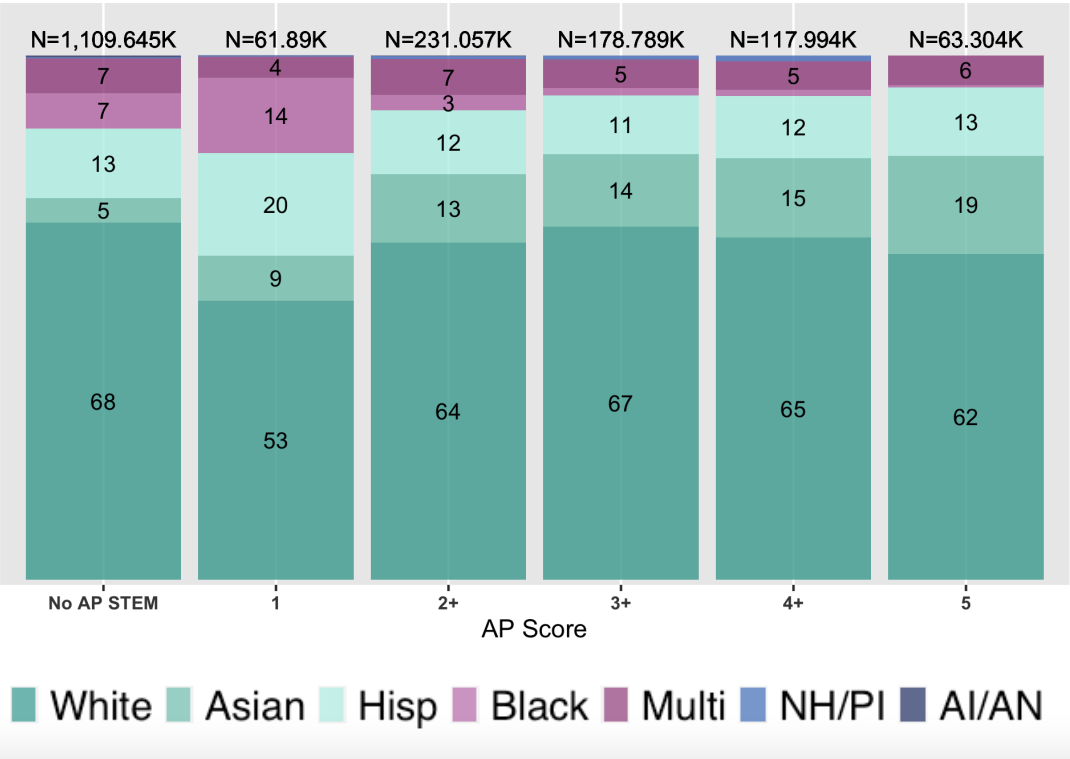
White Asian Hisp Black Multi NH/PI AI/AN

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

Included Prospects by GPA and AP Score



Included Prospects by GPA and AP STEM Score



10 Online Appendix

Table 4: 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 5: 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 6: 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 7: 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 8: 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 9: 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