

A Sociological Analysis of Structural Racism in ‘Student List’ Lead Generation Products

Ozan Jaquette Karina G. Salazar

ABSTRACT

Colleges identify prospective students by purchasing “student lists.” Student list products are selection devices that use search filters to select students. Drawing from the sociology of race, we conceptualize some filters as “racialized inputs,” defined as inputs that are correlated with race because disadvantaged racial groups have historically been excluded from the input. Using a national sample of high school students, we explore the relationship between racialized search filters and the racial composition of included versus excluded students. Using data about actual lists purchased by public universities, we investigate how college administrators utilize racialized search filters. We discuss implications for federal and state policy. We motivate policy research about structural racism embedded in selection devices that allocate students to opportunities.

BIOS

Ozan Jaquette, Ph.D., is an associate professor in the School of Education and Information Studies at the University of California, Los Angeles. His research program analyzes enrollment management and organizational behavior.

Karina G. Salazar, Ph.D., is an assistant professor in the Center for The Study of Higher Education at the University of Arizona. Her research analyzes how the enrollment management practices of public universities shape college access for undererved student populations.

1 Introduction

Racial inequality in college access remains an enduring barrier to social mobility. The share of White and Asian students who attend college remains greater than the share of Black and Hispanic students who attend college (Baker, Klasik, & Reardon, 2018). Based on the High School Longitudinal Survey of 2009, about 83% of Asian, 72% of White, 63% of Hispanic, and 62% of Black students enrolled in college within a year and a half of graduating high school in 2013 (Reber & Smith, 2023). For students who attend college, college selectivity is a strong predictor of outcomes such as degree completion (Shamsuddin, 2016), initial job market participation (Long, 2008), earnings growth over time (MacLeod, Riehl, Saavedra, & Urquiola, 2017), and the ability to repay student loans (Jackson & Reynolds, 2013). However, Black and Hispanic students are less likely to attend selective colleges than White and Asian students (Alon & Tienda, 2007; Posselt, Jaquette, Bielby, & Bastedo, 2012).

In economics, Hoxby (2009, p. 106) conceptualizes the market for college access “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) argues that information costs were historically the primary barrier to efficient matches. Students want to attend the best possible college but they don’t know where they will be admitted or 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. Hoxby (2009, p. 103) notes that “In 1955, 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 the emergence of an efficient, national higher education market. 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. Institutionalization

of the SAT/ACT caused a “dramatic fall” (Hoxby, 2009, p. 102) in the cost of “colleges’ information about students.” Test-takers could send scores to colleges they were interested in, allowing colleges to compare prospective students from disparate places.

However, most colleges cannot survive financially or thrive reputationally solely from prospective students who reach out on their own (EAB, 2018; Hossler & Bontrager, 2014; Ruffalo Noel Levitz, 2018). Rather, colleges must find “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 colleges 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 prospects who reach out on their own – constitute the set of prospects who receive subsequent recruiting interventions (e.g., mail, email) designed to push them toward the application and enrollment stages of the “enrollment funnel.” Ruffalo Noel-Levitz (2022b, p. 5) reports that 86% of public colleges “purchase high school student names to generate inquiries and applicants.” Of these, 80% purchase more than 50,000 names annually.

We argue that student list products exacerbate racial inequality in college access because their design incorporates inequality present in society. We posit two broad mechanisms of racial inequality. 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) that may disadvantage communities of color.

We develop a conceptual framework that draws from the sociology of race. “Selection devices” are routines for making decisions about who receives an intervention or opportunity based on some set of input factors (Hirschman & Bosk, 2020). Student list products are “discretionary” selection devices in that they grant administrators discretion over the set of search filters utilized to select prospects. Norris (2021) defines “racialized inputs” as those

that disadvantage underrepresented students of color because they have been historically excluded from this input. We argue that several student list search filters are racialized inputs (herein racialized search filters), particularly filters for test scores, geography, and geodemography. In turn, administrators utilizing student list products may disproportionately exclude communities of color, either because they purchase lists with reference to admissions criteria that exclude students of color or because of incomplete knowledge about how search filters interact with local patterns of exclusion.

This manuscript analyzes the College Board Student Search Service product. We address three research questions. First, what is the relationship between individual racialized search filters and the racial composition of included vs. excluded students? Second, in what ways do public universities utilize racialized search filters in concert with other search filters when purchasing student lists? Third, what is the racial composition of student list purchases that utilize racialized search filters in concert with other search filters? We address RQ1 by using the High School Longitudinal Survey of 2009 (HSLS) to reconstruct the College Board Student Search Service product. Analyses simulate the racial composition of included versus excluded prospects when individual search filters are utilized. We address RQ2 by analyzing 830 student lists purchased by 14 public universities. We address RQ3 by showing the racial composition of selected student list purchases – simulated and actual – that filter on multiple search filters simultaneously.

Simulations from HSLS:09 show that search filters conceptualized as racialized inputs – taking a college entrance exam, test-score thresholds, or zip codes – are associated with racial inequality in targeted prospects (RQ1). Actual student lists purchased by 14 public universities often utilized one or more racialized search filters (RQ2). Simulations show that using multiple racialized search filters can compound racial disparities between included versus excluded prospects (RQ3). Finally, targeted analyses of actual student list purchases reveal extreme structural inequality, even in purchases that explicitly target underrepresented groups (RQ3).

These results have policy implications. All student list products enable colleges to select prospects based on a set of search filters. Although some colleges may use student list prod-

ucts to increase access for underrepresented groups, the potential for unintended inequities seems high. The question becomes, should policymakers tolerate a product that is likely to do harm because it is capable of doing good? We observe striking similarities between federally regulated “consumer report” products (e.g., Equifax) and student list products. Like consumer reports, student lists systematically lead to the extension of credit vis-a-vis student loans. Federal consumer protection agencies should regulate student list products as they do other consumer reports. We argue that state longitudinal data systems can create a more equitable alternative to private, third-party student lists.

This article has broader implications for scholarship. The sociology of race conceives of critical policy research as being focused on underlying structures that produce inequality. Racial equality cannot be achieved by placing progressive practices atop structurally racist systems. Rather, “the only way to ‘cure’ society of racism is by eliminating its systemic roots” (Bonilla-Silva, 1997, p. 476) by focusing on seemingly neutral structures that are correlated with race because of historical racial exclusion. Student list products are selection devices that allow colleges to target students based on factors that are reflections of structural inequality in educational opportunity.

Our conceptual framework points to selection devices as important objects of critical policy research. Consistent with the research tradition on tracking, selection devices allocate people to different groups. Many contemporary third-party products can be usefully conceptualized as selection devices. Student list products are selection devices that grant administrators discretion over which filters they use to select prospects. Others (e.g., products that identify “at risk” students) make decisions based on a standardized algorithm. Our conceptual framework casts attention to the inputs used by selection devices. Prior research shows that inputs correlated with race yield racial inequality and that administrator discretion usually amplifies racial inequality (Castilla, 2008; Hirschman & Bosk, 2020; Norris, 2021). Therefore, critical policy research should deconstruct selection devices, interrogate seemingly neutral inputs, and examine how these products are used by people on the ground.

The practical importance of critical policy research focused on selection devices is two-fold. First, efforts to defend racially progressive practices may be defeated by conservative legal

challenges. By contrast, critical policy research plays offense by showing that ‘business-as-usual’ practices benefit dominant groups. Thus, critical analyses may sway even conservative courts by showing how products or practices treat people unequally on the basis of race. Second, amidst the data science revolution, third-party products increasingly perform core functions on behalf of schools and colleges (Nichols & Garcia, 2022). If researchers ignore these products, policy research will have a diminishing influence on education. Federal regulatory agencies have become concerned about products that “categorize consumers in ways that can result in exclusion of certain populations” (Federal Trade Commission, 2016a, p. 9). Critical analyses of selection devices can inform regulations that reduce disparities.

The article is organized as follows. We provide background information and review scholarship on recruiting. We develop a conceptual framework based on the sociology of race. We describe methods and present results. We discuss implications for policy and research.

2 Background and Literature Review

This section situates student lists vis-a-vis enrollment management and the recruiting process. Second, we review scholarship on recruiting. Third, we describe student list products.

2.1 Enrollment Management

Enrollment management is the organizational behavior side of college access. 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). Colleges 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 the specific needs of campus constituencies (e.g., College of Engineering needs majors, teams need players) (Stevens, 2007). As an administrative structure, the office of enrollment management typically controls the activities of admissions, financial aid, and recruiting (Kraatz, Ventresca, & Deng, 2010). The enrollment management industry consists of university personnel (e.g., admissions counselors, VP for enrollment management), professional associations (e.g., National Association for College

Admission Counseling), and third-party servicers (e.g., College Board, EAB, PowerSchool). Colleges cannot realize their enrollment goals solely from prospects who find the college on their own. They must incite demand and discover desirable prospects who can be convinced to enroll. The “enrollment funnel” – depicted in Figure S1 – is a conceptual model used by the enrollment management industry to depict broad stages in the process of recruiting students (American Association of Collegiate Registrars and Admissions Officers, 2018; EAB, 2019). 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 the 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). The funnel narrows at each successive stage in order to convey the assumption of “melt” (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 stages. 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 advertisements designed to solicit inquiries and applications (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

Most scholarship at the nexus of enrollment management and college access focuses on the latter stages of the enrollment funnel, particularly the process of deciding which applicants to admit (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 leads, soliciting inquiries, and soliciting applications.

Scholarship on recruiting from economics tends to estimate the effect of specific interventions on the enrollment funnel. For example, one study estimates the effect of a college’s advertising budget on its enrollment rate (Ruffalo Noel-Levitz, 2022b). Another study estimates the effect of a college’s financial aid package on its enrollment rate (Hurwitz, 2012). A third study estimates the effect of a college’s admission policies on its enrollment rate (Bastedo, Howard, & Flaster, 2016).

tions on college access outcomes. Hoxby & Avery (2013) evaluate a nationwide 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 (Castleman & Goodman, 2018; Cunha, Miller, & Weisburst, 2018; Gurantz et al., 2021), but results have been mixed. Another set of studies evaluate interventions by colleges 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 (e.g., Mulhern, 2021). A College Board report by Howell, Hurwitz, Mabel, & Smith (2021) compared SAT test-takers who opted into the College Board Student Search Service – allowing colleges to purchase their contact information – to those who opted out. After controlling for covariates (e.g., SAT score, parental education), 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 $25.3 \left((41.1 - 32.8) / 32.8 \right)$ 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 students (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 bachelor's degree (18.9%). Leveraging a natural experiment in College Board student list purchases, Smith, Howell, & Hurwitz (2022) find that purchasing a prospect profile increases the probability that the student will apply to and enroll at the purchasing college, with larger effects for Black, Hispanic, and low-income students.

Scholarship from sociology tends to document recruiting behavior “in the wild,” often as part of broader analyses of college access or enrollment management (Cottom, 2017; Posecznick, 2017). Holland (2019) probes the “structural holes” between high school counseling and college recruiting efforts from the perspective of 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). 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.

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 have been conceptualized as an indicator of enrollment priorities (Salazar et al., 2021; Stevens, 2007) and/or a network tie indicating the existence of a substantive relationship (Jaquette, Han, & Castaneda, 2023). 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) finds that for-profit colleges found a niche in Black and Hispanic communities 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 the College Board and ACT because their target audience was not recent test-takers. Ironically, Black and Hispanic adult women were vulnerable to marketing from for-profits because they were excluded from the College Board and ACT lists used by traditional colleges.

Reflecting on the recruiting literature, studies from both economics and sociology find that underrepresented student populations are particularly sensitive to recruiting. Economics often evaluates the effect of outreach designed to increase underrepresented student enrollment at selective colleges. By contrast, scholarship from sociology (Salazar et al., 2021; Stevens, 2007) and investigative reporting (MacMilan & Anderson, 2019; Tough, 2019) from the popular press consistently find 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 (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 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. An article on digital advertising by EAB (2018, p. 9) provides insight into 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.

EAB (2018) suggests that, when recruiting college-going high school students, behavioral-based leads are less effective than purchasing names from the 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).

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 the College Board and ACT, which derive student list data from test-takers. Advances in technology yielded new sources of student list data, particularly online college search engines (e.g., Cappex) and college planning software sold to high schools (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 colleges 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 student lists. For public institutions, 20% purchased fewer than 50,000 names annually, 29% purchased 50,000-100,000, 31% purchased 100,000-150,000, and 20% purchased more than 150,000 names annually. Ruffalo Noel-Levitz (2022a) reports that purchasing names was the top expenditure item in the undergraduate 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. Case studies and news articles suggest that larger and more selective institutions purchase more names than smaller and less selective ones (Arcidiacono, Kinsler, & Ransom, 2022; Belkin, 2019; Jaquette et al., 2022).

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 four buckets: academic; geographic; demographic; and student preferences. Academic filters include SAT score, PSAT score, AP score by subject, high school GPA, and 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 on 830 student lists purchased by 14 public universities, Salazar et al. (2022) found that the average list purchase specified 4.44 criteria and 98.8% of purchases specified at least one academic and one geographic filter.

A purchased list 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 limited but can be inferred from search filters.¹

How lists are used. Much like the role of voter files in political campaigns (Culliford, 2020), purchased lists are a 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. For instance, colleges cannot send prospective students brochures and emails without addresses. Purchased lists are combined with student-as-first-contact inquiries and layered with additional data sources, such as consumer data from credit bureaus, historical application/enrollment data, etc. These layered data are the input to predictive models that inform decisions about recruiting interventions designed to push prospects to subsequent stages of the funnel, such as deciding who gets a \$0.50 postcard or 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 student

¹See Salazar et al. (2022) for College Board and ACT student list templates

names. On average, public colleges contact purchased names eight times before giving up. With respect to efficacy, Ruffalo Noel Levitz (2018) asked clients to rate different “first contact” interventions 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 Conceptual Framework

Drawing from the sociology of race, we conceptualize racial inequality due to the architecture of student list products and how these products are utilized by college administrators.

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 on which selection devices differ is individual discretion versus standardization. Discretionary selection processes rely on the judgment of individual administrators, who have discretion about which inputs to consider and/or how to evaluate these inputs. For example, in professional domains such as psychiatric treatment, evaluators exercise judgment based on professional norms about how to evaluate inputs. Another example is targeted advertising of social media users. Facebook classifies users by market segment, demographics, and other variables while granting ad-buyers discretion over which variables they use to target users (Cotter, Medeiros, Pak, & Thorson, 2021). By contrast, in “standardized selection devices,” the relationship between the value of input variables and the outcome is dictated by a mathematical formula that is beyond the discretion of individual administrators using the product (Duncan, Ohlin, Reiss, & Stanton, 1953). In child welfare, for example, the “Structured Decision Making Model” yields recommendations about whether children should be placed in protective care based on an algorithm designed to predict the likelihood of future abuse or neglect (Hirschman & Bosk, 2020).

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, an individual administrator can choose which inputs to filter on and which thresholds to apply to each filter. By contrast, a “direct admissions” policy that admitted applicants based on a function of ACT score and GPA would be considered a standardized selection device; even though the test score and GPA thresholds are chosen by the college or a state policy organization, individual admissions administrators have no discretion over these thresholds.

A long-standing debate in sociology, psychology, and related professions is the extent to which discretionary versus standardized selection devices produce or reduce racial inequality. Following 1970s anti-discrimination legislation, many industries adopted standardized selection devices because “evidence had accumulated that... decision-makers were routinely giving into vague intuitions, personal prejudices, and arbitrary opinions” (Burrell & Fourcade, 2021, p. 22). Reviewing the literature, Hirschman & Bosk (2020) state that standardized selection devices can reduce racial inequality *if* the primary source of inequality is explicit or implicit racial bias from individual decision-makers. However, Bonilla-Silva (1997) criticizes social science disciplines for defining racism as an ideology held by individuals.

Neither standardized nor discretionary selection devices eliminate racial inequality stemming from structural racism. 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 use race as an input, scholarship from sociology argues that selection devices 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, p. 5) defines racialized inputs as “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.” Norris (2021) reconstructs Moody’s city government

credit rating algorithm, which assigns scores to cities based on determinants of loan default. Norris (2021) argues that median family income is a racialized input in that income is correlated with race because of historical wage discrimination. In turn, cities with a greater share of Black residents have lower median income. Once median income is included in Moody's model, the percent of Black residents no longer predicts city credit rating. Thus, through the inclusion of seemingly neutral racialized inputs, "prior disadvantage and racism against Black individuals becomes institutionalized" (Norris, 2021, p. 2), and selection devices yield racially disparate outcomes "in ways that escape legibility/cognition as racially unequal" (p. 5).

Geographic inputs. Geographic borders are the most commonly studied racialized inputs (e.g., Benjamin, 2019; Korver-Glenn, 2022; O'Neil, 2016). 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 (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 using zip code as an input to predict the probability of recidivism for previously incarcerated people. Because zip codes are correlated with race, using zip codes 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 similar audience segments for marketers (Experian, 2023). 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 Mosiac – 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:

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 analytic inputs. Another class of racialized inputs comes from predictive analytics, which have been a focus of scholarship on algorithmic bias (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 which factors predicted default for past cases.

Discretionary selection devices often utilize predictive analytics as another input to filter 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 the “Geomarket” and “Geodemographic Segment” filters. These filters draw geographic borders for including or excluding future prospects based on analyses

of the college-going behavior of past prospects (College Board, 2011).

Whether predictive analytics are used as outcome or input, “predicting the future on the basis of the past threatens to reify and reproduce existing inequalities” (Burrell & Fourcade, 2021, p. 224). Harcourt (2007) refers to this phenomenon as the “ratchet effect,” whereby disproportionately targeted/excluded populations are predicted to have a 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 colleges utilize student list products. We argue that the underlying architecture of student list products produces 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).² Prior research shows that rates of SAT, ACT, and AP test-taking differ by race (Blake & Langenkamp, 2022; 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

²Recently, College Board and ACT, respectively, began allowing non test-takers to opt into student list products by participating in the College Board Big Future (see <https://bigfuture.collegeboard.org/>) or the ACT Encourage (see <https://myoptions.org/>) college search engines.

inputs as search filters. We argue that test score filters (e.g., SAT, PSAT, AP) meet the racialized input criteria of being “theoretically and empirically correlated with historical racial disadvantage” (Norris, 2021, p. 5). 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 considering the historical and contemporary structural inequalities that drive race-based differences in test scores is likely to reproduce racial inequality in educational opportunity.³

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 colleges to target prospects based on where they live. College Board geographic search filters include state, CBSA, county, zip code, geomarket, and geodemographic filters. We conceptualize geographic search filters as racialized inputs because these filters are built on top of historic and contemporary policies and practices promoting racial segregation. Targeting prospective students based on geographic location without considering macro and local structures that produce racial segregation is likely to reinforce historical race-based inequality in educational opportunity.

Scholarship on recruiting finds that for-profit colleges systematically target poor, communities of color (Cottom, 2017; Dache-Gerbino et al., 2018) while selective private and public research universities disproportionately target affluent schools and communities (e.g., Jaquette et al., 2023; Salazar, 2022; Salazar et al., 2021; Stevens, 2007). Newspaper articles report that enrollment managers are intentional about targeting localities based on income. MacMilan & Anderson (2019) reports that “consulting companies may estimate a student’s financial position by checking their zip codes against U.S. Census data for estimated house-

³By 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 et al., 2012).

hold incomes in that area.” Rivard (2013) writes, “College Board does sell zip codes, which are a very good proxy for income levels, meaning colleges and their consultants could use the data to sort out rich and low-income kids.” Rivard (2013) quotes an enrollment management consultant at Scannell & Kurz, who said “‘Everybody wants to go to the magic island of full pay students, but it’s rapidly shrinking real estate.’”

These studies and news articles 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 zip codes in low-income, minority, urban areas. 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 colleges filter on larger localities (e.g., county, state, CBSA) as a 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.

3.3 Utilizing Student List Products

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

Discretion. We highlight three findings from sociological scholarship on product utilization. First, scholars tend to find that administrative discretion over selection devices causes structural inequality to increase (Castilla, 2008; Cotter et al., 2021; Norris, 2022). Discretionary selection devices are sensitive to racialized inputs and allow explicit or implicit individual bias to affect selection decisions (Burrell & Fourcade, 2021; Korver-Glenn, 2018). Korver-Glenn (2018) shows that Houston area homes in white neighborhoods received higher

appraisal values than those in non-white neighborhoods because of appraiser discretion in selecting comparison homes, which is exacerbated by the racialized borders of housing market areas drawn by the real estate board. Second, 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). Third, research shows that Americans dramatically underestimate the magnitude of racial income inequality (Kraus, Onyeador, Daumeyer, Rucker, & Richeson, 2019). Discretionary selection devices that incorporate racialized inputs may produce racial inequality because decision-makers have incomplete knowledge about how these inputs interact with local patterns of racial inequality (Cotter et al., 2021; Korver-Glenn, 2018).

These findings motivate analyses about administrative discretion and racial inequality in student list purchases. Colleges may select academic achievement filters based on admissions standards, which are a function of college stakeholders and macro trends in the admissions profession (Rosinger, Sarita Ford, & Choi, 2021). Until recently, most admissions offices viewed test scores as objective measures of achievement (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.

The utilization literature suggests that admissions standards are not the sole driver of inequality in student list purchases. Scholarship finds that two safeguards against structural inequality in discretionary selection devices are *transparency* – selection criteria are clear to all stakeholders – and *accountability* – consequences for using biased selection criteria (Castilla, 2008; Norris, 2022). The process of purchasing student lists is opaque to most internal and external stakeholders. There can be no accountability without transparency (Norris, 2022). Whereas admissions readers are trained and normed before they evaluate applications (Bastedo, 2016), student list purchases can be executed by any person affiliated with a Title IV institution. Furthermore, we could not find written professional norms about how to purchase lists. Without guardrails against discretion, the utilization of a complicated product is likely to yield unintended consequences, particularly when purchasers select several filters that interact with racial inequality present in society.

Micro-targeting. Filtering on multiple search filters facilitates micro-targeting of desired prospects, which has become a branding strategy for student list products. College Board Student Search promises to “create a real pipeline of best-fit prospects” (College Board, n.d.) while ACT Encoura uses the tagline “find and engage your best-fit students” (Encoura, n.d.). Consultancies encourage colleges to execute multiple student list purchases, each targeting different market segments (e.g., Waxman, 2019).

The flip side of micro-targeting is exclusion (Cotter et al., 2021). Purchased lists do not show how the characteristics of targeted prospects compare to the characteristics 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 less-racialized filter (high school GPA) but adding a racialized input filter (e.g., AP score) may still increase racial inequality. Additionally, filtering on multiple structurally racist inputs (e.g., SAT score and zip code) may compound inequality.

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 and 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 incorporate 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 social inequalities. For example, purchases designed to target “women in STEM” may yield racial or socioeconomic inequality. Purchases that explicitly target underrepresented minority students with high test scores may also yield socioeconomic or geographical inequality by systematically excluding low-income students or excluding students living in predominantly non-white communities.

4 Methods

4.1 Data

Our analyses utilize two data sources. First, the primary data source is the High School Longitudinal Study of 2009 (HSLS09), which we use to run analyses for RQ1 and RQ3. HSLS09 is a nationally representative survey that follows a cohort of more than 23,000 students from more than 940 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.

Of the more than 23,000 respondents included in HSLS09, our unweighted analysis sample consists of the 16,530 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. The survey weight variable W3W2STUTR is designed for respondents who meet these conditions. After weighting, these 16,530 students represent a population of approximately 4.187 million U.S. 9th graders in 2009.

The second data source consists of “order summaries” and their resulting “student lists” for College Board purchases by 14 public universities.⁴ Table S2 in the supplemental online appendix shows selected characteristics of the universities in our data collection sample.⁵ Order summaries were collected as part of a larger project by issuing public records requests to public universities in five states (CA, IL, TX, MN, and AZ) about student lists purchased

⁴See Salazar et al. (2022) for an example of a College Board student list order summary.

⁵Table S2 uses Carnegie 2020 Basic classification to collapse institutions into two institution types. All doctoral universities with highest or high research activity were categorized as “Research Universities. Master’s universities across all programs were categorized as “Master’s Universities”.

from 2016 to 2020. Order summaries from 830 student lists purchased by 14 universities were used to analyze which search filters do public universities use (RQ2).

RQ3 analyzes the racial composition of lists that filter on multiple search filters. We pull selectively from 414 orders – associated with 2,549,085 prospects – in our public records request data where we have both the order summary (i.e., which combinations of filters were used) and the prospect-level data (i.e., the resulting student list).⁶ For more detailed information about data collection, see Salazar et al. (2022).

4.2 Variables

Dependent variable. Our primary dependent variable is prospects' race/ethnicity. For HSLS09, we use 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. For public records request data, College Board data have separate measures for race and ethnicity, which allow students to select more than one option as per changes to census reporting requirements. Similar to HSLS09, we aggregate this into a single race/ethnicity measure with the same seven categories listed above.

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 below in Figure 3). Our conceptual framework restricts the 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 HSLS09 high school transcript file:

⁶See Salazar et al. (2022) for an example of prospect-level data associated with orders linked above.

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 unweighted high school GPA in academic courses (X3TGPAACAD). Consistent with how College Board filters work, **P2** variables are analyzed as categorical rather than continuous variables. To select thresholds for **P2** variables – for example, 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 parsimony.

Proposition **P3** focuses on geographic filters. We create measures for student's school zip code by merging in data from NCES. Next, we attach income data to zip code by merging in data from the American Community Survey (ACS) 2012 5-year estimates.

4.3 Analyses

Analyses utilize simple descriptive statistics. For RQ1, analyses compare the racial composition of included versus excluded HSLS09 prospects when an individual search filter is utilized in isolation. We run (weighted) statistical tests for comparing differences in included versus excluded students by race/ethnicity. Consider a hypothetical purchase of all prospects that 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% ($=90/1800$) of AP STEM test-takers and Black students comprise 11% ($=1560/14720$) 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.

RQ2 asks which search filters public universities utilize. We use counts and proportions of filters used across research (N=9) versus master's (N=5) universities to describe patterns in how racialized search filters are used alongside other filters when purchasing lists.

RQ3 examines the racial composition of student list purchases that utilize multiple search

filters simultaneously. Choices about filters are informed by our theoretical framework. Analyses based on HSLS09 examine included versus excluded prospects by race/ethnicity when multiple filters are utilized. We also analyze RQ3 using data from actual student lists purchased by public universities, selecting from purchases where we are able to obtain the order summary *and* the associated prospect-level data. In contrast to analyses based on HSLS09, analyses based on public records requests are unable to make inferences about the population of student list purchases. We are also unable to make comparative inferences about excluded groups for analyses based on public records requests. We leverage secondary data from NCES and ACS to show how the characteristics of purchased lists compare to relevant groups (e.g., all high school graduates in the metropolitan area).

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 via college entrance (i.e., SAT and ACT) and pre-college entrance (i.e., PSAT and preACT) exams.

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 COVID-19. This undermines the external validity of our findings with respect to current cohorts of high school students. 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 Racial Composition of Individual Filters

We address RQ1 by first describing the racial characteristics of HSLS09 students who completed standardized assessments in comparison to those who did not, which would determine inclusion in the underlying College Board student list database. Figure S2 in the online appendix presents the racial/ethnic composition of prospects included (i.e., completed assessment) and excluded (i.e., did not complete assessment) across standardized tests. More than 1.8 million prospects completed a college entrance exam and would have presumably been included in the College Board student list database. On average, these included prospects were 57% White, 4% Asian, 17% Hispanic, 12% Black, and 8% Multiracial. In comparison, more than 2.3 million prospects did not complete a college entrance exam and would be excluded from the database. Excluded prospects included a lower proportion of White students (47%) and greater proportions of Black (15%) and Hispanic (26%) students. Other standardized assessments such as AP exams and college pre-entrance exams also resulted in similar included prospects that on average were 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. Table S3 in the online appendix shows all differences in proportions between included and excluded students by race/ethnicity were statistically significant, lending support for Proposition P1.

Proposition P2 states that the proportion of underrepresented minority students included in student lists declines 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 HSLS09 students at minimum score thresholds commonly used across student list purchase orders for college entrance, pre-entrance, and AP exams. For

example, Figure 1 presents these results for college entrance (top panel) and pre-entrance assessments (bottom panel). For the top left panel, each bar represents the racial composition of included prospects who completed an entrance exam and scored at the minimum threshold indicated. On the top right panel of Figure 1, each bar represents the racial composition of excluded prospects who did not complete an entrance 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 1 are reported in online appendices (Table S4).

As entrance exam score thresholds increase from less than 1000 to greater than 1400 in Figure 1, proportions of included White and Asian students increase while the proportions of included Hispanic and Black students decrease. For example, White students make up a statistically significant ($p<0.001$) smaller share of included (47%) than excluded (53%) prospects scoring less than 1000 on an entrance exam, 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).

Pre-entrance exam results are also shown in Figure 1 for composite scores that range from 60 to 240 on a PSAT scale.⁷ Similar to entrance exams, as pre-entrance exam 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. Similar racial disparities in included versus excluded prospects are evident across AP score thresholds (see Figure S3 online), providing support for P2.

⁷PSAT exams taken in 2014 or before receive composite scores that range from 60 to 240. PSAT exams taken in 2015 or later range from 320 to 1520 (College Board, 2015). Our lower bound PSAT composite thresholds of 120, 170, 200, and 220 for HSLS students who completed the exam prior to 2014 equate to minimum thresholds of 890, 1220, 1410, and 1510 on the 2015 or later PSAT scale, respectively.

Proposition P3 states that as purchases filter on more affluent zip codes, the proportion of underrepresented minority students included in student lists will decline relative to the proportion who are excluded. 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 states. This approach aligns with common ways in which student list orders filter on zip codes within specific states or metropolitan areas.

Figure 2 presents the racial composition of HSLS09 students included versus excluded in student list purchases when filtering based on their zip code percentile of affluence. 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 31% 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 the 90th percentile or higher, the proportion of Hispanic students declines to 13% ($p<0.001$). Similarly, Black students make up 8% of included prospects relative to making up 15% of excluded prospects ($p<0.001$) within the most affluent zip codes. 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.

5.2 Utilizing Student List Products

To answer RQ2, in what ways do public universities use racialized input search filters, we analyze how often filters were used for College Board student list purchases by the 14 universities in our public records request data collection.

Figure 3 illustrates the prevalence of each individual filter by institution type. We categorize filters within academic, geographic, demographic, and student preference groups. Both research and master's universities commonly used academic filters like GPA, SAT, and PSAT.

However, compared to master’s universities, research universities were less likely to filter on GPA or PSAT scores and more likely to filter on SAT scores. Additionally, AP filters were only used by research universities. Geographic filters used to purchase student lists also differed across institution types. Orders by research universities were more likely to use a state filter whereas master’s universities were more likely to use a zip code filter. Research universities also used filters utilizing predictive analytics to make inferences about the college-going behavior of prospective students living in specific geographic areas (e.g., geomarket, segment). Research universities also used demographic and student preference filters. Filters for race, low socioeconomic status, and gender were the most commonly used demographic filters, whereas college size, major, college type, and location were the most commonly used student preferences.

Given the prevalence of academic filters across universities, Figure 4 shows minimum and maximum thresholds used in SAT score filters by institution type, respectively. The figure shows that research universities tended to specify higher minimum SAT score thresholds and higher maximum SAT score thresholds compared to master’s universities. Similar patterns are evident for PSAT score thresholds (see Figure S4 online).

Student list purchases typically filter on multiple criteria. Table 1 shows the top ten filter combinations used for student list purchases across institution type for the 14 universities in our public records request data collection.⁸ For master’s universities, the top 10 filter combinations account for 97% of all orders. This is a function of nearly half of all orders using a combination of high school graduation class, GPA, PSAT scores, and zip code as filters, and another 35% of orders using the same filters with SAT scores instead of PSAT. For research universities, the top ten filter combinations account for 59% of all orders. The most common filter combination, making up 19% of all orders, included high school graduation class, GPA, SAT, and zip code. The second most common combination included the same filters along with PSAT, high school rank, state, and race. The remaining top orders also used these filter combinations in addition to filters like low-socioeconomic status, AP scores,

⁸While the majority of orders specified multiple filters as “AND” conditions, some orders specified particular filters as “OR” conditions (e.g., SAT score in some range or PSAT score in some range). Additionally, filter combinations are skewed by universities that made large numbers of orders relative to other universities.

Geomarket, Segment, and gender.

5.3 Racial Composition of Multiple Filters

Analyses Based on HSLS09. Our last research question focuses on assessing the racial composition of student list purchases that utilize racialized input search filters in concert with other search filters. We begin by analyzing the racial composition of included versus excluded groups when filtering the HSLS09 sample based on common combinations of filters used by research and master’s universities in Table 1. For example, Figure 5 shows included prospects resulting from combining the two most common academic filters: GPA and college entrance and pre-entrance exams. Drawing on our conceptual framework, high school GPA is a strong predictor of postsecondary student success (Allensworth & Clark, 2020) and is less likely to be theoretically and empirically correlated to racial disadvantage than standardized test scores (Alon & Tienda, 2007; Posselt et al., 2012). On the other hand, colleges may filter on test score thresholds that are consistent with their admissions standards (e.g., more prestigious colleges recruit and enroll students with higher test scores) (Clineinst, 2019) but hold race-based differences driven by structural inequality. We therefore explore the racial composition of prospects included in student lists when filtering on GPA greater than or equal to 3.0 while simulating increases to minimum thresholds for standardized tests.

The left panel of Figure 5 presents the racial composition of prospects included when filtering on GPA greater than or equal to 3.0 across college entrance exam thresholds at increments of 50 beginning at scores just above the HSLS09 sample median of 1010. For space considerations, we do not present the plot for excluded prospect groups across all thresholds but include an “All” column showing the racial composition of the HSLS09 sample. The figure suggests that even at the lowest college entrance exam score threshold, White students make up much larger proportions of included prospects while Black and Hispanic students make up statistically significant smaller proportions when filtering for both GPA and entrance exam scores.⁹ For example, White students make up 72% of included prospects when filtering for GPAs greater than or equal to 3.0 in combination with entrance exam scores

⁹Table S6 within online appendices report statistical tests for proportions between included and excluded students by race/ethnicity for combining GPA with college entrance exams.

greater than 1050, whereas Hispanic and Black students make up 10% and 3%, respectively. Racial disparities only grow as score thresholds increase. Moreover, these racial disparities are greater than when filtering at similar thresholds for entrance exam scores (Figure 1) alone. The right panel of Figure 5 suggest Hispanic and Black students make up decreasing proportions of included students when combining a GPA filter greater than or equal to 3.0 and a pre-college entrance exam filter, whereas the proportions of White and Asian students increase. Although, we do not see the same magnitude of disparities for a college pre-entrance exam filter relative to an entrance-exam filter. Online appendices also show similar racial disparities are evident when filtering on both GPA and AP scores (Figure S5 and Table S7).

In order to assess the characteristics of students lists compiled from combining academic and geographic racialized filters together, Figure 6 adds a zip code affluence filter to the GPA and college entrance/pre-entrance exam order simulations presented above from HSLS09. 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 state. The left panel of Figure 6 presents the racial composition of included students when filtering for GPAs greater than or equal to 3.0, college entrance exam 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 2, the combination of zip code affluence with GPA and entrance exam filters leads to greater disparities even at lower levels of affluence. For example, Figure 6 shows White students make up 72% of included prospects when filtering for GPAs greater than or equal to 3.0 in combination with entrance exam 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 31% of Hispanic and 27% of Black included prospects resulting from only filtering by zip code affluence (Figure 2). Greater and statistically significant racial disparities result from the combination of filters across all levels of zip code affluence in comparison to only filtering by zip code affluence.¹⁰ Similar patterns are evident when combining zip code affluence and

¹⁰Table S8 within online appendices report statistical tests for proportions between included and excluded

GPA filters with a pre-entrance exam filter for composite scores greater than or equal to 150.

Analyses Based on Collected Student List Purchases. By drawing on student list purchases from our public records request data collection, we can explore the potential results of utilizing filters beyond admissions standards. For example, while this study is unable to use HSLS09 to recreate predictive analytics filters that subdivide geographic areas into distinct markets based on past college enrollments to micro-target prospective students, we can analyze the racial composition of student lists using filters like College Board’s Segment services via our project sample. College Board’s segment filter merges demographic, academic, and historical college-going data from geographical areas to create predictive profiles for current college-bound students (College Board, 2011b), which then categorizes every U.S. Census tract into one of 33 “neighborhood clusters” and every high school into one of 29 “high school clusters” based on these predictive profiles. Online appendix Table S9 and Table S10, recreated from College Board (2011b), show the characteristics of Segment neighborhood clusters and school clusters used as filters by one university in our project sample. Drawing on our conceptual framework, Table S9 and Table S10 suggest that predictive analytics used to create clusters are likely drawing on historical and geographical correlations between racial and income demographics. For example, neighborhood cluster EN78 is 26% nonwhite and has a median income of \$134,400 while neighborhood cluster EN:71 is 97% nonwhite and has a median income of \$42,661.

We, therefore, analyze the racial composition of student lists from eight orders by a public research university that utilized Segment filters in concert with academic filters. These eight orders targeted 2019-2023 high school graduating classes, and resulted in 131,562 prospects whose profiles were purchased. All eight Segment orders filtered on GPAs ranging from a low of B- to a high of A+. The orders specified minimum SAT and PSAT scores that ranged from 1220-1240 on an SAT scale, whereas maximum PSAT and SAT scores were filtered at 1450. These student list purchases were also geographically filtered by state, CBSAs, and Segment. All eight orders filtered on the same Segment high school and neighborhood clusters, which are highlighted in online appendices Table S9 and Table S10.

students by race/ethnicity for combining GPA, college entrance exams, and zip code affluence.

Figure 7 compares racial and income characteristics of prospects whose profiles were purchased via Segment to the those of the public high school student population in the top four metropolitan areas where the most prospect profiles were purchased: New York (27,932 prospects, rank #1), Los Angeles (12,307 prospects, rank #2), Philadelphia (9,126 prospects, rank #3), Washington, DC (5,728 prospects, rank #4). For each metropolitan area, we show two figures. The right column shows the racial composition of prospects from the metropolitan area whose profiles were purchased compared to the racial composition of all public high school students in the metropolitan area. The left column shows the income of prospects whose profiles were purchased — defined as average median household income of prospects' home zip codes — compared to overall median income in the metropolitan area.

For New York, Figure 7 shows White and Asian students comprised 58% and 27% of prospects whose profiles were purchased via Segment, respectively, compared to making up 30% and 9% of students in public high schools in the metropolitan area. By contrast, Black and Hispanic students comprised just 1% and 8% of prospects, respectively, compared to 26% and 34% of students in public high schools. Furthermore, prospects whose profiles were purchased lived in zip codes that were much more affluent – an average of \$153,000 – than the overall New York metropolitan area median income of \$91,000. Figure 7 shows similar racial and income patterns in the other three metropolitan areas.

Figure 8 takes a spatial look at the high schools presented in Figure 7 across the four metropolitan areas. The map shows public high schools, with blue markers indicating the location of a school where at least one prospect's profile was purchased and the size of blue markers indicating the number of prospects whose profiles were purchased. Red markers indicate the location of schools where no prospect profiles were purchased. Figure 8 corroborates the findings presented for HSLS09 above. Prospects whose profiles were purchased attend high schools that are largely concentrated in affluent and predominantly White zip codes bordering the central cores of each metropolitan area, whereas schools where zero prospect profiles were purchased are concentrated in the lowest-income communities with larger proportions of People of Color located in the center metropolitan areas. Similar patterns are

evident across maps for all four metropolitan areas.¹¹.

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

We analyze the resulting student lists from these “Women in STEM” orders from one public research university in order to analyze in-depth patterns in the racial and economic characteristics of prospects that result from the combination of achievement, geographic, and gender filters used to target women interested in STEM. Because nearly 85% of prospects from Women in STEM orders were out-of-state prospective students ($N=10,668$), we select four out-of-state metropolitan areas and compare prospective students to the characteristics of public high school women students in those metropolitan areas.¹²

Figure 9 compares the racial and income characteristics of prospective students whose profiles were purchased to the characteristics of all women public high school students in each metropolitan area. For each metropolitan area, we show two figures. The figure on the left column provides the average median household income for the overall metropolitan area, for prospects whose profiles were purchased using AP scores, and prospects whose profiles were purchased using SAT scores. The figure on the right column provides the racial/ethnic composition of all public high school women 12th-grade students in the metropolitan area,

¹¹Interactive versions of metropolitan maps for segment analyses can be accessed at https://mpatricia01.github.io/public_requests_eda/outputs/maps/map_segment.html

¹²The four out-of-state metropolitan areas were selected based on where the greatest number of prospects whose profiles were purchased and based on regional variation: New York (1,663 prospects, rank #1), Atlanta (548 prospects, rank #2), Chicago (485 prospects, rank #3), and Seattle (279 prospects, rank #11).

of prospects whose profiles were purchased using AP scores, and of prospects whose profiles were purchased using SAT scores.

For New York, Figure 9 shows women attending public high schools in metropolitan area are 29% White, 9% Asian, 26% Black, and 34% Hispanic. However, women in STEM prospects from New York who scored a 4 or 5 on an AP STEM exam whose profiles were purchased by the university are 38% White, 44% Asian, 3% Black, 11% Hispanic, and 5% multiracial. Only 2 women in STEM prospects in the New York metropolitan area whose profiles were purchased via AP scores identified as Native American. These racial disparities are most pronounced in orders using SAT scores. Of the 821 prospects from New York that scored a 1200 to 1600 on the SAT and indicated an interest in STEM majors, 54% were White, 43% were Asian, and 3% were multiracial. Only one of the 821 prospects identified as Hispanic and zero prospects identified as Black or Native American. Similar patterns in the racial composition of Women in STEM prospects are evident across Atlanta, Chicago, and Seattle metropolitan areas. Figure 9 also shows Women in STEM prospects are also on average substantially more affluent than the rest of the metropolitan areas where they live.

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. Participating in student list products is positively associated with college access and this relationship is stronger 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 disadvantage underrepresented students.

This manuscript investigates the College Board’s student list search filters, how those filters are utilized, and the racial composition of students who are included in student list purchases. We develop a framework from the sociology of race. Student list products are discretionary

selection devices, enabling college administrators to select prospects by choosing search filters and filter thresholds. 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 national 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) yields racial disparities in students 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 Hispanic students being included in student list purchases. Filtering on more affluent zip codes (P3) is also associated with higher proportions of White students and lower proportions of Black and Hispanic students.

Actual student list purchases select multiple search filters simultaneously. Analyses based on HSLS09 suggest filtering on multiple criteria across score thresholds can compound racial disparities between included versus excluded prospects. Analyses of actual student lists purchased by public universities support findings from the product utilization literature, which suggest that administrator discretion likely increases racial inequality. Relative to the population of public high school students across metropolitan areas, the “Segment” orders and “Women in STEM” orders yielded dramatic disparities in the number of Black and Hispanic prospects included in student list purchases. These disparities may be partially a function of admissions standards. However, student list products are powerful, complicated products that incorporate racialized inputs and have few guardrails on how they are utilized. Therefore, 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.

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. They overcome “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. They 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. These products enable firms to transition from approving/rejecting applicants to the 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. Student list products systematically lead to the extension of credit as follows: at the top of the enrollment funnel (Figure S1), colleges obtain the contact information of prospective students by purchasing student lists; purchased leads are targeted with recruiting interventions designed to maximize the probability of conversion in subsequent stages of the funnel; at the end of the funnel, admits are offered financial aid packages – including loan aid – designed to increase the probability of enrollment. Given the clear link between student lists and the extension of credit, student list products should be classified as consumer credit reports. Considering varying levels of quality amongst higher education institutions, this designation would allow federal consumer protection agencies to ensure that student list products do not systematically funnel students to low-quality institutions along racial lines.

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 (2023) “Dear Colleague” letter broadened the definition of third-party servicers, stating that “most activities and functions performed by outside entities on behalf of an institution are intrinsically intertwined with the institution’s administration of the Title IV programs and thus the entities performing such activities are appropriately subject to third-party servicer requirements.” This definition includes entities that “interact

with prospective students for the purposes of recruiting or securing enrollment.” Although the proposed regulations intend to target online program managers (OPMs), it remains unclear whether they cover other third-party actors in the enrollment management space.

We recommend that the definition of third-party servicers be revised to include student list vendors because student lists are root intermediaries linking students to Title IV institutions and the receipt of Title IV aid. Certain aspects of student list products may violate the anti-discrimination provision of the Higher Education Act (Higher Education Act, 1965). For example, geodemographic filters allow colleges to target schools and neighborhoods with high college-going rates and these characteristics may be tied to historic racial segregation. As test-optional admissions policies erode the number of test-takers, online college search engines and college planning software purchased by high schools have become important sources of student list data. PowerSchool (a K-12 software firm), EAB, and Ruffalo Noel Levitz (enrollment management consultancies) have become student list vendors. These firms sell products that simultaneously identify prospects and deliver recruiting interventions. Meanwhile, the College Board and ACT now sell enrollment management consulting. Because contemporary student list products and vendors directly interact with prospective students, they should be subject to federal third-party servicer regulations. Jaquette et al. (2022) describe these market dynamics in greater detail.

State policy. State policymakers are positioned to develop “public option” alternatives to private, third-party student list 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. Colleges would receive the contact information and academic achievement of students who opt-in for free, which reduces the motivation for fine-grained search filters that enable colleges to micro-target the “right” students.

Even with free names, targeted marketing materials still cost money. Colleges may have an incentive to obtain profiles for a large number of prospects while being selective about which

prospects receive which recruiting interventions. Considering the low marginal cost of email, we believe that colleges would reach out to prospects not currently being contacted under the system where names cost money. This would be an important change because Holland (2019) finds that first-generation and historically underrepresented students of color are particularly sensitive to recruiting overtures from colleges. Because the public option would not contain information about College Board/ACT standardized test scores, it may be less useful for colleges that consider these tests for admissions decisions. Jaquette & Salazar (2022) provide 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) found that adoption by Idaho increased first-time undergraduate enrollment at the campus level and the state level. Five states have adopted direct admissions or are considering adoption (Odle & Delaney, 2022). 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 rely on student lists purchased from third-party vendors. To make sure all eligible students are included, Brown & Burns (2023) recommend that states adopting direct admissions should develop student lists based on state longitudinal data systems.

Institutions. Student lists – and recruiting more generally – become more important for institutional efforts to increase racial diversity following the Supreme Court’s ruling on race and admissions. In August 2023, the Biden Administration issued guidance to institutions, including how to pursue racial diversity while complying with the law. This guidance states that “the Court’s decision in SFFA does not require institutions to ignore race when identifying prospective students for outreach and recruitment” (US Department of Justice & US Department of Education, 2023). Some colleges may respond to this guidance by adding student list purchases that target underrepresented students of color, either directly using race/ethnicity filters or indirectly by targeting particular zip codes.

We recommend a different course of action. Our previous research finds that the off-campus

recruiting visits of public research universities and selective private colleges disproportionately target affluent, predominantly white public and private schools (Jaquette et al., 2023; Salazar, 2022; Salazar et al., 2021). This article shows that student list products systematically exclude students of color and that discretion in student list purchases can unintentionally exclude students of color at alarming rates. Therefore, rather than adding a bit of racially progressive recruiting atop a racially biased recruitment funnel, we recommend that colleges begin by interrogating their recruiting practices. For example, how do the high schools they visit differ racially and socioeconomically from those they skip? How do the prospect profiles they purchase differ racially and socioeconomically from those they do not purchase?

6.3 Implications for Scholarship

Scholarship on recruiting largely assumes that recruiting is done by individual colleges. Prior research has not investigated how the underlying architecture of student list products makes prospects more or less 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 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). The College Board developed several geographic filters that create geographic borders based on historical, 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, the 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.

One topic for future research is demographic search filters, which may be the subject of future legal debate. Demographic filters allow colleges to target prospects by race, ethnicity, sex, and first-generation status. The equity rationale is that these filters may facilitate access for underrepresented populations. 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. Another topic for future research is patterns of segregation in schools and communities, which vary dramatically across and within metropolitan areas. Student list filters seem to interact with segregation patterns in ways that are not easily anticipated by administrators who purchase student lists.

Finally, we hope that this article motivates critical education policy research on selection devices more broadly. Amidst the data science revolution, schools and colleges increasingly use software-as-service products and digital platforms to perform core functions (Nichols & Garcia, 2022). Whether developed internally or by third parties, many products in the broader “edtech” space can be usefully conceptualized as selection devices; they categorize students for some opportunity or intervention based on a set of input factors. In enrollment management, for example, consulting firms develop algorithms that recommend how many names to buy in each zip code and which prospects should be targeted for which recruiting intervention (Fire Engine RED, 2022; James Madison University, 2017; Ruffalo Noel Levitz, 2021). Financial aid optimization products determine which admits receive which aid package (EAB, 2023; Leeds & DesJardins, 2015), while student success products predict which students are “at risk” (Cardona, Cudney, Hoerl, & Snyder, 2022; EAB, 2022).

These selection devices are structures; they structure opportunities for individual students in ways that may promote racial equality or reinforce inequality. Standardized selection device products – in which the outcome is an algorithmic function of the inputs – produce racial inequality by incorporating racialized inputs. Racialized inputs are reflections of historic racial discrimination. Selection devices that incorporate racialized inputs amplify the effects of historic racial discrimination. Discretionary selection devices produce racial inequality

by incorporating racialized inputs and because of bias or incomplete knowledge by the administrator wielding the product. Critical policy research can progress by deconstructing selection devices located at the nexus of education and technology. What inputs do these products incorporate? Are inputs incorporated in ways that counter or amplify historic racial inequality? How are these products utilized on the ground by schools, by colleges, and by the consulting firms they hire? The answers to these questions will inform policy regulations that promote racial equality in educational opportunities.

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. <https://doi.org/10.1177/000312240707200401>
- American Association of Collegiate Registrars and Admissions Officers. (2018). Overlooked inefficiencies in admissions that can impact yield. Retrieved from <https://www.aacrao.org/research-publications/research/eye-on-research-blog/article/aacrao-consulting/2018/05/03/overlooked-inefficiencies-in-admissions-that-can-impact-yield>
- 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/10.1016/j.econedurev.2019.101923>
- Arcidiacono, P., Kinsler, J., & Ransom, T. (2022). Recruit to reject? Harvard and African American applicants. *Economics of Education Review*, 88. <https://doi.org/10.1016/j.econedurev.2022.102255>
- Baker, R., Klasik, D., & Reardon, S. F. (2018). Race and stratification in college enrollment over time. *AERA Open*, 4(1), 2332–8584. <https://doi.org/10.1177/2332858417751896>
- Bastedo, M. N. (2016). Cognitive repairs in the admissions office. Unpublished Manuscript.
- 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*. <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. Retrieved from <https://sheeoed.medium.com/the-link-between-student-lists-and-direct-admissions-pros-cons-and-policy-implications-166882543e39>
- Burrell, J., & Fourcade, M. (2021). The society of algorithms. *Annual Review of Sociology*, 47, 213–237. <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. <https://doi.org/10.1177/0038038506067507>
- Campbell, A. (2017). *Higher education marketing: How to master your admissions funnel*. Hop Online. Retrieved from <https://hop-online.com/blog/higher-education-marketing-admissions-process/>
- Cardona, T., Cudney, E. A., Hoerl, R., & Snyder, J. (2022). Data mining and machine learning retention models in higher education. *Journal of College Student Retention: Research, Theory & Practice*, 0(0), 51–75. <https://doi.org/10.1177/1521025120964920>
- 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
- 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.
- Clinedinst, M. (2019). *State of college admission*. National Association for College Admission Counseling. Retrieved from <https://www.nacacnet.org/state-of-college-admission-report/>
- 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. (2015). *PSAT/NMSQT: Understanding scores*. College Board. Retrieved from <https://satsuite.collegeboard.org/media/pdf/2015-psat-nmsqt-concordance-tables.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*. The New Press.
- 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. <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. <https://doi.org/10.1353/rhe.2018.0033>
- Doyle, W. R. (2010). Changes in institutional aid, 1992-2003: The evolving role of merit aid. *Research in Higher Education*, 51(8), 789–810. <https://doi.org/10.1007/s11162-010-9177-0>
- Duncan, O. D., Ohlin, L. E., Reiss, A. J., & Stanton, H. R. (1953). Formal devices for making selection decisions. *American Journal of Sociology*, 58(6), 573–584. Retrieved from <http://www.jstor.org/stable/2772558>
- Dynarski, S., Libassi, C., Michelmore, K., & Owen, S. (2021). Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students. *American Economic Review*, 111(6), 1721–1756. <https://doi.org/10.1257/aer.20200451>
- EAB. (2018). *Making your digital ads count: 15 lessons on new and emerging techniques in undergraduate recruitment marketing*. EAB.
- EAB. (2019). The 5 key stages of college enrollment—and which metrics to track during each. Retrieved from <https://eab.com/insights/daily-briefing/enrollment/the-5-key-stages-of-college-enrollment-and-which-metrics-to-track-during-each/>
- EAB. (2022). Starfish: Scale your student success efforts. Retrieved from <https://eab.com/products/starfish/>
- EAB. (2023). Financial aid optimization. Retrieved from <https://eab.com/products/financial-aid-optimization/>
- Encoura. (n.d.). Encoura. ACT. Retrieved from <https://encoura.org/>
- Equifax. (2023). Drive customer growth: Best practices for financial services firms to acquire the right customers and grow share. Retrieved from <https://assets.equifax.com/marketing/US/assets/drive-customer-growth-ebook.pdf>
- Experian. (2023). *Mosaic: The best data powers the most robust segmentation*. Retrieved from <https://www.experian.com/content/dam/marketing/na/assets/ems/marketing-services/documents/product-sheets/mosaic-product-sheet.pdf>
- 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/>

- Green III, P. C., Baker, B. D., & Oluwole, J. O. (2021). School finance, race, and reparations. *Washington 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/10.1002/pam.22262>
- Harcourt, B. E. (2007). *Against prediction: Profiling, policing, and punishing in an actuarial age*. 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>
- Higher Education Act, 79. Stat. 1219., Pub. L. No. 89-329. (1965). Retrieved from <https://www.govinfo.gov/content/pkg/COMPS-765/pdf/COMPS-765.pdf>
- 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*. Jossey-Bass.
- Hossler, D., & Bontrager, B. (2014). *Handbook of strategic enrollment management*. Jossey-Bass. Retrieved from <http://www.wiley.com/WileyCDA/WileyTitle/productCd-1118819489.html>
- 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. 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. <https://doi.org/DOI: 10.1257/jep.23.4.95>
- 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
- Jackson, B. A., & Reynolds, J. R. (2013). The price of opportunity: Race, student loan debt, and college achievement. *Sociological Inquiry*, 83(3), 335–368. <https://doi.org/10.1111/soin.12012>
- James Madison University. (2017). Commonwealth of virginia standard contract, contract

- no. UCPJMU4648.
- Jaquette, O., Han, C., & Castaneda, I. (2023). The private school network: Recruiting visits to private high schools by public and private universities. [*Unpublished Manuscript*].
- Jaquette, O., & Salazar, K. G. (2022). *Student list policy: Problems, regulations, and a solution*. The Institute for College Access and Success. Retrieved from https://ticas.org/wp-content/uploads/2022/09/Student-list-policy_problems_regulations_a_solution.pdf
- Jaquette, O., Salazar, K. G., & Martin, P. (2022). *The student list business: Primer and market dynamics*. The Institute for College Access and Success. 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*. Princeton, N.J.: Princeton University Press.
- Killgore, L. (2009). Merit and competition in selective college admissions. *Review of Higher Education*, 32(4), 469–488. <https://doi.org/10.1353/rhe.0.0083>
- 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*. 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. <https://doi.org/10.5465/amj.2010.57319260>
- 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/10.1080/03085149900000013>
- Long, M. C. (2008). College quality and early adult outcomes. *Economics of Education Review*, 27(5), 588–602. <https://doi.org/10.1016/j.econedurev.2007.04.004>
- MacLeod, W. B., Riehl, E., Saavedra, J. E., & Urquiola, M. (2017). The big sort: College reputation and labor market outcomes. *American Economic Journal: Applied Economics*, 9(3), 223–261. <https://doi.org/10.1257/app.20160126>

- MacMilan, D., & Anderson, N. (2019). Student tracking, secret scores: How college admissions offices rank prospects before they apply. *Washington Post*, (October 14). Retrieved from <https://www.washingtonpost.com/business/2019/10/14/colleges-quietly-rank-prospective-students-based-their-personal-data/>
- 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.
- 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>
- 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*. <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.). 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. <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. <https://doi.org/10.1353/rhe.2015.0038>
- Poon, M. (2007). Scorecards as devices for consumer credit: The case of Fair, Isaac & Company Incorporated. *Sociological Review*, 55, 284–306. <https://doi.org/10.1111/j.1467-954X.2007.00740.x>
- Posecznick, A. (2017). *Selling hope and college: Merit, markets, and recruitment in an unranked school*. Ithaca, NY: Cornell University Press.
- Posselt, J. R. (2016). *Inside graduate admissions: Merit, diversity, and faculty gatekeeping*. 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. <https://doi.org/10.3102/0002831212439456>

- 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>
- Shamsuddin, S. (2016). Berkeley or bust? Estimating the causal effect of college selectivity on bachelor's degree completion. *Research in Higher Education*, 57(7), 795–822. <https://doi.org/10.1007/s11162-016-9408-0>
- 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*. Cambridge, MA: Harvard University Press.
- Tiako, M. J. N., South, E., & Ray, V. (2021). Medical schools as racialized organizations: A primer. *Annals of Internal Medicine*, 174(8), 1143–1144. <https://doi.org/10.7326/m21-0369> %m 34058105
- Tough, P. (2019). What college admissions officers really want. *New York Times Magazine*, (Sept. 10). Retrieved from <https://www.nytimes.com/interactive/2019/09/10/magazine/college-admissions-paul-tough.html>
- US Department of Education. (2023). *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 Justice, & US Department of Education. (2023). *Questions and answers regarding the Supreme Court's decision in Students for Fair Admissions, Inc. v. Harvard College and University of North Carolina*. US Department of Justice; US Department of Education. Retrieved from https://www2.ed.gov/about/offices/list/ocr/docs/ocr-questionsandanswers-tvi-20230814.pdf?utm_content=&utm_medium=email&utm_name=&utm_source=govdelivery&utm_term=
- Waxman, B. (2019). The power of list segmentation part I: Are you already doing this? instead, *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. <https://doi.org/10.1257/jep.13.1.13>

8 Tables

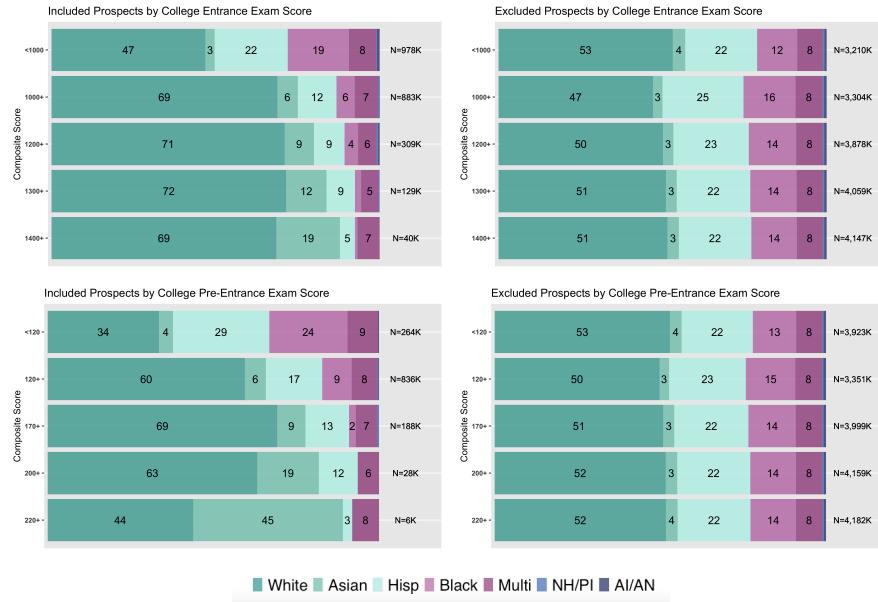
Table 1: Top filter combinations used in College Board orders purchased purchased by research vs. master's

Research			M.A.		
Filters	Count	Percent	Filters	Count	Percent
HS grad class, GPA, SAT, Zip code	99	19%	HS grad class, GPA, PSAT, Zip code	143	47%
HS grad class, GPA, SAT, PSAT, Rank, State, Race	39	7%	HS grad class, GPA, SAT, Zip code	107	35%
HS grad class, SAT, State	38	7%	HS grad class, GPA, SAT, PSAT, Zip code	28	9%
HS grad class, PSAT, State	28	5%	HS grad class, SAT, Geomarket	6	2%
HS grad class, GPA, SAT, State	23	4%	HS grad class, GPA, SAT, PSAT, County	4	1%
HS grad class, GPA, PSAT, State, Race	20	4%	HS grad class, GPA, SAT, County	4	1%
HS grad class, PSAT, State, Low SES	20	4%	HS grad class, SAT, Geomarket, College type	2	1%
HS grad class, GPA, PSAT, State	19	4%	HS grad class, PSAT, Geomarket	2	1%
HS grad class, GPA, AP score, Geomarket	15	3%	HS grad class, GPA, SAT, State, Segment, Major, Edu aspiration	1	0%
HS grad class, GPA, SAT, PSAT, State, Segment, Gender	13	2%	HS grad class, GPA, PSAT, State, Geomarket, Edu aspiration	1	0%

NOTES: Top filter combinations are based on 830 College Board student list purchases made by research (N=9) and master's (N=5) universities from 2016-2020. Geodemographic and Segment filters are proprietary filters created by College Board that use geodemography to predict the college-going behaviors of students within specific geographic areas.

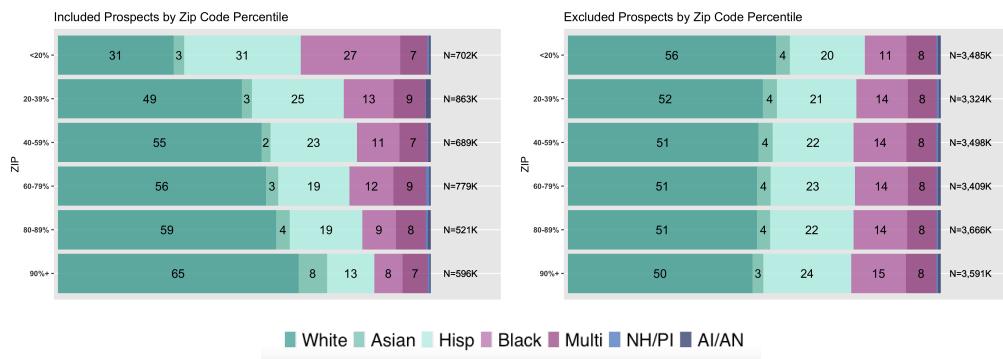
9 Figures

Figure 1: College Entrance and Pre-Entrance Exam Filters Across Thresholds



SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS09).

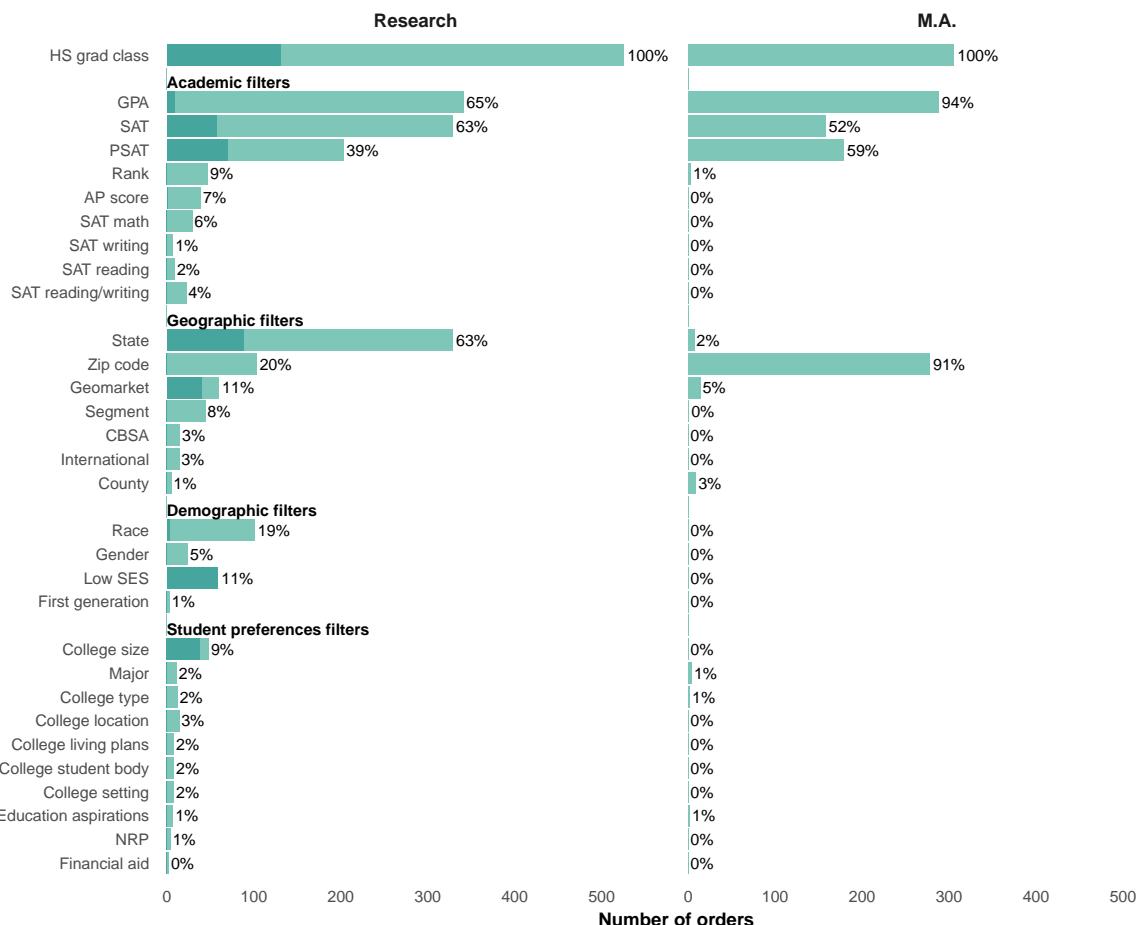
Figure 2: Zip Code Filter Across Affluence Percentiles



SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS09).

\end{landscape}

Figure 3: Filters used in College Board Orders Purchased by 14 Public Universities



NOTE: One research university placed a large amount of orders, which influenced the overall filters used for all research universities. The contribution of this particular university is shown in the darker color in the figure.

Figure 4: SAT Filter Used by Research vs. Master's Public Universities

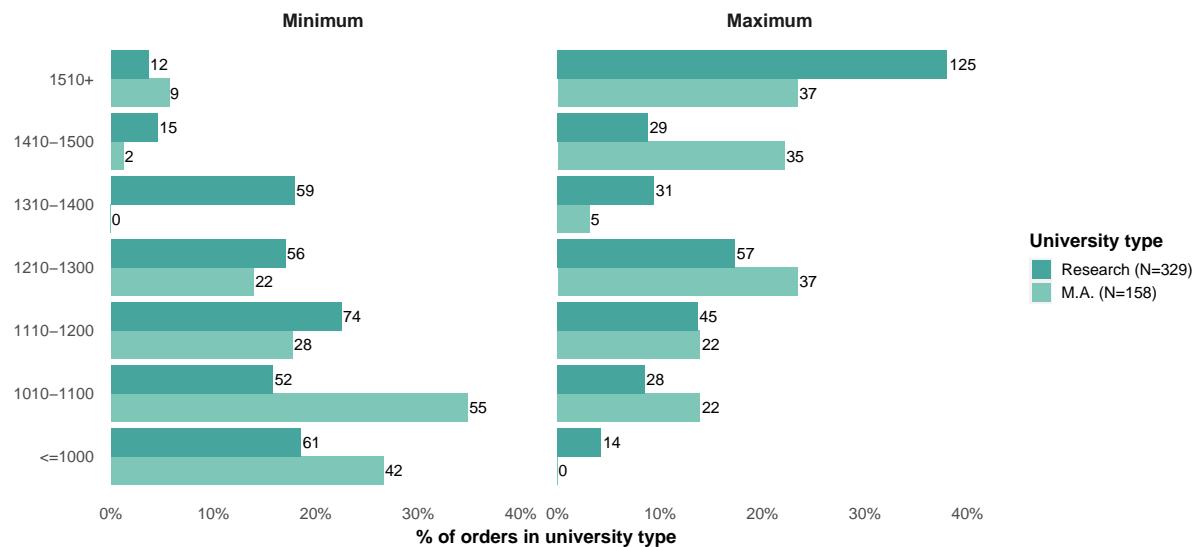
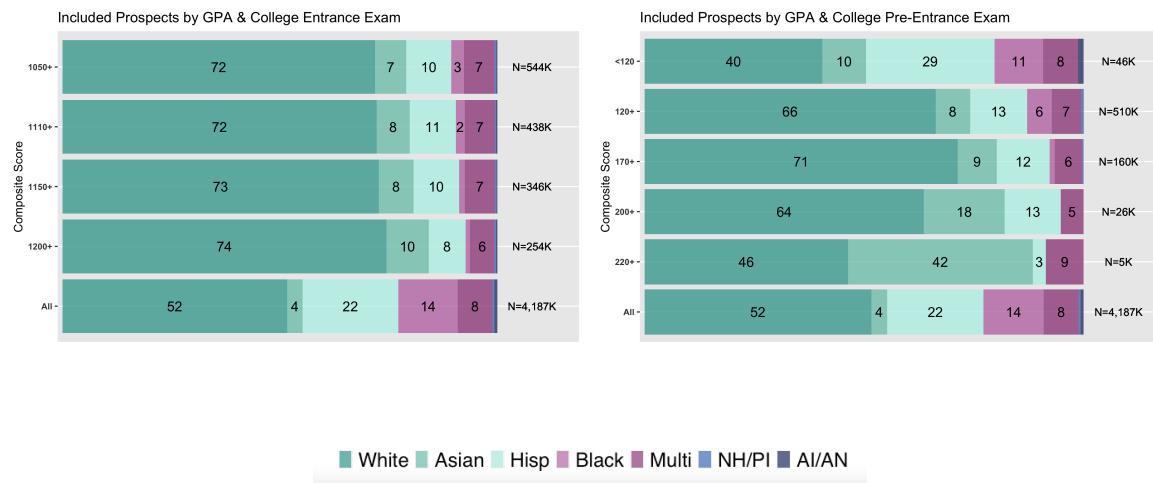
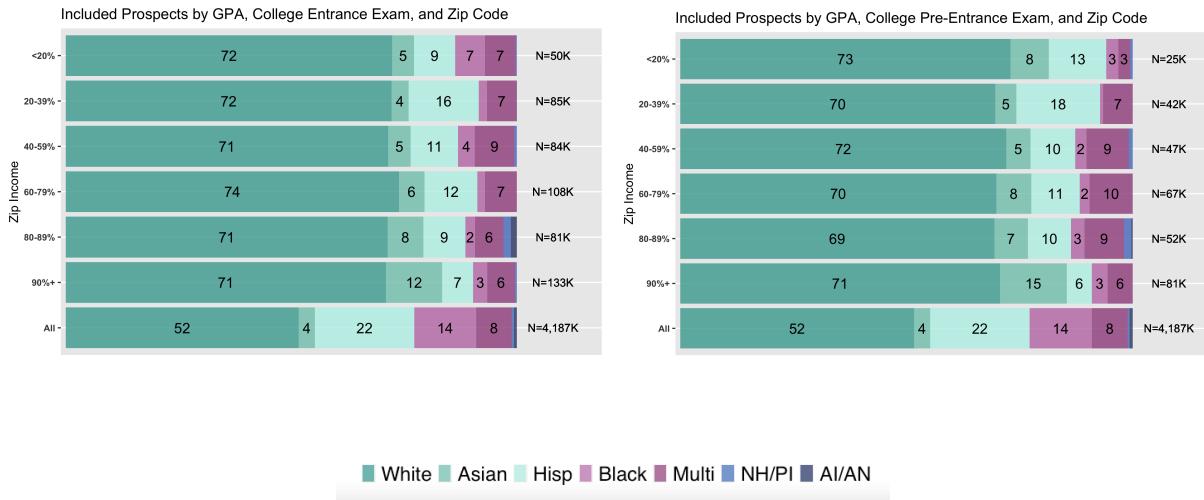


Figure 5: Academic Combinations: GPA (3.0+) and College Entrance or Pre-Entrance Exams (across score thresholds)



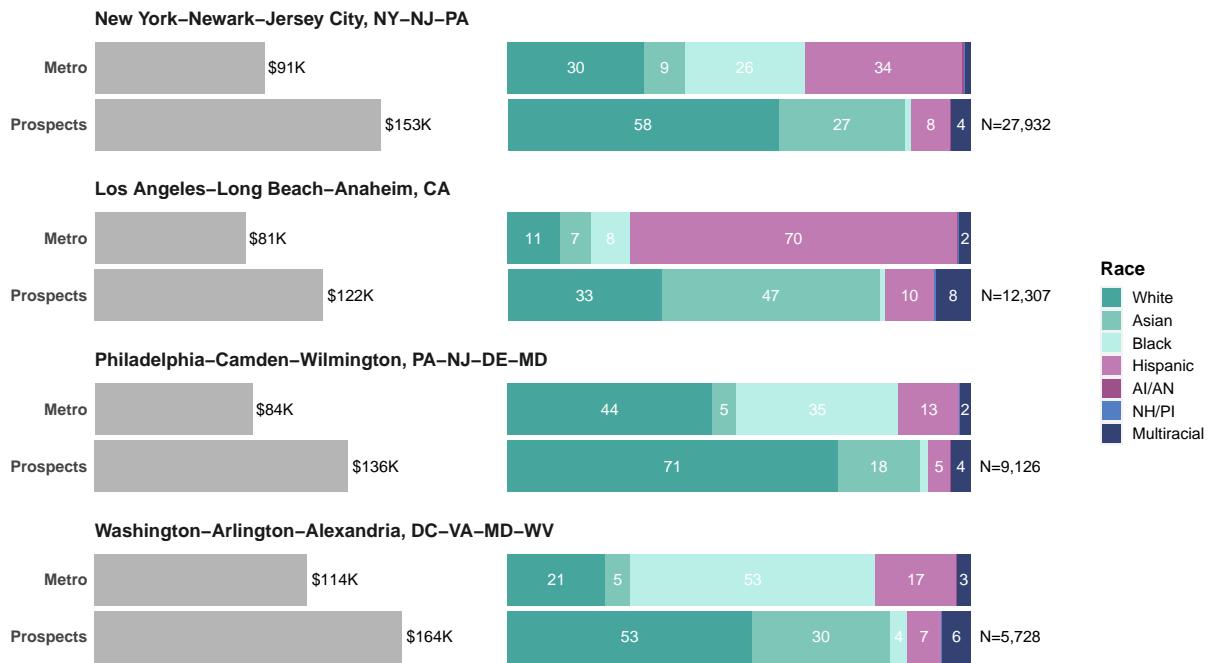
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS09).

Figure 6: Academic and Geographic Combination: GPA (3.0+), College Pre-Entrance (150+) or Entrance (1050+) Exams, and Zip (across income thresholds)



SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS09).

Figure 7: Segment Filter Prospects by Metropolitan Area



Note: Filters used across these orders include HS Class, GPA (B-A+), State (in-state vs. out-of-state), AP STEM (3 min for in-state; 4 min for out-of-state) or SAT (1200 minimum for in-state; 1300 minimum for out-of-state) with STEM major interest. Metro population source is the National Center for Education Statistic's Common Core of Data and includes all students attending a public high school in the metropolitan area.

Figure 8: Maps of Segment Filter Prospects by Metropolitan Area

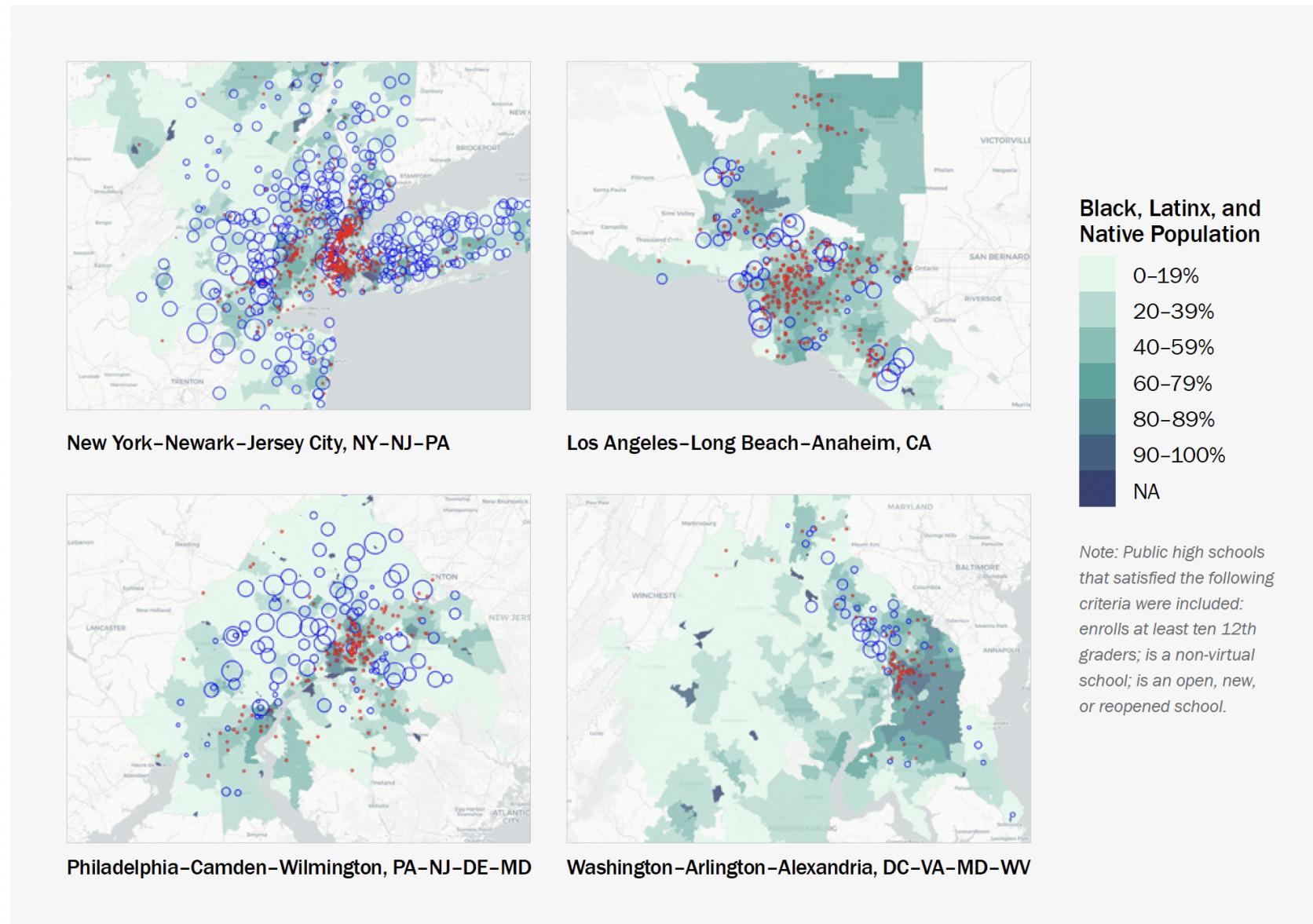
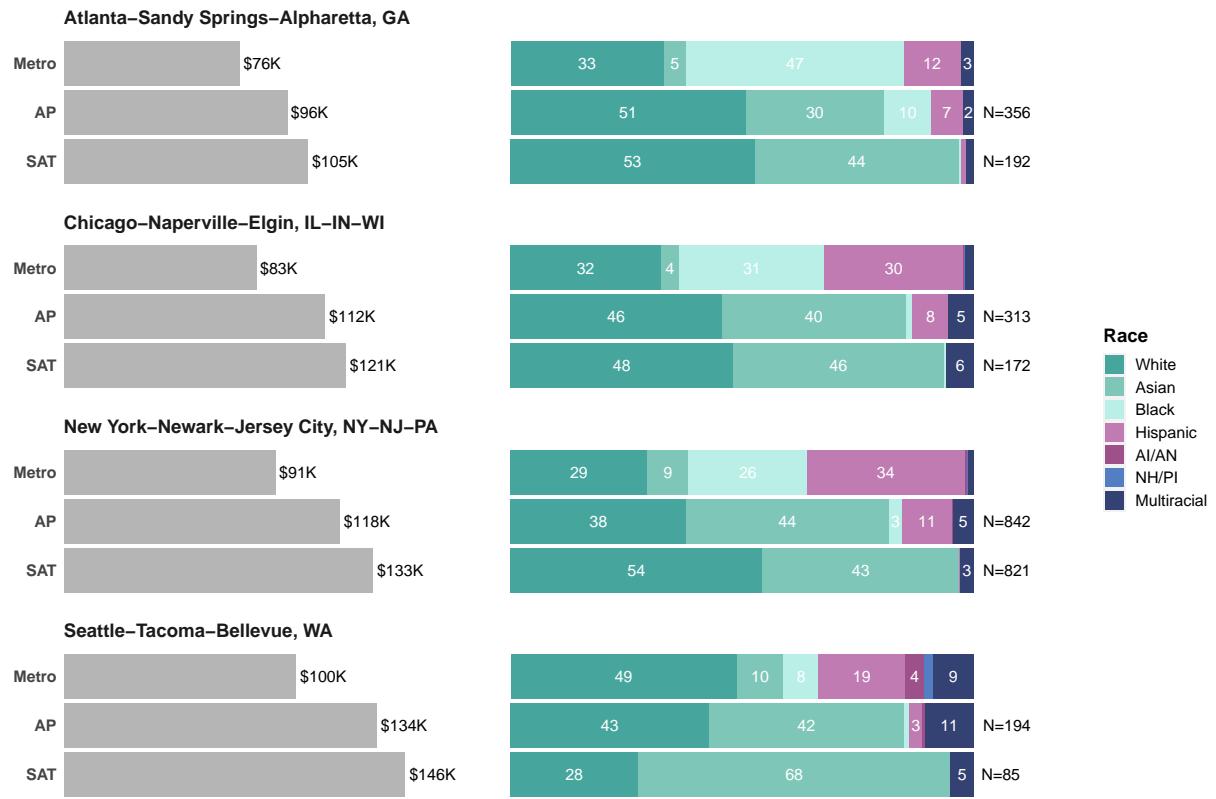


Figure 9: Women in STEM Prospects by Metropolitan Area



Note: Filters used across these orders include HS Class, Segment, GPA (B-A+), PSAT/SAT (1220-1450); State/CBSAs. Metro population source is the National Center for Education Statistic's Common Core of Data and includes all students attending a public high school in the metropolitan area.

10 Supplemental Online Appendix

Figure S1: The Enrollment Funnel

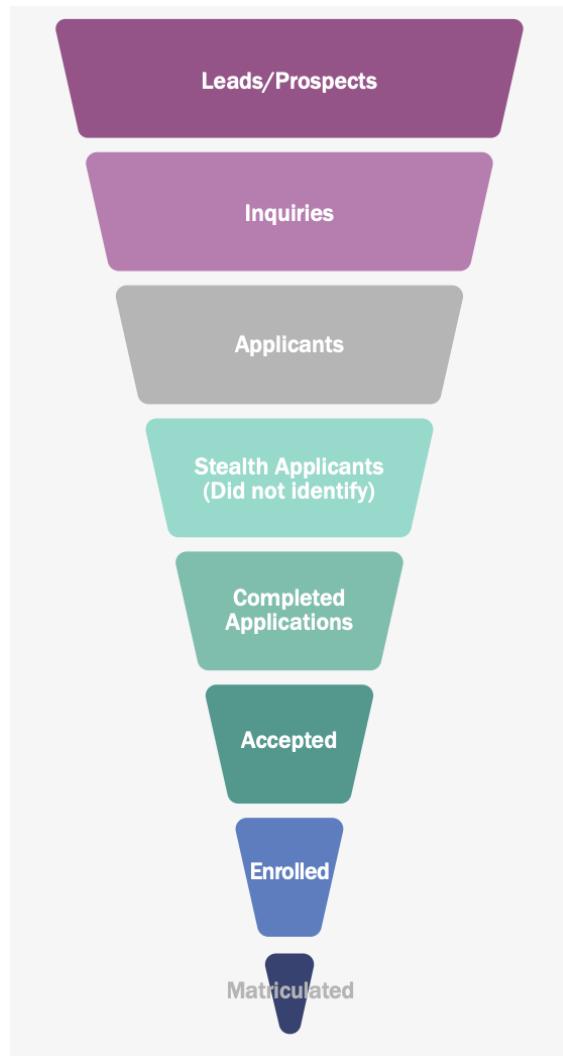
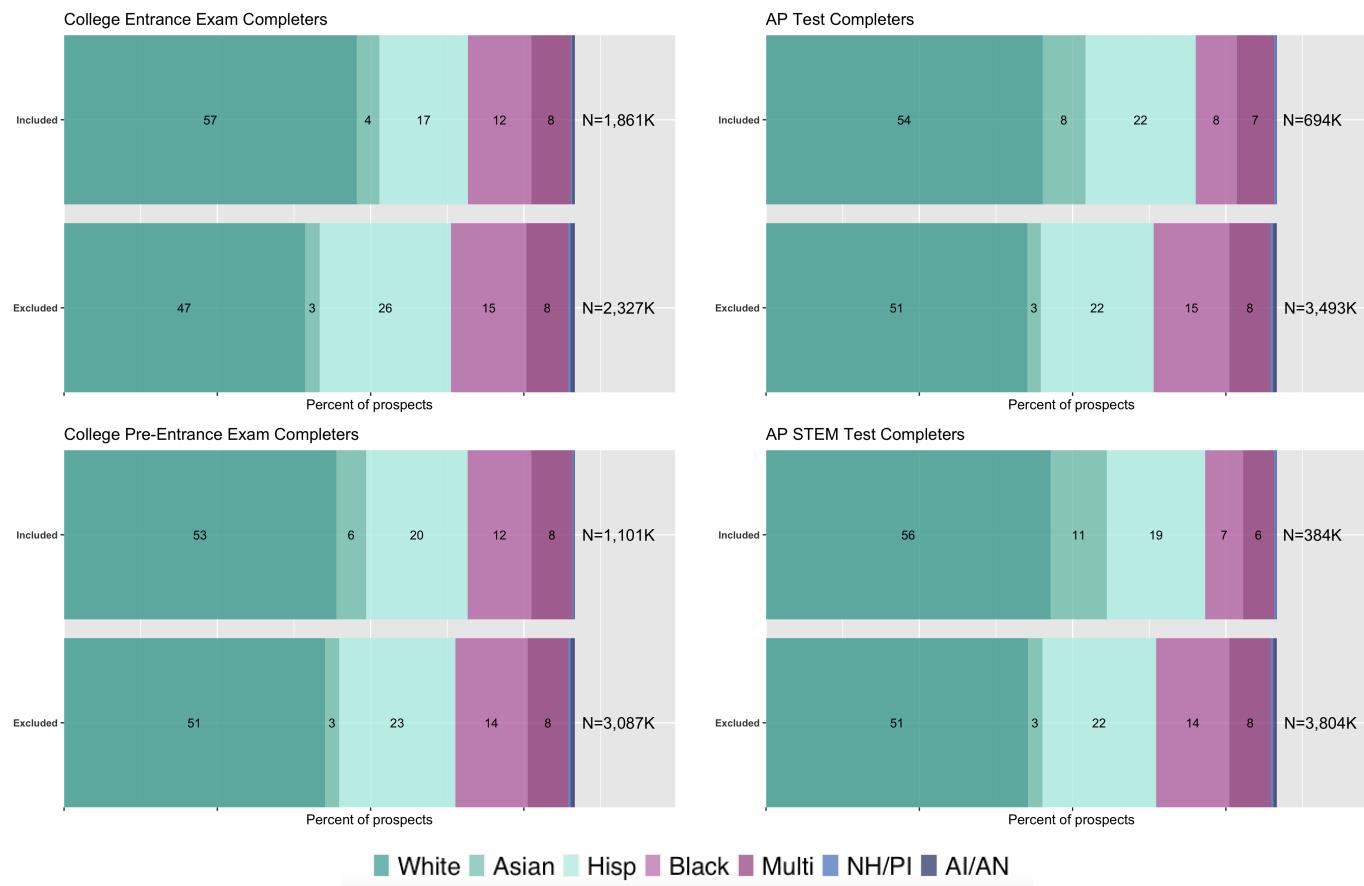


Figure S2: Test Takers Across College Entrance, College Pre-Entrance, and AP Assessments



Δ2

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS09).



SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS09).

Figure S4: PSAT Filter Used by Research vs. Master's Universities

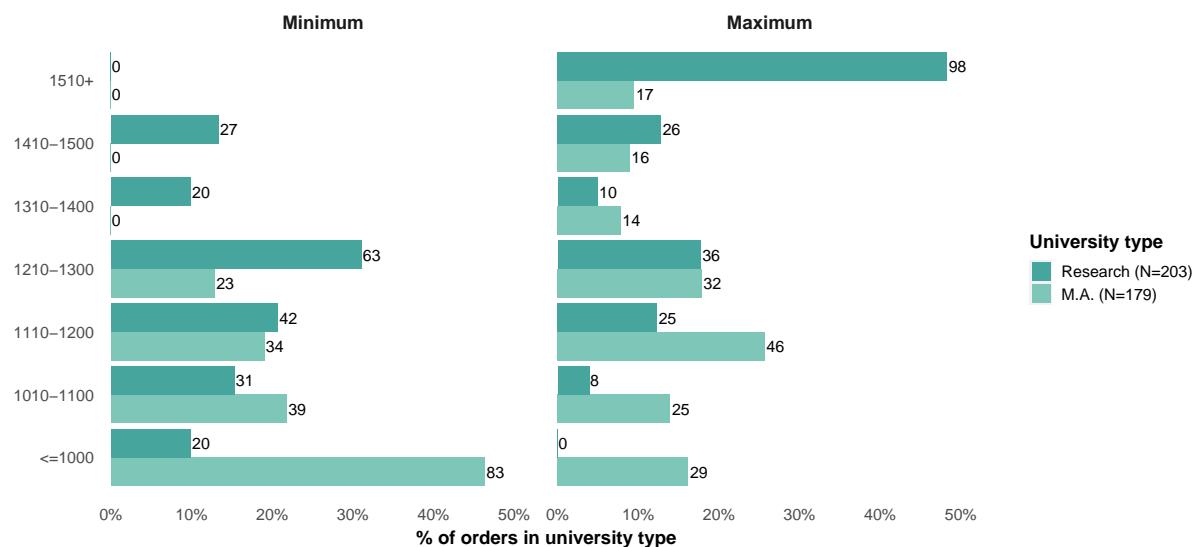
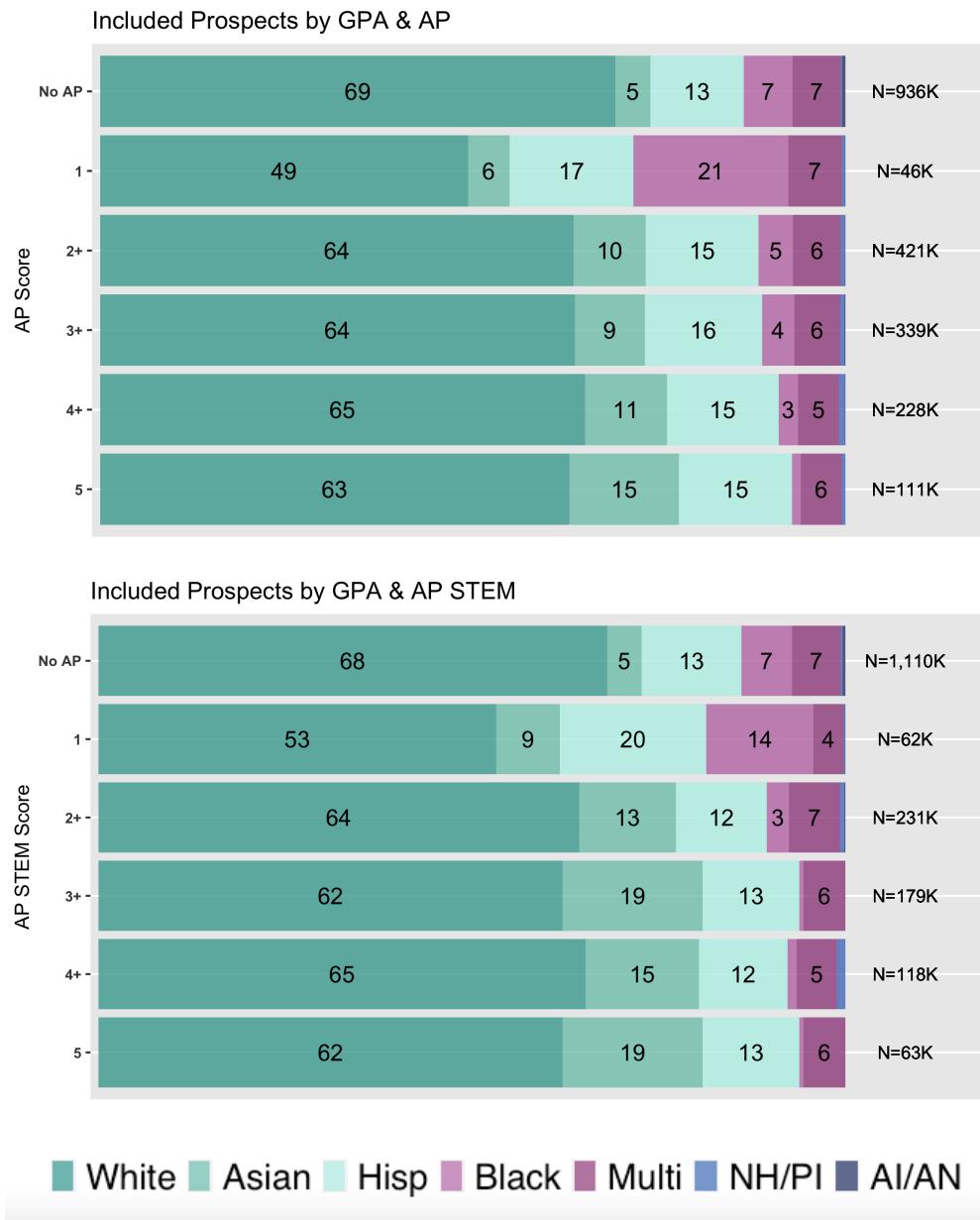


Figure S5: Academic Combination: GPA (3.0+) and AP (across score thresholds)



SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS09).

Table S1: HSLS09 Descriptive Statistics

	Unweighted	N	SE	Pct
Race/Ethnicity				
White	9,390	2,163,043	45,293	51.7
Asian	1,370	150,222	15,373	3.6
Hisp	2,520	920,384	41,451	22.0
Black	1,660	574,370	36,346	13.7
Multi	1,410	332,043	12,921	7.9
NH/PI	70	18,784	5,241	0.4
AI/AN	110	28,519	6,288	0.7
Academic Filters				
College Entrance Exam Completer	7,910	1,860,677	54,277	55.6
College Pre-Entrance Exam Completer	4,780	3,086,739	51,247	73.7
AP test-taker (any)	2,990	694,359	33,918	16.6
AP test-taker (STEM)	1,800	383,669	23,721	9.2
Academic GPA	16,480	4,177,402	6,863	99.8
Missing Academic GPA	40	9,964	6,562	0.2

NOTE: Unweighted sample sizes rounded to nearest 10 per NCES restricted data license regulations.

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS09).

Table S2: Characteristics of universities in public records request sample

Type	Tuition & Fees, In-State	Total Enrollment	Undergraduate Freshmen								
			% Non-Resident	% Pell	% Asian	% Black	% Hispanic	% Native American	% Multiracial	% White	
Arizona State University-Tempe	Research	\$10,792	54,161	40	29	9	4	23	1	5	50
Northern Arizona University	Research	\$10,289	29,609	38	35	2	3	26	2	7	57
University of California-Davis	Research	\$14,419	31,437	19	33	25	2	23	0	5	21
University of California-Irvine	Research	\$13,738	28,947	26	38	38	2	26	0	4	11
University of California-San Diego	Research	\$14,018	29,613	26	31	34	1	23	0	5	18
University of Illinois at Chicago	Research	\$15,027	19,476	8	59	22	8	40	0	3	22
University of Illinois at Urbana-Champaign	Research	\$17,293	35,476	26	24	20	7	13	0	3	41
Illinois State University	Research	\$12,353	20,115	4	30	2	10	12	0	3	71
Northeastern Illinois University	Master's	\$9,638	9,260	2	64	6	30	44	0	2	11
University of Illinois at Springfield	Master's	\$11,523	3,500	10	38	2	18	18	1	5	52
Texas A&M University-Texarkana	Master's	\$6,963	1,953	16	42	1	16	23	0	4	46
Stephen F Austin State University	Master's	\$7,716	12,149	1	41	1	16	24	0	5	53
Tarleton State University	Master's	\$7,367	12,853	2	40	1	6	21	0	3	67
Texas A & M University-College Station	Research	\$10,294	53,515	6	21	9	3	26	0	3	58

SOURCE: U.S. Department of Education, *Integrated Postsecondary Education Data System, 2017-18*.

Table S3: Assessment Completer Differences in Proportion

	Included	Excluded	Difference	Lower CI	Upper CI
College Entrance Exam					
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
College Pre-Entrance Exam					
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

NOTE: * $p<0.05$, ** $p<0.01$, *** $p<0.001$

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS09).

Table S4: Score Threshold Proportion Differences in Included vs. Excluded across Race/Ethnicity

	White	Asian	Hisp	Black	Multi	NH/PI	AI/AN
College Entrance Exam							
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
College Pre-Entrance Exam							
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***

NOTE: *p<0.05, **p<0.01, ***p<0.001

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS09).

Table S5: Zip Code Affluence Proportion Differences in Included vs. Excluded across Race/Ethnicity

	White	Asian	Hisp	Black	Multi	NH/PI	AI/AN
Affluence Percentile							
Less than 20%	-0.248***	-0.009***	0.112***	0.156***	-0.008***	0.001***	-0.003***
20-39%	-0.030***	-0.011***	0.034***	-0.005***	0.008***	-0.004***	0.008***
40-59%	0.035***	-0.014***	0.015***	-0.028***	-0.005***	-0.002***	0.000
60-79%	0.051***	-0.004***	-0.035***	-0.024***	0.011***	0.004***	-0.003***
80-89%	0.079***	0.000*	-0.028***	-0.054***	0.001*	0.002***	0.000***
Greater than 90%	0.151***	0.047***	-0.108***	-0.071***	-0.013***	0.001***	-0.004***

NOTE: *p<0.05, **p<0.01, ***p<0.001

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS09).

Table S6: GPA and College Entrance or Pre-Entrance Exam Score Threshold Proportion Differences in Included vs. Excluded across Race/Ethnicity

	White	Asian	Hisp	Black	Multi	NH/PI	AI/AN
College Entrance Exam							
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***
1200+	0.243***	0.066***	-0.144***	-0.135***	-0.025***	-0.001***	-0.004***
College Pre-Entrance Exam							
<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

NOTE: * $p<0.05$, ** $p<0.01$, *** $p<0.001$

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS09).

Table S7: 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

NOTE: * $p<0.05$, ** $p<0.01$, *** $p<0.001$

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS09).

Table S8: GPA, College Entrance/Pre-Entrance Exam, by Zip Code Affluence Proportion Differences in Included vs. Excluded across Race/Ethnicity

	White	Asian	Hisp	Black	Multi	NH/PI	AI/AN
College Entrance Exam (1050+)							
Less than 20%	0.210***	0.012***	-0.129***	-0.072***	-0.011***	-0.002***	-0.007***
20-39%	0.210***	0.001*	-0.065***	-0.122***	-0.015***	-0.005***	-0.006***
40-59%	0.201***	0.014***	-0.117***	-0.102***	0.009***	0.002***	-0.007***
60-79%	0.228***	0.021***	-0.106***	-0.123***	-0.009***	-0.005***	-0.007***
80-89%	0.200***	0.044***	-0.129***	-0.117***	-0.017***	0.012***	0.006***
Greater than 90%	0.200***	0.091***	-0.156***	-0.110***	-0.019***	0.000	NA
College Pre-Entrance Exam (150+)							
Less than 20%	0.215***	0.049***	-0.093***	-0.111***	-0.054***	0.002**	-0.007***
20-39%	0.181***	0.011***	-0.036***	-0.131***	-0.014***	-0.005***	-0.007***
40-59%	0.206***	0.017***	-0.122***	-0.113***	0.014***	0.005***	-0.007***
60-79%	0.184***	0.042***	-0.115***	-0.118***	0.017***	-0.005***	-0.007***
80-89%	0.180***	0.039***	-0.125***	-0.109***	0.008***	0.012***	-0.004***
Greater than 90%	0.193***	0.114***	-0.167***	-0.104***	-0.024***	NA	NA

NOTE: * $p<0.05$, ** $p<0.01$, *** $p<0.001$

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS09).

Table S9: Filter by neighborhood segments

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

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

Table S10: Filter by high school segments

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

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