

1 Introduction

Colleges and universities (herein universities) identify prospective students by purchasing “student lists” from College Board, ACT, and other vendors. A Student list contains the contact information of prospective students who meet the search filter criteria (e.g., test score range, high school GPA, zip codes) specified by the university. Purchased lists are a fundamental input for undergraduate recruiting campaigns, which target individual prospects by mail, email, and on social media (EAB, 2018).

Recent research suggests that student lists are surprisingly important for college access – and degree completion – for millions of students each year. Howell et al. (2021) compared SAT test-takers who opted into the College Board Student Search Service – allowing accredited institutions to “licence” their contact information – to those who opted out.¹ Figure 1 reproduces the main results. After controlling for covariates (e.g., SAT score, parental education, high school), 41.1% of students who participated in Search attended a 4-year college compared to 32.8% of students who opted out, an 8.3 percentage point difference and a 25.3 $((41.1 - 32.8) / 32.8)$ percent change in the relative probability. Participating in Search was associated with larger change in the relative probability of attending a 4-year college for students who identified as Black (24.5%) and Latinx (34.4%) than it was for students who identified as White (21.6%), and this change was larger for students whose parents did not attend college (40.6%) than it was for students whose parents had a BA (18.9%).

A report series from *TICAS* argues that student list products systematically exclude underrepresented students in two ways (Jaquette & Salazar, 2022; Jaquette, Salazar, & Martin, 2022; Salazar, Jaquette, & Han, 2022). The first is which prospective students are in the underlying database. Historically, student list products sold by College Board and ACT exclude non test-takers, but rates of test-taking differ by race and class [CITE]. Second, several search filters (e.g., AP score, geodemographic segment) used to control which prospect profiles are purchased facilitate the exclusion of students from communities of color and low-income communities. Prior scholarship on recruiting (e.g., Salazar, Jaquette, & Han, 2021; Stevens, 2007) ignores student list products because this literature assumes that recruiting is something done by individual colleges and universities. However, Jaquette et al. (2022) show that university recruiting behavior is structured by products sold by vendors and consultancies from the edtech/enrollment management industry.

Scholarship from critical data studies unpacks the business models and algorithms of digital platforms (e.g., Coursera). “Rentier capitalism” – as in rent tenants pays a landlord –

¹For a similar analysis of ACT’s Educational Opportunity Service see Moore (2017).

generates profit by charging monetary rent and data rent (Sadowski, 2019, 2020). Data rent is the “digital traces” created by users interacting with the platform, which often become the basis for new products. Student list data are extracted from user-data created by students laboring on platforms (e.g., taking the SAT, searching for college on Naviance) and these data are packaged to universities for monetary rent. Other scholarship shows how platforms reinforce racial inequality by embedding structural inequality within platform algorithms (e.g., Benjamin, 2019; Noble, 2018). In student list products, several search filters used to target prospects (e.g., zip code, text score) reflect historical inequality in educational opportunity. The nascent “platform studies in education” literature (Nichols & Garcia, 2022) observes that platforms increasingly perform core functions in education and calls for critical scholarship to inform policy regulations about edtech. However, the literature has not investigated how third-party platforms structure educational opportunity along the dimensions of race, class, and geography.

This study investigates the College Board Student Search Service product. We ask, what is the relationship between student list search filters and the characteristics of students who are included versus excluded in student lists purchased from College Board [KS - RQ HERE OR BEFORE ABOVE PARAGRAPH]?

We develop a conceptual framework by drawing from scholarship about algorithmic products from the sociology of race and critical data studies. Student list products are algorithmic selection devices that — similar to Google Ads or Facebook — allow advertisers to control the prospective customers through the use of search filters. Structural racism is systematic racial bias in which processes viewed as neutral or common-sense systematically advantage dominant groups (Bonilla-Silva, 1997; Ray, 2019). Structurally racist inputs are determinants of a selection device that are correlated with race because non-white people have been historically excluded from the input (Hirschman & Bosk, 2020). We conceptualize several student list search filters as structurally racist inputs. We develop propositions about the relationship between structurally racist search filters and racial exclusion.

We assess propositions using a nationally representative sample of high school students from the High School Longitudinal Study (HSL:09) and also data about student lists purchased by public universities, which we collected via public records requests. We reconstruct search filters and filter thresholds from the College Board Student Search Service product. We simulate student list purchases using theoretically motivated and commonly observed search filters with the goal of understanding how combinations of search filters and filter thresholds yield racial and socioeconomic inequality included versus excluded prospects.

The manuscript is organized as follows. First, we provide background on student list products, situating them vis-a-vis the process of recruiting students and summarizing recent dynamics in the market for student list data. Second, we review empirical scholarship on recruiting, focusing on scholarship from sociology. Third, we develop a a conceptual framework and propositions. Next, we describe methods and present results. Finally, we discuss implications for policy and scholarship. Several search filters may satisfy all “unfair practices” criteria of the FTC Act. Because of the systematic link between student lists and student loans, student list vendors may be classified as Consumer Reporting Agencies, thereby falling under the regulatory jurisdiction of the Consumer Finance Protection Bureau. Our broader contribution is to policy research on college access. Extant scholarship analyzes students, schools, or universities, often in relation to federal or state policies. We propose a critical, empirical research program on college access that focuses on third-party organizations and products in the edtech sector, which now dwarfs the market for for-profit Title IV institutions. As courts challenge progressive policies like affirmative action, policy research should go on the offensive. Use theories of structural inequality to investigate structural racism by edtech and target this research at regulatory agencies with relevant jurisdiction.

2 Background: Student List Products

2.1 Situating Lists Vis-a-vis Recruiting

Student lists are a match-making intermediary connecting universities to prospective students. The U.S. higher education market can be conceived as a national voucher system, whereby tuition revenue – composed of household savings and grants and loans from federal, state, and private sources – follow students to whichever institution they enroll in. Students want to attend college but do not know all their options, where they would be admitted, and how much it will cost. Universities have a financial incentive to provide access to students. Additionally, universities pursue some mix mix of broad enrollment goals (e.g., academic profile, racial diversity), while also meeting the needs of various campus constituencies (e.g., College of Engineering needs majors, marching band needs players) (Stevens, 2007). Universities cannot realize these goals solely from prospects who contact the university on their own. They must find prospects who can convinced to apply. However, universities don’t know who they are, where they are, or how to contact them. Student lists overcome the problem faced by universities, providing the contact information of prospects who satisfy their criteria.

The “enrollment funnel” – depicted in in Figure 2 – is a conceptual model used in the enrollment management industry to describe stages in the process of recruiting students. The

funnel begins with a large pool of “prospects” (i.e., prospective students) that the university would like to convert into enrolled students. “Leads” are prospects whose contact information (or “profiles”) has been purchased. “Inquiries” are prospects that contact your institution and consist of two types: first, inquiries who respond to an initial solicitation (e.g., email) from the university; and second, “student as first contact” inquiries who reach out to the university on their own (e.g., sending ACT scores, taking a [virtual tours](#) that records IP address). Applicants consist of inquiries who apply plus “stealth applicants” who do not contact the university before applying. The funnel narrows at each successive stage in order to convey the assumption of “melt” at each stage (e.g., a subset of “inquiries” will apply). Practically, the enrollment funnel informs interventions that increase the probability of “conversion” from one stage to another (Campbell, 2017). For example, financial aid packages are used to convert admits to enrolled students (e.g., McPherson & Schapiro, 1998).

At the top of the enrollment funnel, universities identify leads by buying student lists. The sum of purchased leads plus student-as-first-contact inquiries constitutes the set of all prospects the university has contact information for who may receive targeted recruiting interventions. Based on data provided by university clients, Ruffalo Noel Levitz (2020) reported that 28% of public universities purchased less than 50,000 names annually, 44% purchased 50,000-100,000 names, 13% purchased 100,000-150,000 names, and 15% purchased more than 150,000 names. Ruffalo Noel Levitz (2018) asked clients to rate different “first contact” interventions (e.g., off-campus recruiting visit) as sources of inquiries and enrolled students. For the median public university, 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%).²

2.2 The Market for Student List Data

Historically, the student list business has been dominated by College Board and ACT. In the 21st Century, student lists have been central a surprising level of dynamism in the broader enrollment management industry. Jaquette et al. (2022) describes key dynamics that shaped the contemporary market for student list data. The first dynamic is the centrality of enrollment management consulting firms to the student list business. Although universities are the paying customers of student list products, many universities outsource student list

²For private non-profit institutions, 34% of private institutions purchased fewer than 50,000 names, 24% purchased 50,000-100,000 names, 23% purchased 100,000-150,000 names, and 18% purchased more than 150,000 names (Ruffalo Noel Levitz, 2020). Additionally, student list purchases were the highest source of inquiries, accounting for 32% of inquiries and were tied with off-campus recruiting visits as the highest source of enrolled students, accounting for 18% of enrolled students (Ruffalo Noel Levitz, 2018).

purchases to enrollment management consulting firms. Furthermore, student lists are an essential input to the predictive models and recruiting interventions (e.g., emailing prospects) the consultancies provide.

The second dynamic is competition followed by concentration. In the 2000s, advances in technology yielded new sources of student list data, creating opportunities for new vendors. Start-ups entered the student list market by creating college search engines that asked students to submit information in order to receive recommendations about colleges and scholarships. Another new data source is college planning software (e.g., Naviance) sold to high schools and used by high school students and guidance counselors. In the 2010s, the enrollment management industry experienced a surge in horizontal followed by vertical acquisitions. Horizontal acquisitions occurred when one enrollment management consulting firm acquired a competitor (e.g., e.g., RuffaloCODY acquired Noel-Levitz in 2014) [CITE]. Vertical transformations transformed the student list business. For example, K-12 software provider PowerSchool entered the student list business by acquiring the Naviance college planning and Intersect student recruiting platforms from Hobsons. Enrollment consultant EAB entered the market for student list data through acquisitions (e.g., Cappex college search engine) and by becoming the exclusive reseller of the Intersect. As in the past, however, all student list products derive student list data from the user-data of students laboring on platforms.

Third, incumbents College Board and ACT attempted to retain their competitive advantage amidst the test-optional movement. Both organizations embraced data science by developing new search filters (e.g., ACT's predicted probability of enrolling) based on statistical models. Additionally, organizations leveraged their oligopoly position in the student list business to sell enrollment management consulting, offering clients information about prospects that is not included in purchased lists. However, the test-optional movement poses an existential threat to the college entrance exam. As fewer prospective students take College Board and ACT assessments, their competitive advantage in the coverage of college-going high school students is eroding, and private equity edtech firms such as EAB and PowerSchool are positioned to acquire market share ceded by College Board and ACT. Whereas College Board and ACT historically sold names at a price per-prospect (e.g., \$0.50 in 2021), for-profit edtech firms maximize profit by wrapping a large proprietary database of prospects within a software-as-service product (e.g., EAB's Enroll360) that universities must purchase in order to obtain access to these prospects.

3 Scholarship on Recruiting from Sociology

We position our study vis-a-vis scholarship on recruiting. Most scholarship on enrollment management focuses on latter stages of the enrollment funnel, particularly which applicants get admitted and financial aid leveraging to convert admits to enrolled students. Fewer studies investigate the earlier “recruiting” stages of identifying prospects, acquiring leads, and soliciting inquiries and applications. Whereas scholarship from economics tends to analyze the effects of specific recruiting interventions (e.g., Andrews, Imberman, & Lovenheim, 2020; Miller & Skimmyhorn, 2018; Smith, Howell, & Hurwitz, 2021), scholarship from sociology tends to observe how recruiting happens “in the wild” so to speak. Our review focuses on scholarship from sociology. We identify a key blind spot, one that is shared by scholarship from economics and the broader interdisciplinary field of education research.

Scholarship from sociology has analyzed recruiting from the perspective of students, high schools, and postsecondary institutions (e.g., Cottom, 2020; Holland, 2019; Posecznick, 2017; Salazar et al., 2021; Stevens, 2007). This literature primarily utilizes ethnographic or case-study designs, and often analyzes recruiting as part of a broader analysis of college access or enrollment management. Holland’s (2019) analysis of pathways from high school to college exemplifies scholarship that engages with recruiting from the perspective of high school students (e.g., McDonough, 1997). Students from groups underrepresented in higher education were drawn to colleges that made them feel wanted because they felt “school counselors had low expectations for them and were too quick to suggest that they attend community college” (Holland, 2019, p. 97). These students were strongly influenced by marketing material and high school recruiting visits, including small-group representative visits and instant decision events. By contrast, affluent students with college-educated parents were less taken by such overtures and more concerned with college prestige.

Several studies analyze connections between colleges and high schools from an organizational perspective. Off-campus recruiting visits are often conceptualized as an indicator of enrollment priorities and/or a network tie indicating the existence of a substantive relationship (Jaquette, Han, & Castaneda, forthcoming). Stevens (2007) provides an ethnography of enrollment management at a selective, tuition dependent private college sensitive about rankings. The college valued recruiting visits to (affluent) high schools as a means of maintaining relationships with guidance counselors at feeder schools. Khan (2011) analyzes the other side of the coin, showing how guidance counselors at an elite private school get under-qualified applicants into top colleges by exploiting colleges’ desire for information about which applicants will enroll if admitted. 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. Salazar (2022) analyzed by public research universities to out-of-state metropolitan areas, finding that universities engage in “recruitment redlining – the circuitous avoidance of predominantly Black and Latinx communities along recruiting visit paths” [p. X]. In contrast to branding about the commitment to racial diversity (Holland & Ford, 2021), these studies find that the recruiting efforts of selective institutions prioritize affluent, predominantly white schools and communities.

A smaller set of studies analyze recruiting at open-access institutions that target working adults (e.g., Cottom, 2017; Posecznick, 2017). Cottom (2017) is simultaneously an ethnography of enrollment management by for-profits and an analysis of political economy. For-profits found a niche in Black and Latinx communities precisely because traditional colleges ignored these communities. They systematically targeted women of color and generated profit by encouraging these students to take on federal and private loans. This business model exemplifies “predatory inclusion,” the logic of “including marginalized consumer-citizens into ostensibly democratizing mobility schemes on extractive terms” (Cottom, 2020, p. 443).

Collectively, empirical scholarship on recruiting from sociology – and other disciplines – assumes that recruiting is something done by individual colleges and universities. As a consequence, the recruiting literature ignores the role of third-party products and vendors. This blind spot has two root causes. First, scholarship on recruiting has not considered scholarship from critical data studies, which shows that digital platforms increasingly perform core organizational functions (e.g., Sadowski, 2019, 2020) and that digital platforms utilize algorithms that reinforce racial inequality (e.g., Benjamin, 2019). Second, scholarship on recruiting ignores the broader enrollment management industry that surrounds colleges and universities.

Jaquette et al. (2022) provide a conceptual analysis of the market for student list data in relation to the enrollment management industry. Although universities make choices about which names to purchase, these choices are structured by the algorithmic architecture of student list products — which prospects are included in the product, the targeting behaviors allowed by the product, the targeting behaviors encouraged by the product. Furthermore, many universities are uninformed about which prospective students they target because they outsource student list purchases to enrollment management consultancies (Salazar et al., 2022). These student list purchases substantially determine which prospective students will receive recruiting interventions at subsequent stages of the enrollment funnel. Thus, products and services sold by third-party vendors structure the recruiting behavior of individual universities and, in turn, college access opportunities for students.

Salazar et al. (2022) issued public records request to collect data about student lists purchased by public universities in four states. Their analyses sought to investigate the College Board student list product, rather than the behavior of universities purchasing the product. The primary research question was, what is the relationship between student list filter criteria and the characteristics of prospects included in purchased lists? For example, an analysis of several “women in STEM” purchases – which filtered on a combination of SAT/AP score, GPA, state, and intended major – showed that the racial and socioeconomic composition of targeted prospects differed dramatically from their surrounding metro area. Because of data limitations, however, Salazar et al. (2022) could not determine which filters and filter thresholds were driving this exclusion. This paper advances beyond Salazar et al. (2022) in two ways. First, we develop theoretically motivated propositions about search filters that are likely to yield racial inequality. Second, we test propositions using a nationally representative sample of high school students. These data allow us to examine who is included and excluded as filters and filter thresholds are changed and the analyses yield inferences that generalize to student populations of interest to policymakers.

4 Conceptual Framework

We develop a conceptual framework in two parts. First, we introduce concepts at the nexus of algorithms and structural racism, drawing largely from the sociology of race, but also from critical data studies and critical geography. Second, we apply these concepts to student list products and develop propositions about the relationship between search filters and exclusion.

4.1 Algorithms and Structural Racism

Algorithms and actuarialism. Algorithms are “sets of instructions written as code and run on computers” (Burrell & Fourcade, 2021, p. 215). Sociologists observe that algorithmic products utilize actuarial methods and are based on the logic of actuarialism (Burrell & Fourcade, 2021; Hirschman & Bosk, 2020; Simon, 1988). Pioneered by the insurance industry, actuarial methods proceed in two steps. First, apply statistical techniques to previous cases in order to identify factors positively and negatively associated with an outcome of interest. Second, apply these results to future cases in order to make predictions and to assign levels of risk to each case. Actuarialism is the ideology that equates fairness with risk, as determined by predicted probabilities. Under the logic of actuarialism, individuals or businesses that have characteristics associated with loan default should be charged higher interest rates.

Actuarial methods standardize decision-making by replacing individual judgment with de-

cisions based on a formula.³ Hirschman & Bosk (2020) observes that actuarial methods can promote racial equity if the primary source of inequality is racial bias of individual decision-makers. For example, E. Korver-Glenn (2018) shows that homes in predominantly white neighborhoods received higher appraisal values than those in non-white neighborhoods because appraisers have discretion in selecting similar comparison homes (“comps”) for the valuation.

However, scholarship from the sociology of race argues that actuarial methods may not reduce racial inequality due to structural racism. Drawing from Bonilla Silva’s (1997) concept of “racialized social systems,” structural racism is defined as “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. Hirschman & Bosk (2020, pp. 352–353) states that “actuarialism tends to bake [racial] inequality into the decision-making process, transmuting social disadvantages into seemingly objective measures of individual riskiness.” We discuss two mechanisms of structural racism in algorithmic products: (1) structurally racist inputs; and (2) market segmentation and micro-targeting.

Structurally racist inputs Actuarial products predict future outcomes by modeling the determinants of the outcome using historical data. Burrell & Fourcade (2021, p. 224) state that “predicting the future on the basis of the past threatens to reify and reproduce existing inequalities of treatment by institutions.” Even when actuarial products do not include race as a determinant, they often include determinants that are highly correlated with race. *Structurally racist inputs* are determinants of an outcome that systematically disadvantage non-white people and are correlated with race because non-white people have been historically excluded from this input (Harcourt, 2015; Hirschman & Bosk, 2020).

Obermeyer, Powers, Vogeli, & Mullainathan (2019) provide an empirical example of structurally racist inputs. They found that an algorithm used by hospital systems to predict patient health care needs under-predicted the needs of Black patients because the algorithm used health care costs as a proxy for health needs, but Black patients tend to receive less care relative to their health needs than other patients. An example by Norris (2021) reconstructed Moody’s city government credit rating algorithm, which assigns credit scores to cities based on determinants thought to predict loan default. The algorithm does not include the percent of residents who are Black, but it does include median household income, which is correlated with percent black because of historic wage discrimination. Once household

³Burrell & Fourcade (2021) observes that the adoption of actuarial methods across many industries was buoyed by concerns about racial equity following antidiscrimination legislation in the 1970s.

income is included in the model, percent Black is not longer a significant predictor of city credit rating. Thus, household income is a “racialized input” Norris (2021), defined as a seemingly neutral, structurally racist input that masks the structural racism of an algorithm by “explaining away” the relationship between race and the outcome.

Geographic structurally racist inputs emerge from algorithms that make predictions based on where people live (Benjamin, 2019). Because targeting by race can be profitable, platforms often capitalize on residential segregation by using geography as an input to circumvent laws prohibiting race as an input. Thus, Benjamin (2019, p. 147) writes, “racialized zip codes are the output of Jim Crow policies and the input of New Jim Code practices.”

The concepts “space” and “place” from critical geography (Agnew, 2011; Bell, 2007) provide insight about geographic structurally racist inputs. Place denotes a holistic understanding of a geographic location that incorporates the “history, peoples, and purposes within the political, social, and economic landscape” (Bell, 2007, p. 317). By contrast, space simply refers to a physical location which can be described in terms of quantifiable spatial features (e.g., distance, demographics, economic activity). Geospatial analyses typically adopt this view of space “as a location on a surface where things ‘just happen’” (Agnew, 2011, p. 318).

Market research conceives of geography as space and exploits racial segregation as a means of identifying customers (Benjamin, 2019; Noble, 2018). For example, geodemography is a branch of market research that estimates the behavior of consumers based on where they live. College Board (2011, p. 1), which develops geodemographic segment search filters, states that “the basic tenet of geodemography is that people with similar cultural backgrounds, means, and perspectives naturally gravitate toward one another or form relatively homogeneous communities; in other words, birds of a feather flock together.” By contrast, scholarship on racial segregation conceives of geography as place. Racial segregation is a function of historic and contemporary structurally racist laws, policies, and practices promoting residential segregation (Harris, 1993; J. Howell & Korver-Glenn, 2018; E. Korver-Glenn, 2018; e.g., Rothstein, 2017). Algorithms that categorize people based on geographic location (space) without considering historic and contemporary structures (place) that produce residential segregation are to reproduce historical race-based inequality in opportunity.

Market segments and micro-targeting. A second source of structural racism in products utilizing actuarial methods are the related processes of market segmentation and micro-targeting. Market segmentation categorizes customers into groups (e.g., “married sophisticates,” “rural everlasting”) that are useful for advertisers (Federal Trade Commission, 2014). Micro-targeting is the process of using data to precisely identify and target granular

segments of society (Cotter, 2022).

In sociology, Fourcade & Healy (2013) define “classification situations” – which encompasses market segmentation – as the use of actuarial techniques to categorize consumers into different groups. Historically, classifications were binary; consumers with “good” credit were offered loans and those with bad credit were not. Advances in data analytics (e.g., individual credit scores) enabled finer classifications; customers are classified into many groups, or along a continuum, alongside the emergence of tiered products targeting different consumer groups with different benefit levels. For example, “payday loans” charge high interest rates to consumer groups that were previously denied credit altogether. Thus, at one end of the continuum, these classifications are similar to Cottom’s (2017, 2020) concept “predatory inclusion,” while at the other end of the continuum marginalized populations are excluded from attractive product offerings.

Scholarship from critical data studies shows that racial exclusion is a predictable consequence market segmentation (Benjamin, 2019; Cotter, Medeiros, Pak, & Thorson, 2021; Noble, 2018). The process of developing a classification system requires developers to make a series of inherently subjective decisions (e.g., who is in the dataset, which measures to utilize, which categories to identify), creating opportunities for individual biases of developers and structurally racist inputs to enter the algorithm (Noble, 2018). Furthermore, classification systems are developed to optimize profit. In turn, “audiences are treated as a commodity to be bought and sold. When audience segments are under-valued in the market, demand among advertisers for the ability to reach them will be relatively low, which decreases the likelihood that a corresponding segment will be produced” (Cotter et al., 2021, p. 3).

Critical data studies has also investigated micro-targeting (e.g., Benjamin, 2019; Cotter, 2022; Cotter et al., 2021). For example, Cotter et al. (2021, p. 1) state that micro-targeting by Facebook “is driven not by a goal of making all users available to advertisers, but of making the ‘right’ individuals available. Facebook advises that advertisers ‘Implement a targeting strategy that focuses on reach and precision and eliminates waste.’” A theme from scholarship on micro-targeting in politics is that these technologies *could* be used to to increase outreach to marginalized groups, but in practice they are not. Instead, “campaigns routinely ‘redline’ the electorate, ignoring individuals they model as unlikely to vote, such as unregistered, uneducated, and poor voters (Kreiss, 2012, p. 74-75).⁴

⁴WE PULLED THIS QUOTE FROM COTTER 2021

4.2 Mechanisms of Exclusion in Student List Products

Conceptualizing student-list products. Student list products allow universities to choose prospective students from within some database of prospects by selecting on search filters. Salazar et al. (2022) categorize the filters available in the College Board Student Search Service product into the four buckets of geographic, academic, demographic, and student preferences (e.g., desired campus size, intended major), as shown in Table X.

Student list products have similarities and differences to algorithmic products that have been analyzed within sociology and critical data studies. Sociologists often study algorithmic products that assign scores based on the value of input determinants. For example, Moody’s algorithm assigns credit scores to cities based on inputs correlated with default in previous cases (Norris, 2021). Similar to these products, student lists utilize search filters that can be conceptualized as structurally racist inputs. For example, the College Board product enables universities to filter prospects based on 5-digit zip code, but zip codes are racially segregated.

Algorithmic products that make decisions purely based on scores (e.g., most credit offers) remove racial inequality caused by explicit or implicit individual decision-making bias (Hirschman & Bosk, 2020). By contrast, student list products are similar to purchasing ads from Google or Facebook in that advertisers (universities) choose prospective customers by selecting on search filters. Universities purchasing lists may be thoughtful about avoiding structurally racist search filters. However, this individual discretion raises the possibility of racial disparities due to individual bias or lack of knowledge about the products.

Market segmentation and micro-targeting are at the core of student list products. Universities are encouraged to execute multiple student list purchase, each targeting different market segments (Waxman, 2019). Purchases target granular populations by simultaneously filtering on several filters. As in political advertising (Cotter, 2022), micro-targeting has become a branding strategy for student list products. For example, College Board Student Search promises to “create a real pipeline of best-fit prospects” College Board (n.d.) while ACT Encoura uses the tag-line “find and engage your best-fit students” Encoura (n.d.). Recently, student list products have developed new search filters that utilize market segmentation and micro-targeting approaches that have been analyzed within critical data studies (Cotter et al., 2021; Noble, 2018). For example, College Board (2011) utilized market segmentation methodologies to create “geodemographic segment” filters, which classify each high school and each neighborhood to a group based on the past college-going behaviors of students.⁵ Other new filters are designed to increase precision, for example College Board’s “interest

⁵In 2021, College Board added three new geodemographic filters (College Board, 2021).

in my peers” filter and ACT’s predicted probability of enrollment filter. To the extent that universities want to target students from affluent schools and communities (Salazar et al., 2021), contemporary student list products facilitate this goal with great efficiency.

4.2.1 Predicting Exclusion

Our analyses focus exclusion due to structural inequality embedded in the underlying student list product as opposed to exclusion that emerges from the individual bias of people purchasing lists. We posit two broad sources of structural inequality in student list products: (1) who is included in the underlying database; and (2) utilizing structurally racist inputs as search filters for selecting prospects from the underlying database. This section draws from the theoretical discussion above to develop propositions about the relationship between search filters and exclusion, focusing on academic and geographic search filters.

Academic filters. College Board academic filters include high school graduating class, SAT score, PSAT score, AP score by subject, high school GPA, and high school class rank.

The first source of structural inequality in student list products is which prospective students are included in the underlying database. 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).⁶ Prior research shows that rates of SAT/ACT test-taking differ by race and class (e.g., Bastedo & Jaquette, 2011). Similarly, the percentage of students who take AP exams vary across race, particularly for STEM exams (Kolluri, 2018), and Black students are more likely than white students to attend a high school with few AP course offerings (Rodriguez & McGuire, 2019). These findings motivate the following proposition, which we analyze separately by assessment (SAT, PSAT, AP) and for taking any assessment.

P1: The condition of taking standardized assessments (X) is associated with racial disparities in who is included versus excluded (Y)

Second, search filters that condition on test scores thresholds are a source of exclusion that builds on differences in test-taking. Average standardized test scores differ by race and by class [CITE]. Prior research finds that access to test preparation varies by race and class (Park & Becks, 2015). Furthermore, prior research finds that SAT question items are racially and socioeconomically biased (Freedle, 2003; e.g., Santelices & Wilson, 2010). Therefore, we conceptualize test score filters as a structurally racist input in student list products. By contrast, prior research suggests that high school GPA is a less biased measure of performance (Alon & Tienda, 2007; Bastedo & Jaquette, 2011; Posselt, Jaquette, Bielby, & Bastedo, 2012)

⁶footnote: recently added college search engine blah blah

and that GPA is a strong predictor of postsecondary student success (Allensworth & Clark, 2020; Niu & Tienda, 2010). We test the following proposition separately by assessment (SAT, PSAT, AP) and for GPA:

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

Geographic filters. Geographic search filters enable universities to target prospects based on where they live. College Board Student Search Service offers two kinds of geographic search filters: filters based on known geographic borders (e.g., state, CBSA, county, zip code); and filters based on new geographic borders that College Board creates using historic data about test-takers (e.g., geomarket, geodemographic segment). At present we do not have access to data on borders created by College Board.

We conceptualize geographic search filters as structurally racist inputs because these search filters are built on the back of historic and contemporary policies and practices promoting racial segregation in American schools and communities. Drawing from critical geography (e.g., Agnew, 2011), targeting prospective students based on geographic location (space) without consideration to the historic and contemporary structures (place) that produce residential segregation is likely to reinforce historical race-based inequality in educational opportunity.

We expect that utilizing finer geographic filters (e.g., zip code rather than county) is associated with greater racial and socioeconomic disparities in student list purchases because American residential segregation occurs at fine-grained geographic levels (Elizabeth Korver-Glenn, 2022).

P3. Filtering on smaller geographic localities is associated with greater racial disparities in included vs. excluded than filtering on larger geographic localities.

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

P4. As purchases filter on higher levels of zip-code affluence, the proportion of underrepresented minority students included in student lists declines relative to the proportion who are

excluded

Filtering on academics and geography. Actual student list purchases filter on several criteria rather than one. Salazar et al. (2022) analyzed data on 830 student lists purchased by 14 public universities. The average purchase specified X.X criteria and XX% of purchases simultaneously specified at least one academic and one geographic filter. Filtering on multiple search criteria facilitates micro-targeting of desired prospects. The flipside of micro-targeting is exclusion. We suggest that filtering on multiple structurally racist inputs has a compounding effect on racial inequality in which prospects are included versus excluded. To assess this claim, we draw on the Salazar et al. (2022) sample of orders placed by public research universities and select several orders that utilized common, but potentially problematic filter criteria. We analyze the racial composition of students included versus excluded from these purchases. Next, we simulate marginal changes to order criteria to gain insight about how structurally racist inputs drive exclusion.

5 Data and Methods

5.1 Data

We use data from the High School Longitudinal Study of 2009 (HSL09), which is a nationally representative survey that follows a cohort of more than 20,000 students from 944 schools entering the ninth grade in Fall 2009. Surveys collected during the base year include responses from students' parents, teachers, school administrators, and school counselors. Five subsequent data collection waves capture students' schooling experiences through 11th grade, high school graduation and postsecondary plans, and enrolling in college or entering the labor market. HSL09 provides the extensive student-level demographic, geographic, and academic variables needed to recreate academic and geographic filters for this study to simulate.

6 Discussion

What these platforms do with student data is opaque, because obfuscation is a deliberate strategy to avoid regulation (Cottom, 2020). Obfuscation is not an excuse for not doing the research, but rather the reason that policy research must take on this challenge

Although filtering on specific zip codes may be instrumentally rational from the perspective of university enrollment goals, there is no equality of opportunity rationale for a product that enables universities to target students living in one zip code and exclude students from the

neighboring zip code [?MOVE SENTENCE TO DISCUSSION?].

7 References

- Agnew, J. A. (2011). Space and place. In J. A. Agnew & D. N. Livingstone (Eds.), *The SAGE handbook of geographical knowledge* (pp. 316–330). Los Angeles: SAGE.
- Allensworth, E. M., & Clark, K. (2020). High school GPAs and ACT scores as predictors of college completion: Examining assumptions about consistency across high schools. *Educational Researcher*, 49(3), 198–211. <https://doi.org/10.3102/0013189X20902110>
- Alon, S., & Tienda, M. (2007). Diversity, opportunity, and the shifting meritocracy in higher education. *American Sociological Review*, 72(4), 487–511. Journal Article. Retrieved from <Go to ISI>://000248696500001
- Andrews, R. J., Imberman, S. A., & Lovenheim, M. F. (2020). Recruiting and supporting low-income, high-achieving students at flagship universities. *Economics of Education Review*, 74, 101923. <https://doi.org/https://doi.org/10.1016/j.econedurev.2019.101923>
- Bastedo, M. N., & Jaquette, O. (2011). Running in place: Low income students and the dynamics of higher education stratification. *Educational Evaluation and Policy Analysis*, 33(3), 318–339.
- Bell, C. A. (2007). Space and place: Urban parents' geographical preferences for schools. *The Urban Review*, 39(4), 375–404. <https://doi.org/10.1007/s11256-007-0059-5>
- Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new jim code*. Medford, MA: Polity.
- Bonilla-Silva, E. (1997). Rethinking racism: Toward a structural interpretation. *American Sociological Review*, 62(3), 465–480. <https://doi.org/10.2307/2657316>
- Burrell, J., & Fourcade, M. (2021). The society of algorithms. *Annual Review of Sociology*, 47, 213–237. Journal Article. <https://doi.org/10.1146/annurev-soc-090820-020800>
- 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/>
- College Board. (n.d.). College board search solutions. The College Board. Retrieved from <https://cbsearch.collegeboard.org/solutions>
- College Board. (2011). *Segment Analysis Service: An educationally relevant geodemographic tagging service*. College Board. Retrieved from <https://secure-media.collegeboard.org/mSSS/media/pdf/segment-analysis-service-overview.pdf>
- College Board. (2021). Introducing environmental attributes. *YouTube*. Retrieved from <https://www.youtube.com/watch?v=VmTU9sb4ZiY>
- Cotter, K. (2022). Selling political data: How political ad tech firms' discourses legitimate microtargeting. In *17th international conference on information for a better world - shaping the global future (iConference)* (Vol. 13192, pp. 195–208). CHAM: Springer

- International Publishing Ag. https://doi.org/10.1007/978-3-030-96957-8_18
- Cotter, K., Medeiros, M., Pak, C., & Thorson, K. (2021). "Reach the right people": The politics of "interests" in facebook's classification system for ad targeting. *Big Data & Society*, 8(1), 16. <https://doi.org/10.1177/2053951721996046>
- Cottom, T. M. (2017). *Lower ed: The troubling rise of for-profit colleges in the new economy*. New Press, The.
- Cottom, T. M. (2020). Where platform capitalism and racial capitalism meet: The sociology of race and racism in the digital society. *Sociology of Race and Ethnicity*, 6(4), 441–449. <https://doi.org/10.1177/2332649220949473>
- EAB. (2018). *Making your digital ads count: 15 lessons on new and emerging techniques in undergraduate recruitment marketing*. EAB.
- Encoura. (n.d.). Encoura. ACT. Retrieved from <https://encoura.org/>
- 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>
- Fourcade, M., & Healy, K. (2013). Classification situations: Life-chances in the neoliberal era. *Accounting Organizations and Society*, 38(8), 559–572. Journal Article. <https://doi.org/10.1016/j.aos.2013.11.002>
- Freedle, R. (2003). Correcting the SAT's ethnic and social-class bias: A method for reestimating SAT scores. *Harvard Educational Review*, 73(1), 1–43.
- Harcourt, B. E. (2015). Risk as a proxy for race: The dangers of risk assessment. *Federal Sentencing Reporter*, 27(4), 237–243. <https://doi.org/10.1525/fsr.2015.27.4.237>
- Harris, C. I. (1993). Whiteness as property. *Harvard Law Review*, 106(8), 1707–1791. <https://doi.org/10.2307/1341787>
- Hirschman, D., & Bosk, E. A. (2020). Standardizing biases: Selection devices and the quantification of race. *Sociology of Race and Ethnicity*, 6(3), 348–364. <https://doi.org/10.1177/2332649219844797>
- Holland, M. M. (2019). *Divergent paths to college: Race, class, and inequality in high schools*. Rutgers University Press. <https://doi.org/10.36019/9780813590288>
- Holland, M. M., & Ford, K. S. (2021). Legitimizing prestige through diversity: How higher education institutions represent ethno-racial diversity across levels of selectivity. *Journal of Higher Education*, 92(1), 1–30. <https://doi.org/10.1080/00221546.2020.1740532>
- Howell, Jessica, 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

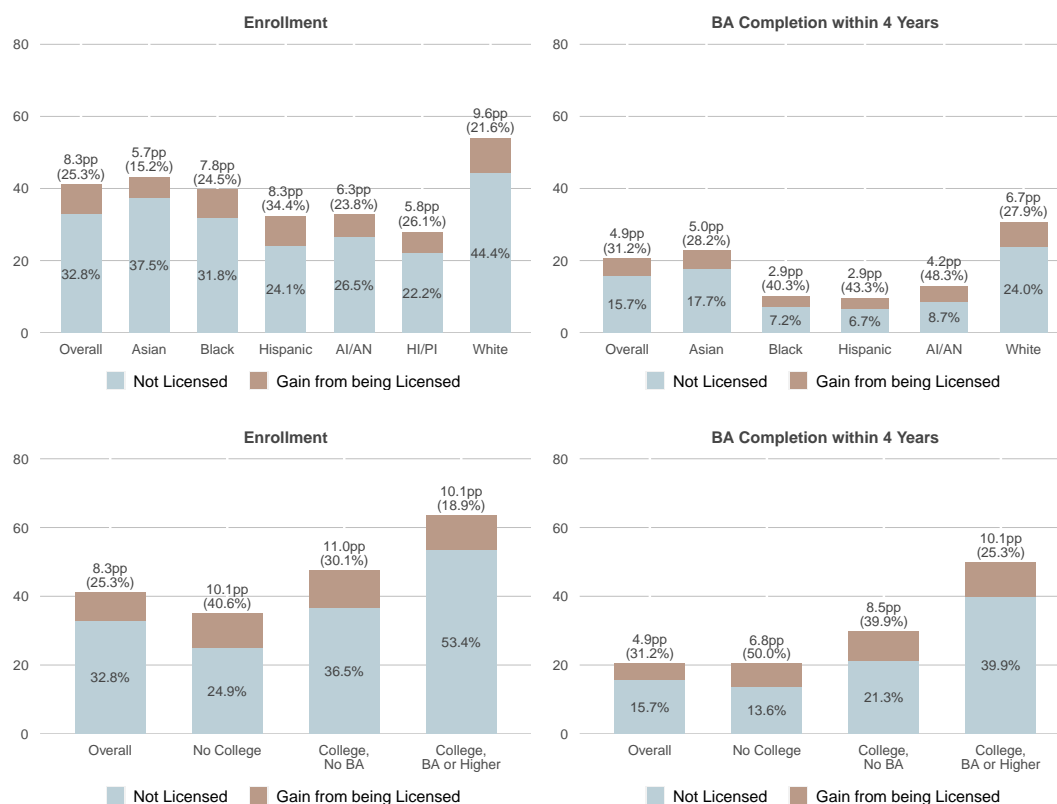
- Howell, J., & Korver-Glenn, E. (2018). Neighborhoods, race, and the twenty-first-century housing appraisal industry. *Sociology of Race and Ethnicity*, 4(4), 473–490. <https://doi.org/10.1177/2332649218755178>
- Jaquette, O., Han, C., & Castaneda, I. (forthcoming). The private school network: Recruiting visits to private high schools by public and private universities. In S. Burd (Ed.), *Lifting the veil on enrollment management: How a powerful industry is limiting social mobility in american higher education*. Book Section, Cambridge, MA: Harvard Education Press.
- Jaquette, O., & Salazar, K. G. (2022). *Student list policy: Problems, regulations, and a solution*. Washington, DC: TICAS. Retrieved from #
- Jaquette, O., Salazar, K. G., & Martin, P. (2022). *The student list business: Primer and market dynamics*. Washington, DC: TICAS. Retrieved from https://ticas.org/wp-content/uploads/2022/09/The-Student-List-Business_-Primer-and-Market-Dynamics.pdf
- Khan, S. R. (2011). *Privilege: The making of an adolescent elite at st. Paul's school* (p. 232). Princeton, N.J.: Princeton University Press.
- 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, Elizabeth. (2022). *Race brokers: Housing markets and racial segregation in 21st century urban america* (pp. pages cm). New York, NY: Oxford University Press.
- McDonough, P. M. (1997). *Choosing colleges: How social class and schools structure opportunity* (pp. xi, 174 p.). Albany: State University of New York Press.
- McPherson, M. S., & Schapiro, M. O. (1998). *The student aid game*. Princeton, NJ: Princeton University Press.
- Miller, B. J., & Skimmyhorn, W. L. (2018). I want you! Expanding college access through targeted recruiting efforts. *Education Finance and Policy*, 13(3), 395–418. https://doi.org/10.1162/edfp_a_00232
- 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>
- 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>
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453. <https://doi.org/10.1126/science.aax2342>
- Park, J. J., & Becks, A. H. (2015). Who benefits from SAT prep?: An examination of high school context and race/ethnicity. *Review of Higher Education*, 39(1), 1–23. Retrieved from <Go to ISI>://WOS:000360777300001
- Posiecznick, A. (2017). *Selling hope and college merit, markets, and recruitment in an unranked school*. Ithaca ; London: Cornell University Press. Retrieved from <https://muse.jhu.edu/book/52616>
- Posselt, J. R., 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.
- Ray, V. (2019). A theory of racialized organizations. *American Sociological Review*, 84(1), 26–53. <https://doi.org/10.1177/0003122418822335>
- Rodriguez, A., & McGuire, K. M. (2019). More casses, mre acess? Understanding the effects of course offerings on back-white gaps in advanced placement course-taking. *Review of Higher Education*, 42(2), 641–679. Retrieved from <Go to ISI>://WOS:000454537100016
- Rothstein, R. (2017). *The color of law: A forgotten history of how our government segregated America*. Liveright Publishing.
- Ruffalo Noel Levitz. (2018). *2018 marketing and student recruitment report of effective practices*. Ruffalo Noel Levitz. Retrieved from http://learn.ruffalonl.com/rs/395-EOG-977/images/RNL_2018_Student_Recruitment_Marketing_Report_EM-19.pdf
- Ruffalo Noel Levitz. (2020). *2020 marketing and recruitment practices for undergraduate students report*. Ruffalo Noel Levitz. Retrieved from https://learn.ruffalonl.com/rs/395-EOG-977/images/2020_Marketing_Recruitment%20Practices_Undergraduate_Students.pdf
- Sadowski, J. (2019). When data is capital: Datafication, accumulation, and extraction. *Big Data & Society*, 6(1), 12. <https://doi.org/10.1177/2053951718820549>
- Sadowski, J. (2020). The internet of landlords: Digital platforms and new mechanisms of rentier capitalism. *Antipode*, 52(2), 19. <https://doi.org/10.1111/anti.12595>

- Salazar, K. G. (2022). Recruitment redlining by public research universities in the los angeles and dallas metropolitan areas. *The Journal of Higher Education*, 1–37. <https://doi.org/10.1080/00221546.2021.2004811>
- Salazar, K. G., Jaquette, O., & Han, C. (2021). Coming soon to a neighborhood near you? Off-campus recruiting by public research universities. *American Educational Research Journal*, 58(6), 1270–1314. <https://doi.org/10.3102/00028312211001810>
- Salazar, K. G., Jaquette, O., & Han, C. (2022). *Geodemographics of student list purchases: A first look*. Washington, DC: TICAS. Retrieved from https://ticas.org/wp-content/uploads/2022/09/Geodemographics-of-Student-List-Purchases_A-First-Look.pdf
- Santelices, M. V., & Wilson, M. (2010). Unfair treatment? The case of freedle, the SAT, and the standardization approach to differential item functioning. *Harvard Educational Review*, 80(1), 106–133.
- Simon, J. (1988). The ideological effects of actuarial practices. *Law & Society Review*, 22(4), 771–800. Journal Article. <https://doi.org/10.2307/3053709>
- Smith, J., Howell, J., & Hurwitz, M. (2021). The impact of college outreach on high schoolers' college choices: Results from over one thousand natural experiments. *Education Finance and Policy*, 1–25. https://doi.org/10.1162/edfp_a_00334
- Stevens, M. L. (2007). *Creating a class: College admissions and the education of elites* (p. 308). Cambridge, MA: Harvard University Press.
- 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
- Waxman, B. (2019). The power of list segmentation part i: Are you already doing this? *intead, Global & Local Academic Branding*. Retrieved from <https://services.intead.com/blog/the-power-of-list-segmentation-part-i-are-you-already-doing-this>

8 Figures

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



Notes: AI/AN = American Indian or Alaska Native. HI/PI = Hawaiian or Pacific Islander. Sample for enrollment outcomes is all SAT takers in the 2015–2018 high school graduation cohorts. Sample for completion outcomes is students in the 2015–2016 cohorts. Results are estimated from regressions that include student-level controls for: sex, race/ethnicity, SAT score, parental education level, last Student Search Service opt-in status, graduation cohort, and high school fixed effects. All differences between licensed versus non-licensed students are statistically significant at the 1% level.

Figure 2: The enrollment funnel



Figure 3: Women in STEM purchases compared to metro, average income and racial composition

