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 that purchases the list. In turn, student lists are a fundamental input for undergraduate recruiting "campaigns," which target individual prospects by mail, email, and on social media.

Recent research suggests that student lists are surprisingly important for college access – and degree completion – for millions of students each year. Jessica Howell, Hurwitz, Mabel, & Smith (2021) compared SAT test-takers who opted into the College Board Student Search Service – allowing accredited institutions to "licence" their contact information – and test-takers who opted out, after controlling for covariates (for a similar analysis of ACT's Educational Opportunity Service see Moore (2017)). Figure 1 reproduces the main results. For students with the same values of SAT score, parental education, race/ethnicity, sex, graduation year, and who attended the same high school, 41.1% of students who participated in Search attended a 4-year college compared to 32.8% of students who opted out, representing an 8.3 percentage point difference and a 25.3 (`=(41.1-32.8)/32.8`) percent change in the relative probability of attending a 4-year college. Participating in Search was associated with a larger percent change in the probability of attending a 4-year institution 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 percent change was also larger for for students whose parents did not attend college (40.6%) than it was for students whose parents had a BA (18.9%).

Despite the importance of student lists for college access, recent reports by *TICAS* argue that student list products systematically exclude underrepresented student populations in two ways [CITE]. The first source of exclusion 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. Second, several "search filters" (e.g., zip code, AP score) used to control which prospect profiles are purchased facilitate the effecient exclusion of students from communities of color and low-income communities.

Prior scholarship assumes that recruiting is something done by individual universities [CITE], motivating policies that incentivize or regulate (e.g., nonresident enrollment caps) university behavior. In reality, the recruiting behavior of universities is structured by products sold by third-party vendors and consultancies in the enrollment management industry (Jaquette, Salazar, & Martin, 2022). The nascent "platform studies in education" literature observes that digital platforms increasingly perform core functions of schools and universities (Komljenovic, 2022a; Nichols & Garcia, 2022). This literature calls for critical, empirical scholarship that informs policy regulations about the role of big tech and edtech in education (Kerssens & Dijck, 2022; Ben Williamson, Gulson, Perrotta, & Witzenberger, 2022). In contrast to the robust scholarship about racial discrimination within the field of critical data studies (e.g., Benjamin, 2019; Noble, 2018), few studies within education research investigate how platforms that structure educational opportunity exclude along the dimensions of race, class, and geography. This study analyzes student list products sold by College Board. We ask, what is the relationship between student list search filters (e.g., test score range, zip code) and the characteristics of students who are included vs. excluded in student lists purchased from College Board?

We develop a conceptual framework about the relationship between search filters and exclusion by drawing from recent scholarship in the sociology of race about algorithmic products. Algorithms are instructions written in code (Burrell & Fourcade, 2021). Student list products are algorithmic selection devices that – similar to Google Ads or Facebook – allow advertisers to control the prospective customers through the use of search filters. Structural racism is "systematic racial bias embedded in the 'normal' functions of laws and social relations" (Tiako, South, & Ray, 2021, p. 1143), whereby processes viewed as neutral or common-sense systematically advantage dominant groups. Structurally racist inputs are determinants of a selection device that are correlated with race because non-white people have been historically excluded from the input. We conceptualize several "geographic" and "academic" search filters as structurally racist inputs. For example, prospects may be filtered by zip code, but zip codes are correlated with race because of residential segregation. Prospects may be filtered by AP test scores, but who attends schools with access to AP curricula? [say geodemographic?]

We address the research question using data from the High School Longitudinal Study (HSLS:09), which follows a nationally representative sample of 9th graders from 2009. We reconstruct the search filters and filter thresholds from the College Board Student Search Service product. We then simulate simulate student list purchases commonly observed and theoretically motivated search filters with the goal of understanding the racial, socioeconomic, and geographic characteristics of students who are included versus excluded from student list purchases. [RESULTS SHOW. . .]

The discussion section addresses policy implications and scholarly contribution. We argue that several search filters satisfy the criteria of "unfair practices" of the Federal Trade Commission (FTC) Act. Additionally, because of the systematic link between student lists and student loans, student list vendors may meet the criteria of Consumer Reporting Agencies, which are regulated by the FTC [TRUE?] and the Consumer Finance Protection Bureau (CFPB).

Our broader contribution is to policy research on college access. Extant research analyzes students, schools, or universities, often in relation to local, state, or federal policies. Although federal higher education policies often focus on for-profit colleges, third-party for-profit vendors now dwarf direct providers. We propose a critical, empirical research program on college access that focuses on organizations and products in the edtech sector. Like student list data, most digital platforms in education are derived from the user-data of students laboring on platforms. What these platforms do with student data is opaque, because obfuscation is a deliberate strategy to avoid regulation (Cottom, 2020). Scholarship from critical data studies shows that structural racism in digital platforms is "a feature, not a bug" because racial exploitation is the defining feature of capitalism (Benjamin, 2019). As courts challenge progressive college access policies like affirmative action, policy research should go on the offensive by applying theory about structural mechanisms to investigate structural racism by third-party products and vendors. Given the narrow scope of the Department of Education and the Higher Education Act (HEA), this research should target the FTC, the CFPB, and other agencies that serve equality of opportunity for consumers.

2 Background: The Student List Business

2.1 Situating Student 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., by sending ACT scores). Applicants consist of inquiries who apply plus "stealth applicants" who do not contact the university before applying. The funnel narrows at each successive stage in order to convey the assumption of "melt" at each stage (e.g., a subset of "inquiries" will apply). Practically, the 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 (e.g., taking a "virtual tours" that records IP address) constitutes the set of all prospects the university has contact information for, who can receive targeted recruiting interventions via mail, email, social media, etc. The majority of BA granting public and private non-profit institutions purchase student lists annually. Based on data provided by university clients, Ruffalo Noel Levitz (2020) reported that 28% of public universities purchased less than 50,000 names, 44% purchased 50,000-100,000 names, 13% purchased 100,000-150,000 names, and 15% purchased more than 150,000 names. 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 (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, student list purchases were the highest source of inquiries, accounting for 26% of inquiries, and accounted for 14% of enrolled students, which ranked fourth after "application as first contact" (19%), campus visit (17%), and off-campus visit (16%). For the median private non-profit university, 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.

2.2 The Market for Student List Data

Although the student list business has historically been dominated by College Board and ACT, in the 21st century student lists have been central to a surprising level of dynamism in the enrollment management industry. Drawing from Jaquette et al. (2022), this section summarizes key dynamics that have have 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 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-up firms entered the student list market by creating college search engines, which asked students to submit information in order to receive recommendations about colleges and scholarships. Another new source of student list data comes from college planning software that is 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, the K-12 information system provider PowerSchool entered the student list business by acquiring the edtech/enrollment management firm Hobsons, which operated the Naviance college planning software and Intersect student recruiting software. 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 recruiting platform.

Third, incumbents College Board and ACT attempted to retain their competitive advantage – through new products and features – amidst the test-optional movement. Both organizations conspicuously embraced data science by developing new search filters based on statistical models that promise to help universities make "efficient" name buys that target "right-fit" students. For example, ACT allows universities to filter prospects based on their predicted probability of enrolling, while College Board developed "geodemographic" search filters that target prospects based on the characteristics/behavior of their high school and their neighborhood. While EAB has become a supplier of names, both College Board and ACT 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 will erode. Several for-profit firms are positioned to acquire market share ceded by College Board and ACT.

3 Literature Review

We position our scholarly contribution as a bridge between two literatures in sociology. First, we review scholarship that engages with the recruiting side of enrollment management (Cottom, 2020; Megan M. Holland, 2019; e.g., Stevens, 2007). Second, we review scholarship from the sociology of race that focuses on the nexus between structural racism and the digital economy (e.g., Cottom, 2020; Hirschman & Bosk, 2020; Norris, 2021) [THIS OPENING PARAGRAPH NEEDS TO BE REVISED]

3.1 Platform Studies in Education

An emerging literature examines digital platforms in education (Kerssens & Dijck, 2022; Komljenovic, 2022a; e.g., Nichols & Garcia, 2022; B. Williamson, 2021), drawing from a broader multidisciplinary, transnational set of literatures around “platform studies” (Benjamin, 2019; e.g., Noble, 2018; Sadowski, 2019, 2020). Digital platforms (e.g., Uber, Coursera, Naviance) are intermediaries for exchange coordinate market transactions and create new markets (Sadowski, 2020). Platforms are also the “ground on which all the user activity happens, allowing the platform to record everything happening in it” (Komljenovic, 2021, p. 322).

One thread of platform studies deconstructs “platform capitalism,” the business models utilized to generate profit from digital platforms. Platform capitalism (often called “rentier capitalism) generates profit by charging customers “rent” – as in the rent a tenant pays a landlord – for the right to use the platform without transferring ownership rights to the customer (Sadowski, 2020). Monetary rent refers to money a customer pays to an organization for access to digital products, for example a university pays annual subscription fees to Elsevier for access to academic journals (Komljenovic, 2021). Data rent refers to "digital traces" that platform users create by interacting with the platform (e.g., personal information they submit, interactions on the platform) (Komljenovic, 2021). Digital platforms gain ownership over user data via terms-of-use agreements. drawing from Marx (1978), Sadowski (2019) develops the concept "data as capital," to describe how platforms monetize user data, which may be used to improve the platform or may become the basis for a new platform.

Another thread of platform studies examines how digital platforms reproduce structural inequality (Benjamin, 2019; e.g., Noble, 2018; O'Neil, 2016). Noble (2018) shows that the results of search algorithms reflect racist ideologies of people on the internet and the profit imperative of adverters that capitalize on these ideologies. Benjamin (2019) develops *race critical code studies* and attendant concepts. "Discriminatory design" is the process embedding structural inequality in platform algorithms, for example, by scoring customers based on an input that people of color have been excluded from. Through "technological determinism," the usage of these biased algorithms affects society, often amplifying the effects of structurally racist inputs. Because targeting by race is profitable, platforms capitalize on residential segregation by using geography as an input to circumvent laws prohibiting race as an input. As Benjamin (2019, p. 147) writes, "racialized zip codes are the output of Jim Crow policies and the input of New Jim Code practices."

Student list products are exemplars of platform capitalism that reproduce structural inequality. Student list data are extracted from the user-data of students laboring on platforms to prepare for college (e.g., taking the SAT) or search for college. Terms-of-use agreements grant platforms ownership over these data. Following Sadowski (2019), College Board monetizes this commodity by licensing names to universities for roughly $0.50 per prospect. New entrants to the market (e.g., EAB, PowerSchool) wrap proprietary databases of prospects within software-as-service products that recruit these prospects (e.g., Intersect, Enroll360), which are then sold to universities for an annual subscription. Student list products reproduce structural inequality because the search filters used to target particular prospects are themselves products of historical inequality in educational opportunity. Examples of problematic search filters include zip code, AP test scores, and whether a prospect indicated an interest in a "peer" university.

The Summer 2022 special issue of *Harvard Educational Review* sets the commitments and direction of the nascent "platform studies in education" literature. Nichols & Garcia (2022) reviews extant scholarship on technology within education research, observing that most scholarship focuses on technical questions about student learning outcomes and instructional practice. Napier & Orrick (2022, p. 207) states that, "platform studies scholars urge us to go beyond pedagogical and technical questions toward social, political, and economic critiques. Consistent with this call, a growing transnational literature examines the economic business models of platform capitalism in the education sector (e.g., Komljenovic, 2022b, 2022a; B. Williamson, 2021). Big tech and edtech companies profit by developing software systems – sold for an annual subscription – that perform core functions of education systems [E.G.,]. Other studies observe that, as education systems outsource core functions, digital platforms exert influence on organizational governance and education policy (Napier & Orrick, 2022; e.g., Ben Williamson et al., 2022).

In contrast to the robust scholarship on discrimination within critical data studies (e.g., Noble, 2018), the nascent literature on platform studies in education does now show how platforms that structure educational opportunity exclude along racial and class dimensions. Additionally, while scholarship argues that digital platforms influence education policy, extant scholarship does not conduct concrete empirical analyses that show how platforms should be regulated. Thus, our analyses contributes to platform studies in education literature, first, by showing how College Board student list products discriminate and, second, by showing how problematic search filters can be regulated by existing federal policy.

3.2 Sociological Scholarship on Recruiting

[NOW THAT THE FIRST PART OF LIT REVIEW IS NOT JUST ABOUT SOC OF ED, IT IS LESS DEFENSIBLE TO ONLY REVIEW SOCIOLOGY OF RECRUITING. YOU MIGHT BE ABLE TO GET AWAY WITH IT BY SAYING THAT MOST RESEARCH ON RECRUITING COMES FROM WITHIN SOCIOLOGY AND WHAT WE SAY ABOUT WHERE THE SOC OF ED LITERATURE FALLS SHORT IS ALSO TRUE OF ECON AND INTERDISICPLINARY RESEARCH IN HIGHER ED] Considering the "enrollment funnel" depicted in Figure 2, scholarship from the sociology of education has focused more on the latter stage of which applicants get admitted (Killgore, 2009; e.g., Posselt, 2016) [and financial aid? check] than earlier "recruiting" stages of identifying prospects, acquiring leads, and soliciting inquiries and applications. However, a growing body of research substantively analyzes recruiting from the perspective of students, high schools, and postsecondary institutions (e.g., Cottom, 2020; Megan M. Holland, 2019; Posecznick, 2017; Salazar, Jaquette, & Han, 2021; Stevens, 2007), often utilizing ethnographic or case-study designs and often 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)][CITE]. First-generation students and underrepresented students of color often reported that "school counselors had low expectations for them and were too quick to suggest that they attend community college" (Megan M. Holland, 2019, p. 97). This trust vacuum created an enrollment opportunity because these students were drawn to colleges that made them feel wanted. Megan M. Holland (2019) found that high school recruiting visits -– including college fairs, instant decision events, and small-group representative visits -– influenced where students applied and where they enrolled, but this finding was strongest for first-generation students and underrepresented students of color. 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. These studies often conceptualize off-campus recruiting visits as an indicator of enrollment priorities and 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 private college that depends on revenue from tuition and donations and is sensitive about acceptance and yield rates. The College valued recruiting visits to local high schools as a means of maintaining relationships with guidance counselors at feeder schools and tended to visit the same set of largely affluent private and public schools year after year. Analyzing the other side of the coin, Khan (2011) shows how private school guidance counselors exploit colleges' desire for trustworthy information about applicant enrollment intentions to get less-qualified students into top colleges. Salazar et al. (2021) analyzed off-campus recruiting visits by public research universities. 12 of the 15 universities made more visits to out-of-state schools than in-state schools, and these out-of-state visits concentrated focused on affluent public and private schools located in predominantly white communities. Salazar (2022) analyzes recruiting visits to Los Angeles and Dallas by out-of-state public research universities, finding that universities engage in "recruitment redlining – the circuitous avoidance of predominantly Black and Latinx communities along recruiting visit paths" [p. X]. Thus, in contrast to branding by about the commitment to racial diversity (M. M. Holland & Ford, 2021), scholarship consistently finds 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., CITE]. Cottom's (2017) analysis of the for-profit sector is simultaneously an ethnography of enrollment management and a work of political economy. For-profits found a niche in Black and Latinx communities because traditional colleges and universities disregarded these communities. They systematically sold low-quality programs to women of color, generating 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)

[MODIFY THIS PARAGRAPH/SECTION SO THAT IT ARGUES THAT SCHOLARSHIP ON RECRUITING HAS NOT TAKEN THE LEAD OF CRITICAL DATA STUDIES/PLATFORM STUDIES IN FOCUSING ON PLATFORMS/PRODUCTS RATHER INSTEAD OF ASSUMING THAT RECRUITING IS BASED ON INDIVIDUAL ORG BEHAVIOR] Upon reflection, scholarship assumes that recruiting is something done by individual colleges and universities. But university enrollment management behaviors are increasingly structured by software and services purchased from third-party vendors. Scholarship on enrollment management must analyze the products being sold to universities and the vendors that create these products. For most universities, student list purchases largely determine which prospective undergraduate students will receive recruiting interventions. Although universities make choices about which names 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, and the targeting behaviors encouraged by the product. Cottom (2020) argues that algorithmic products are not race neutral. Rather, scholarship on the sociology of race shows that algorithmic products reproduce racial inequality by incorporating seemingly neutral inputs that systematically exclude non-white people. Therefore, we review key advances from the sociology of race in order to conceptualize how student list products reproduce racial inequality.

4 Conceptual Framework

4.1 Structural Racism and Algorithms

Structural racism. Contemporary sociology is largely concerned with a structural analysis of race. Bonilla-Silva (1997) argues that most social sciences define racism as an ideology held by individuals (e.g., explicit or implicit racial bias). This approach measures societal racism by examining the attitudes of individuals and excludes the possibility that institutions can be racist. Bonilla-Silva (1997) argues for a focus on the underlying social structure rather than the ideology associated with it. He defines "racialized social systems" as "societies that allocate differential economic, political, social, and even psychological rewards to groups along racial lines" (p. 474). Racial groups are a social construction of a racialized social system. For example, in the U.S. "black people" is a construct that emerges from the slave trade and its legacy. "Racial ideology" – commonly known as racism – is the ideological component of a racialized social system, which includes individual bias and also institutions that that benefit dominant racial groups. Bonilla-Silva (1997, p. 476) argues "that the only way to 'cure' society of racism is by eliminating its systemic roots." As the practices that produce racial inequality become increasingly covert, The sociology of race defines structural racism as a "a form of systematic racial bias embedded in the 'normal' functions of laws and social relations" (Tiako et al., 2021, p. 1143), whereby processes viewed as neutral or common-sense systematically advantage dominant groups and disadvantage marginalized groups.

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 decisions based on a formula.[1] Hirschman & Bosk (2020) argues that actuarial methods can promote racial equity the source of racial inequality is explicit or implicit racial bias of individual decision-makers. In residential real estate, for example, homes in predominantly white neighborhoods received higher appraisal values than those in non-white neighborhoods because appraisers have discretion in selecting similar comparision homes ("comps") for the valuation (Korver-Glenn, 2018).

Structurally racist inputs However, scholarship from the sociology of race argues that actuarial methods reinforce structural racism through two mechanisms. The first is structurally racist inputs. Actuarial products predict future outcomes by modeling the determinants of

[1]Burrell & 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.

the outcome using historical data. Even when these models do not explicitly include race, they often include structurally racist inputs, defined as 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. For example, Obermeyer, Powers, Vogeli, & Mullainathan (2019) found that a commercial algorithm used by hospital systems under-predicted the health care 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. Thus, 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."

Norris (2021) develops the concept "racialized input" by reconstructing Moody's city government credit rating algorithm, which assigns credits scores to cities based on determinants thought to predict the probability of loan default. When cities issue bonds, these credit scores affect bond interest rates. While Moody's does not use percent of residents who are Black as an input, Norris (2021) shows that %Black is negatively associated with city credit rating. This negative relationship persists after controlling for all determinants (e.g., revenue divided by expenditures, total liabilities as a percent of revenue, etc.) with the exception of median household income. Household income is correlated with %Black because of centuries of wage discrimination. Once the model controls for household income, %Black is no longer a significant predictor city credit rating. Thus, Norris (2021) argues that household income is a "racialized input," a structurally racist but seemingly neutral input that masks the structural racism of an algorithm by "explaining away" the relationship between race and the outcome.

Market segments and micro-targeting. A second source of structural racism in actuarial methods is market segmentation, whereby customers are categorized into groups (e.g., "married sophisticates," "rural everlasting") for advertisers. Fourcade & Healy (2013) defines "classification situations" as the use of actuarial techniques by organizations to categorize consumers into different groups. Historically, classifications were binary; consumers with "good" credit were offered loans and consumers with bad credit were not. Advances in data analytics (e.g., individual credit scores) enabled finer classifications, whereby customers are classified into many groups – or along a continuum – alongside the emergence of tiered products targeting different consumer groups with different benefit levels. Fourcade & Healy (2013) (p. 566) offer a quote from a banking trade publication: "Stop trying to lend at low margin to accountants, lawyers and civil servants who are reliable but earn the bank peanuts. Instead, find the customers who used to be turned away; by using modern techniques, in credit scoring and securitization, they can be transformed into profitable business." For example, the "payday" loan industries targets consumers groups that were previously denied credit altogether, but charges excessive interest rates. Cottom (2020) defines "predatory inclusion" as the inclusion of "including marginalized consumer-citizens into ostensibly democratizing mobility schemes on extractive terms." Predatory inclusion is exemplified by the for-profit college industry (Cottom, 2017), but other examples include sub-prime mortgage schemes and the "gig economy" Cottom (2020).

Scholars within critical data studies argues 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. Classification systems developed to optimize profit treat audiences "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).

Related to market segmentation, scholarship from critical data studies also investigates "micro-targeting" approaches to customer identification, which promise to reach granular segments of society with great precision (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. Rather, micro-targeting practices raise concerns about "political redlining," whereby "Campaigns routinely 'redline' the electorate, ignoring individuals they model as unlikely to vote, such as unegistered, uneducated, and poor voters (Kreiss, 2012, p. X)[QUOTED FROM COTTER 2021].

4.2 Mechanisms of Exclusion in Student List Products

Conceptualizing student-list products. Sociologists have studied products that standardize decision-making based on an algorithm (Hirschman & Bosk, 2020; e.g., Norris, 2021). After the algorithm has been created, these products eliminate racial inequality caused by individual decision-making bias, but may cause racial inequality because structural racism is baked into the underlying product (e.g., structurally racist inputs, predatory inclusion). For example, Moody's algorithm assigns credit scores to cities based on analyses of the determinants of default in previous cases (Norris, 2021). By contrast, student list products are similar to purchasing ads from Facebook or Google in that student list products allow universities to choose prospective students by selecting search filters. When purchasing student lists, individual universities may be thoughtful about avoiding structurally racist search filters. However, in contrast to purely standardized algorithms, individual discretion in student list purchases raises the possibility of racial disparities due to individual bias (explicit or implicit) and due to lack of knowledge about these products.

This research focuses on exclusion due to structural inequality embedded in the underlying product. We argue that twp sources of structural exclusion in student list products are (1) who is included in the underlying database and (2) utilizing structurally racist inputs as search filters for selecting prospects from the underlying database. First, students who do not complete a College Board assessment are generally excluded from the underlying database.[2] Test-taking rates differ by race and by class [CITE]. Thus, we expect College Board student list databases to exhibit racial inequality even before search filters are selected. Second, College Board search filters have long utilized search filters (e.g., zip code, AP score) that are structurally racist because communities of color have been historically excluded from these inputs.

In the era of digital platforms, student list products have developed new algorithmic search filters that utilize market segmentation and micro-targeting approaches. 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. In 2021, College Board added three new geodemographic filters (College Board, 2021). More generally, student list products have incorporated new filters designed to increase precision in targeting desired prospects. Examples include ACT's predicted probability of enrollment filter, and College Board's "interest in my peers" filter. To the extent that universities want to target students from affluent schools and communities, student list products facilitate this goal with great efficiency. 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.). Enrollment management consulting firms

[2] footnote: recently added college search engine blah blah

promise precision when marketing list buying services. For example, Ruffalo Noel Levitz states the "RNL Student Search and Engagement" product enables universities to "target the right students in the right markets" by making "the most efficient name purchases using predictive modeling" (Ruffalo Noel Levitz, 2021).

4.2.1 Predicting Exclusion

Salazar, Jaquette, & Han (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). Table X lists selected filters by category. We conceptualize a search filter as a structurally racist input if it is correlated with race due to historic exclusion from the input. This section draws from the sociology of race and related fields to develop expectations [?predictions?] about the relationship between search filters and exclusion, focusing on geographic search filters and academic search filters.

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, zip code); and filters based on geographic borders that College Board constructs using historic data about test-takers (e.g, geomarket, geodemographic filters).

The concepts "space" and "place" from critical geography (Agnew, 2011; Bell, 2007) describe alternative approaches to conceptualizing geographic location, which are useful for developing predictions about the relationship between geographic filters and exclusion. 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).

Marketing conceives of geography as space. Market research exploits racial segregation as a means to identify and target prospective customers (Benjamin, 2019; Noble, 2018). For example, geodemography (now referred to as "spatial big data") 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; Korver-Glenn, 2018; e.g., Rothstein, 2017). Geographic search filters are built on the back of racial segregation. 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.

With respect to filters based on known geographic borders (e.g., county, zip code), we expect that utilizing finer geographic filters is associated with greater racial and socioeconomic dispairities in student list purchases because American residential segregation occurs at fine-grained geographic levels [CITE]. Prior research on recruiting, which 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), implies that some 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.

In addition to filters for known geographic borders, College Board uses data on test-takers from previous admissions cycles to create new geographic borders for the purpose of filtering prospective students. "Geomarkets" divide metropolitan areas into smaller pieces. For example, the San Francisco Bay Area is divided into eight geomarkets, including CA10 the "City of San Jose" and CA11 "Santa Clara County excluding San Jose" [CITE]. Geodemographic segment filters utilize cluster analysis allocate each high school and each neighborhood (census tract) to a market segment based on based on past college-going behaviors of students from that school or neighborhood (College Board, 2011). The resulting classification is highly correlated with race because communities of color that have been historically excluded from higher education are more likely to be lumped together. More generally, we argue that filtering on geographic borders created from past education data is associated with racial exclusion because these borders likely reflect historic disparities in educational opportunity. Further, these filters increase the effects of historic place-based inequality because they enable universities to discriminate between prospects based on previously unknown geographic borders.[BUT CAN'T ANALYZE THESE CUZ WE DON'T HAVE/NEED TO CREATE THE BORDERS]

Academic filters. [KARINA + OZAN REVISE THIS SECTION AFTER SOME PRE-

LIMINARY WORK] 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.

Broadly, the first source of racial and socioeconomic exclusion is that the underlying student list database excludes students who have not taken College Board assessments (SAT, AP, PSAT). Prior research shows that test-taking rates differ by race and class. Thus, we expect that who is included in the underlying database exhibits racial and socioeconomic exclusion prior to conditioning on a particular assessment or a particular score.

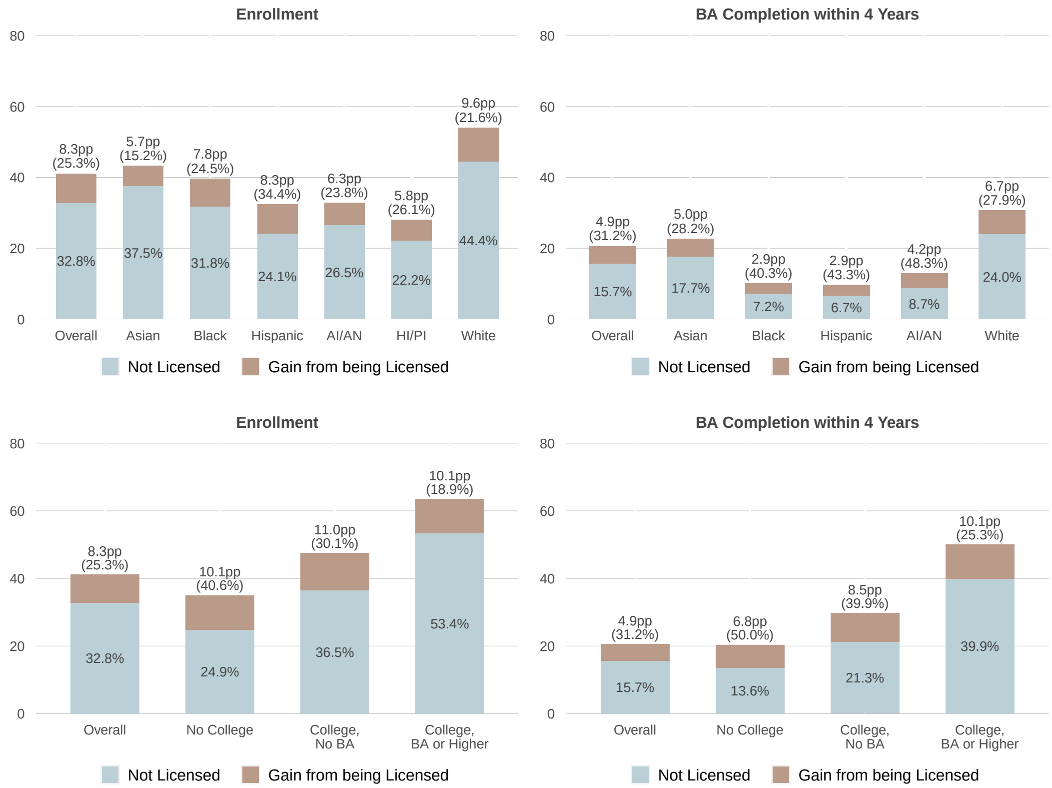
A related source of exclusion within the population of test-takers is who takes which assessment. Students attending schools in affluent communities tend to have better access to AP curricula. Research shows that students from underrpresented populations are less likely to take the PSAT [CITE] or AP tests [TRUE? CITE?] [note that CB did not release AP participation rates by race]. Therefore, conditioning on these assessments may increase exclusion. Additionally, College Board Search enables universities to target high school students early in the search process (e.g., sophomore PSAT takers) but students from underrepresented populations are more likely to take assessments later in high school (e.g., take SAT senior year). We expect that filtering for prospects early in high school is positively associated with racial and socioeconomic exclusion.

A second broad source of exclusion comes from test score thresholds utilized on filters. Test scores differ by race and class as a function of differential access to test preparation and . . . . [READ SOME STUFF AND THEN WRITE; STUFF FROM STRUCTURAL RACISM IN STANDARDIZED TESTING BY TIAKO AND RAY?]. We expect a positive relationship between test score thresholds and racial and socioeconomic exclusion. As an alternative to test scores, universities may filter on high school GPA or high school class rank [RESEARCH BY TIENDA AND FOLKS?]. We expect a weaker relationship between these filters and racial/socioeconomic exclusion.

5 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 reasearch, but rather the reason that policy research must take on this challenge

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.

6 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.

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

Figure 2: The enrollment funnel

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

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

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

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, Megan 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). Legitimating 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., & 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

Kerssens, N., & Dijck, J. van. (2022). Governed by edtech? Valuing pedagogical autonomy in a platform society. *Harvard Educational Review*, *92*(2), 284–303. https://doi.org/10.17763/1943-5045-92.2.284

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

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

Komljenovic, J. (2021). The rise of education rentiers: Digital platforms, digital data and rents. *Learning Media and Technology*, *46*(3), 320–332. https://doi.org/10.1080/17439884.2021.1891422

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

Komljenovic, J. (2022b). Where is value in digital higher education: From commodities to assets. *International Higher Education*, (111), 9–11.

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

Marx, K. (1978). Capital. In R. C. Tucker (Ed.), *The marx-engels reader* (second). New York: W.W. Norton; Company.

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.

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

Napier, A., & Orrick, A. (2022). The economic, social, and political dimensions of platform studies in education. *Harvard Educational Review*, *92*(2), 206–208. https://doi.org/10.17763/1943-5045-92.2.206

Nichols, T. P., & Garcia, A. (2022). Platform studies in education. *Harvard Educational Review*, *92*(2), 209–230. https://doi.org/10.17763/1943-5045-92.2.209

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

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy* (First edition., pp. x, 259 pages). New York: Crown.

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

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

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

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

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

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

Simon, J. (1988). The ideological effects of actuarial practices. *Law & Society Review*, *22*(4), 771–800. Journal Article. https://doi.org/10.2307/3053709

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

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

Williamson, Ben, Gulson, K. N., Perrotta, C., & Witzenberger, K. (2022). Amazon and the new global connective architectures of education governance. *Harvard Educational Review*, *92*(2), 231–256. https://doi.org/10.17763/1943-5045-92.2.231