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 (EAB, 2018), which target individual prospects by mail, email, and on social media.

Research suggests that student lists are 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), 41.1% of students who participated in Search attended a 4-year college compared to 32.8% of students who opted out, an 8.3 percentage point difference and a 25.3 (`(41.1-32.8)/32.8`) percent change in the relative probability. Participating in Search was associated with a larger change in the relative probability of attending a 4-year college for Black students (24.5%) and Latinx student (34.4%) than White students (21.6%), and a larger change for students whose parents did not attend college (40.6%) than those whose parents had a BA (18.9%).

However, a series of *TICAS* reports argue that student list products systematically exclude underrepresented students in two ways (Jaquette & Salazar, 2022; Jaquette, Salazar, & Martin, 2022; Salazar, Jaquette, & Han, 2022). First, student list products sold by College Board and ACT have historically excluded non test-takers, but rates of test-taking differ by race and class (Blake & Langenkamp, 2022; Hyman, 2017). 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 assumes that recruiting is something done by individual universities (e.g., Salazar, Jaquette, & Han, 2021; Stevens, 2007), but university recruiting

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

behavior is structured by third-party products from vendors and consultancies in the enrollment management industry (Jaquette et al., 2022). The nascent "platform studies in education" literature observes that third-party platforms increasingly perform core functions in education and calls for critical scholarship to inform policy regulations about edtech (Nichols & Garcia, 2022). 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 racial composition of students who are included versus excluded in student lists purchased from College Board?

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). Scholarship from critical data studies shows how platforms reinforce racial inequality by embedding structural inequality within platform algorithms (e.g., Benjamin, 2019; Noble, 2018). 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 (e.g., zip code, AP exam score) as structurally racist inputs that reflect historical inequality in educational opportunity. 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 (HSLS: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 inequality in 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. We argue that particular College Board search filters may satisfy the "unfair practices" criteria of the FTC Act. Our broader contribution is to scholarship on education policy. In concert with scholarship on digital platforms in education (e.g., Nichols & Garcia, 2022), we propose a critical policy literature that examines third-party products and vendors in education with the goal of informing regulatory a. As conservative courts challenge progressive policies like affirmative action, policy research should go on the offensive, using theories of structural inequality to investigate structural racism by edtech.

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 – including household savings and financial aid – 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 pursue some mix of broad enrollment goals (e.g., tuition revenue, academic profile, racial diversity), while also meeting the needs of various campus constituencies (e.g., College of Engineering needs majors). 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.

Figure 2 depicts the "enrollment funnel," 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). 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.

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 a survey of their 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]For 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).

2.2 The Market for Student List Data

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. First, enrollment management consulting firms are central to the student list business. 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 in the 2000s followed by concentration in the 2010s. Scholarship on platform capitalism defines data rent as "digital traces" created by users ineracting with the platform (Sadowski, 2019). Student list data are data rent extracted from the user-data of students laboring on platforms (e.g., taking the SAT). Historically, the student list business has been dominated by College Board and ACT. 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). 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.

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 "Enrollment Predictor") based on statistical models. Additionally, both 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. 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. 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

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. We review scholarship on recruiting from sociology, identifying a blind spot shared by scholarship from other disciplines. Scholarship from sociology primarily utilizes ethnographic or case-study designs, and often analyzes recruiting as part of a broader analysis of college access or enrollment management. This literature has analyzed recruiting from the perspective of students, high schools, and postsecondary institutions (e.g., Holland, 2019; McDonough, 1997; Posecznick, 2017; Salazar et al., 2021; Stevens, 2007). For example, Holland's (2019) analysis of pathways from high school to college exemplifies scholarship that engages with recruiting from the perspective of high school students.

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 liberal arts college. 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].

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 universities 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 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 perform core organizational functions (e.g., Sadowski, 2019, 2020) based on algorithms that reinforce racial inequality (e.g., Benjamin, 2019). Second, scholarship on recruiting ignores the enrollment management industry that surrounds 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. Nevertheless, 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. However, because of data limitations – a non-random sample of student list purchases, Salazar et al. (2022) could not determine which filters were driving this exclusion.

This paper advances beyond Salazar et al. (2022) in two ways. First, we develop theoretically motivated propositions about which search filters 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 as filters and filter thresholds are changed. The analyses yield inferences that generalize to student populations of interest to policymakers.

4 Conceptual Framework

We introduce concepts at the nexus of algorithms and structural racism, drawing from the sociology of race and also from critical data studies and critical geography. Next, we apply these concepts to develop propositions about the relationship between student list search filters and racial 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). Actuarial methods – pioneered by the insurance industry – 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, entities 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. Selection devices are "standardization tools designed for making categorical decisions about individuals" (Hirschman & Bosk, 2020, p. 349) "solely on the basis of their assigned scores" (Duncan, Ohlin, Reiss, & Stanton, 1953, p. 572) on one or more selection instruments, such as a test. Examples of selection devices include student list products or an algorithm that uses GPA and SAT score to determine which applicants are accepted.

Hirschman & Bosk (2020) states that actuarial methods can promote racial equity *if* the primary source of inequality is racial bias of individual decision-makers. For example, Korver-Glenn (2018) shows that homes in white neighborhoods received higher appraisal values than those in non-white neighborhoods because appraisers have discretion in selecting comparision homes ("comps") for the valuation. The adoption of actuarial methods across many industries was buoyed by concerns about racial equity following anti-discrimination legislation in the 1970s (Burrell & Fourcade, 2021).

However, the sociology of race argues that actuarial methods do not reduce racial inequality stemming from 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 historically dominated racial/ethnic groups because these groups have been historically excluded from this input (Harcourt, 2015; Hirschman & Bosk, 2020).

Obermeyer et al. (2019) provide an empirical example of structurally racist inputs. An algorithm that hospital systems used to predict patient health care needs under-predicted the needs of Black patients because the algorithm used healthcare costs as a proxy for needs, but Black patients tend to receive less care than others relative to their needs. In another example, Norris (2021) reconstructed Moody's city government credit rating algorithm, which assigns credits scores to cities based on determinants thought to predict loan default. The algorithm does not include the percent of residents who are Black, but does include median household income, which is correlated with percent black because of historic wage discrimination. Once the model includes household income, percent Black is no longer a significant predictor of city credit rating. Thus, household income is a "racialized input," 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.

Algorithmic selection devices often use geography as an input (Benjamin, 2019). Targeting by race can be profitable but is often illegal. Structurally racist geographic inputs capitalize on residential segregation to circumvent laws prohibiting race as an input. Thus, "racialized zip codes are the output of Jim Crow policies and the input of New Jim Code practices" (Benjamin, 2019, p. 147).

The concepts "space" and "place" from critical geography (Agnew, 2011; Bell, 2007) provide insight about structurally racist geographic 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) develops geodemographic segment search filters, stating 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" (p. 1). By contrast, structural analyses of racial segregation conceive of geography as place (e.g., Harris, 1993; Korver-Glenn, 2018; Rothstein, 2017). These analyses view segregation as a consequence of historic and contemporary laws, policies, and practices promoting residential segregation. Algorithmic selection devices that categorize people based on geographic location (space) without considering structures that produce segregation (place) are likely to reproduce historical race-based inequality in opportunity.

Market segments and micro-targeting. Another source of structural racism in algorithmic products 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 granular segments of society (Cotter, 2022).

Sociology conceives of market segmentation as an example of "classification situations." Related to the selection devices concept, Fourcade & Healy (2013) define classification situations 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, classifying customers into many groups, or along a continuum. These classifications are tied to tiered products that targets different consumer groups with different levels of benefits and costs. For example, "payday loans" charge high interest rates to consumer groups that were previously denied credit altogether. Thus, at one end of the continuum, contemporary classification situations produce "predatory inclusion" (Cottom, 2017, 2020). 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 (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). Because classification systems are developed to optimize profit, Cotter et al. (2021, p. 3) argue, "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".

Scholarship on micro-targeting from critical data studies raises similar concerns. In their analysis of Facebook, Cotter et al. (2021, p. 1) state that micro-targeting is "driven not by a goal of making all users available to advertisers, but of making the 'right' individuals available. [Therefore] 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 (Cotter, 2022). Instead, according to Kreiss (2012, pp. 74–75), "campaigns routinely 'redline' the electorate, ignoring individuals they model as unlikely to vote, such as unegistered, uneducated, and poor voters."

4.2 Mechanisms of Exclusion in Student List Products

Conceptualizing student-list products. Student list products are algorithmic selection devices that enable universities to select prospective students from within some database of prospects by choosing search filters.

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 analyses, student lists utilize search filters that can be conceptualized as structurally racist inputs. For example, College Board Student Search Service 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 central to student list products. Consulting firms encourage universities to execute multiple student list purchase, each targeting different market segments (e.g., 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.). Student list products have developed new search filters based on predictive analytics. 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. College Board added three new geodemographic filters in 2021 (College Board, 2021). Other new filters are designed to increase precision, for example 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. Search filters in the College Board Student Search Service can be categorized into the four buckets of academic, geographic, demographic, and student preferences (e.g., desired campus size, intended major). Drawing from theory, we 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).[3] Prior research shows that rates of SAT/ACT test-taking differ by race and class (e.g., Bastedo & Jaquette, 2011; Blake & Langenkamp, 2022; Hyman, 2017). Similarly, the

[3]footnote: recently added college search engine blah PATRICIA - ADD.

percentage of students who take AP exams vary across race, particularly for STEM exams (Kolluri, 2018), and Black students are more likely than white students to attend a high school with few AP course offerings (Rodriguez & McGuire, 2019). These findings motivate the following proposition, which we analyze separately by assessment (SAT, PSAT, AP) and for taking any assessment.

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

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 (Alon, 2009; Alon & Tienda, 2007). 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. 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) 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 geographic search filters include state, CBSA, county, and zip code. Additionally, the geomarket filter and geodemographic filters create new borders based on historical College Board data, but we do not currently have access to these borders.

We conceptualize geographic search filters as structurally racist inputs because these filters are built on top of historic and contemporary policies and practices promoting racial segregation. Targeting prospective students based on geographic location (space) without consideration to the structures that produce racial segregation (place) is likely to reinforce historical race-based inequality in educational opportunity.

Prior research on recruiting consistently finds that selective private and public research universities disproportionately target affluent schools and communities (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 structures of racial segregation often prohibit people of color from living in affluent neighborhoods.

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

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

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

Filtering on multiple filters. 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 4.44 criteria and 98.8% of purchases simultaneously specified at least one academic and one geographic filter. Table 1 shows the top 20 filter combinations. 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 universities and select several orders that utilized common, but potentially problematic filter criteria. seeWe 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 Methods

5.1 Data

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

Our analysis sample includes students who meet all of the following conditions: completed Spring 2012 follow-up 1 survey; completed 2013 update survey; and obtained high school transcript data. Of the 23,503 respondents included in HSLS09, our unweighted analysis sample consists of the 16,525 students who meet all conditions. The survey weight variable `W3W2STUTR` is designed for respondents who meet these conditions. After weighting, these 16,525 students represent the population of approximately 4.187 million U.S. 9th graders in 2009 [STATE 95% CI FOR THIS POINT ESTIMATE?]. The analysis sample is smaller for analyses that utilize variables that have missing values for some respondents. [? SAY MIN ANALYSIS SAMPLE SIZE IS THIS?].

The second data source consists of the "order summaries" of student lists that public universities purchased from College Board. These data are used to inform hypothetical student list purchases in the final set of analyses. As described in Salazar et al. (2022), we collected these data by issuing public records requests to all public universities in five states (CA, IL, TX, MN, and one anonymous state) about student lists purchased from 2016-2020. Salazar et al. (2022) analyzed 804 College Board orders, which yielded XXXX prospect profiles. These orders were placed by 14 public universities. Salazar et al. (2022) shows that some universities placed many more orders and/or purchased many more prospect profiles than others. Figure 3 shows the filters utilized in these orders, separately for research universities and ma/doctoral universities.

5.2 Variables

Our research question is, what is the relationship between student list search filters and the racial composition of students who are included versus excluded from College Board student list purchases? In turn, our dependent variable measures student demographic characteristics and our independent variables are measures of student list filters, which can be utilized in isolation or in combination with one another. Descriptive statistics for analysis variables are shown in Table X [KARINA NEEDS TO CREATE TABLE].

Dependent variable. Following our conceptual framework, this manuscript is primarily concerned with the racial composition of prospects who are included in student list purchases compared to the racial composition of prospects who are excluded. Our primary dependent variable is the student race/ethnicity composite variable `X2RACE`, which includes the following seven categories: American Indian/Alaska Native, non-Hispanic; Asian, non-Hispanic; Black/African-American, non-Hispanic; Hispanic; More than one race, non-Hispanic; Native Hawaiian/Pacific Islander, non-Hispanic; and White, non-Hispanic.[4] We also conducted analyses that utilized parental education (`X2PAREDU`) and family income (`X2FAMINCOME`) as dependent variables, but we exclude these analyses because of manuscript space limits.

Independent variables. Independent variables are measures of student list filters. Choices about independent variables were based on our conceptual framework and the set of student list filters observed in our public records request data collection, shown in Figure 3. Our conceptual framework restricts analytic focus to academic filters and geographic filters, leaving demographic filters and student preferences filters for a future analysis [?EXCEPT INTENDED MAJOR FOR WOMEN IN STEM?].

Propositions **P1** and **P2** focus on academic filters. **P1** is concerned about which students take standardized assessments, which determines inclusion in the underlying College Board student list database. **P2** is concerned with test score thresholds utilized to filter prospects.

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

We create. For **P1**, we create dichotomous measures for each of the following assessments (input variables in parentheses) based on test score variables from the high school transcript file: PSAT/PreACT (`X3TXPSATCOMP`); SAT/ACT (`X3TXSATCOMP`); any AP exam (variables with names that start with `X3TXAP`); and any STEM AP exam. For **P2**, we use these same input variables to create test score measures for PSAT/PreACT; SAT/ACT; highest AP exam score; and highest AP STEM exam score. We also create a measure of high school GPA in academic courses (`X3TGPAACAD`), which is a question asked in the pre-test questionnaire of College Board assessments. Consistent with how College Board filters work, **P2** variables are analyzed as categorical rather than continuous variable. To select thresholds for **P2** variables – for example, an SAT score thresholds of 1000+, 1100+, 1200+, etc. – we considered what the product allows, what we observed in orders collected via public records requests, and the goal of parsimony.

A limitation of measures created for **P1** and **P2** is that HSLS does have separate measures for SAT and ACT. Instead, SAT and ACT test scores are converted to the same scale, but we do not know which students took which assessments. The same is true for PSAT and PreACT assessment.

Propositions **P3** and **P4** focus on geographic filters filters. Drawing from Figure 3, we create measures for student home state (`X2GSTATE`), county (`X2GCNTY`), zip code (`X2GZIPCD`), and CBSA (based on crosswalk with home zip code). Next, we attach income data to localities by merging in data from the American Community Survey (ACS) YYYY 5-year estimates. We do not create independent variables for geomarket filter or geodemograhic segment filter because these filters utilize geographic borders that College Board created using proprietary, historical data about test-takers.

5.3 Analyses

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

Consider a hypothetical purchase that that all prospects took an AP STEM exam. Conceptually, two types of comparisons are possible. Comparison type one compares the proportion of students from a particular race/ethnicity group (e.g. Black) who are included to the proportion of students not from that particular race who are included. For example, using unweighted sample sizes, `91/1,655=5.5%` of Black students took an AP STEM exam and `1,712/14,870=11.5%` of non-Black students took an AP STEM exam. The test for difference in proportions compares the 5.5% of Black students who are included to the 11.5% of non-Black students who are included, and this test is run separately for each race/ethnicity group.

Comparison type two compares the racial composition of the included group to the racial composition of the excluded group. For example, Black students comprise 5.05% (`=91/1803`) of AP STEM test-takers and Black students comprise 10.6% (`=1564/14722`) of students who do not take an AP STEM exam. The test for difference in proportions compares whether the proportion of included prospects who identify are Black differs from the proportion of excluded prospects who identify as Black, and this test is run separately for each race/ethnicity group.

While the significance tests from comparison type one and comparison type two are mathematically equivalent, the two comparisons differ conceptually. The first comparison analyzes the probability of being targeted from the student perspective; that is, do students who identify as Black have a higher/lower probability of being included than students who do not identify as Black? The second comparison focuses on the racial composition targeting from the university perspective; that is, what is the racial composition of prospects who are targeted by a particular set of filters versus the racial composition of prospects who are excluded by these filters? This manuscript focuses on comparison type two because we are interested in how student list products structure the racial composition of university recruiting efforts.

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

5.4 Limitations

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

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

6 Results

6.1 Academic Filters

TEXT

6.2 Geographic Filters

TEXT

6.3 Academic and Geographic Filters

TEXT

7 Discussion

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

Results for proposition **P1** .... [KARINA WRITE HIGH-LEVEL SUMMARY OF RESULTS; ONE OR TWO PARAGRAPHS]

These results have policy implications for federal regulatory agencies concerned with consumer protection and equality of opportunity. Consider zip code filters. Given the history of racial segregation, there is no equality of opportunity rationale for products that enable universities to target particular zip codes. Over the last decade, the Federal Trade Commission (FTC) has become concerned about algorithmic products that "categorize consumers in ways that can result in exclusion of certain populations" (Federal Trade Commission, 2016, p. 9). The FTC enforces the FTC Act, which applies to all organizations engaged in interstate commerce. Section 5 of the FTC Act prohibits "unfair" practices, defined as practices that meet three criteria: (1) causes substantial harm to consumers; (2) harm cannot be reasonably avoided; and (3) harm not outweighed by benefits to other consumers and to competition (FDIC, 2018). Zip code filters may cause substantial harm to consumers (criterion #1) because students who live in nearby non-targeted zip codes are excluded from college access opportunities. Consumers cannot reasonably avoid the injury (criterion #2) because they cannot easily move to a different zip code. The benefit to targeted consumers may not outweigh the harm to excluded consumers (criterion #3).

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

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

In addition to empirical analyses, legal scholarship informs how regulatory agencies interpret the law. For example, Lawler & Dold (2021) argues that the Consumer Financial Protection Bureau (CFPB) has regulatory authority over for-profit colleges because organizations that provide financial advisory services to consumers seeking loans – the activities of any financial aid office – are "covered persons" under the Consumer Financial Protection Act (CFPA). Similarly, legal scholarship can inform how the FTC and the CFPB intepret regulatory authority over student list vendors and products. One issue is whether and how student list vendors can be regulated as "consumer reporting agencies," which are regulated by the Fair Credit Reporting Act and the CFPA. A consumer reporting agency is an entity that sells information about prospective consumers that leads to the extension of credit (i.e., loans) (15 U.S.C §1681a). Student list vendors may qualify as consumer reporting agencies because of the systematic link between student lists and student loans. That is, the first stage of the enrollment funnel is to identify "leads" by purchasing student lists and the last stage is convert admits to enrolled students by offering financial aid packages.

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

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

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

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

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

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

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

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

The payoff for developing this critical policy literature is great, and so is the cost of inaction. Third-party providers do not want to be the object of research because scrutiny from scholars will lead to scrutiny from regulators, which may disrupt profitable practices (Cottom, 2020).

Increasingly, third-party providers perform core functions of schools and universities (Komljenovic, 2022; Nichols & Garcia, 2022). If researchers continues to ignore these products, then policy research will have diminishing influence on core functions of schools and universities. As conservative courts challenge progressive education policies like affirmative action, policy research should go on the offensive by applying theory about structural mechanisms to investigate structural racism by third-party products and vendors.

In higher education, third-party providers now dwarf the for-profit college market. Nevertheless, the Higher Education Act (HEA) – enforced by the US Department of Education – regulates for-profit Title IV institutions, but remains agnostic about "third-party servicers" (aside from lenders and guaranty agencies). For example, responding to concerns about incentive-based compensation for online program management (OPM) companies, the Department of Education (2011) argued that tuition sharing with third-party vendors is not problematic because enrollment goals are determined by the institution, not the vendor.[5]

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

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

8 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. (2009). The evolution of class inequality in higher education: Competition, exclusion, and adaptation. *American Sociological Review*, *74*(5), 731–755. Journal Article. Retrieved from <Go to ISI>://WOS:000270863800003

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

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

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

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.

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

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

Burrell, J., & Fourcade, M. (2021). The society of algorithms. *Annual Review of Sociology*, *47*, 213–237. Journal Article. https://doi.org/10.1146/annurev-soc-090820-020800

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Hirschman, D., & Bosk, E. A. (2020). Standardizing biases: Selection devices and the quantification of race. *Sociology of Race and Ethnicity*, *6*(3), 348–364. https://doi.org/10.1177/2332649219844797

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

Howell, J., Hurwitz, M. H., Mabel, Z., & Smith, J. (2021). *Participation in student search service is associated with higher college enrollment and completion*. College Board. Retrieved from https://cbsearch.collegeboard.org/pdf/college-outreach-and-student-outcomes.pdf

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

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

Jaquette, O., 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

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

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

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

Kreiss, D. (2012). Yes we can (profile you): A brief primer on campaigns and political data. *Stanford Law Review, Online*, *64*(February 2), 70–74. Retrieved from https://review.law.stanford.edu/wp-content/uploads/sites/3/2016/08/64-SLRO-70_0.pdf

Kuhn, T. S. (1962). *The structure of scientific revolutions* (pp. xv, 172 p.). Chicago: University of Chicago Press. Retrieved from file:///C:/Documents%20and%20Settings/Ozan%20Jaquette/My%20Documents/academic%20papers/philosophy-higher-education/kuhn-structure-scientific-revolutions.doc

Lawler, M., & Dold, M. (2021). For-profit schools as covered persons under the CFPA. *Berkeley Law, Center for Consumer Law & Economic Justice*. Retrieved from https://protectborrowers.org/wp-content/uploads/2021/03/SBPC_UCB_For-Profits.pdf

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.

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

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

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

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., 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

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

Rodriguez, A., & 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

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

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.

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

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

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

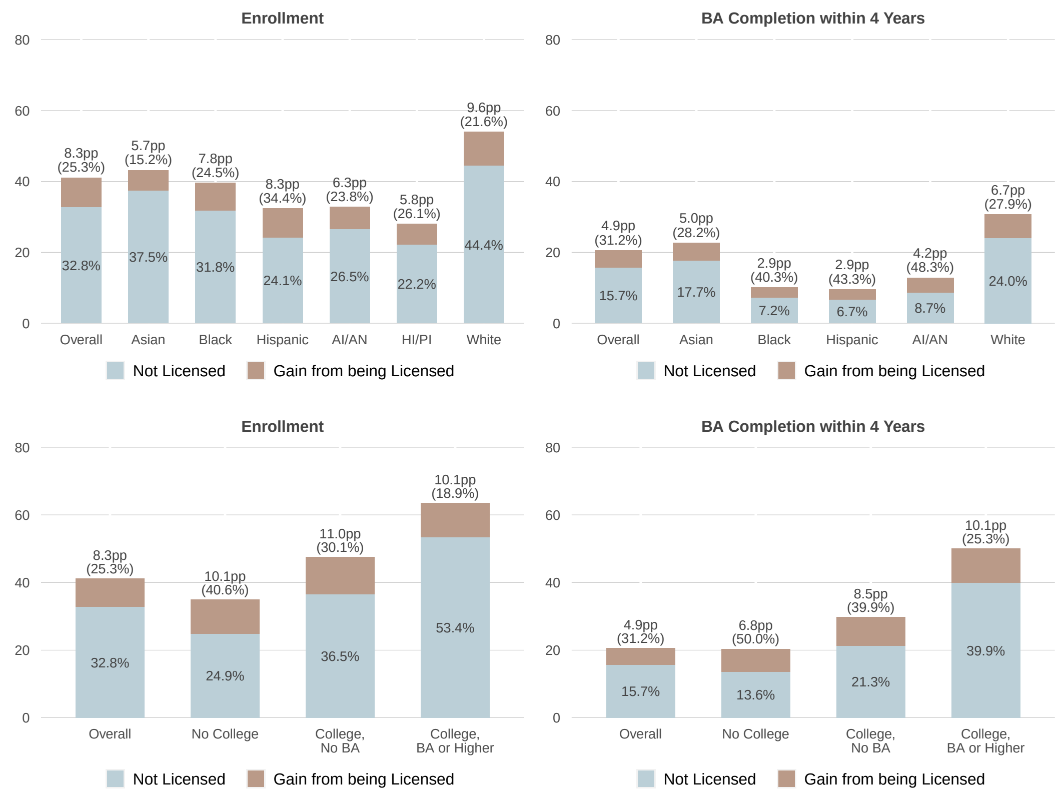
Waxman, B. (2019). The power of list segmentation part i: Are you already doing this? *intead, Global & Local Academic Branding*. Retrieved from https://services.intead.com/blog/the-power-of-list-segmentation-part-i-are-you-already-doing-this

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

9 Figures

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



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

Figure 2: The enrollment funnel

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

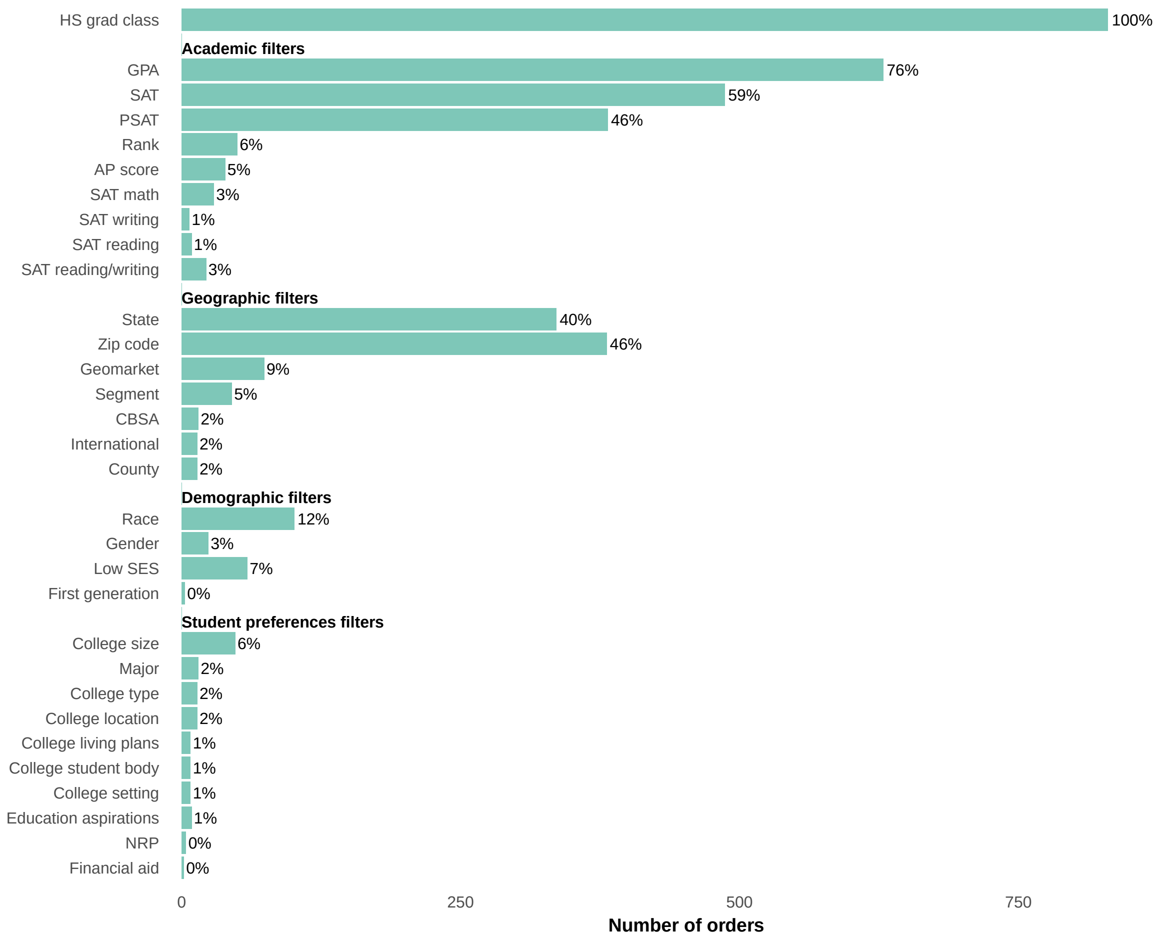


Table 1: Top filter combinations used in College Board orders purchased purchased by 14 public universities

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