Structuring College Choice: The Curious, Mundane Case of College Board Geomarkets

Ozan Jaquette (UCLA)   Karina Salazar (University of Arizona)

1 Introduction

Quarto enables you to weave together content and executable code into a finished document. To learn more about Quarto see https://quarto.org.

the bal bla Reber & Smith (2023)

2 Literature Review

Scholarship on college access can be categorized as focusing on demand-side versus supply-side factors. Broadly, economics defines the demand-side as factors that affect consumer demand for a product (consumer income, population trends, prices, consumer tastes) whereas the supply-side refers to factors that affect production of a product (e.g., cost of inputs and other production costs, technological advancement, government regulation). The majority of research on college access focuses on demand-side factors. In economics, demand-side studies include how students respond to tuition and financial aid, student academic preparation (e.g., racial and socioeconomic "test score gaps"), and information interventions. In sociology, demand-side topics include analyses of college-going culture [cite Bourdieu, McDonough]

3 Background

According to its bylaws,

> The College Board is a not-for-profit membership corporation operating for educational purposes …Members of the College Board are secondary and postsecondary institutions, districts and systems, and nonprofit educational associations, organizations, and agencies serving secondary and/or postsecondary education

Historically, College Board partnered with Members to develop projects that later became revenue-generating products for College Board. For example, in the 1970s, Carleton College became a leader in applying market research to student recruiting under the direction of Larry Litten, then director of the Office of Institutional Research. In the late 1970s Carleton College partnered with College Board to develop the "Carleton College-College Board Six-Market Study" about the college choice decisions of households, with College Board providing funding and data for the project Litten, Sullivan, & Brodigan (1983) [and Litten became Associate Director of the Consortium of Financing Higher Education (COFHE). This research project was the basis for the College Board Admitted Student Questionnaire (ASQ), a revenue-generating product that provides colleges with information about admitted students – those who enrolled and who did not enroll. Similarly, in 1979 The College Board partnered with Robert Zemsky, a professor at UPenn, to develop the *Comprehensive Undergraduate Enrollment Planning Project* (Zemsky & Oedel, 1983). This project was the basis for the College Board Enrollment Planning Service (EPS), the market research and enrollment consulting division [diff word?] of College Board.

3.1 The Market Segment Model

In 1983, The College Board published *The Structure of College Choice* by Robert Zemsky and Penney Oedel. This book, based on the *Comprehensive Undergraduate Enrollment Planning Project* develops the Market Segment model, which introduced "geo-markets" and became the basis for EPS.

Zemsky & Oedel (1983) describes the Market Segment model as an effort to develop a systematic, quantitative description of which institutions are likely to be considered by which institutions, based on the shared stories and wisdom of college admissions officers. Zemsky & Oedel (1983, p. 9) write that,

> Admissions officers invariably are tellers of stories – about the colleges they represent, about the colleges they attended, about each other, and about the often vagabond life of college recruiting .... We have begun with this celebtration of storytelling for two reasons. First, we believe that the institutions of admissions officers actually comprise a remarkably systematic body of knowledge about the college selection process. Second, much of that knowledge has been captured in the admissions officers' tales, which exhibit all the hallmarks of a classic folklore: language, symbol, and system. Our research thus is based on listening carefully to what admissions officers have to say."
>
> From the beginning, we have sought to mold our research to these concepts, to capture and quantify the phenomenon behind the folklore. Thus, it is the admissions officers' notion of admissions pools that dictates the geographic units of our analysis." [p. 11]

As such, the first step in developing the Market Segment model was to identify p. 11 "three types of boundaries – region, state, and community" – as follows: first, regions were defined as New Egland, Middle States, and the South; Second, states are states; and, third, each state was divided into "as few as two and as many as thirty community-based enrollment markets or pools, for a total of 143 separate markets" [in New England, Middle States, and the South]. These community-based enrollment markets – later called "geomarkets" by The College Board Enrollment Planning Service – are the analytic focus of this manuscript. Enrollment markets are intended to be consistent with the conception of a catchment market from the perspective of admissions counselors. Zemsky & Oedel (1983) write that,

> In many cases, the market boundaries match formal political and educational divisions, reflecting natural channels of communication. Each major metropolitan area is composed of several markets, usually corresponding to the inner city, a first ring of suburbs, and an outer ring of suburbs. In more sparsely populated areas, communities are sometimes combined in order to make the analysis meaningful."

Having defined these geographic boundaries, each SAT score a student submits to an institution can be defined as "local" (institution is in the same local market that the student lives in), "in-state" (institution is located in different local market but the same state the that the student lives in), "regional" (institution is located in a different state but the same regional market that the student lives in), or "national" (institution is located in a different regional market than the student lives in).

Next, high school students are categorized into found different *segments* – local, in-state, regional, and national – based on their score-sending behavior. A "local" student submits more SAT scores to local institutions than they do to in-state, regional, or national institutions. An "in-state" student submits more SAT scores to in-state institutions than they do to local, regional, or national institutions, etc.

The Market Segment Profile. The two primary outputs of the Market Segment Model are the (1) Market Segment Profile and (2) the Institutional Profile. Zemsky & Oedel (1983) state that "the Market Segment Profile is each market's signature, a detailing of student choice and institutional standing within each segment of that market. In a single-page format, this information is presented for each of the 143 markets currently included in the Comprehensive Undergraduate Enrollment Planning Project." Table 1 reproduces a partial, simplified version of the bottom panel of Zemsky & Oedel (1983, fig. 2.1) which represents the Market Segment Profile for "Connecticut Market 3: Fairfield County."

Table 1 has separate columns for each market segment (local, in-state, regional, national) and rows show the number of test-takers and characteristics of test-takers. For example, there were 550 "local" students in Fairfield county and these students submitted SAT scores to 2.8 postsecondary institutions on average. By contrast, there were 1,664 "regional" students who submitted test scores to 4.8 institutions on average. For local students, 11.2% had family income greater than $35,000 and 9% had both parents with a BA. For regional students, 41.9% had family income greater than $35,000 and 34.0% had both parents with a BA. Each Market Segment Profile also present information about the institutions that students from each segment sent scores to.[1]

INSERT Table 1 ABOUT HERE

The Institutional Profile. For each postsecondary institution, the Market Segment Model produces several Institutional Profiles, which describe the students who sent scores to the institution and which majors they are interested in. Institutional profiles are created separately for students from a particular local market, for all students in a state, or all students in a region. Table 2 reproduces a partial, simplified version of Zemsky & Oedel (1983, fig. 2.3), which is the institutional profile of an anonymous institution for students from Fairfield County, CT. Table 2 shows that 58 students submitted SAT scores to the institution. These 58 students represented 4.8% of the total 1,199 in-state students from Fairfield County. 69 regional students sent scores to the institution, representing 4.1% of all 1,664 regional test-takers. Of these 69 regional students, 35 expressed interest in majoring in the liberal arts. These 35 students represent 7.0% of all regional students from Fairfield County who expressed interest in the liberal arts.

INSERT Table 2 ABOUT HERE

[1] Table 3 recreates the top panel of Zemsky & Oedel (1983, fig. 2.1) for the Fairfield County, CT local market. For example, of the 550 local students, 277 (50.4%) sent scores to institution #1, which was a private master's granting institution. Of the 1,199 in-state students, 757 (63.1) sent scores to institution #1, a public doctoral granting institution, and 515 (43.0%) sent scores to institution #2, a public master's granting institution.

Zemsky & Oedel (1983) argue that the Institutional Profile and the Market Segment Profile enable admissions officers to know where to look for students and which institutions are competing for those students (p. 25):

> The Institutional Profile and the Market Segment Profile quantify the admission officers' intuitive grasp of market structure. Structure here carries a dual meaning, connoting both the structure of student choice and the structure of institutional competition…This two-sided interpretation furnishes the essential framework for planning by individual colleges and universities…To draw effectively on its own natural constituency, a college not only must contact the "right" kind of students — that is, students who are predisposed toward that type of institution – but also must persuade them of its special character. This means knowing the competition as well as the clientele.

Enrollment Planning Service. The College Board Enrollment Planning Service (EPS) was created in 1984 to provide institution-specific market research that would inform the recruiting strategy of member institutions (College Board, 2012).

Just as Litten et al. (1983) and the "six market area study" – underwritten by College Board – was basis for the College Board Admitted Student Questtionnaire product, Zemsky & Oedel (1983) and the the *Comprehensive Undergraduate Enrollment PlanningProject* – underwritten by College Board – was the basis for EPS (Takamiya, 2005). Although information is spotty, the original EPS product seemed to be based fundamentally on the Market Segment Model developed by Zemsky & Oedel (1983). Each institution purchasing EPS information services would obtain access to the Market Segment Report for each local market. They also received the Institutional report for each state and for each local market. Collectively, these reports enabled institutions to determine where attractive students were located, which students were considering – or might consider – the institution, and which institutions were likely to be in competition for these students.

The appendix of Zemsky & Oedel (1983) maps the local markets for New England, Middle States, and the South, for a total of 143 local markets. At some point, College Board EPS created geo-markets for the remaining US states. College Board (2023) shows the College Board geo-markets circa 2023. Incredibly, the geomarkets for New England, Middle States, and the South remain [?ALMOST?] exactly the same.

WHAT ELSE TO SAY ABOUT ENROLLMENT PLANNING SERVICE???

PX: think about how geomarkets are used More generally as part of recruitment planning by EPS NEED TO GET SOME CONCRETE INFORMATION ABOUT HOW GEOMARKETS ARE USED In student list purchases, geomarkets are used as a filter alongside other filters

/https://secure-media.collegeboard.org/homeOrg/content/pdf/welcome_to_the_college_board_fullversi
This document is from 2012 "Enrollment Planning Service (EPS®) is an online database using a wide range of criteria, including geography, high schools, score-sending patterns and international regions. EPS allows staff to analyze data and find prospective students both nationally and internationally who are likely to succeed at their institution. This allows for better and more targeted recruiting and communications based on a wide range of criteria, including geography, demographics, academic preparation and collegiate aspirations. EPS also provides comprehensive reporting and fully integrates with Student Search Service.

This is interesting https://www.pdffiller.com/28049652-fillable-college-board-enrollment-planning-service-form?mode=view

EPS "is the analysis tool that pinpoints the schools and geomarkets where your best prospects are most likely to be found. You can research your existing (or new) markets using selection criteria and locate your top prospects in various ways – by states, geomarkets, counties, Zip codes, high schools and international regions. EPS provides you with comprehensive reports on your markets, your position in those markets, and your copetition."

Market research by college board https://www.tandfonline.com/doi/pdf/10.1080/00091383.1984.10570121 https://eric.ed.gov/?id=ED238341 Druesne, B., Harvey, J. & Zavada, M. (Summer 1980). College mailings: What works. The College Board Review, 116,12-17.

ENROLLMENT PLANNING SERVICE From https://www.jstor.org/stable/pdf/40195699.pdf [1986 RIHE] Other Erollment Planning SerVIce EPS and Descriptor PLUS are data serVIces that make use of College Board infornMation to provide Institutions with intelligence on trends and conditions that directly affect their enrollment management strategy

Student Search Service.

SAY BASIC SHIT ABOUT STUDENT SEARCH SERVICE. COME BACK TO THIS LATER.

HOW GEOMARKETS ARE USED IN STUDENT LIST PRODUCTS

4 Conceptual Framework

We develop a conceptual framework by drawing from *Discriminating Data: Correlation, Neighborhoods, and the New Politics of Recognition* by Wendy Hui Krong Chun (2021). *Discriminating Data* investigates how the correlation and assumptions about homophily have been utilized to create a future that amplifies the power of past discrimination. In particular, we draw from chapter one, "Correlating Eugenics," and chapter two "Homophily, or the Swarming of the Segregated Neighborhood."

4.1 Correlations

Correlation measures the extent to which two or more variables move together. According to the Pearson correlation coefficient, if an observation x_i greater than the mean $\bar{x}$ is associated with an observation y_i greater than the mean $\bar{y}$ then variables x and y are positively correlated.

Chun (2021) traces the origins of correlation to 19th century statisticions and "biometric eugenicists" Francis Galton and and his student Karl Pearson. These scholars developed the concepts of correlation and linear regression while conducting research on plants, animals, and peoples about the relationships between breeding and genetic inheritence. Biometric eugenicists viewed selective breeding as the most important tool for human progress. Eugenics, Pearson stated is "the science of improving stock, not only by judicious mating, but by all the influences which gave the more suitable strains a better chance" [quoted in Chun page 59]. Eugenicists believed that efforts to improve conditions – for plants, animals, people – would only improve performance for a single generation whereas selective breeding would improve performance for multiple generations. Therefore, Pearson stated, "Give educational facilities to all, limit the hours of labour to eight-a-day…give a minimum wage with free medical advice, and yet you will find that the unemployables, the degenerates and the physical and mental weaklings increase rather than decrease" [quoted from Chun page 65]. The key to human progress, is to keep unfit bloodlines separate from pure bloodlines and to stop unpure from breeding.

The predictive analytics embedded in most algorithmic products are based on correlation. The creation of predictions proceeds in two steps: Using "training" data, apply statistical models to previous cases to determine the predictors of an outcome; second, apply the results of these analyses to predict the outcome for future cases [give brief explanation of training data and then predict for subsequent datasets, can draw from your EEPA chapter] Correlations can be used to predict outcomes without requiring knowledge about causal relationships. Chun (2021) provides the example of Kosinski, Stillwell, & Graepel (2013), who develop a method to predict sensitive personal attributes based on Facebook Likes. Their approach was subsequently utilized by political consulting firm Cambridge Analytica to identify and target subgroups that could be swayed to vote for their clients, such as Ted Cruz, who were seeking elected office.

In Zemsky & Oedel (1983), chapter 3 "A Sense of Place: Students, Families, and Communities" is fundamentally about correlations. Having categorized each high school student as local, in-state, regional, or national based on their score-sending behavior, chapter 3 analyzes which social and economic characteristics are associated with these behaviors and the extent to which these relationships hold across states and local markets. The data analyzed by Zemsky & Oedel (1983) can be seen as training data. In actual recruiting applications, colleges would have historic data on the number of students in each segment in a local market, but they would not know a current high school student is a local, in-state, regional, or national prospect. Therefore, knowing which observable student characteristics are associated with score sending outcomes can be useful for knowing which prospects an institution should target. [NOTE: THEY WOULD ALSO KNOW THE NUMBER OF STUDENTS IN EACH SEGMENT FRONM PAST COHORTS; CUZ THIS IS IN THE MARKET SEGMENT REPORT]

Zemsky & Oedel (1983) identify four variables that can be used individually or in combination to predict score sending behavior. "These four attributes – educational aspirations, parental education, scholastic aptitude, and family income – reflect the basic social patterns of the nation. It would have been surprising if these were not the four social variables that best explained the patterns of college choice" (p. 33). Commenting on parental education, Zemsky & Oedel (1983, p. 33) state that, "these data allow us to say with considerable confidence that local and in-state students are not likely to come from families in which both parents have received college educations." Commenting on family income, they write that we could predict that all local students would come from moderate-income or low-income families and be wrong only 5.5 percent of the time" (p. 33). Further analyses reveal that the SAT score is the most important predictor, followed by parental education, family income, and educational aspirations.

Many scholars of critical data studies have shown the negative consequences of using predictive analytics to make decisions [CITE SAFIYA, AMONGST OTHERS]. Proxy variables are variables that are highly correlated to a variable of interest (e.g., EXAMPLE). O'Neil (2016) states that "weapons of math destruction" use proxy variables to make predictions in the absence of desired data. For example, the COMPAS algorithm for predicting re-arrest utilizes home zip-code as an input, which results in higher predicted probabilities of re-arrest for Black people because of racial segregation in housing. Weapons of math destruction often exclude race as an input and use inputs that are correlated with race, while simultaneously ignoring historical structural racism that produced the correlation in the first place. For example, Norris [CITE] shows that Moody's algorithm for assigning city credit scores uses median income as an input, but not racial composition. Given historic wage discrimination, the resulting system penalizes cities with large Black populations while appearing to be race-neutral [or colorblind?]. Another consequence of using historic correlations to make future decisions is that this amplifies the effects of past discrimination that produced these correlations [some redundancy in this paragraph]. Thus, "Correlations, do not simply predict actions; they also form them" (Zemsky & Oedel, 1983, p. 58).

We argue that the designation and usage of local markets in the Market Segment Model exemplify the negative consequences of correlations. Zemsky & Oedel (1983, p. 44) describes a hypothetical college that wants to target regional and national students in New England and is considering whether to expend limited recruiting efforts in the Boston geomarket – which in 1980, was relatively low-income – or in more affluent geomarkets nearby:

> Where would you concentrate your energies? Ideally, you would seek communities with a high proportion of students already predisposed toward institutions such as your own. The Market Segment Model would provide this information through segment percentages for the community in question. Further classification of students by social attributes allows you to identify a group for mailings or recruiting…if you were to recruit in Boston, only about two out of every ten

> students with fewer than two attributes would likely listen, while slightly less than half of the students with two or more attributes would be receptive.... Your efforts would surely be better directed toward three of the four communities in the bottom band, Manchester, Hartford, and Fairfield County. Simply by knowing a little bit about the students' backgrounds and academic records you could quickly focus your attention on those most likely to consider your kind of institution. Indeed, in Fairfield County alone you could reach more than 40 percent of your "primary target" population – that is, students with a greater than 75-percent probability of concentrating their college choices among institutions like your own."

In this example, historical data on the number of students by Segment is used to divert recruiting energies from lower-income to higher-income local markets. It also shows how geomarkets can be utilized in concert and within student list products. First, decide which geomarkets to concentrate on, then filter students in those geomarkets based on socioeconomic and academic characteristics.

SOMEWHERE: THIS QUOTE FROM CHUN PAGE 57 IS PRETTY SOLID; BUT MAYBE YOU ALREADY HAVE TOO MANY QUOTES?

> As this example makes clear, these models not only "discover" the effects of discrimination; they also automate and perpetuate them for they exploit, rather than remedy inequalities. These correlations are at the heart of what communications scholar Oscar Gandy, writing in 2009..... "identified as"technologies of rational discrimination": unless there is a clear determination not to discriminate, Gandy explained, these technologies perpetuate inequality by creating and comparing "analytically generated groups in terms of their expected value or risk." That homophily drives these groups and correlations "that work" is no accident.

4.2 Homophily and Social Networks

In network science, actors (nodes) are connected to one another via network ties. The first stage in social network analyis is to decide who are the actors and what are the ties that link them. For example, in Zemsky & Oedel (1983) we can think of students sending SAT scores to colleges as a "two-mode" social network in which students (node 1) send SAT scores – the network tie – to colleges (node 2). The second, theorization, stage of social network analysis makes makes statements about which actors are likely to form direct or indirect ties with one another. Chun (2021, p. 90) observes that this stage "builds models that reproduce the abstractions produced in the first. Whatever repeats the initial mapping is true or causal: truth within these mathematical or logical systems as Arendt points out is consistency." For this reason, Chun (2021) argues that network science is biased in the direction of its theoretical models and its assumptions – which are previously proven theoretical models – in that behavior consistent with the model is defined as "truth" and behavior inconsistent with the model is cast away as "noise."

Homophily has become a core assumption of network science (McPherson, Smith-Lovin, & Cook, 2001). Homophily means that actors who share common characteristics are likely to form connections with one another, an idea captured by the phrases "birds of a feather flock together" and "similarity breeds connection" (McPherson et al., 2001, p. 415). By contrast, heterophily is the idea that actors with different characteristics are likely to form connections with one another. According to the review by (McPherson et al., 2001, p. 417), "the classic citation [on homophily] in the sociological literature" was the Lazarsfeld & Merton (1954) study of friendship formation in Hilltown and Craftown, two mixed-race housing developments. As survey methods and social network methods advanced in the 1970s and 1980s, evidence in favor of homophily proliferated (McPherson et al., 2001), to the point where homophily became an assumption rather than a finding (Chun, 2021).

Chun (2021) problematizes the idea that homophily is a naturally occurring phenomenon.

In commercial social networks, such as Facebook, Twitter/X, and TikTok, homophily is programmed into the algorithms that create connections between users. Thus, "echo chambers are not unfortunate errors, but deliberate goals" because "homophily is used to create agitated clusters of individuals whose angry similarity and overwhelming attrication to their common object of hatred both repel them from one another and glue them together" [p. 82].

Chun (2021) also problematizes academic scholarship about homophily in two ways. First, network science defines behavior consistent with the theoretical model to be "truth" and behavior inconsistent with the theoretical model to be noise. Chun (2021) re-analyzed data about friendship patterns in Hilltown and Craftown and found substantial evidence supporting heterophily in the Black residents, but Lazarsfeld & Merton (1954) focus on friendship formation for white residents. Second, drawing heavily from rational choice theory, network science treats homophily as the result of "voluntary" actions by individuals; "it erases historical contingencies, institutional discrimination, and economic realities" (Chun, 2021, p. 95) that underlie behavior consistent with homophily.

In Zemsky & Oedel (1983), Chapter 4 – "The Company We Keep: Colleges and Their Competition" – conducts a social network analysis to identify which institutions are in competition with one another. A network tie between two institutions occur if "15 percent of the students who sent SAT scores to the first institution also sent a score to the second that 15 percent of students who sent an SAT score to the second institution also sent one to the first" (Zemsky & Oedel, 1983). Next, the authors develop "Tinker Toy" diagrams that show which institutions are connected to one another. These diagrams are drawn separately for each student segment – local, in-state, regional, and national – and separately for each geomarket, such that the analyses convey which institutions compete with one another for which student segments in each local market.

Homophily is the core finding of these network analyses. Zemsky & Oedel (1983, p. 42) state that "On the basis of this analysis, we draw a fundamental conclusion about the structure of college choice: collegiate competition occurs principally between like institutions."[2] Subsequent analyses investigate competitive overlap across institutional type – e.g., private flagship, private selective college, public flagship – overlaying tuition price and student student socioeconomic status (SES). The analyses reveal substantial overlap between private selective colleges and private flagship universities, charge the highest prices and enroll students with much higher SES than other institutional types. In conclusion, Zemsky & Oedel (1983, p. 72) states,

> We now know why students so seldom speak of their own social or family backgrounds in explaining how they go about choosing a college. They have no need to. Students describe themselves socially simply by telling us the colleges and universities in which they are interested. The layering of collegiate competition is primarily a socioeconomic layering. The hierarchical structure of collegiate competition largely reflects the stratified social and economic dimensions of the communities from which colleges draw their students. Competition among colleges, as admissions officers have told us for so long, is in fact, a matter of keeping company with one's peers.

Chun (2021) discusses the negative consequences arising from homophily. First, when we assume homophily is the naturally occurring behavior of individuals rather than a consequence of societal structures, then we ignore the structures that give rise to the behaviors we observe. Second, when we "engineer homophily" – through algorithms or policies – we make it less likely for heterophily to emerge, eliminating possibilities – of friendships, of matches – that may have occurred in the absence of such engineering.

The Market Segment Model manifests both negative consequences of homophily. First, Zem-

[2]For example, describing the Figure 4.4 "Structure of Fairfield County Regional Market," (Zemsky & Oedel, 1983, p. 54) state that "competitive overlap, moreover, is often confined to institutions belonging to the same [Carnegie] type as well sector. For example, public flagships compete primarily with other public flagships; private standard colleges, with other private standard colleges; Catholic institutions, with other Catholic institutions."

sky & Oedel (1983) observes hierarchical hierarchical matching between the socioeconomic status of students and the socioeconomic status of colleges, but interprets these findings as naturally occurring phenomena rather than the consequences of historic structural inequality in the opportunity structure of the U.S. Second, the Market Segment Model not only describes the structure of college choice, it influences the future structure of college choice by becoming the basis for products and practices. The Market Segment Model is incorporated into the College Board Enrollment Planning Service (EPS) and Student Search Service. This product amplifies the effect of historic structural inequalities on future opportunity structure by encouraging colleges to pursue local markets – and student segments within these local markets – in ways that are consistent with the patterns observed by Zemsky & Oedel (1983).

4.3 Geomarkets and Structural Inequality

Geomarkets and socioeconomic inequality. We argue that geomarkets – and the Market Segment model more broadly – are built on structural inequality, exhibit structural inequality, and are used in ways that amplify structural inequality.

The Market Segment Model is explicitly a function of socioeconomic status. Zemsky & Oedel (1983) state that socioeconomic status drives whether students belong to the local, in-state, regional, or national segment. In turn, geomarkets are described by the relative numbers of local, in-state, regional, or national students, which is substantially a function of socioeconomic status. Colleges prioritize geomarkets depending on the number of students from desired students. When purchasing student lists, Zemsky & Oedel (1983) suggest that affluent coleges should consider targeting affluent geomarkets. Within geomarkets targeted by a student list purchase, colleges may use additional filters to capture student market segment and these filters may systematically exclude low-income students. Finally, in large metropolitan areas, the borders of geomarkets may be drawn in ways that separate affluent from less affluent communities. Zemsky & Oedel (1983, pp. 11–12) state that "each major metroplitan area is composed of several markets, usually corresponding to the inner city, a first ring of suburbs, and an outer ring of suburbs." We present the following hypothesis.

H1. In large metropolitan areas that contain multiple geomarkets, we expect significant socioeconomic inequality (e.g., income, parental education) between geomarkets.

Geomarkets and racial inequality. The relationship between geomarkets and racial inequality is less obvious because race is not an explicit component of the Market Segment Model and Zemsky & Oedel (1983) make no mention of race. The question of whether geomarket borders are correlated with race is a question about structural racism. Structural racism is "a form of systematic racial bias embedded in the 'normal' functions of laws and social relations" (Tiako, South, & Ray, 2021, p. 1143), whereby processes viewed as neutral or common sense systematically advantage dominant groups and disadvantage marginalized groups.

Colorblind racism refers to behaviors and ideology that ignore the significance of race and racial discrimination [CITE LATER WORK BY BONILLA-SILVA]. The Market Segment Model can be seen as an example of colorblind racism because Zemsky & Oedel (1983) develop an analysis of the structure of college choice that completely ignores race. In turn, enrollment management professionals interpet the Market Segment Model and geomarkets as race-neutral recruiting tools. Race-neutral racism is racial inequality that results from using ostensibly race-neutral inputs that are systematically correlated with race [CITE] (e.g., Benjamin, 2019; Chun, 2021; Norris, 2021). DAN HIRSCHMAN CALLS IT A Dan calls geomarket colorblind racist market device.

We argue that geomarkets are seemingly race-neutral inputs that are correlated with race. Geographic borders are the most commonly studied racialized inputs (e.g., Benjamin, 2019; Korver-Glenn, 2022; O'Neil, 2016). These studies build on the fact that American communities and schools are racially segregated as a consequence of historic and contemporary laws, policies, and practices promoting racial segregation (Chun, 2021; Harris, 1993; Korver-Glenn, 2018; Rothstein, 2017). Selection devices that categorize people based on geographic location without considering structures that produce segregation are likely to reproduce historical race-based inequality in opportunity. Although previous research has focused racial inequality stemming from the use of zip codes as an input (e.g., O'Neil, 2016), College Board geomarkets are much larger than individual zip codes.

Nevertheless, we argue that geomarket borders may be correlated with race for two reasons. First, borders may be drawn in a racicialized way that follows the countours of racial segregation in residential housing. Examples include: the "South and South Central Los Angeles" geomarket (CA21); the "City of Oakland" geomarket (CA07), which is surrounded by the "Alameda County excluding Oakland" geomarket (CA08); and the "Wayne County Detroit" geomarket (MI01), which is surrounded by the "Detroit's Northern Suburbs" (MI02) and "Ann Arbor" (MI03) geomarkets. Second, geomarket borders drawn along class divides may exhibit racial inequality because of race-based inequality in income and wealth [CITE]. We present the following hypothesis:

H2. In large metropolitan areas that contain multiple geomarkets, we expect significant racial inequality (e.g., income, parental education) between geomarkets.

MOTIVATE HYPOTHESIS ABOUT INCREASE/DECREASE OVER TIME IN BETWEEN-GEOMARKET RACIAL INEQUALITY. POTENTIAL FACTORS MIGHT BE GENTRIFICATION; RISING INCOME INEQUALITY

Product utilization and racial inequality. In addition to geomarket borders being correlated with class and race, we argue that geomarkets are likely to be used in ways that amplify socioeconomic and racial inequality in college access. Student lists are purchased by university enrollment personnel or by enrollment management consultants acting as agents on behalf of a university. Student list products are designed to filter on multiple search filters simultaneously, and they grant administrative discretion over which filters to select, which filter thresholds to select, and how many purchases to execute.

Compared to standardized selection devices, scholars tend to find that administrative discretion over selection devices causes structural inequality to increase (Castilla, 2008; Cotter, Medeiros, Pak, & Thorson, 2021; Norris, 2022). Discretionary selection devices allow explicit or implicit individual bias to affect selection decisions (Burrell & Fourcade, 2021; Korver-Glenn, 2018). Even without implicit bias, discretionary selection devices can increase structural inequality when decision-makers have incomplete knowledge about how the inputs they choose are correlated with race. As such, discretionary selection devices are particularly sensitive to racialized inputs that are perceived to be race-neutral. For example, research shows that Americans dramatically underestimate the magnitude of racial income inequality (Kraus, Onyeador, Daumeyer, Rucker, & Richeson, 2019).

Student list purchases that filter on College Board geo-markets do not not inherently exclude communities of color. Indeed, a university could filter on geomarekts in order to explicitly target an area with a high number of low-SES students or students of color. As such, the incorporation of geomarkets into the student list product creates new ways to target underrepresented communities and also creates new ways to discriminate against these communities. For example, an enrollment professional can reason, there are a lot of 'good' students in the "Contra Cost County" (CA06) and the "Alameda County excluding Oakland" (CA08) geomarkets, but not so many in the "City of Oakland" (CA07) geomarket.

We conceive of geomarkets as a "colorbind racist market device" in that geomarkets are often correlated with race, but geomarkets are likely perceived as race-neutral by people making decisions about student list purchases.[3] Student list purchases do not show how the characteristics of targeted prospects compare to the characteristics of the surrounding community. From this perspective, racial inequality from the use of geomarkets in student list purchases is likely to be an unintentional result of using a seemingly race-neutral product. Additionally, purchases that specify multiple filters (e.g., test scores, grades, intended major) including geomarkets can yield unintended racial inequality because administrators have

[3] Colorblind racism is the result of ideology that attemps to minimize race; using geomarkets one doesn't have to minimize race, because race has already been eliminated from geomarkets and Market Segment Model.

incomplete knowledge about how the intersection of these filters interact with local patterns of segregation.'

We argue that unintentional racial inequality can arise because the Market Segment Model – and the enrollment management profession more generally – are oriented towards realizing enrollment goals rather than equality of opportunity for students. For example, Zemsky & Oedel (1983) recommends that instead of targeting the Boston geo-market, a (hypothetical) college targets Manchester-NH, Harford-CT, Burlington-VT, and Fairfield County-CT. The rationale is that students from these local markets are more likely to have socioeconomic characteristics associated with attending this college. By contrast, in 1980, the Boston geomarket was relatively low income and – though ignored by Zemsky & Oedel (1983) – was more racially diverse. Zemsky & Oedel (1983, p. 44) comment on this recommendation:

> On occasion, senior spokespersons for the profession worry that students outside the main market areas remain forgotten and hence, unchallenged. Inevitably, the increasing competition for students, the expense of travel and mailings, and internal political constraints compel institutions to concentrate their efforts where they will do the most good. The result is a natural reinforcing of the basic socioeconomic patterns that gave shape in the first place to the structure of college choice.

Consistent with Zemsky & Oedel (1983), prior research finds that selective private colleges and universities target affluent schools and communities (Jaquette, Han, & Castaneda, online first; Jaquette & Salazar, 2018). The same is true for the out-of-state recruiting efforts of public research universities (Salazar, 2022; Salazar, Jaquette, & Han, 2021). We argue that student list purchases that target affluent geomarkets of large metropolitan areas are likely to exclude more students of color compared to student list purchases that target the entire metropolitan area or target less-affluent geomarkets in the MSA. We test this argument by analyzing actual student list purchases of public reearch universities. We examine student list purchases that filter on geomarket. In student list purchases that did not filter on geomarket, we simulate what would happen if they had.

HX: STATE HYPOTHESIS???

5 References

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Medford, MA: Polity.

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

Castilla, E. J. (2008). Gender, race, and meritocracy in organizational careers. *American Journal of Sociology*, *113*(6), 1479–1526. https://doi.org/10.1086/588738

Chun, W. H. K. (2021). *Discriminating data: Correlation, neighborhoods, and the new politics of recognition* (pp. xi, 327 pages). Cambridge, Massachusetts: The MIT Press.

College Board. (2012). *Welcome to the college board*. College Board. Retrieved from https://secure-media.collegeboard.org/homeOrg/content/pdf/welcome_to_the_college_board_fullversion.pdf

College Board. (2023). *Detailed connections geography maps: Reach more students where they are*. College Board. Retrieved from https://cbsearch.collegeboard.org/media/pdf/connections-geographies.pdf

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

Harris, C. I. (1993). Whiteness as property. *Harvard Law Review*, *106*(8), 1707–1791. https://doi.org/10.2307/1341787

Jaquette, O., Han, C., & Castaneda, I. (online first). The private school network: Recruiting visits to private high schools by public and private universities. *Research in Higher*

Education.

Jaquette, O., & Salazar, K. (2018). Colleges recruit at richer, whiter high schools. Retrieved from https://www.nytimes.com/interactive/2018/04/13/opinion/college-recruitment-rich-white.html

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

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

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proc Natl Acad Sci U S A*, *110*(15), 5802–5805. https://doi.org/10.1073/pnas.1218772110

Kraus, M. W., Onyeador, I. N., Daumeyer, N. M., Rucker, J. M., & Richeson, J. A. (2019). The misperception of racial economic inequality. *Perspectives on Psychological Science*, *14*(6), 899–921. https://doi.org/10.1177/1745691619863049

Lazarsfeld, P. F., & Merton, R. K. (1954). Friendship as a social process: A substantive and methodological analysis. In M. Berger (Ed.), *Freedom and control in modern society* (pp. 18–66). New York: Von Nostrand.

Litten, L. H., Sullivan, D. J., & Brodigan, D. L. (1983). *Applying market research in college admissions* (pp. xxii, 303 p.). New York: College Entrance Examination Board.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415–444.

Norris, D. (2021). Embedding racism: City government credit ratings and the institutionalization of race in markets. *Social Problems*. https://doi.org/10.1093/socpro/spab066

Norris, D. (2022). The illusion of transparency: The political double standard in city credit ratings. *Socio-Economic Review*. https://doi.org/10.1093/ser/mwac016

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

Reber, S., & Smith, E. (2023). College enrollment disparities: Understanding the role of academic preperation. *The Brookings Institution*. Retrieved from https://www.brookings.edu/research/college-enrollment-disparities/

Rothstein, R. (2017). *The color of law: A forgotten history of how our government segregated America*. Liveright Publishing.

Salazar, K. G. (2022). Recruitment redlining by public research universities in the los angeles and dallas metropolitan areas. *The Journal of Higher Education*, *93*, 585–621. 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

Takamiya, T. (2005). *Mechanisms for marketing higher education information services: The case of the college board*. University of Pennsylvania.

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

Zemsky, R., & Oedel, P. (1983). *The structure of college choice / robert zemsky, penney oedel*. New York: College Entrance Examination Board.

Table 1: Simplified market segment profile, Connecticut Market 3: Fairfield County

Characteristic	Local	In-state	Regional	National
Total test takers	550.0	1199.0	1664.0	3766.0
Avg SAT (verbal + math)	770.0	850.0	970.0	980.0
Avg # scores sent per test taker	2.8	3.5	4.8	5.3
Percent in top 20% of HS class	27.8	26.1	44.7	45.7
Percent aspiring to more than BA	30.6	41.5	54.5	62.2
Percent family income more than $35,000	11.2	20.6	41.9	43.0
Percent both parents with BA	9.0	16.3	34.0	37.1

Table 2: Simplified sample institutional profile for Connecticut Market 3: Fairfield County

	Local	In-state	Regional	National	Total
Total number of scores received	1.0	58.0	69.0	109.0	237.0
Pct of all test-takers in segment	0.2	4.8	4.1	2.9	3.3
lib_arts_num	0.0	25.0	35.0	61.0	121.0
lib_arts_share	0.0	8.4	7.0	5.3	5.9
engineering_num	0.0	2.0	5.0	3.0	10.0
engineering_share	0.0	3.0	5.8	0.8	1.8

Table 3: Top 5 institutions in terms of number of scores sent by segment, Connecticut Market 3: Fairfield County

	Local (N=550)			In-state (N=1,199)			Regional (N=1,664)			National (N=3,766)		
	Num	Pct	Type	Num	Pct	Type	Num	Pct	Type	Num	Pct	Type
1	277	50.4	priv ma	757	63.1	pub doct	610	36.7	pub doct	1226	32.6	pub doc
2	261	47.5	priv ma	515	43.0	pub ma	348	20.9	priv doct	371	9.9	priv doct
3	183	33.3	priv ma	438	36.5	pub ma	272	16.3	priv doct	327	8.7	priv res
4	103	18.7	pub doct	183	15.3	pub ma	248	14.9	pub doct	312	8.3	priv doct
5	100	18.2	pub ma	177	14.8	pub ma	197	11.8	pub doct	308	8.2	priv doct