# Structuring College Access: The Market Segment Model and College Board Geomarkets

Ozan Jaquette (UCLA) Karina Salazar (University of Arizona)
Crystal Han (Stanford University)

**ABSTRACT**

*The Structure of College Choice* (Zemsky and Oedel 1983) created "Geomarkets" and the "Market Segment Model" (MSM). Geomarkets carve states and metropolitan areas into smaller geographic units to define local recruiting markets. MSM predicts how student demand for a college varies by Geomarket, based on household socioeconomic characteristics. Geomarkets became an input for two College Board recruiting products. First, the Enrollment Planning Service informs which Geomarkets and high schools a college recruits from. Second, the Student Search Service sells prospective student contact information – or "student lists" – which colleges filter by Geomarket to determine which prospect profiles to purchase. Drawing from scholarship on quantification, particularly discussions of correlation and homophily, we conceptualize how recruiting products incorporate Geomarkets. We address two research questions: What is the socioeconomic and racial variation between Geomarkets? How does the socioeconomic and racial composition of included versus excluded prospects vary when student lists filter on Geomarkets? We answer RQ1 by analyzing Census data from 1980, 2000, and 2020. RQ2 analyzes data on student lists purchased by public universities, collected through public records requests. Using a quantitative case study design with metropolitan areas as cases, descriptive statistics and interactive maps suggest Geomarkets mirror deeply-rooted patterns of racial and class segregation. Moreover, student lists that exclude low-income Geomarkets systematically overlook non-White students, first-generation students, and especially first-generation non-White students.

# 1 Introduction

On January 9, 2019 a public research university purchased College Board Student Search Service order no. 448922, thereby obtaining the contact information of out-of-state prospective students, who would subsequently receive brochures, postcards, emails, and targeted social media. This was a large student list purchase, yielding $122,426$ "names." At \$0.45 per name, the order cost \$$55,091.70$. The order specified three "search filters" to control which names were purchased: first, the 2020 high school graduating class; second, under the heading "College Board Exams," PSAT score range of 1070 - 1180; and third, under the heading "Geography", the order filtered on 67 "Geomarkets," identified by titles like "IL08 - Northwest Suburbs" and "IL09 - North Shore."[1] The Geomarket filter carves states and metropolitan areas into smaller geographic units. Upon closer inspection, this student list purchase selected all Geomarkets in the Chicago-land area except for "IL10 - City of Chicago," which is significantly poorer and has a higher share of Black residents. Similarly, in the DC/Maryland/Virginia metro area, the order selected the affluent "MD02 - Montgomery Metropolitan," "MD03 - Central Maryland (w/o Baltimore)," "VA01 - Arlington and Alexandria," and "VA02 - Fairfax County" Geomarkets, but excluded the "MD05 - Prince Georges Metropolitan," "MD07 - Baltimore (Urban)," and "DC01 - District of Columbia" Geomarkets.

Sociology offers several perspectives to describe the process of sorting students into colleges as a process that tends toward social reproduction (Domina et al. 2017; Stevens et al. 2008). On the student demand side, the status attainment model argues that student postsecondary destination is a function of family socioeconomic background, particularly parental education (e.g., Blau and Duncan 1967; Sewell and Shah 1967; Sewell 1971; Fishman 2020; Karen 2002). The cultural capital model argues that bourgeois households bestow the pedigree, social networks, and information necessary to claim spots at selective colleges (e.g., Bourdieu 1984, 1988; Huang 2023). On the supply-side, the credentialism literature recognizes that colleges have a financial incentive for educational credentials to be required to claim advantageous social and economic positions (Weber 1948), and this race for credentials benefits affluent households (e.g., Collins 1979; Labaree 1997). Scholarship on enrollment management offers an agentic depiction of college behavior, showing

[1] On the same day, the university purchased two additional orders that filtered on PSAT score ranges of $1190-1260$ and $1270-1520$, respectively, but selected the same HS graduating class and Geomarkets.

that colleges expend considerable resources recruiting desirable prospects (Stevens 2007; Author XXXXa; Cottom 2017; Holland 2019).

We argue that these perspectives ignore an important mechanism of social reproduction: third-party vendors create products that sort students on behalf of college. These products often amplify historic inequality in educational opportunity by using historical snapshots of student demand to make recommendations about where colleges should prioritize recruiting efforts. In the contemporary era of private equity funded platform capitalism, third-party recruiting products have become increasingly powerful, granular, and ubiquitous. However, a watershed moment in their development was the creation of College Board Geomarkets.

In 1983, Robert Zemsky and Penny Odell authored *The Structure of College Choice*. College Board underwrote the project – providing data and funding – and published the book as part of their nascent efforts "to help colleges estimate their enrollment potential, efforts which had faltered largely because the smallest geographic unit used in these analyses was the state" (Zemsky and Oedel 1983, x). Conceptualized as an effort to "capture and quantify" (p. 11) the knowledge of admissions officers, Zemsky and Oedel (1983) developed the Market Segment Model by analyzing the SAT score-sending behavior of high school seniors in 1980. The thesis of the Market Segment Model is that student demand for your institution is a function of social class and geography. Therefore, colleges should recruit from territories that contain large populations of your target social class. Geomarkets are geographic borders meant to define local recruiting markets. The left panel of Figure 1 shows Geomarkets in the Chicago-land area. As a planning tool for colleges, the Market Segment Model predicts how demand for a particular college varies by Geomarket, based on characteristics of households in the Geomarket.

Geomarkets have been incorporated into the supply-side structure of college access in three ways. Anecdotally, Geomarkets became an organizing principle for how college admissions offices allocate admissions recruiters to territories. For example, Figure 1 shows that Chicago Geomarkets define recruiting territories for University of Chicago admissions recruiters. Second, Geomarkets are the basis for the College Board Enrollment Planning Service (EPS). EPS software recommends which Geomarkets a college should recruit from and which schools/communities they should prioritize within targeted Geomarkets.

Third, Geomarkets were incorporated into College Board's student list product, Student Search Service. Student lists contain the contact information of prospective students and have been the primary source of "lead generation" in U.S. higher education since 1972, when College Board began selling names (Belkin 2019; Author XXXXb; EAB 2018). Colleges control which prospect profiles they purchase by selecting search filters, such as SAT score, AP score, and state. Geomarket filters enable colleges to include or exclude prospects from particular Geomarkets.

Scholarship on enrollment management has focused on the behavior of colleges (e.g., Cottom 2017; Author XXXXa, XXXXc; Stevens 2007), contributing to the belief that recruiting is a function of individual colleges. However, scholarship has failed to investigate the broader enrollment management industry – consisting of third-party vendors, their products, their consulting solutions – as a set of mechanisms that structure college access. These mechanisms, which lie upstream of the behavior of individual colleges, are incorporated by individual colleges in ways that funnel certain kinds of students to certain kinds of institutions. This omission is surprising because sociology is concerned with social sorting (Domina et al. 2017), particularly sorting into college.

This manuscript analyzes College Board Geomarkets as a case study of quantification. In particular, we draw from the discussions of correlation and homophily from Chun (2021). Zemsky and Oedel (1983) identified the correlates of 1980 SAT score-sending behavior and concluded that student demand for higher education is primarily a function of social class. The Market Segment Model argues that homophily – actors that share characteristics form connections – is the organizing principle of competition and college choice, stating that "the hierarchical structure of collegiate competition largely reflects the stratified social and economic dimensions of the communities from which colleges draw their students" (Zemsky and Oedel 1983, 72). Scholarship on quantification demonstrates that making recommendations based on past correlations amplifies the effects of historic structural inequality (Burrell and Fourcade 2021). The snapshot of student demand in 1980 – itself a consequence of historic structural inequality – was programmed into recruiting products that colleges utilize to identify and target prospective students. The result is a supply-side that amplifies structural inequalities observed on the demand-side.

Our analyses address two research questions, which speak to how Geomarkets are utilized within EPS software (RQ1) and within the Student Search Service student list product (RQ2). First, what is the socioeconomic and racial variation between Geomarkets in metropolitan areas and how does this variation change over time? We address this question by spatially joining Geomarket shapefiles to Census data about socioeconomic and racial characteristics. Second, how does the socioeconomic and racial composition of included versus excluded prospects vary when student list purchases filter on particular Geomarkets? This research question is motivated by Zemsky and Oedel (1983), which recommend that selective colleges target affluent Geomarkets, and by prior research on enrollment management behavior, which finds that selective colleges focus recruiting visits on high schools in affluent, predominantly White communities (Stevens 2007; Author XXXXa). We analyze this research question using data from actual student list purchases, which we collected using public records requests. We analyze student list purchases that include all Geomarkets in a metropolitan area in order to investigate which prospects would be included or excluded had the student list purchase filtered on particular Geomarkets.

The following section introduces background and scholarship about enrollment management and introduces salient concepts from scholarship on quantification. Second, we introduce the case, describing how Zemsky and Oedel (1983) developed Geomarkets and the Market Segment Model and how these concepts were incorporated into College Board recruiting products. Third, we describe data and methods. Fourth, we present results. Findings for RQ1 reveal that Geomarkets mirror deep-rooted patterns of racial and class segregation, with Black and Latinx communities historically and often still concentrated in the poorest Geomarkets, though these dynamics vary by metropolitan area. Findings for RQ2 indicate student list purchases that exclude low-income Geomarkets are likely to systematically overlook high-achieving Black, Hispanic, and Asian prospective students, especially those who are first-generation college students, thereby highlighting how third-party recruitment tools serve as mechanisms of social reproduction by sorting students in ways that reinforce historic inequalities in educational opportunity.

Finally, we discuss implications for scholarship. Third-party vendors have structured college access for decades and their influence is growing. The contemporary enrollment management industry is characterized by private-equity owned firms selling software-as-a-service platforms that utilize the same logic as Zemsky and Oedel (1983), for example EAB's "pipeline analytics" machine learning product (Vista Equity Partners 2025). We propose two fruitful approaches to studying the role of third-party vendors in enrollment management. First, case studies – like this manuscript and Hamilton et al. (2024) – utilize primary data collection to conduct granular analyses of particular products or vendors. Second, a growing number of subscription databases (e.g., Orbis, Pitchbook, LSEG) serve the information needs of the investment community by providing data on private firms and acquisitions or investments in private firms. These data sources may provide the basis for a scalable "knowledge infrastructure" (Hirschman 2021) that investigates the role of third-party vendors in education.

## 2 Enrollment Management and Quantification

### Enrollment Management

Enrollment management is simultaneously a profession, an administrative structure, and an industry. As a profession, enrollment management (EM) integrates techniques from marketing and economics in order to "influence the characteristics and the size of enrolled student bodies" (Hossler and Bean 1990, xiv). As an administrative structure, the EM office typically controls the activities of admissions, financial aid, and recruiting (Kraatz et al. 2010).

Figure 2 depicts the "enrollment funnel," which modifies the traditional marketing funnel to depict broad stages in the process of recruiting students (Litten et al. 1983; EAB 2019). The funnel begins with a large pool of "prospects" (i.e., prospective students) that the university would like to enroll. "Leads" are prospects whose contact information has been obtained. "Inquiries" are prospects that contact the institution, including those that respond to an initial solicitation (e.g., email) and those that reach out on their own (e.g., sending SAT scores). The purpose of the enrollment funnel is to inform recruiting interventions that target one or more stages. These interventions seek to increase the probability of "conversion" across stages (Campbell 2017). At the top of the enrollment funnel, purchasing student lists is the primary means of converting prospects to leads (Author XXXXb). Purchased leads are served emails, brochures, and targeted social media designed to solicit inquiries and applications (Ruffalo Noel-Levitz 2022b).

Scholarship at the nexus of EM and college access can be categorized by which part(s) of the enrollment funnel it speaks to. The majority of scholarship focuses on the admissions stage, analyzing which admissions criteria are utilized and/or which applicants are admitted (e.g., Hirschman et al. 2016; Taylor et al. 2024; Posselt 2016; Killgore 2009; O. Poon and Bastedo 2022). Scholarship from the economics of education often investigate financial aid leveraging, which seeks to convert admits to enrolled students (e.g., Hurwitz 2012; Leeds and DesJardins 2015).

A growing literature analyzes the earlier "recruiting" stages of identifying leads, soliciting inquiries, and soliciting applications. Author (XXXXa) conceptualize recruiting behavior as an indicator of college enrollment priorities. Ethnographies by Stevens (2007) and Khan (2011) identify connections between private school guidance counselors and college admissions officers as a mechanism for social reproduction. In turn, recruiting visits to high schools – and subsequent phone calls – are a means of maintaining ties with guidance counselors at feeder schools and establishing relationships with prospective students (Stevens 2007; Ruffalo Noel-Levitz 2022b). Quantitative case-studies of off-campus recruiting visits by public research universities and by selective private universities reveal a preference for visiting private schools and affluent, predominantly White public schools (Author XXXXa, XXXXc, XXXXd). From the student perspective, Holland (2019) finds that underrepresented students were drawn to colleges that made them feel wanted, often attending institutions with lower graduation rates and requiring larger loans than other college options. Cottom (2017) shows that for-profit colleges found a niche in Black and Latinx communities because traditional colleges ignored these communities.

By focusing on the behavior of colleges, scholarship on EM implicitly assumes that recruiting is something done by individual colleges. The EM industry includes relevant stakeholders in the organizational field (DiMaggio and Powell 1983), including professional associations (e.g., National Association for College Admission Counseling) and third-party servicers that supply products and consulting solutions to colleges (e.g., College Board, EAB). We argue college access is structured by third-party servicers and products that interact with direct-providers (colleges). Although sociologists have hinted at the ways the broader EM industry contributes to inequality in college access (e.g., Kraatz et al. 2010), scholarship has failed to make third-party servicers and products the object of empirical analysis.

Drawing from scholarship on organizational theory, EM processes require colleges to make many "make or buy" (Scott and Davis 2007; Coase 1937) decisions about whether to perform a given task in-house or outsource it to a third-party vendor (Author XXXXb). The most widely-known class of EM vendor is EM consulting firms, which provide advice and implementation in the areas of marketing, recruiting, pricing and financial aid, and student success (Marcus 2024). As "creating a class" has become complicated and high-stakes (Stevens 2007), many colleges hire EM consulting firms to develop and/or implement recruiting campaigns. The two largest firms – Ruffalo Noel Levitz and EAB – claim to serve more than 3,000 colleges and universities collectively (Ruffalo Noel Levitz 2023; EAB, n.d.).[2]

While the contemporary EM industry is characterized by software-as-a-service products sold to colleges by private equity backed firms (Author XXXXb), College Board products played a pivotal role in transforming recruiting from an in-house process to a process structured by third-party products. Every year, colleges must figure out which prospects should receive promotional material and which high schools to visit. The College Board Student Search Service – copied by ACT's Educational Opportunity Service – became the ubiquitous means of converting prospects into leads that could be served targeted recruiting interventions. The College Board Enrollment Planning Service (EPS), launched in 1984, is an early software-as-a-service product that provides recommendations about which Geomarkets colleges should recruit from and which high schools they should visit within targeted Geomarkets. Both products are applications of quantification that sort students into colleges. Student Search Service includes or excludes prospects from student list purchases by colleges based on the value of selected filters like Geomarket while EPS encourages colleges to target Geomarkets and high schools based on the characteristics of households. The multi-disciplinary literature on quantification helps conceptualize how third-party products utilize Geomarkets and the potential consequences of this utilization on the inequitable sorting of students into colleges.

**Quantification**

Espeland and Stephens (2008, 402) define quantification as "the production and communication of numbers." Sociological studies on quantification have predominantly explored how numbers are produced, used, and interpreted across institutional contexts such as administration, democratic rule, economics, personal life, and education, with a particular emphasis on rankings (Mennicken

[2] In our data collection, attempting to obtain data about student list purchases from all public universities in four states, at least half of these universities outsourced student list purchases to an EM consultancy (Author XXXXb, XXXXe).

and Espeland 2019; Espeland and Sauder 2016; McArthur and Reeves 2022). In the field of critical data studies, quantification is central to scholarship that explores how computation and algorithms structure social opportunities (e.g., Noble 2018; O'Neil 2016). Our study is grounded in the systematic investigation of "interactions between different quantification regimes" across sociology and critical data studies, as called for by Mennicken and Espeland (2019) [p. 239], to explore the role of third-party servicers and products in the sorting of students into colleges. In particular, the discussions of correlation and homophily by Chun (2021) show how social reproduction is encoded in algorithms that direct future resources. We explore how correlation and homophily were incorporated into the construction and application of Geomarkets within the Market Segment Model (Zemsky and Oedel 1983).

**Correlation**. Broadly, correlation measures the extent to which two or more variables move together. Predictive algorithms are based on correlation and are developed in two steps. First, apply statistical techniques to previous cases (training data) in order to identify factors positively and negatively associated with an outcome of interest. Second, apply these results (e.g., regression coefficients) to future cases in order to make predictions and/or to assign levels of risk to each case. Chun (2021) provides the example of Kosinski et al. (2013), who developed a method to predict sensitive personal attributes (e.g., gender, political party) based on Facebook Likes. These models predict outcomes based on correlations without requiring knowledge about underlying causal relationships. Chun (2021, 50) writes that, "correlation grounds big data's so-called revolutionary potential. As Wired editor Chris Anderson infamously declared..., big data proved that 'correlation supersedes causation, and science can advance even without coherent models, unified theories, or really any mechanistic explanation at all.' "

Due to data limitations, predictive analytics often utilize "proxy" variables – variables highly correlated with a variable of interest (O'Neil 2016; Chun 2021). For example, "e-scores" utilize proxy variables to identify "people like you" and then predict your buying behavior based on the past buying behavior of people like you. O'Neil (2016) [p. 146] states that "the modelers for e-scores have to make do with trying to answer the question 'How have people like you behaved in the past?' when ideally they would ask, 'How have you behaved in the past?' " We observe similar behavior in market research about college access (e.g., College Board 2011b), when enrollment managers use ‘which colleges did students near you consider’ as a proxy for, ‘which colleges would you consider?’

Many studies show that predictions based on correlations reproduce structural inequality (for a review see Burrell and Fourcade 2021). The correlations observed during the training data stage are a snapshot of relationships between variables at a particular point in time. Observed correlations may be a function of enduring structural inequality, but underlying causes are not considered by applications of predictive models. Reviewing scholarship about algorithms, Burrell and Fourcade (2021, 224) state that “predicting the future on the basis of the past threatens to reify and reproduce existing inequalities.” Disproportionately targeted/excluded populations are predicted to have a higher risk of an outcome, which amplifies subsequent targeting/exclusion. This phenomenon has been termed the “ratchet effect” (Harcourt 2015) and “pernicious feedback loops” (O’Neil 2016).[3] As we discuss below, Zemsky and Oedel (1983) analyzed 1980 SAT score-sending behavior (training data) and inferred that demand for college is a function of class. They then recommended that selective colleges recruit from localities with many affluent, college-educated households.

**Homophily**. Homophily is the idea that actors who share common characteristics are likely to form connections with one another, or “birds of a feather flock together” (McPherson et al. 2001). Homophily is a core concept of network science, in which actors (nodes) are connected to one another directly and indirectly via network ties (edges). In *The Company We Keep: Colleges and Their Competition*, Zemsky and Oedel (1983, 4) conduct a network analysis whereby two colleges are defined as competitors if a large number of students submit SAT scores to both colleges. Chun (2021) observes that network science models often draw from rational choice theory, which leads to the assumption that homophily is the result of voluntary action by individuals.

Chun (2021) problematizes the idea that homophily is a naturally occurring phenomenon. Because network science models – like correlational models above – typically do not observe how historical structures affect relationships, they “erase historical contingencies, institutional discrimination, and economic realities” (Chun 2021, 95) that cause behavior consistent with homophily. Second, in commercial social networks, homophily is more than an assumption; rather, it is programmed

[3] An often cited example is the LSI-R recidivism model which predicts a prisoner’s chances of re-arrest and is used by 24 states (O’Neil 2016). Because the algorithm uses zip code as an input, people who live in highly policed neighborhoods have a higher predicted probability of being arrested, which leads to more policing in those neighborhoods, which perpetuates racialized inequality in arrests.

into algorithms that create connections between users. Thus, “networks create and spawn the reality they imagine; they become self-fulfilling prophecies” (Chun 2018, 66). We observe similar phenomena in our case. Based on a network analysis of 1980 SAT score-sending, Zemsky and Oedel (1983) conclude that like-colleges compete for like-students (i.e., students of similar social origin). In turn, Zemsky and Oedel (1983) reason that colleges should target the Geomarkets and high schools targeted by peer colleges. This logic is subsequently programmed into EPS software that recommends which Geomarkets and high schools to recruit from. Thus, homophily observed on the demand side of college choice – which is a function of historic structural inequality – is programmed into the supply-side.

Homophily is central to market research products that categorize customers. Geodemography emerged in the 1970s as a branch of market research that estimates the behavior of consumers based on where they live (Burrows and Gane 2006). Market segments are subgroups within a larger market that have similar consumer demand. Early geodemographic classifications of consumers (e.g., PRIZM by Claritas Corporation) were derived from publicly available Census data, which disaggregated data to the zip code level. The Claritas Corporation had a financial incentive to argue that people living near one another share similar consumer preferences because geographic localities could then be categorized into market segments that would be useful for direct mail marketing campaigns (McKelvey 2022). Later, the development of individual credit scores (e.g., FICO score) enabled merchants to classify consumers into many, fine-grained groups (M. Poon 2007). Fourcade and Healy (2013) introduce the concept “classification situations” to describe the expansion of actuarial techniques to categorize customers into many, ordinally ranked groups. Merchants and lenders began tying these classifications to tiered products that target different consumer groups with different levels of benefits and costs (Fourcade and Healy 2024).[4] Classification situations engender markets where a vertical hierarchy of products are matched to a vertical hierarchy of consumers. We observe similar processes in our case study. Zemsky and Oedel (1983) categorize students into four market segments – local, in-state, regional, and national – based on their score-sending behavior and then evaluate the attractiveness of each Geomarket based on how many “regional” and “national” students live there.

[4]For example, “payday loans” charge high interest rates to consumer groups that were previously denied credit.

We make targeted contributions to scholarship on quantification in the sociology of education. As Stevens (2007) demonstrates in the chapter *Numbers*, EM is fundamentally concerned with quantification. However, sociological analyses of EM focus on the behavior of colleges and their agents, creating the sense that inequalities created by EM are a function of individual colleges reacting to macro structures in their environment such as rankings and financial pressures (e.g., Espeland and Sauder 2016; Stevens 2007). To our knowledge, the sociology of education has not considered producer-facing products that quantify students on behalf of colleges. Market research and investigative journalism indicates that the vast majority of colleges purchase third-party EM products to identify prospects and to decide which schools and communities to target (Marcus 2024). Recruiting products grounded in the logic of predictive analytics take a snapshot of student demand – without considering historical structures that produce inequality – and then recommend that colleges divert recruiting resources to localities with strong student demand. Based on a snapshot of existing social stratification, market research matches vertically categorized consumers to vertically categorized producers, thereby amplifying the effect of initial stratification on subsequent stratification. Third-party EM products have been sold since the 1980s and have become more abundant with the ubiquity of software-as-a-service platforms, but have remained invisible to sociology.

This manuscript analyzes the Market Segment Model (Zemsky and Oedel 1983), which categorized high school students into vertical market segments – local, in-state, regional, national – and simultaneously created local Geomarkets that could be evaluated based on their composition of student market segments. The Market Segment Model and Geomarkets became the basis for the College Board Enrollment Planning Service (EPS), which advises colleges which Geomarkets to target. Later, Geomarkets were incorporated into the College Board student list product named Student Search Service. Unlike the analysis of UK school league tables by McArthur and Reeves (2022), we cannot show the effect of quantification on social reproduction. However, by demonstrating which students are included versus excluded in student list purchases that filter on particular Geomarkets, our analyses provide novel insight into the underlying mechanism by which quantification can reproduce historical class-based and race-based inequality in the sorting of students into colleges.

## 3 The Market Segment Model and College Board Geomarkets

**Creating Geomarkets and the Market Segment Model.**

In 1978, set against the backdrop of impending college-age demographic decline, University of Pennsylvania professor Robert Zemsky was tasked by the President to figure out, " 'Who thinks about Penn?' " and " 'What other institutions do they think about when they think about us?' " (Zemsky and Oedel 1983, x). To answer these questions, Zemsky began working with the Market Research Committee of the Consortium on Financing Higher Education (COFHE), a consortium of 30 selective private universities founded in the mid-1970s. The project goal became the creation of the Market Segment Model, which had the lofty ambition of predicting student demand for any college from any locality. However, "to gain a truly comprehensive view of the collegiate enrollment market, we needed a database that described most institutions and most students" (Zemsky and Oedel 1983, x). The group approached College Board. "Coincidentally, the Board was reviewing its own efforts to help colleges estimate their enrollment potential" (Zemsky and Oedel 1983, x). Although the analyses of Zemsky and Oedel (1983) are based on the score-sending behavior of SAT test-takers, the authors had great respect for admissions officers and conceived of the project as an effort to "capture and quantify" (p. 11) their knowledge:

> 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 celebration of storytelling…[because] we believe that the intuitions of admissions officers actually comprise a remarkably systematic body of knowledge about the college selection process…Our research is based upon listening carefully to what admissions officers have to say (pp. 9-10).

The "initial task" was to define geographic boundaries "in a manner consistent with admissions officers' intuitive understanding of student pools" (Zemsky and Oedel 1983, 4). Quoting an admissions officer, Zemsky and Oedel (1983) [p. 11] write, " 'There are only three kinds of college-bound students: those who want to live at home, those who want to live on campus but bring their laundry home, and those who want to go far enough from home that Mom and Dad can't visit without calling first.' " As such, Zemsky and Oedel (1983, 11) created "three types of boundaries

– region, state, and community.” The initial regions in the project were New England, Middle States, and the South. Next, they “divided each state into as few as two and as many as thirty community-based enrollment markets or pools, for a total of 143 separate markets” (p. 11). These enrollment markets, later called Geomarkets, were intended to be consistent with the conception of a catchment market from the perspective of admissions counselors. Zemsky and Oedel (1983, 11–12) described the creation of Geomarket borders only briefly:

> 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 geography, the Market Segment Model sought to predict how student demand for a particular college varies by Geomarket, based on the characteristics of households. The model created a classification system that grouped SAT test-takers into one of four different *market segments* – local, in-state, regional, and national – based on SAT score-sending behavior. For a given student, each college that receives a score from the student can be defined as “local” (college located in the the same Geomarket as the student), “in-state” (same state but different Geomarket), “regional” (same region but different state), or “national” (different region). In turn, a test-taker is categorized in the “local” market segment if they submit 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. Zemsky and Oedel (1983, 4) comment that the Market Segment Model “was nothing more than a set of simple rules for disaggregating high school seniors into similar groups. The model worked because students, once so disaggregated, appeared to behave in remarkably consistent ways.”

The two primary outputs of the Market Segment Model are the (1) Market Segment Profile and (2) the Institutional Profile. The Online Appendix describes these outputs in more detail. Both outputs are created separately for each Geomarket. For each market segment (local, in-state, regional, national), the Market Segment Profile shows the number of students, their average SAT score, the percent aspiring to a BA+, percent with family income greater than $35,000, percent with both parents having a BA, etc.[5] The Institutional Profile shows the number of students who send scores to a particular institution – separately by market segment – and which majors these students are interested in. A university could obtain this for their own institution or for a competitor.[6] Taken together, the Market Segment Profile shows a college which Geomarkets possess attractive student market segments while the Institutional Profile shows the extent to which students in a particular Geomarket are interested in your college or a competitor college. These outputs became the basis for the Enrollment Planning Service (EPS) software.

**Correlations**. Zemsky and Oedel (1983, 3), *A Sense of Place: Students, Families, and Communities*, investigates the student characteristics correlated with being in the local, in-state, regional, or national market segment. The analyses identify four variables – educational aspirations, parental education, scholastic aptitude, and family income – that predict score-sending behavior, both individually and in combination. Because these four variables "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" (Zemsky and Oedel 1983, 33).

The thesis of this chapter is that student demand for higher education is a function of social origin. For example, the authors state (p. 33) 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" and that "the implication is simply that college-educated parents instill in their children more wide-ranging educational aspirations." Commenting on family income, Zemsky and Oedel (1983, 33) 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." They conclude (p. 42) that "our research has simply demonstrated what everyone has always known: communities with high levels of family income and parental education are also communities in which students have higher than average SATs and more far-reaching aspirations."

The authors also find that these four variables predict student score-sending behavior at the Geomarket-level. However, Geomarkets differ in the relative abundance of students with particular

[5]For example, the Market Segment Profile in the Online Appendix shows how many students in "CT3 – Fairfield County" are defined as regional or national based on their SAT score-sending behavior.

[6]The Online Appendix reproduces a partial, simplified version of Zemsky and Oedel (1983, fig. 2.3), the Institutional Profile of an anonymous college for students from Fairfield County, CT. This table shows that 69 Regional students and 109 National students from Fairfield County sent SAT scores to the college.

socioeconomic characteristics, which has practical implications for recruiting. Therefore, Zemsky and Oedel (1983, 44) recommend that colleges target Geomarkets with desirable compositions of socioeconomic characteristics in order to reach students from desired student market segments:

> On occasion, senior spokespersons for the profession worry that students outside the main [Geo]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.

**Homophily**. *The Company We Keep: Colleges and Their Competition* (Zemsky and Oedel 1983, 4) conducts a network analysis to determine which institutions are in competition with one another.[7] Based on these analyses, Zemsky and Oedel (1983) [p. 46] state that competition between colleges is characterized by homophily: “we draw a fundamental conclusion about the structure of college choice: collegiate competition occurs principally between like institutions.” Subsequent analyses investigate the tuition price and the socioeconomic composition of institutions in competition with one another. Private selective colleges and private flagship universities compete directly for students, charge the highest prices, and enroll students with the highest socioeconomic status. The authors argue that like-colleges compete for like-students as defined by socioeconomic characteristics. Zemsky and Oedel (1983, 72) describe observed patterns as a natural process in which a vertical socioeconomic hierarchy of students is matched to a vertical hierarchy of universities:

> 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

[7]This is a “two-mode” social network in which students (mode 1) send SAT scores – the network tie – to colleges (mode 2). The authors turn this into a one-mode college network that defines two institutions as being in competition with one another – the network tie – if at least 15% of students who sent SAT scores to one institution also sent scores to the other institution and vice-versa. 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. For example, describing the Figure 4.4 “Structure of Fairfield County Regional Market,” Zemsky and Oedel (1983, 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.”

> 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.

The discussion of competition by Zemsky and Oedel (1983) exemplifies the concerns about correlation and homophily described by Chun (2021). A correlational analysis of 1980 SAT score-sending patterns finds competition between colleges is defined by socioeconomic homophily. This homophily is presented as a naturally occurring phenomenon. Given these findings, Zemsky and Oedel (1983) recommend that colleges should target Geomarkets that contain a critical mass of students interested in peer colleges, information that can be discerned from the Institutional Profiles (see Online Appendix). In itself, Zemsky and Oedel (1983)'s Market Segment Model is a social science depiction of student demand – akin to the status attainment model – that does not consider historic, structural inequalities that cause observed patterns. However, by inscribing the Market Segment Model and Geomarkets into the EPS software, College Board amplified structural inequalities that contributed to homophily observed in 1980 SAT score-sending behavior. Drawing from Chun (2021), commodification of the Market Segment Model and Geomarkets engineers homophily.

**Enrollment Planning Service**. In 1984, College Board created the Enrollment Planning Service (EPS), based on the Market Segment Model (College Board 2012; Takamiya 2005). EPS was an early software-as-a-service platform that recreated the analyses of Zemsky and Oedel (1983). For each Geomarket, colleges could obtain Market Segment Report for each local market and the Institutional Profile – their own and that of competitors. Based on background conversations with EM professionals, EPS software also provided information about the score-sending behavior of individual high schools within each Geomarket. Therefore, colleges used EPS software to decide which Geomarkets to recruit from and which high schools to visit within targeted Geomarkets. Typical College Board (2005) marketing material describes EPS as,

> The marketing software that pinpoints the schools and Geomarkets where your best prospects are most likely to be found. With the click of a mouse, EPS provides you with comprehensive reports on your markets, your position in those markets, and your competition. Focus your valuable time and resources on the right prospects.

Whereas Zemsky and Oedel (1983) identified 143 Geomarkets covering the New England, Middle States, and South region, EPS created Geomarkets for the remaining U.S. states, with 304 Geomarkets in total. College Board (2023b) shows the contemporary Geomarkets. Documentation and promotional material suggest that geomarket borders were chosen based on a combination of formal geographic borders (e.g., counties) as well as proprietary College Board data designed to identify geographic areas with different college-going behaviors.[8] However, Geomarkets for New England, Middle States, and the South are identical to those developed by Zemsky and Oedel (1983).

Quantification has reactivity, discipline, and authority effects to the extent that stakeholders care about the numbers (Berman and Hirschman 2018). Reactivity is the idea that salient quantitative measures cause people and organizations to change their behavior. Market research suggests that EPS was highly salient. Noel-Levitz (1998) reports that in 1995, 37% of 4-year publics and 49% of 4-year privates used EPS, while 41% of 4-year publics and 16% of 4-year privates used ACT's market analysis service product.

Quantification "disciplines" (Espeland and Stephens 2008) actors to react in particular ways. EPS software may discipline colleges to approach recruiting in a manner consistent with the Market Segment Model. Drawing from promotional literature (e.g., College Board 2005, 2011a), Takamiya (2005), and background conversations with enrollment managers, the practical purpose of EPS was to inform the "travel schedule" of admissions recruiters. EPS promotional material and user guides encourage users to begin by identifying which Geomarkets they will recruit from.[9] Second, EPS users decide which high schools they will visit within selected Geomarkets. Based on the principle of homophily, the Market Segment Model suggests that selective colleges should target Geomarkets with large numbers of affluent, college-educated households, while low-income communities are left to local four-year and community colleges. Anecdotally, the salience of Geomarkets and EPS software affected the decision-making authority of admissions officers and the territories that

[8] College Board (n.d.) states that "geomarkets are areas within a state that represent a further segmentation of a population. Students from California don't all share the same college-going behaviors. We have accounted for this variance by segmenting the 50 states into 304 geomarkets to provide further insight into student behaviors within particular areas of individual states."

[9] Software documentation by Oracle (n.d.) states that, "EPS market codes are proprietary market codes owned by the College Board and are used to categorize external organizations and people into geographical areas.... Some admissions offices use EPS market codes to focus their recruiting efforts in geographic areas in which they believe they will be the most successful."

admissions officers were assigned to.[10]

Because Geomarkets are foundational to EPS software, our first research question is interested in the extent to which this geographical building block is associated with race and class. We ask (RQ1), what is the socioeconomic and racial variation between Geomarkets and how does this variation change over time? Zemsky and Oedel (1983) viewed demand for higher education as a function of class. They developed 143 Geomarkets borders with an eye towards identifying geographic areas that differed from one another in terms of class composition, while (College Board n.d.) created the remaining Geomarket borders with an eye towards distinguishing geographies based on college-going behaviors. Considering the extent of class-based residential segregation in the U.S. (Reardon and Bischoff 2011), we expect substantial socioeconomic inequality between Geomarkets in large metropolitan areas.

We also expect substantial racial inequality between Geomarkets within metropolitan areas. Interestingly, the Market Segment Model is explicitly based on socioeconomic stratification, but Zemsky and Oedel (1983) do not mention race. U.S. cities are characterized by extreme historic and contemporary residential racial segregation (Korver-Glenn 2022). Unless designers intentionally consider racial segregation, selection devices that categorize people based on geographic location are likely to reproduce historical race-based inequality in opportunity (Chun 2021). Second, although geomarket borders may have been drawn along class divides, a strong correlation exists between race and wealth (Kraus et al. 2019). Third, Geomarket borders may have been drawn in a way that follows the contours 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 sur-

[10] Quantification often reduces the decision-making authority of local actors (Espeland and Stephens 2008). The Market Segment Model sought to replicate the aggregate knowledge of local admissions officers. Once this knowledge was quantified and commodified onto a CD-ROM, the local expertise of admissions officers becomes less valuable. The EPS product increases the ability of a college admissions leader – working with College Board staff or an EM consultant – to plan recruiting efforts centrally. In background conversations, EM professionals told us that EPS software enabled colleges to plan travel without relying on admissions officers having strong local knowledge of their territories. However, EPS software often recommended visiting the same sets of affluent, high-achieving high schools that were receiving visits from other colleges. We were told that savvier, quantitatively-adept admissions offices used EPS to visit schools that EPS recommended ignoring because there would be less competition for the good students who attended these high schools. Finally, Geomarkets were created to mirror the territories of admissions officers (Zemsky and Oedel 1983). In turn, as Geomarkets became more salient, admissions offices often structured their territories around Geomarkets. On background, admissions officers and enrollment consultants told us it remains common parlance to hear an admissions officer say something like, "I recruit 'PA 2'," which refers to the "Chester County, PA Geomarket." In states a college recruits heavily from, admissions officers are often assigned specific Geomarkets as their territory. This can be seen in the Chicago-land recruiting territories of University of Chicago in Figure 1.

rounded 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.

**Geomarket Filter in Student Search Service**. Geomarkets also function as a student list search filter for the College Board's Student Search Service. A student list contains the contact information of prospective students who meet the search filter criteria (e.g., test score, GPA) specified by the university. Student lists are the fundamental input for undergraduate recruiting campaigns because purchased names – alongside prospects who reach out on their own – constitute the set of prospects who receive subsequent recruiting interventions (e.g., mail, email) designed to push them toward the application and enrollment stages of the "enrollment funnel." Ruffalo Noel-Levitz (2022b) reports that 86% of public colleges and 87% of private colleges purchase student lists.[11] Historically, dominant list vendors are the College Board and ACT, which derived student list data from test-takers, but the test-optional movement, advances in technology, and surging private equity investments have contributed to new sources of student list data (Author XXXXb).

Student lists are positively associated with student outcomes. A College Board research report by Howell et al. (2021) compared SAT test-takers who opted into the College Board Student Search Service – allowing colleges to purchase their contact information – to those who opted out. After controlling for covariates (e.g., SAT score, parental education, school fixed effects), 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 percent change – (41.1–32.8)/32.8 – in the relative probability.[12] Howell et al. (2021) also found that participating in Search was strongly associated with obtaining a BA in four years.[13]

However, Author (XXXXf) argue that student list products may exacerbate racial inequality in college access. The authors conceptualize student list products as "selection devices" (Hirschman and

[11]For public universities that purchased lists, 80% purchase more than 50,000 names annually. Ruffalo Noel-Levitz (2022a) reports that student lists were the top expenditure item in the undergraduate recruiting budget for both private and public institutions in 2022, with the average public institution allocating 15% of its budget to names.

[12]Leveraging a natural experiment in College Board student list purchases, Smith et al. (2022) find that purchasing a prospect profile increases the probability that the student will apply to and enroll at the purchasing college, with larger effects for Black, Hispanic, and low-income students.

[13]20.6% of students who participated in Search obtained a BA in four years compared to 15.7% of students who opted out, representing a 31.2% (=(20.6-15.7)/15.7) increase in the relative probability of graduation.

Bosk 2020) that enable colleges to select which prospective students they target by incorporating search filters (e.g., high school graduating class, state). Norris (2021) defines "racialized inputs" as ostensibly race-neutral inputs that are systematically correlated with race because marginalized racial/ethnic groups have historically been excluded from the input. Author (XXXXf) argue that several frequently utilized student list filters (e.g., zip code, AP test score, SAT score) meet the criteria of racialized inputs. Using a national sample of high school students and using data from actual student lists purchased by public universities, Author (XXXXf) show that racialized search filters have a strong negative relationship with the selection of Black and Latinx prospects.

Because of the extent of residential segregation in the U.S., geographic borders are a commonly studied racialized input in scholarship about algorithmic bias (Benjamin 2019; O'Neil 2016; Harcourt 2007). Student list products offer many geographic filters. Some geographic filters are based on known geographic borders (e.g., zip code, county, CBSA, state). College Board also created geographic borders that subsequently became filters in the Student Search Service, such as Geomarkets and Geodemographic Segment filters.[14] In their sample, Author (XXXXf) found that 11% of student lists purchased by public research universities filtered on Geomarkets. Drawing from Norris (2021), we argue that geomarkets satisfy the two criteria of racialized inputs. First, they are ostensibly race-neutral inputs in that neither Zemsky and Oedel (1983) – nor subsequent EPS promotional material (e.g., College Board 2011a) – mention race. Second, Geomarkets may be systematically correlated with race. For example, Geomarket "CA07" is "City of Oakland" and "CA08" is "Alameda County, excluding Oakland." We extend Author (XXXXf) by examining how the racial composition of actual student list purchases changes if universities filter on particular Geomarkets.

Zemsky and Oedel (1983, 42–44) discuss how the Market Segment Model and Geomarkets should be utilized in the context of purchasing student lists. The authors ask,

> Imagine that you recruit for a college that draws most of its students from regional or national segments. Where would you concentrate your energies? Ideally, you would seek communities [Geomarkets] with a high proportion of students already predisposed

[14] Geodemographic segment filters utilize cluster analysis to allocate each census tract and each high school into different categories based on past college enrollment and other factors (College Board 2011b).

> 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...Your efforts would surely be better directed toward...Manchester, Hartford, and Fairfield County...Indeed, in Fairfield County alone you could reach more than 40 percent of your "primary target" population.

In both mailings and off-campus recruiting visits, Zemsky and Oedel (1983) recommend focusing on Geomarkets with large populations of students who are deciding between "institutions like your own" (p. 44). Selective colleges primarily draw from students in the "regional" and "national" segments. Zemsky and Oedel (1983) suggest that the student list purchases of selective colleges should focus on Geomarkets with large numbers of affluent, highly educated households. By contrast, the student list purchases of local state colleges and community colleges, which rely on enrollment from students in the "local" and "in-state," should focus on nearby Geomarkets with large numbers of low-SES students. Our analyses attempt to recreate the strategy recommended by Zemsky and Oedel (1983) by showing who is included versus excluded when student list purchases filter on particular Geomarkets. Considering the correlation between race and wealth in segregated America, we suggest that a class-based list-buying strategy will increase racial stratification in college access.

# 4 Data and Methods

## 4.1 Data and Variables

**RQ1**. RQ1 primarily utilizes Census Data while RQ2 primarily utilizes data about student list purchases. Research question 1 is about socioeconomic and racial variation between Geomarkets in metropolitan areas and how such variation changes over time. To answer this question, we use census tract data from the U.S. Census Bureau. Utilizing a small geographic area is important for implementing a spatial merge between Census data and Geomarket shape files (described below). Census tracts are the smallest geographic area for which measures of both race/ethnicity and socioeconomic characteristics are available. Defined as statistical subdivisions of counties, census tracts boundaries are reassessed every decennial census to reflect relatively homogeneous demographic communities, typically ranging in population size from 2,500 to 8,000 residents (U. S. Census Bureau 2014, 1994). Compared to zip codes, which are fixed spatial units designed primarily to demarcate route efficiency for postal services, census tracts are considered more reliable for population-based analyses.

We use data from the 1980 and the 2000 Decennial Census as well as 5-year estimates from the 2020 American Community Survey (ACS). For the 1980 Decennial Census, we draw from Summary Tape File 1 (STF1) for variables about race and Hispanic origin collected via "short form" questions answered by all households. We draw from Summary Tape File 3 (STF3) for variables about socioeconomic characteristics (income, education, and poverty) collected via the "long form" questionnaire completed by a sample of roughly 20% of all households. The 1980 Decennial Census data collection period closely matches the data collection period of SAT score-sending behavior that Zemsky and Oedel (1983) utilized to create Geomarkets and the Market Segment Model. For the 2000 Decennial Census, we similarly use the Summary File 1A (SF1A) for census-tract measures of race and Hispanic origin and the Summary File 3A (SF3A) for census-tract measures of income, education, and poverty.[15] Beginning in 2010, the Decennial Census no longer collected a "long form" questionnaire. These questions were replaced by ACS. Therefore, we also utilized the 2020 5-year ACS for both race and socioeconomic characteristics at the census tract level, which includes data collected from 2016-2020.[16] This data collection period mirrors the period for our primary data collection about student list purchases used to analyze RQ2, as described below.

Across these data sources, we created measures of race and ethnicity, specifically the number and proportion of the census-tract population in the following categories: White, non-Hispanic; Black, non-Hispanic; Hispanic; Asian, non-Hispanic; Native Hawaiian and Pacific Islander, non-Hispanic; American Indian and Alaskan Native, non-Hispanic; and two or more races, non-Hispanic. Measures of Asian, Native Hawaiian and Pacific Islander, American Indian and Alaskan Native, and two or more races (all non-Hispanic) were not available for the 1980 Decennial Census. We also created socioeconomic measures that approximate the predictors of market segment identified by Zemsky

[15] Data for the 1980 and 2000 Decennial Census were retrieved from the IPUMS National Historical Geographic Information System, created by Manson et al. (2024).

[16] Data for the 2020 5-year ACS were retrieved from the tidycensus R package, created by Walker and Herman (2024).

and Oedel (1983). These variables are median household income, percent of households below the poverty line, and people age 25+ with a BA. All measures were created at the census tract-level and at the Geomarket-level by aggregating tract-level data to Geomarkets.

We assigned each census tract to a College Board Geomarket, which we constructed through manual digitization of maps published by The College Board in 2023 marketing materials promoting their Enrollment Planning Service (EPS) software products (College Board 2023a). Specifically, we georeferenced the College Board's Geomarket maps by overlaying PDF images onto spatial coordinates aligned with census tract boundaries. We then extracted vector data by manually tracing the Geomarket boundaries and generating individual shapefiles for each Geomarket. For instance, Figure 3 shows The College Board's map of Geomarkets in the Chicago metropolitan area on the left and our created shapefiles on the right.

This approach to constructing Geomarkets aligned precisely with 2020 census tract boundaries from ACS 5-year estimates; that is, each census tract corresponded uniquely to a Geomarket generated from the College Board images. For example, Figure 4 zooms into the northern Geomarkets of the Chicago metropolitan area from Figure 3. The center map showcases how our created Geomarket shapefiles (delineated in purple) align with the spatial coordinates of 2020 census tract boundaries (delineated in grey). However, this process was less precise when applied to the 1980 and 2000 Decennial Census tract boundaries, likely a factor of census tract delineations shifting over time based on demographic changes and our use of map images from 2023. The map on the right in Figure 4 delineates 1980 census tract boundaries, which are visibly larger and more spatially aggregated than the smaller, more granular 2020 tracts shown in the center map.

Some census tract boundaries from the 1980 and 2000 Decennial Censuses intersected multiple Geomarkets (i.e., tracts whose boundaries spanned two or more Geomarkets), resulting in spatial misalignment. In such cases, we assigned each tract to the Geomarket that encompassed the majority of its spatial area. When the spatial majority was relatively balanced across multiple Geomarkets, we assigned tracts in a way that maintained Geomarket boundaries that most smoothly followed Geomarket delineations sourced from The College Board. While this approach minimized ambiguity, it introduced minor variations in Geomarket shapes across decennial censuses. We illustrate this process in Figure 4, using a 1980 census tract in Chicago as an example, where our constructed shapefiles for the Chain of Lakes and North Shore Geomarkets do not align with the tract boundary (circled in red). In this case, because the majority of the tract's spatial area fell within the Chain of Lakes Geomarket, we assigned the tract to Chain of Lakes rather than to North Shore. This assignment, along with others like it, slightly alters the overall boundaries of Geomarkets when compared to those based on 2020 census tract delineations. To assess the robustness of our approach to spatial misalignment, we tested alternative methods for resolving ambiguous tract assignments to Geomarkets and found overall patterns in the results remained relatively consistent across approaches.[17]

**RQ2**. Our second research question asks, how does the racial and socioeconomic composition of included versus excluded prospects vary when student list purchases filter on Geomarkets? We answer this question using actual student lists purchased by public universities. In February 2020, we began issuing public records requests to public universities about student lists purchased from 2016 through 2020. We narrowed the scope of our request to student lists purchased from College Board, ACT, and the National Research Center for College and University Admissions (NRCCUA), the three largest student list vendors at the time. For each student list purchased from 2016 through 2020, we requested two related pieces of data: (1) the order summary, which specifies search criteria for the student list purchase; and (2) the de-identified prospect-level list produced from these search criteria.

The data collection sample for this project was all public universities in California, Illinois, Minnesota, and Texas. We chose these states because they were priorities for the two foundations

[17]An alternative approach to dealing with spatial misalignment included creating shapefiles for Geomarkets from a 2012 R-bloggers post written by a data scientist working in the enrollment management space. Essentially, the data file was a Geomarket-zip code crosswalk that held one observation per zip code and a column that assigned each zip code to one Geomarket. Using these zip code-level data, we utilized the aggregate() function to create shapefiles for each Geomarket. Next, we utilized the st_intersection() function to assign each census tract to a Geomarket, based on the intersection of their associated spatial shapefiles. This spatial merge kept Geomarket boundaries consistent across all three Decennial Census and split census-tract shapefiles that intersected multiple Geomarkets into smaller shapefiles, each entirely contained within a single Geomarket. We replaced each variable in these shapefiles with its original value scaled by the share of the tract's land area that lay within the Geomarket. For a particular census tract, imagine that 60% of its land area was contained in Geomarket A, 40% of its land area was contained in Geomarket B, and the census tract reports 100 people who identify as Asian, non-Hispanic. The partial spatial merge splits this census tract into two observations. The observation assigned to Geomarket A reports $.60 * 100 = 60$ Asian, non-Hispanic people and the observation assigned to Geomarket B reports $.40 * 100 = 40$ Asian, non-Hispanic people. Even under this most conservative approach to addressing spatial misalignment, overall patterns in the findings remained consistent. However, we opted to georeference the manually digitized maps directly sourced from the College Board, as the original source of the Geomarket-zip code crosswalk cited in the R-bloggers post could not be verified.

that funded the data collection. Utilizing public records requests to obtain quantitative data is a painstaking process, made more difficult by the onset of the COVID-19 Pandemic. Initially, the majority of universities ignored or denied our request. We established a client relationship with a civil rights legal organization and leveraged this relationship to obtain pro bono representation from four law firms. We also collected data from Arizona State University and Northern Arizona University, because we were able to obtain pro-bono legal representation for these two universities. However, we were unable to obtain legal representation for Texas because firms interested in representing us had conflicts of interest in Texas.

Firm representation substantially increased the success of data collection. Nevertheless, data collection remained difficult. Some universities provided records that were not usable for quantitative analyses (e.g., summary statistics across multiple orders; or data did not contain important fields). Some universities did not provide records based on legitimate legal grounds (e.g., data not in university possession; not required to create records that do not currently exist). We learned that many universities outsourced student list purchases to third-party consulting firms, particularly EAB. Unfortunately, we were rarely able to obtain usable data from these universities. Several universities denied requests based on questionable legal rationale, but we lacked the resources to litigate.

The primary data source for our second research question is student lists purchased from College Board. Originally collected from a larger project sample (Author, BLINDED), the data for this study's sample includes a total of 414 orders associated with 2,549,085 prospects made by 11 public universities located in California, Illinois, Texas, and Arizona. Although these lists were purchased by individual universities, the set of prospects included in each list is a function of the search criteria specified for that student list purchase. The resulting student lists could hypothetically extend to any university placing student lists orders based on the same filter criteria. In other words, any observed patterns in the findings below would remain constant for any university selecting the same filters.[18] Therefore, we utilize these data not to analyze the behavior of individual universities, but to identify which prospective students are included when a particular set of search criteria are selected.

[18]The caveat to this statement is that most purchases select to exclude prospect profiles that were previously purchased by the university.

In particular, we answer the second research question by analyzing particular student list purchases that did *not* utilize the Geomarket search filter and purchased prospects from all Geomarkets in a particular metropolitan area.[19] For each prospect, we know their high school, and the exact longitude and latitude of their high school, which we use to assign high school students to Geomarkets. Our analyses then show which students would have been included or excluded from the student list purchase had the purchase filtered on particular Geomarkets.

The variables of interest are Geomarket and prospect-level characteristics, which are derived from the pretest questionnaire administered to College Board test-takers that are included in the de-identified student lists we received (e.g., HS code, HS GPA range, intended major). College Board asks students separate questions about ethnicity (Cuban, Mexican, Puerto Rican, other Hispanic, non-Hispanic, ethnicity non-response) and race (American Indian or Alaska Native, Asian, Black, Native Hawaiian or other Pacific Islander, White, race non-response), allowing students to check as many boxes as they want (College Board 2016). From these responses, a College Board "derived aggregate race/ethnicity" was re-created based on the U.S. Department of Education reporting guidelines that includes the following categories: no response; American Indian/Alaska Native; Asian; Black; Hispanic/Latino; Native Hawaiian or Other Pacific Islander; White; other; two or more races, non-Hispanic (College Board 2016).[20] Information about student socioeconomic status is limited. However, we create a measure of whether the prospect is a first-generation college student based on their parents' level of education that was self-reported on the College Board questionnaire.

### 4.2 Methods

**Research design**. We utilize a multiple, quantitative case study design. Metropolitan areas are cases. Case studies deepen understanding of phenomena by comparing and contrasting across different settings (Yin 2009). Our research questions explore the socioeconomic and racial variation between Geomarkets, how such variation changes over time, and how Geomarket landscapes shape the composition of included versus excluded prospects. Our goal is not to make generaliza-

[19]For analyses of Dallas, Detroit, Long Island, Northern New Jersey, and Philadelphia, we utilize data from student list purchases that did filter on Geomarkets but included all Geomarkets in the region.

[20]Any student who selects a Hispanic ethnicity category is defined as Hispanic/Latino, regardless of the race categories they select. Additionally, non-Hispanic students who check "American Indian or Alaska Native" and another race group are defined as "two or more races, non-Hispanic."

tions about Geomarket variation nationally, but to explore how different geographical models of metropolitan areas, shaped by their unique demographic characteristics, influence the composition of prospect pools. Such aims align with the purpose of case studies, which drive the exploration of research questions focused on the "how" and "why" of phenomena in rich and nuanced contexts. We, therefore, select a small number of metropolitan areas in order to provide sufficient analytic depth while situating each case within its specific historic and contemporary sociodemographic characteristics. The Online Appendix presents results for a larger number of metropolitan areas.

The central thesis of Zemsky & Oedel's (1983) Market Segment Model suggests that student demand for various colleges is a function of social class and geography, and that colleges strategically target recruiting in high socioeconomic areas. However, because Geomarkets vary significantly in their racial and class characteristics, and in how colleges target these Geomarkets, a multiple case study enables us to compare how these dynamics unfold across different metropolitan areas. By treating each metropolitan area as a separate case, we explore both within-case processes (e.g., how homophily structures recruitment in a single metropolitan area) and cross-case patterns (e.g., how the same recruitment strategy results in different racial/class outcomes across metropolitan areas).

Case selection was informed by purposeful sampling to identify "information-rich" cases that can provide knowledge about variation across Geomarkets and over time (Yin 2009; Patton 2002). We considered metropolitan areas as information rich cases across four broad dimensions: 1) urban model structures, 2) geographical variation, 3) variation across racial/ethnic and socioeconomic characteristics, and 4) availability of student list orders and de-identified list data. Based on Zemsky and Oedel (1983), we identified classic urban model structures that reflected "major metropolitan area[s] …composed of several markets, usually corresponding to the inner city, a first ring of suburbs, and an outer ring of suburbs" (p. 11-12), such as Boston, Chicago, Detroit, Baltimore, Denver, Houston, Miami, and Philadelphia. To compare and contrast with classic urban model structures, we also identified alternative urban models such as bi-modal or metroplex models that contain two centralized cores surrounded by the more classic suburban ring (e.g., Dallas-Fort Worth and San Francisco) and multi-nuclei models where urban growth develops around multiple, sometimes uncentralized and independent, metropolitan areas (e.g., Los Angeles and New York City). The third consideration addresses geographical variations in demographic populations by intentionally selecting metropolitan cases across different U.S. regions (West, Midwest, Northeast, South) to ensure resident and prospect pools mirror broader populations. Related to all previous selection dimensions, we are also interested in variation of racial and socioeconomic populations between Geomarkets within a metropolitan area (RQ1) and how such variation shapes which prospective students are targeted by student list purchases (RQ2). We therefore considered whether or not metropolitan areas provided sufficient and varied racial and socioeconomic diversity for our research purposes.

Lastly, choice of metropolitan cases was informed by student list data availability. Although Census data are available for all metropolitan areas nationally (RQ1), we do not have good candidate student list data for all metropolitan areas (RQ2). In order to show which prospective students would have been included or excluded from the student list purchase had the purchase filtered on particular Geomarkets, the student list purchases we analyze must include all Geomarkets in a metropolitan area. This lead us to select on orders that filtered for an entire state (e.g., State filter is set to Pennsylvania), an entire metropolitan area without a Geomarket filter (e.g., CBSA filter is set to the Philadelphia metropolitan area), or an entire metropolitan area with a Geomarket filter inclusive of all Geomarkets (e.g., CBSA filter is set to the Philadelphia metropolitan area and Geomarket filter is set to all five Geomarkets in the Philadelphia area). Although we would like to analyze the D.C./Maryland/Virginia (DMV) area, the student list purchases we collected exclude "MD 5 – Prince George's Metropolitan" and "MD 7 – Baltimore (Urban)," which are substantially poorer and more Black than other Geomarkets in the DMV. Similarly, student list purchases targeting Cleveland exclude "OH 4 – City of Cleveland (East)."

In order to isolate the impact of Geomarkets on the racial/ethnic and socioeconomic characteristics of prospect pools, we did not consider purchases that utilized student preference (e.g., campus size, intended major), low-income household, or first-generation status filters. Because nearly all orders we collected (Author, BLINDED) filter on some academic criteria, we only considered orders that included assessment threshold filters (e.g., PSAT, SAT) across a wide score range to explore whether and to what extent Geomarket contributions to prospect pools are consistent across all score ranges.

Our case selection process resulted in three metropolitan cases: Chicago, Dallas-Fort Worth, and

Los Angeles.

**Analyses**. Analyses are simultaneously descriptive and spatial. We answer RQ1 by producing Geomarket-level tables and graphs that show how the Geomarkets in a selected metropolitan area vary on racial and socioeconomic characteristics and how they vary over time. We also produce interactive maps at the census tract-level to show more granular variation within and between Geomarkets.

We answer RQ2 by analyzing student list purchases that encompassed all Geomarkets in a selected metropolitan area. Descriptive tables and graphs describe the racial and socioeconomic characteristics of prospects that would have been included/excluded had the purchase filtered on particular Geomarkets. We also create interactive maps of prospect pools by Geomarket contributions to provide a more granular visualization of analyses.

# 5 Results

Results are presented below for Chicago, Dallas, and Los Angeles. The Online Appendix provides a more convenient means of viewing results. The online appendix also contains interactive maps, additional figures and tables not contained in the main text, and results for additional metropolitan areas (e.g., Bay Area, Philadelphia). The table of contents for navigating between metro areas can be accessed by clicking on the three horizontal lines on the bottom left of the screen.

## 5.1 Chicago

### Shifting Demographics, Persistent Segregation in Geomarket Landscape

Figure 5 presents the seven Geomarkets in the Chicago metropolitan area by total population across the three census periods. The maps in the figure show how Chicago Geomarkets reflect the classic urban geographical structure of an inner-city core surrounded by a ring of suburbs as described by Zemsky and Oedel (1983, 12). Based on the color gradient scale representing total population density, Figure 5 indicates northern suburb Geomarkets in the metropolitan area – Chain of Lakes, Northwest Suburbs, Northshore, Evanston and Skokie – experienced relatively little population growth from 1980 through 2020. In contrast, the most significant population increases occurred in the Western Suburbs (1.0 Million to 1.4 Million) and the South and Southwest Suburbs (less than 1 Million to 1.2 Million) Geomarkets. The City of Chicago Geomarket, forming the metropolitan area's central core, was the only locality to experience a population decline, from 3.5 Million in 1980 to 3.2 Million in 2020, despite remaining the most populous Geomarket.

These modest population shifts belie substantial change in the racial/ethnic composition within and across Geomarkets over time. Figure 6 illustrates the racial and ethnic composition of each Geomarket across the three census periods. In 1980, nearly all Geomarkets were predominantly White, with proportions ranging from 84% in the South and Southwest Suburbs to approximately 96% in North Shore. In contrast, the City of Chicago had a substantially more diverse population, with White residents comprising only 52% of its total population. One data caveat is that the following race/ethnicity categories were not available in 1980: Asian, non-Hispanic; Two+ races, non-Hispanic; NHPI, non-Hispanic; AIAN non-Hispanic. By 2020, several Geomarkets that were previously predominantly White shifted to a more heterogeneous racial/ethnic distribution, including Chain of Lakes, Northwest Suburbs, Evanston and Skokie, Western Suburbs, and South and Southwest Suburbs. For many Geomarkets, this balance in racial/ethnic composition was the result of large increases in the Latinx population from 1980 to 2020, which ranged from a 7 percentage point increase for the Evanston and Skokie Geomarket to a 20 percentage point increase for Chain of Lakes. Some Geomarkets also experienced a substantial increase in Black populations (South and Southwest Suburbs) and in Asian populations (Northwest Suburbs and Evanston and Skokie). Figure 6 also indicates that, although the North Shore Geomarket underwent some changes in racial and ethnic composition, its overall demographic patterns remained relatively the same across the three census periods. City of Chicago experienced relatively modest declines in its Black population (35% in 1980 to 26% in 2020), alongside increases in its Hispanic (13% in 1980 to 30% in 2020) and Asian populations (4% in 2000 to 7% in 2020), further substantiating the Geomarket as the most racially/ethnic diverse in the metropolitan area.

The interactive maps in the Online Appendix can show these shifts at the census tract-level. For example, in 1980 Black people were concentrated in southern tracts of the City of Chicago Geomarket, as described by scholarship on gentrification in Chicago (Ewing et al. 2024; Snidal et al. 2022). However, in 2000 and 2020 Black populations shift towards northeast tracts of the South and Southwest Suburbs Geomarket

We explore socioeconomic demographics – median income, percent living below poverty, and educational attainment – within and across Geomarkets over time in Figure 7. The primary takeaways are the City of Chicago – and to a somewhat lesser degree the South and Southwest Suburbs – stands out as having lower median income, higher poverty rates, and lower BA attainment than other Geomarkets. By contrast, the North Shore stands out as being particularly affluent and highly educated. Overall, median income both within and across Geomarkets in the Chicago metropolitan area remained relatively stable across the three census periods. For instance, in 1980, median household income across the Chicago metropolitan area ranged from a high of $165,000 in the North Shore Geomarket to a low of $68,000 in the City of Chicago Geomarket. However, median incomes in all other Geomarkets were closer to the upper thresholds of affluence, such as $125,000 in Evanston and Skokie, $117,000 in the Northwest Suburbs, $108,000 in the Western Suburbs, $104,000 in Chain of Lakes, and $99,000 in the South and Southwest Suburbs. While overall income patterns across Geomarkets remained consistent through 2020, a slight median income increase for the City of Chicago Geomarket ($71,000) and decrease in the South and Southwest Suburbs Geomarket ($84,000) narrowed the income gap between the lowest and second-lowest Geomarkets to just $13,000. Figure 7 also shows that from 1980 to 2020, the percentage of residents living in poverty declined in the City of Chicago Geomarket while it increased in all other Geomarkets, resulting in a narrowing of poverty gaps across the Chicago metropolitan area. Similar patterns are evident in educational attainment across Geomarkets. In 1980, the City of Chicago and South and Southwest Suburbs had the lowest percentage of residents with a bachelor's degree or higher (14%). However, by 2020, the City of Chicago experienced the largest percentage point increase in college degree attainment, reaching nearly 40%.

Overall, demographic shifts transformed the Chicago-land Geomarket landscape from largely a binary structure into a more tiered dynamic that retains substantial racial/ethnic and socioeconomic inequality between Geomarkets. The original Geomarket boundaries delineated significant racial and income segregation in the 1980s. This included Black, Hispanic, and Asian residents primarily in the City of Chicago Geomarket and White residents in Geomarkets forming the suburban ring of the metropolitan area. Income, poverty, and educational attainment also created a static socioeconomic divide in 1980, with southern Geormarkets (City of Chicago and South and Southwest Suburbs) characterized by lower affluence and educational attainment compared to the higher socioeconomic standing of the northwestern suburbs. However, declining White populations, Hispanic and Asian population growth, and Black out-migration, as well as narrowing of education and poverty gaps within some Geomarkets, have contributed to a more stratified Geomarket dynamic in the metropolitan area with significant disparities remaining between Geomarkets. The City of Chicago Geomarket remains the critical geographical hub across all Geomarkets for racial/ethnic diversity; however, other Geomarkets now also include substantive numbers of residents across racial/ethnic groups (such as South and Southwest Geomarket for Black residents, Chain of Lakes for Hispanic residents, and Northwest Suburbs for Asian residents). While some stark socioeconomic gaps within Geomarkets have narrowed since 1980, significant geographical income segregation remains between Geomarkets despite a more clearly defined middle-to-high socioeconomic Geomarket group (Chain of Lakes, Northwest Suburbs, Western Suburbs).

**Composition of Student Lists by Geomarkets**

To examine how the socioeconomic and racial composition of included versus excluded prospects varies when student list purchases filter on particular Geomarkets, we analyze six orders – placed by a research university – that target prospects across the entire state of Illinois. The first two orders filter on the same SAT thresholds (ranging from 1020 to 1150), the same GPA thresholds (ranging from B- to A+), and filter on the high school graduating classes of 2019/2020 and 2020/2021, respectively. The remaining four orders used the same GPA and high school class filters but differed in SAT thresholds. Two orders targeted students with relatively middle-range scores (1160 to 1300) and two orders targeted those with relatively high scores (1310 to 1600). Because all six orders included an Illinois state filter, we can utilize prospects from the Chicago metropolitan area within the resulting student lists to examine Geomarket contributions by simulating the application of a Geomarket filter across low-, middle-, and high-range SAT threshold orders.

Figure 8 presents the racial/ethnic composition of purchased student profiles within Geomarkets in the Chicago metropolitan area, notably revealing that the City of Chicago stands out distinctly from all other Geomarkets in terms of the substantial racial and ethnic diversity of its purchased profiles.

At the lowest SAT score thresholds (top panel, 1020-1150) the racial/ethnic distribution of prospects within Geomarkets most closely reflects the demographic makeup of the metropolitan area in 2020 (see Figure 6). For example, more than 7,000 of the 15,964 total prospects (whose race/ethnicity is known) in this SAT range order were concentrated in the Western Suburbs (59% White, 9% Asian, 6% Black, 23% Hispanic, 4% Multiracial) and South and Southwest Suburbs (54% White, 2% Asian, 19% Black, 21% Hispanic, and 3% Multiracial), both of which yield majority White prospect pools comparable to the racial/ethnic makeup of their overall Geomarket populations. The North Shore (70% White, 12% Asian, 4% Black, 11% Hispanic) and the Northwest Suburbs (59% White, 11% Asian, 2% Black, 24% Hispanic) Geomarket pools exhibit the greatest proportions of White prospects. Unlike any other Geomarket, nearly eight of every 10 prospects from the City of Chicago Geomarket were Hispanic (47%), Black (24%), Asian (5%), or Multiracial (3%).

As SAT score ranges increase, the City of Chicago Geomarket becomes the only Geomarket with substantial shares of Black and Hispanic prospects. Moving from the top to the bottom panels in Figure 8, all Geomarkets experience growth in the shares of White and Asian prospects, while Black and Hispanic prospect shares decline. For the highest SAT score range, the City of Chicago is the only Geomarket where White prospects comprise a slim majority (56%), with the remaining prospect pool made up of 16% Asian, 6% Black, 18% Hispanic, and 5% Multiracial students. Prospect pools from all other Geomarkets in this high SAT order are predominantly White and Asian, with Black and Hispanic shares ranging from 0.3% (North Shore) to 4% (South and Southwest Suburbs) of prospects and 5% (North Shore) to 10% (South and Southwest Suburbs) of prospects, respectively.

We also explore contributions across Geomarkets to overall prospect pools by race/ethnicity. Figure 9 shows the contributions of each Geomarket to the pool of prospects from the two orders targeting middle-range SAT scores (1160-1300) by race/ethnicity. The top panel presents the full distribution of the 13,729 total prospects (whose race/ethnicity is known) included in these orders by Geomarkets: 10% of all purchased prospect profiles reside in Chain of Lakes, 11% in Northwest Suburbs, 6% in North Shore, 4% in Evanston and Skokie, 23% in City of Chicago, 29% in Western Suburbs, and 17% in South and Southwest Suburbs. These overall distributions are compared to those in the lower panels of Figure 9, which illustrate the proportional representation of each race/ethnicity group across Geomarkets.

Figure 9 shows which Geomarkets send the bulk of students for a given race/ethnicity. The City of Chicago accounts for 23% of all prospects but 51% of Black students. Similarly, South and Southwest Suburbs accounts for 17% of all prospects but 26% of all Black prospects. The upshot is that if the student list purchases had excluded these two Geomarkets, 77% of the nearly 1,000 Black students currently included in these purchases would be excluded. On the other hand, all other Geomarkets contributed substantially smaller shares of Black prospects. For example, the Northwest Suburbs accounts for about 11% of all prospects but only about 3% of Black prospects. The City of Chicago Geomarket also contributed larger shares of the more than 2,500 Hispanic students (45% versus 23% for all prospects), while all other Geomarkets contributed comparable or relatively smaller proportions of Hispanic prospects than their overall representation among all included prospects.

In contrast, these representational patterns by Geomarket are reversed for White and Asian students. Figure 9 shows nearly all Geomarkets contributed to the pool of White prospects at rates proportional to their overall contributions to all included prospects (except City of Chicago). Similar patterns are observed for Asian students, with proportional contributions by Geomarket largely consistent with their overall distribution, except for Northwest Suburbs (17% versus 11% for all prospects) and Evanston and Skokie (9% versus 4% for all prospects) contributing disproportionately larger shares and the South and Southwest Suburbs (6% versus 17% for all prospects) contributing smaller shares relative to their overall contribution to the larger prospect pool.

These contribution patterns by race/ethnicity demonstrate that student list orders targeting prospects in the Chicago metropolitan area using Geomarket filters that exclude the City of Chicago – as observed in some orders collected via our public requests – result in the loss of substantially disproportionate shares of Black and Hispanic students from the resulting prospect pools. The Online Appendix shows that this pattern holds across both lower-range (1020-1150) and higher-range (1310-1600) SAT score filters. In both cases, the City of Chicago contributes disproportionately larger shares of Black and Hispanic prospects compared to all other Geomarkets.

We also examine the contributions of each Geomarket to the pool of prospects based on first-generation college student status. Similar to race/ethnicity figures above, Figure 10 presents the Geomarket distribution for the 13,848 total prospects (whose first-generation status is known) included in the middle-range SAT score orders (1160-1300). Among the 1,750 prospects whose parents did not attend college ("no college"), the City of Chicago stands out as the only Geomarket contributing a disproportionately larger share of first-generation college students – 43% in comparison to 23% for all included prospects in this SAT range. In contrast, all other Geomarkets contributed disproportionately smaller shares of first-generation college students whose parents did not attend college relative to all included prospects, with such disparity ranging from a one-percentage-point difference (Northwest Suburbs, South and Southwest Suburbs) to a ten-percentage-point difference (Western Suburbs). For prospects whose parents attended some college but did not complete their degrees ("some college"), Figure 10 indicates both City of Chicago and South and Southwest Suburbs Geomarkets contribute a disproportionately larger share of first-generation students. These representational patterns are reversed among prospects that are not first-generation college students. The City of Chicago Geomarket contributed a substantially smaller share of not-first generation prospects (18%) relative to all prospects (23%) in this SAT range, whereas all other Geomarkets contributed comparable shares.

Finally, we examine Geomarket contributions to prospect pools by both first-generation college student status and race/ethnicity in the Chicago metropolitan area. Figure 11 presents the Geomarket distribution of the 13,059 total prospects (whose first-generation status and race/ethnicity is known) included in the middle-range SAT (1160-1300) orders. Figure 11 shows the City of Chicago Geomarket consistently contributes a disproportionately larger share of first-generation college students whose parents did not attend college across all racial/ethnic categories. For instance, among all Asian prospects (n=1,370) whose contact information was purchased in this SAT score range, approximately 12% were first-generation college students whose parents did not attend college, 16% were first-generation college students whose parents attended some college but did not complete their degree, and 73% had parents with a BA. By contrast, for 292 Asian students from the City of Chicago Geomarket, 30% had parents who did not attend college and approximately 50% had parents with a BA.

For Hispanic prospects, the City of Chicago stands out as the Geomarket contributing the largest overall number and share of Hispanic first-generation college students across those whose parents did not attend college (44% versus 38% overall) and those whose parents attended some college but did not complete their degrees (29% versus 26% overall). Other Geomarkets such as Northwest Suburbs and Chain of Lakes also contributed substantially larger shares. By contrast, Hispanic students from the more affluent North Shore Geomarket and Evanston and Skokie Geomarket were more likely to have parents with a BA.

Although Black prospects who are first-generation college students make up a relatively smaller portion of the overall pool, Figure 11 suggests that such prospects are primarily contributed by the Evanston and Skokie, City of Chicago, and South and Southwest Suburbs Geomarkets. Similarly, Geomarkets such as Chain of Lakes, North Shore, Evanston and Skokie, and Western Suburbs contribute a disproportionately larger share of White prospects who are not first-generation college students, whereas City of Chicago and South and Southwest Suburbs contribute larger shares of White first-generation college students.

In summary, the enduring segregated structure of the Chicago metropolitan area from 1980 to 2020 plays a significant role in how Geomarkets shape the racial/ethnic and socioeconomic composition of prospect pools. Reflecting the larger metropolitan area dynamics, Geomarkets that are predominantly comprised of low-income communities of color such as the City of Chicago and South and Southwest Suburbs consistently contribute disproportionately larger shares of Black, Hispanic, and first-generation college prospects, especially at lower SAT score thresholds. On the other hand, predominantly White and affluent Geomarkets in the suburban ring of the metropolitan area like the Chain of Lakes, North Shore, and Evanston and Skokie contribute larger proportions of White and Asian prospects, as well as non-first generation college students. However, the City of Chicago stands out in playing a crucial role in fostering college access for first-generation Asian and Hispanic college students, contributing such prospects to the overall pool at significantly higher, disproportionate rates than any other Geomarket. Given the broader racial/ethnic distribution of the Chicago metropolitan area, student list orders that use Geomarket filters excluding the City of Chicago and including more affluent Geomarkets in targeting high-achieving students may still capture some Asian and Hispanic prospects. However, these students are far less likely to be first-generation students of color.

### 5.2 Dallas-Fort Worth

#### Suburban Geomarket Growth Eases Racial Disparities, Concentrated Wealth Persists

The six Geomarkets within the Dallas-Fort Worth metropolitan area are illustrated in Figure 12 by total population across the three census periods. The Dallas-Fort Worth region represents a unique "metroplex" case formed by the convergence of two separate metropolitan areas, which creates a unique urban spatial structure that differs from the classic urban core and suburban ring model showcased by Chicago. As depicted in Figure 12, two distinct inner-core Geomarkets (City of Dallas and City of Fort Worth) are separated by a centrally located suburban Geomarket encompassing Irving, Arlington, and Grand Prairie. These central Geomarkets are surrounded by the more conventional outer ring of suburban Geomarkets (Dallas County, Collin and Rockwall Counties, and West of Dallas/Fort Worth Metroplex).

From 1980 to 2020, the population of the overall Dallas-Fort Worth metropolitan area increased dramatically from approximately 2.7 Million to 6.8 Million. Figure 12 shows all six Geomarkets experienced population growth over the three census periods. However, the most pronounced increases occurred in the West of Dallas/Fort Worth Metroplex and Collin and Rockwall Counties Geomarkets, which grew from approximately 350,000 to 1.6 Million and 160,000 to 1.1 Million, respectively. These population trends, as shown in Figure 12, suggest population growth has been partly driven by urban sprawl from the metropolitan centers – City of Fort Worth and City of Dallas – into the surrounding suburban and rural areas of the metropolitan areas. This outward expansion may also account for the comparatively larger geographic size of the West of Dallas/Fort Worth Metroplex Geomarket from 1980 to 2000.

Figure 13 presents the racial and ethnic composition of each Geomarket within the Dallas-Fort Worth metropolitan area across the three census periods. In 1980, all Geomarkets were predominantly White, with White populations ranging from 61% in City of Dallas to 93% in West of Dallas/Fort Worth Metroplex. By 2000, however, Geomarkets located in the central part of the metropolitan area exhibited a more balanced racial/ethnic distribution, including Dallas County, Irving, Arlington, and Grand Prairie, and City of Fort Worth. In contrast, Collin and Rockwall Counties and West of Dallas/Fort Worth Metroplex Geomarkets remained predominantly White in 2000 at 77% and 80%, respectively. By 2000, the City of Dallas was no longer a predominantly White Geomarket, with a population that was approximately 25% Black, 36% Hispanic, and 2% Asian. Over the next two decades, the Hispanic population grew steadily, contributing to greater diversity. By 2020, the City of Dallas Geomarket had the lowest proportions of White and Asian residents across the Dallas-Fort Worth Metropolitan area, with approximately 23% Black and 42% Hispanic residents.

These shifts in racial and ethnic composition also became more pronounced for other Geomarkets between 2000 and 2020. The Collin and Rockwall and West of Dallas/Fort Worth Metroplex Geomarkets experienced continued declines in their share of White residents. However, both Geomarkets remained predominantly White (57% and 63%, respectively). In contrast, all other Geomarkets experienced continued increases in the shares of Black and Hispanic residents, resulting in a diminished previously White-majority populations for Dallas County (31% White, 25% Black, 34% Hispanic; 8% Asian), Irving, Arlington, and Grand Prairie (29% White, 19% Black, 38% Hispanic; 11% Asian), and City of Fort Worth (42% White, 16% Black, 35% Hispanic; 4% Asian).

Socioeconomic demographics of median income, percent living below poverty, and educational attainment over time are presented in Figure 14 for the Dallas-Fort Worth metropolitan area. Overall, median income both within and across Geomarkets remained relatively stable across the three census periods. In 2020, both City of Dallas and City of Fort Worth continued to rank at the lowest levels of affluence within the metropolitan area at $68,000 and $74,000 median household incomes, respectively. By contrast, Collin and Rockwall Counties remained the most affluent Geomarket with a median household income of $124,000, followed by West of Dallas/Fort Worth Metroplex ($104,000). Figure 14 also illustrates changes in the percentage of residents living in poverty and educational attainment within Geomarkets over time, revealing similar dynamics across Geomarkets to those observed in median household income. Together, these trends underscore the growing racial/ethnic diversity but widening economic inequality within the evolving Geomarket landscape of the Dallas-Fort Worth metroplex.

**Composition of Student Lists by Geomarkets**

We analyze how the racial/ethnic and socioeconomic composition of included versus excluded prospects varies when filtering on particular Geomarkets within the Dallas-Fort Worth metropolitan area. Specifically, we analyze three list orders placed by a research university targeting prospects across 87 different Geomarkets nationwide, including all six Geomarkets in the Dallas-Fort Worth Metropolitan area. Each order also filtered for prospects in the 2020 graduating class and on PSAT scores across three ranges: 1070-1180, 1190-1260, and 1270-1520.

Figure 15 presents the racial/ethnic composition of purchased student profiles within Geomarkets in the Dallas-Fort Worth area, revealing a comparatively more distributed pattern of racial/ethnic diversity across Geomarkets unlike Chicago's concentrated diversity within a single Geomarket. This distribution is most notable at the lowest PSAT score threshold. For instance, at the 1070-1180 PSAT score range, the central suburban Geomarkets of Dallas County (37% White, 17% Asian, 15% Black, 27% Hispanic) and Irving, Arlington, and Grand Prairie (34% White, 13% Asian, 15% Black, 33% Hispanic) have the greatest proportions of non-White prospects, which is largely attributed to their greater shares of Asian and Black prospects. While the urban central Geomarkets of City of Dallas and City of Fort Worth are the most racially/ethnically diverse in the larger metropolitan area (see Figure 13), their prospect pools in Figure 15 includes larger proportions of Hispanic prospects but lower proportions of Asian and Black students. The Collin and Rockwall Counties (58% White, 7% Black, 17% Hispanic, 13% Asian) and West of Dallas/Fort Worth Metroplex Geomarkets (60% White, 7% Black, 18% Hispanic, 9% Asian) both yield the largest but predominantly White prospect pools. Prospect pools within Geomarkets from the middle range PSAT score (1190-1260) to the high PSAT score range (1270-1520) become disproportionately more Asian, while the representation of Hispanic and Black prospects decline substantially.

Figure 16 presents the race/ethnicity of all prospects targeted in the middle PSAT score range across Geomarket contributions. The bars for "All (race known)" show that Collin and Rockwall (30%) and West of Dallas/Fort Worth (27%) contribute many more students than the other four Geomarkets. Disparities in Geomarket contributions were evident but differed for Asian, Black, and Hispanic students. For instance, The Dallas County (19% versus 12% for all prospects) and Collin and Rockwall Counties (39% versus 30% for all prospects) Geomarkets contributed a disproportionately larger share of Asian students, whereas nearly all other Geomarkets contributed disproportionately smaller shares.

Representational patterns across Geomarkets are reversed for Black and Hispanic prospects in Figure 16. All Geomarkets besides Collin and Rockwall Counties and West of Dallas/Fort Worth Metroplex contributed nearly equal or larger proportions of Black and Hispanic prospects relative to their contributions to the overall prospect pool in the Dallas-Fort Worth metropolitan area. However, the City of Dallas Geomarket contributed the largest disproportionate share of Hispanic prospects (19% versus 12% for all prospects), whereas the Irving, Arlington, and Grand Prairie Geomarket contributed the largest disproportionate share of Black prospects (15% versus 7% for all prospects).

Contributions of each Geomarket to the pool of prospects based on first-generation college student status from the Dallas-Fort Worth metropolitan area are presented in Figure 17. Only 303 prospects of the 3,926 total (whose race/ethnicity is known) in this middle PSAT score range pool were first-generation college students whose parents did not attend college. Again, all Geomarkets besides Collin and Rockwall Counties (12% versus 30% for all prospects) and West of Dallas/Fort Worth Metroplex (16% versus 27% for all prospects) contributed equal or larger proportions of first-generation college students (no college) relative to their contributions to the overall prospect pool. Disproportionately larger contributions range from five percentage points by the City of Fort Worth (16% versus 11% for all prospects) to 10 percentage points by the City of Dallas Geomarket (22% versus 12% for all prospects). Contributions become more proportional for the pool of first-generation college students whose parents attended but did not complete their degree. Proportional balances across Geomarkets are also evident for the 3,067 prospects that are not first-generation college students.

Lastly, Figure 18 presents Geomarket contributions to prospect pools by both first-generation college student status and race/ethnicity, which we view as the most interesting results for the Dallas-Fort Worth metropolitan area. Among all White prospects (n=2,093) whose contact information was purchased in this middle PSAT score range, approximately 3% were first-generation college students whose parents did not attend college, 11% were first-generation college students whose parents attended some college but did not complete their degree, and 86% were not first-generation college students. Figure 18 shows almost all Geomarkets contribute a nearly proportionate share of prospects by first-generation college students relative to the overall pool, the only modest exception being the Irving, Arlington, and Grand Prairie Geomarket.

By contrast, for non-White groups, a larger share of prospects were first-generation students. Furthermore, first generation status differed substantially across Geomarkets, especially for Hispanic and Asian students. For Hispanic students, more than 50% of prospects from the City of Dallas, City of Fort Worth, Irving, Arlington and Grand Prairie, and Dallas County are first-generation college students. By contrast, for the more affluent Collin and Rockwall Geomarket and the West of Dallas/Fort Worth Metroplex Geomarkets, less than 30% of prospects were first-generation. For Asian prospects, City of Fort Worth and Irving, Arlington, and Grand Prairie Geomarkets contributed disproportionately larger shares of first-generation college students across both parents with no college and some college. By contrast, the more affluent Collin and Rockwall sent more Asian prospects (234) than any other Geomarket, but about 89% of these prospects were not first-generation college students.

### 5.3 Los Angeles

#### Decentralized Diversity in a Multi-Nuclei Landscape

Due to the large number of Geomarkets in urban Southern California, we distinguish Los Angeles (11 Geomarkets) – results shown here – from Orange County (four Geomarkets) and San Diego (three Geomarkets), results for which are shown in the Online Appendix. The 11 Geomarkets within the Los Angeles metropolitan area are illustrated in Figure 19 by total population across the three census periods. The Los Angeles metropolitan area is one of the only multi-nuclei models in the country that is formed when urban growth occurs around multiple centers that serve different functions rather than expanding from a single central business district core (Harris and Ullman 1945). As depicted in Figure 19, Geomarkets closely follow this overall metropolitan model. Hollywood and Wilshire, South and South Central Los Angeles, and East Los Angeles Geomarkets are centrally located in the metropolitan area followed by an outer core ring made up of the Glendale and Pasadena, San Fernando Valley-East, West Los Angeles and West Beach, South Bay, and Long Beach Geomarkets. The San Fernando Valley-West Geomarket is located in the northwest part of the region, whereas Covina and West Covina as well as the Riverside, San Bernardino, and Ontario Geomarkets are located inland, comprising the Inland Empire.

The population growth in the Los Angeles metropolitan area increased from approximately 8 Million in 1980 to more than 12 Million in 2020. Figure 19 showcases population growth by Geomarkets over the three census periods, reflecting the urban sprawl forming the multi-nuclei model. The most pronounced increases occurred in the South and South Central Los Angeles Geomarket, as well as the outer/inland Geomarkets of San Fernando Valley-West and Riverside, San Bernardino, and Ontario. The South and South Central Los Angeles Geomarket grew from approximately 1.1 to 1.6 Million from 1980 to 2020. In contrast, the San Fernando Valley-West and Riverside, San Bernardino, and Ontario Geomarkets, respectively, grew from approximately 580,000 to nearly 1 Million and from 1.1 Million to 3.3 Million during this time.

Figure 20 presents the racial and ethnic composition of each Geomarket within the Los Angeles metropolitan area from 1980 to 2020. As mentioned previously, a data caveat here is that the Asian race/ethnicity category was not available in 1980. However, over the three decennial census periods, the proportion of White residents declined across all Geomarkets, while Asian and Hispanic populations generally increased. In 1980, all Geomarkets – except for East Los Angeles, South and South Central Los Angeles, and Hollywood and Wilshire – were predominantly White, with percentages ranging from 63% White in Covina and West Covina to 86% in the San Fernando Valley-East Geomarket. By 2020, however, only the West Los Angeles and West Beach (51%) and San Fernando Valley-East (59%) Geomarkets remained predominantly White. The remaining Geomarkets exhibited a more racially/ethnic diverse population by 2020, including San Fernando Valley-West (40% White, 4% Black, 40% Hispanic; 13% Asian), Glendale and Pasadena (35% White, 4% Black, 41% Hispanic; 17% Asian), South Bay (32% White, 10% Black, 32% Hispanic; 22% Asian), Long Beach (27% White, 10% Black, 47% Hispanic; 13% Asian), Covina and West Covina (19% White, 4% Black, 57% Hispanic; 19% Asian), and Riverside, San Bernardino, and Ontario (28% White, 7% Black, 53% Hispanic; 8% Asian). However, the South and South Central Los Angeles and Hollywood and Wilshire Geomarkets experienced substantial declines in the share of Black residents from 1980 to 2020, dropping from 44% to 16% and 22% to 11%, respectively.

Figure 21 presents trends in income, poverty, and educational attainment in the Los Angeles metropolitan area from 1980 to 2020, demonstrating the South and South Central Los Angeles, East Los Angeles, and Hollywood and Wilshire Geomarkets remained at the lowest socioeconomic levels across these indicators. For instance, the median household income in the South and South Central Los Angeles Geomarket increased modestly from $51,000 in 1980 to $56,000 in 2000, reaching $61,000 by 2020. In comparison, East Los Angeles and Hollywood and Wilshire recorded slightly higher median incomes in 2020 at $74,000 and $70,000, respectively. These figures remain substantially lower than that of the most affluent Geomarket, West Los Angeles and West Beach, where the median household income reached $117,000 in 2020.

### Composition of Student Lists by Geomarkets

To examine how the socioeconomic and racial composition of included versus excluded prospects varies when student list purchases filter on particular Geomarkets in the Los Angeles metropolitan areas, we analyze six orders placed by a research university that filtered for prospects across the entire state of California, in the 2020 through 2022 high school graduating classes, and by PSAT scores. Two orders indicated PSAT score thresholds ranging from 1070-1180, another two orders filtered for a 1190-1260 PSAT range, and the remaining two orders filtered for scores ranging from 1270-1520.

Figure 22 presents the racial/ethnic composition of purchased student profiles. For the nearly 15,000 prospects in the resulting student lists at the low PSAT score thresholds (1070-1180), South and South Central Los Angeles (n=1,035) and East Los Angeles (n=1,076) Geomarket pools stand out as having an overwhelming majority of students of color, 98% and 95% respectively. Relative to other Geomarkets in the Los Angeles metropolitan area, such patterns are attributed to both South and South Central Los Angeles (88%) and East Los Angeles (61%) having substantially greater proportions of Hispanic prospects, as well as a larger proportion of Black prospects for South and South Central Los Angeles (5%) and larger share of Asian prospects for East Los Angeles (31%). Even among higher PSAT score ranges, South and South Central Los Angeles remains the most prominent Geomarket with significant proportions of Hispanic, Asian, and Black student prospects.

Figure 23 shows how much each Geomarket accounts for the share of prospects in the middle PSAT score range by race/ethnicity. South and South Central Los Angeles accounts for only 3%

of all purchased prospect profiles but accounts for 8% of Black prospects. The Riverside, San Bernardino, and Ontario Geomarket accounts for 24% of all prospects but 29% of Black prospects. For Hispanic prospects, South and South Central Los Angeles (9% versus 3% for all prospects) and Riverside, San Bernardino, and Ontario (30% versus 24% for all prospects) also contributed a disproportionately larger proportion of Hispanic prospects relative to their contributions to the overall pool in the Los Angeles metropolitan area. For Asian students, the Glendale and Pasadena (20% versus 15% for all prospects), East Los Angeles (14% versus 7% for all prospects) and Covina and West Covina (12% versus 9% for all prospects) Geomarkets contributed a disproportionately larger share of Asian prospects relative to their contributions to the overall prospect pool. For the more than 2,300 White prospects in this pool, the San Fernando Valley - West (17% versus 12% for all prospects), West Los Angeles and West Beach (9% versus 6% for all prospects), and Glendale and Pasadena (17% versus 15% for all prospects) Geomarkets contribute disproportionately larger shares of White prospects.

These results demonstrate that following the Zemsky and Oedel (1983) strategy of filtering on high-income Geomarkets and excluding low-income Geomarkets – such as the South and South Central Los Angeles Geomarket and the Riverside, San Bernardino, and Ontario Geomarket – would result in the disproportionate exclusion of Black and Hispanic students from the resulting prospect pools. Similarly, using a Geomarket filter excluding East Los Angeles would result in the loss of disproportionate shares of Asian students.

Figure 24 presents Geomarket contributions by first-generation college student status, demonstrating the disproportionately larger share of first-generation college students contributed to the overall pool by East Los Angeles, South and South Central Los Angeles, and Riverside, San Bernardino, and Ontario Geomarkets. For students whose parents did not attend college, East Los Angeles (18% versus 7% of all prospects), South and South Central Los Angeles (11% versus 3% of all prospects), and Riverside, San Bernardino, and Ontario (27% versus 24% of all prospects) Geomarkets provided a disproportionately larger contribution of first-generation college students.

Lastly, Figure 25 presents Geomarket contributions to prospect pools by both first-generation college student status and race/ethnicity in the Los Angeles metropolitan area for the middle range SAT orders. Across all Hispanic prospects, 26% were first-generation college students whose parents did not attend college, 26% were first-generation college students whose parents attended some college but did not complete their degrees, and 49% were not first-generation college students. By contrast, for the 175 Hispanic prospects from the South and South Central Geomarket, 51% had parents who never attended college, 28% had parents who attended some college, and only 21% had parents with a BA. Similarly, Hispanic prospects from East Los Angeles and from Riverside, San Bernardino, and Ontario were more likely to have parents without a BA compared to all Hispanic prospects in the Los Angeles metro area.

Among all Asian prospects whose contact information was purchased, approximately 11% were first-generation college students whose parents did not attend college, 22% were first-generation college students whose parents attended some college but did not complete their degree, and 67% were not first-generation college students. Figure 25 highlights the East Los Angeles Geomarket as a key contributor of Asian prospects (n = 276), a significant proportion of whom are first-generation college students (34%) whose parents did not attend college. The Hollywood and Wilshire Geomarket also contributed a modest number of Asian first-generation college students (some college) at a disproportionately larger rate (34%).

While the pool of Black prospects is relatively small (n=182) in Figure 25, some patterns emerge by first-generation status and Geomarket. Overall, approximately 6% of Black prospects were first-generation college students whose parents did not attend college, 21% were first-generation college students whose parents attended some college, and 73% were not first-generation. The South and South Central Los Angeles Geomarket stands out for contributing a disproportionately high share (13%) of Black first-generation college students whose parents did not attend college, despite representing a relatively smaller overall number. Additionally, the Long Beach Geomarket is notable for its higher proportion (42%) of Black first-generation students whose parents attended college but did not complete a degree.

# 6 Discussion

The discussion of homophily by Chun (2021) describes not only the logic of the Market Segment Model but also a broader institutional logic of the enrollment management industry. The logic of homophily assumes that there is a natural cultural match between students and colleges, suggesting colleges should find prospective students who resemble students who enrolled at their institution in the past and focus recruiting efforts on the same schools and communities that are popular with peer colleges. We suggest that homophily is the operating logic most consistent with college recruiting behavior observed in the wild. For example, Author (XXXXd) find great overlap in the sets of private schools visited by the University of South Carolina and the University of Alabama. As Chun (2021) argues, observed homophily is not the result of voluntary action; rather, homophily is programmed into algorithms that create connections.

The Market Segment Model begins with the status attainment model idea that student demand for college is a function of parental education and parental income. Zemsky and Oedel (1983) illustrated this argument by demonstrating a correlation between student SAT score-sending behavior and measures of class. The thesis of the Market Segment Model is that households of specific class characteristics and college aspirations are likely to live in particular communities and that these geographic territories can be meaningfully captured by Geomarkets. Although Geomarkets appear coarse by contemporary standards, they were a landmark innovation in Geographic segmentation. The Market Segment Model recommends that colleges should identify their core student market segment (i.e., local, in-state, regional, or national), identify Geomarkets that contain large numbers of households from this market segment, and then target high schools and communities within these Geomarkets. This information is contained in the Market Segment Profile (Online Appendix), a standard output in EPS software. When considering new Geomarkets to target, EPS recommends colleges utilize the logic of homophily by identifying Geomarkets that are popular with peer colleges. This information is contained in the Institutional Profile (Online Appendix).

Research question 1 investigates racial and socioeconomic variation between and within Geomarkets over time. We find that Geomarkets are highly correlated with race and class. Findings illustrate that in 1980, when Geomarkets were being created, Black and Hispanic residents tended to be highly concentrated in the poorest Geomarket in the Chicago, Dallas-Fort Worth, and to some extent the Los Angeles metropolitan areas. These Geomarket patterns extend to other metropolitan areas in the country, such as CA7 - City of Oakland, MA 6 - Boston and Cambridge, OH 4 - City of Cleveland (East), TX17 - City of Houston (East), and PA5 - Philadelphia County. In 2020

Chicago, Black residents remain concentrated in the poorest Geomarket but Gentrification also pushed Black residents to the South and Southwest Suburbs Geomarket. In Dallas-Fort Worth, Black and Hispanic communities are distributed across multiple Geomarkets; however, significant socioeconomic disparities persist between Geomarkets. The Los Angeles metropolitan area presents another distinctive Geomarket landscape in which Black, Hispanic, and Asian communities across a range of socioeconomic levels are distributed across multiple Geomarkets rather than concentrated in one or two.

Research question 2 analyzes the racial and socioeconomic composition of student list purchases. We examine purchases that include all Geomarkets in a metropolitan area in order to assess the potential consequences of including or excluding particular Geomarkets from the purchase. This question is motivated by the Zemsky and Oedel (1983) recommendation that selective colleges that enroll students from the regional and national market segment should target affluent Geomarkets, while community colleges and non-selective 4-year colleges should focus on middle- and working-class Geomarkets. Findings across metropolitan cases demonstrate that if student list purchases exclude working- and middle-class Geomarkets – such as IL11 - City of Chicago, TX19 - City of Dallas, TX20 - City of Fort Worth, CA19 - East Los Angeles, and CA21 - South and South Central Los Angeles – the resulting student list purchases would disproportionately exclude Black and Hispanic students who meet the requisite academic achievement criteria. Although these purchases would still capture some Black, Hispanic, and Asian prospects, these prospects are much less likely to be first-generation students.

This article adds a new perspective to existing explanations about how students are sorted into colleges. The status attainment model argues that college destination is a function of parental education and occupation (Sewell and Shah 1967, 1968b, 1968a). We conceive of the Market Segment Model as the status attainment model applied to geodemographic market research, akin to the rise of credit-scoring in UK retail banking (Leyshon and Thrift 1999). Fishman (2020) finds that Asian American students whose parents are immigrants tend to have high educational achievement even if their parents do not, while the educational achievement of later-generation Asian Americans conforms to status attainment theory. In our analyses of prospects included in student list purchases, Asian American prospects from poor Geomarkets tended to be first-generation students while Asian American prospects from affluent Geomarkets tended to have parents with a BA. Building on Fishman (2020), Asian students living in low-income Geomarkets may tend to have parents who immigrated to the U.S. whereas Asian students living in affluent Geomarkets are more likely to have parents born in the U.S.

The cultural capital model explains a process by which upper and upper-middle class families sort themselves into selective colleges by providing their children with the pedigree (academic opportunities, extracurriculars) valued by selective colleges, information about how to navigate the admissions gauntlet, and social networks that provide an inside track (Bourdieu 1984, 1988; Huang 2023; McDonough 1997). This model explains how affluent households maintain a disproportionate enrollment share at selective institutions in an era of holistic admissions ostensibly designed to increase racial and class diversity (Huang 2023). Both the status attainment model and the cultural capital model are demand-side explanations for why education is a "social sieve" (Stevens et al. 2008; Jencks and Riesman 1968) that allows for a modicum of mobility while maintaining a larger flow of intergenerational class transmission (Labaree 1997). Chetty et al. (2020) – who obtained parental income data from federal income tax returns for every U.S. college – show that the disproportionate enrollment share of high income families at selective private colleges – and even at most public research universities – is staggering. These patterns are consistent with the observation by Weber (1948, 241–42) that educational credentials are property that "support their holders'…claims to monopolize socially and economically advantageous positions."

Supply-side explanations of credentialism and more recent scholarship on enrollment management complement the cultural capital model of sorting students to colleges. The credentialism literature recognizes that colleges have a financial incentive for educational credentials to determine the competition for socially and economically advantageous positions (Collins 1979; Labaree 1997; Larson 1977; Brint and Karabel 1989) and this credentials arms race favors affluent households. Scholarship on enrollment management describes which students colleges want and how colleges go about attracting these students at different stages of the enrollment funnel (Stevens 2007; Holland 2019; Killgore 2009; Khan 2011; Author XXXXa; Karabel 2005, 1984; Cottom 2017). Stevens (2007) and Khan (2010) describe a tacit arrangement between high school guidance counselors on the demand side and college admissions officers on the supply side. Counselors at well-resourced – especially private – schools are motivated to give their students a competitive advantage in admissions. Meanwhile, college admissions officers are motivated to enroll students who can afford full tuition price and are likely to donate in the future. These mutually beneficial desires are consummated by recognizing that upper and upper-middle class applicants satisfy the extracurricular needs of the college. The orchestra needs oboists. The lacrosse team needs players. Quantitative analyses of visits from colleges to high schools are consistent with Stevens (2007); selective private colleges and public research universities devote most of their recruiting resources on courting students from privileged schools and communities (Author XXXXd, XXXXa). However, these analyses conceive of recruiting as something that is done solely by individual colleges.

Existing explanations of how students are sorted into colleges miss an important supply-side mechanism – third-party vendors that sort students on behalf of colleges. Zemsky and Oedel (1983) created the Market Segment Model and Geomarkets based on a snapshot of student demand from 1980. Zemsky and Oedel (1983) argue that demand for higher education is correlated with class, ignoring the historical structural barriers that produced class- and race-based inequality in 1980 student demand. In itself, the Market Segment Model is a social science depiction. However, the College Board inscribed Geomarkets and the Market Segment Model into EPS software that tells colleges which Geomarkets and high schools to target. Technologies that target customers based on geography inevitably leverage racial and class segregation (Benjamin 2019; O'Neil 2016; Chun 2021). Furthermore, technologies that use the logic of predictive analytics – identifying correlates in past cases to make predictions for future cases – are prone to "pernicious feedback loops" (O'Neil 2016). The structural barriers (e.g., segregation, slavery, Jim Crow) that caused historic place-based inequality in student demand are amplified because EPS recommends that colleges focus recruiting resources on localities that already have high student demand for peer colleges. Geomarkets were subsequently incorporated into the Student Search Service product. We show that excluding low-income, non-White Geomarkets from student list purchases results in the disproportionate exclusion of first-generation, non-White students with strong academic achievement.

### 6.1 Enrollment Management Industry

The growing salience of third-party vendors warrants a reconfiguration of the organizational field salient to college access. DiMaggio and Powell (1983) [p. 148] define organizational fields as "those organizations that, in the aggregate, constitute a recognized area of institutional life: key suppliers, resource and product consumers, regulatory agencies, and other organizations that produce similar services or products" and that "the virtue of this unit of analysis is that it directs our attention…to the totality of relevant actors." Sociological literatures salient to college access devote substantive attention to schools, colleges, families, and communities. College access scholarship from economics often analyzes the effects of district, state, and federal policies (e.g., Fuller et al. 2023). These literatures ignore the role of third-party vendors, which are often owned by private equity and offer software-as-a-service platforms that perform core functions for schools and colleges.[21]

Author (XXXXb) describes four key dynamics in the enrollment management (EM) industry, with a focus on the market for student list data. First, EM consulting firms are central to the creation and implementation of recruiting campaigns (Marcus 2024). The top two EM consultancies are Ruffalo Noel Levitz, which claims to serve 1,900 colleges annually (Ruffalo Noel Levitz 2025), and EAB, which claims to serve 2,100 colleges (EAB 2025). College reliance on EM consultancies is partially explained by the high level of burnout and employee turnover in the EM profession (Hoover 2023). When we issued records requests to public universities about their student list purchases, at least 50% of universities indicated that they outsourced student list purchases to an EM consultancy. Often, the university was buying names each year but no employee had knowledge about which vendors they were buying from or which search filters they were utilizing.

Second, technological advances create new means of identifying and serving prospects. College Board and ACT leveraged data on test-takers to enjoy a near duopoly in the market for student list data, one that lasted for several decades. This business can be described as list-based lead generation based on the direct-mail model (Singer 1988). "Free" college search engines (e.g., Cappex, Niche) yielded new sources of student list data. Another data source is college planning software purchased by high school districts and utilized by high school students and guidance counselors. The most widely-known product is Naviance, which claims to be used by more than 10 million K-12

[21] Huang (2023) studies independent admissions consultants, third-party vendors on the demand-side.

students and by 40% of US high schools (PowerSchool 2021). Naviance college planning software feeds into the Intersect recruiting platform – an example of behavioral-based targeting – which allows colleges to target Naviance users while they are on the platform.

The third dynamic in the EM industry has been the growth of the test-optional admissions movement, which reduces both the number of paying test-takers and the coverage of student list products sold by College Board and ACT. ACT responded to uncertainty in their core testing business by attempting to become an edtech company (Blumenstyk 2019), failing (Molnar 2020), and being acquired by Nexus Capital in 2024 (Knox 2024). Over the last two decades, the testing companies developed new search filters that explicitly incorporated predictive analytics – for example College Board's Geodemographic Segment filters (College Board 2011b) – developed free college search engines, and entered the enrollment management consulting market more aggressively (Author XXXXb). College Board has weathered the test-optional movement thanks to robust revenues from the Advanced Placement product, but the decline in PSAT/SAT/PreACT/ACT test-takers has undermined their oligopoly in the market for student list data.

The fourth – and most important – dynamic is the transformation of EM from an owner-operated to private-equity owned industry and the resulting eroding distinction between consultancy and software vendor. The 1980s and 1990s was a period of market entry in EM, when college admissions professionals or professors hung up a shingle (Marcus 2024). The 2010s were a period of acquisitions and concentration. Drawing from resource dependence theory, Author (XXXXb) describe the proliferation of horizontal acquisitions to grow market share and vertical acquisitions designed to make customers more dependent on particular recruiting products. After regulatory scrutiny pushed private-equity interests out of the for-profit college market (Eaton 2022a), private equity investors found value in acquiring firms that provided services to direct providers. Direct providers are subsidized by state and federal funds (Title IV financial aid) and these funds are then paid to third-party vendors. Therefore, private equity ownership of direct providers and third-party vendors are both examples of government-subsidized rent-seeking.

EAB exemplifies the trend towards private equity ownership and industry concentration. EAB is the preeminent EM consultancy, offering software-as-a-service products along the domains of recruitment, pricing and financial aid, student success, and advancement (EAB 2025). The origins of EAB trace to Bill Royall, who founded Royall & Company in 1983 to provide direct marketing and fundraising for Republican political campaigns (Jump 2020). By 1995 EM became their primary focus. Royall was acquired by the health tech firm Advisory Board Company for $850 million in 2015 and then by Vista Equity Partners for $1.5 billion in 2017, renaming it EAB (Hansen 2017).

As an owner-operated consultancy, Royall was the market leader in purchasing names from student list vendors on behalf of clients. Under Vista, EAB entered the market for student list data by completing a series of acquisitions that targeted the "leads" and "inquiries" stages of the enrollment funnel. In 2019, EAB acquired YouVisit, a leading provider of virtual tours. YouVisit is an engine for "inquiries" in that taking a virtual tour shows interest in a college and leaves a trail of contact information for subsequent recruiting interventions. In 2020, EAB acquired Cappex, a leading free college search engine, which is an engine for leads in that students provide contact details and other information. In 2021, EAB and PowerSchool – the leading provider of K12 student information systems and a subsidiary of Vista – jointly acquired the EM consultancy/software vendor Hobsons. PowerSchool acquired Hobson's Naviance product and its sister recruiting product Intersect and EAB paid PowerSchool to become the exclusive re-seller of Intersect. EAB integrated these acquisitions to create the Enroll360 recruiting platform.[22]

Enroll360 exemplifies a new kind of student list product that leverages proprietary data to sell software. Historically, College Board and ACT sold prospect contact information at a price-per-prospect (e.g., $0.50 per name). The goal was to generate revenue from the sale of names. By contrast, Enroll360 wraps several proprietary databases of names (e.g., Cappex, Intersect) into a software-as-a-service platform that recruits those prospects at different stages of the enrollment funnel. Here, the goal of lists is to create demand for software. Colleges that are interested in recruiting these prospects must buy an Enroll360 subscription. Like College Board's Student Search Service, contemporary software-as-a-service recruiting products incorporate search filters that enable colleges to choose which prospects they recruit. For example, Feathers (2022) shows how the Intersect recruiting platform allows colleges to send paid advertisements to Naviance users.

[22]The Enroll360 press release states, "We spent the last couple of years creating a connected recruitment ecosystem that allows enrollment leaders to keep pace with students pursuing increasingly digital journeys to college. This work led us to join forces with several leading companies: Cappex, Intersect, Wisr, and YouVisit…Individually, each solution can solve important challenges at various stages of the enrollment funnel…By bringing these capabilities together, our vision is to reinvent how enrollment leaders reach their goals" (Koppenheffer 2021).

Intersect customers (colleges) control which Naviance users will receive recruiting messages by filtering on filters such as “academic ability,” intended majors, and whether students used Naviance to “research competitor institutions” (Feathers 2022). A University of Utah procurement document states the Intersect subscription is necessary because “there is a unique group of prospective students who are only in the PowerSchool Naviance platform” (Sole Source 2022).

An under-studied and under-regulated third-party EM industry has negative consequences for colleges and for students. Across multiple student list products, we see a growing number of filters that recommend which prospects to target based on the behavior of previous cohorts of students. The ACT/Encoura “Enrollment Predictor” filter enables colleges to filter prospects based on the predicted probability of enrollment, which is created based on analyzing the enrollment decisions of previous cohorts of students (Schmidt 2019). College Board Geodemographic Segment filters categorize each census tract into one of about two dozen clusters based on the college-going behavior of previous cohorts (College Board 2011b) and is based on the Claritas/PRIZM market segmentation system that categorized zip codes into groups useful for merchants (McKelvey 2022). An investor-facing Vista Equity Partners (2025) video promoting EAB’s “Pipeline Analytics” product states, “we are using artificial intelligence and machine learning to help schools build lists of students that may be more likely to enroll and persist in their university.” The product “identifies students that would be a good match for a school based on their similarity to students who have succeeded there before and then provides them [the students] information about those institutions so that they choose to apply and ultimately enroll.” Each of these products is based on the underlying logic of correlations and homophily discussed by Chun (2021) and applied by Zemsky and Oedel (1983). These products identify correlations in a historic snapshot of students and apply those relationships to make recommendations about which prospects to target in the present. As such, these technologies amplify the effect of previous inequality in student demand, which is itself a consequence of structural inequality in educational opportunity.

Market dynamics in the EM industry raise policy concerns about tuition prices, student choice, and competition. To the extent that colleges must buy expensive software to recruit prospects contained in an associated database, this expenditure will ultimately be passed on to students in the form of higher prices or reduced expenditure on education. Unfettered horizontal and vertical acquisitions in the EM industry cause the EM industry to become more concentrated. Smaller consultancies that do not control vertical inputs may be unable to compete with companies that have acquired proprietary pools of prospects. Fewer EM consultancies means less competition and higher consulting fees, which will be passed on to consumers. Finally, the blurring of the lines between consultancy and student list vendor suggests that large consultancies will funnel prospects only to those colleges that pay for subscriptions, raising policy concerns about student choice.

### 6.2 Developing a Knowledge Infrastructure

Regulation of the higher education industry should be informed by research that investigates the "totality of relevant actors" (DiMaggio and Powell 1983, 148) in the organizational field. Unfortunately, scholarship salient to education policy and regulation have ignored two, related economic transformations. First, third-party vendors offer software platforms that perform core tasks of organizations, including schools and colleges. Second, private firms comprise a growing share of economic activity (Davis 2016; Kalemli-Özcan et al. 2024). Most firms in EM and the broader edtech sector are private and face fewer disclosure requirements than publicly traded firms. The lack of research about third-party vendors that structure college access is partially due to difficulty obtaining data for empirical analyses. In order to hold private interests accountable for education policy goals, the policy community needs the research community to figure out how to systematically investigate the role of private firms and investors in education.

Knowledge infrastructures are "systems of observation and measurement" (Hirschman 2021, 743) that provide the foundation for scalable research. They collect, process, and distribute data in ways that "enable certain kinds of knowledge production while simultaneously channeling researchers away from questions not readily answerable within…that infrastructure" (Hirschman 2021, 742). The knowledge infrastructure of education research largely consists of data about students, schools, and colleges. It has been shaped by the ascendance of economics (Berman 2022), which incorporates district, state, and federal administrative data to evaluate the effects of schools and policies on student outcomes (e.g., Fuller et al. 2023). The research community mobilized to investigate for-profit colleges (e.g., Eaton 2022b, 2022a) because for-profit colleges and students were included in ongoing data collection. By contrast, researchers ignored the growth of private, for-profit, third-party vendors (e.g., PowerSchool, Parchment) because these organizations are excluded from data collections known to education researchers.

We argue that sociologists and adjacent scholars should expand the domain of the education research knowledge infrastructure to include third-party vendors in enrollment management and edtech more broadly. This is a challenging task. Student unit record data owned by district, state, and federal agencies is simultaneously granular (can evaluate a particular policy) and macro (can describe broad trends in the education sector). By contrast, obtaining granular data about third-party products is difficult because these data are owned by private interests that do not want researchers to interrogate their products (Cottom 2020; Pasquale 2015). We see two avenues for research: case studies that require primary data collection and tend to have a granular focus on particular products or vendors; and secondary data analysis, which may have a more macro focus.

Although case studies often rely on interview and ethnographic data, another approach is to issue records requests to public entities that contract with third-party entities. For example, Hamilton et al. (2024) analyze public university contracts with online program managers (OPMs). The present manuscript is an example of case study research. We issued public records requests to obtain tabular data about student list purchases. However, this process was labor intensive, requiring the pro bono efforts of several law firms. Public records requests seem better suited to obtaining contracts. Qualitative or computational text analysis of publicly available websites and social media provides another avenue for data collection. Here, data collection is efficient but data processing may be laborious. Like investigative reporting (Feathers 2022), case studies can produce granular analyses of particular products (e.g., Intersect, Enroll360) that yield insight into the mechanisms of how third-party vendors structure college access. However, case study data collections do not meet the standard of repeated, ongoing data collection that can be the basis for scalable research.

Subscription databases have the potential to create a knowledge infrastructure for scalable research on the role of private firms and private equity funders in education, and also their role in other domains of interest to sociology (e.g., Eaton 2025). A growing set of data providers serve the information needs of the investment community by providing data on private firms and deals involving private firms. These databases tend to include three broad "types" of data: (1) *firm-level data* (e.g., location, employees, financials); (2) *acquisitions*, encompassing who acquired whom and resulting ownership structure; and (3) *investments*, which include amounts/valuations/stake percentages of private equity investments. Some providers provide all three types of data (e.g., *S&P Capital IQ*, *LSEG Workspace*). Other providers specialize. For example, *PitchBook* is a leading provider of data about investments in private companies by private equity firms, while *Orbis* is a leading provider of firm-level data for non-U.S. private firms. These databases are often utilized in journals of finance or management that consider private equity markets or patterns of acquisitions (e.g., Humphery-Jenner et al. 2017; Kalemli-Özcan et al. 2024). Interestingly, the database subscriptions operate similar to contemporary student list products; an interactive, user-facing platform is wrapped around several sources of proprietary data. University libraries provide full or partial access to some subscriptions, while others must be purchased.

Although scholars cannot share subscription data, they can share code to process and analyze these data. We recommend that scholars develop panel datasets examining change over time in the firms operating in particular industries and acquisitions and private equity investments in those industries. For example, empirical scholarship can examine change over time in concentration in the enrollment management industry and its transformation from an owner-operated to a largely private-equity owned industry. Subsequent analyses could investigate the portfolio of large owners, their ties to other industries – such as the direct-provider for-profit college market – and the vertical inputs/technologies they are investing in.

# 7 References

Author. XXXXb.

Author. XXXXe.

Author. XXXXc.

Author. XXXXd.

Author. XXXXa.

Author. XXXXf.

Belkin, Douglas. 2019. "For Sale: SAT-Takers' Names. Colleges Buy Student Data and Boost Exclusivity." *The Wall Street Journal*, November. https://www.wsj.com/articles/for-sale-sat-takers-names-colleges-buy-student-data-and-boost-exclusivity-11572976621.

Benjamin, Ruha. 2019. *Race After Technology: Abolitionist Tools for the New Jim code*. Polity.

Berman, Elizabeth Popp. 2022. *Thinking Like an Economist: How Efficiency Replaced Equality in u.s. Public Policy*. Princeton University Press.

Berman, Elizabeth Popp, and Daniel Hirschman. 2018. "The Sociology of Quantification: Where Are We Now?" *Contemporary Sociology-a Journal of Reviews* 47 (3): 257–66. https://doi.org/10.1177/0094306118767649.

Blau, Peter Michael, and Otis Dudley Duncan. 1967. *The American Occupational Structure*. Wiley.

Blumenstyk, Goldie. 2019. "Think ACT Is Just a College-Admissions Test? Think Again." *The Chronicle of Higher Education*, February 5. https://www.chronicle.com/article/think-act-is-just-a/245610.

Bourdieu, Pierre. 1984. *Distinction: A Social Critique of the Judgment of Taste*. Harvard University Press.

Bourdieu, Pierre. 1988. *Homo Academicus*. Polity Press. file://C%3A%2FUsers%2Fozanj%2FDocuments%2Facademic%20papers%2Fphilosophy%20of%20social%20sciences%2Fhomo-academicus-Bourdieu.doc.

Brint, Steven G., and Jerome Karabel. 1989. *The Diverted Dream: Community Colleges and the Promise of Educational Opportunity in America, 1900-1985*. Oxford University Press.

Burrell, Jenna, and Marion Fourcade. 2021. "The Society of Algorithms." *Annual Review of Sociology* 47: 213–37. https://doi.org/10.1146/annurev-soc-090820-020800.

Burrows, Roger, and Nicholas Gane. 2006. "Geodemographics, Software and Class." *Sociology-the Journal of the British Sociological Association* 40 (5): 793–812. https://doi.org/10.1177/0038038506067507.

Campbell, Alec. 2017. Higher Education Marketing: How to Master Your Admissions Funnel. Hop Online. https://hop-online.com/blog/higher-education-marketing-admissions-process/.

Chetty, Raj, John N. Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2020. "Income Segregation and Intergenerational Mobility Across Colleges in the United States." *The Quarterly Journal of Economics* 135 (3): 1567–633. https://doi.org/10.1093/qje/qjaa005.

Chun, Wendy Hui Kyong. 2018. "Queerying Homophily." In *Pattern Discrimination*, edited by Clemens Apprich, Wendy Hui Kyong Chun, Florian Cramer, and Hito Steyerl. Meson press. https://doi.org/10.25969/mediarep/12350.

Chun, Wendy Hui Kyong. 2021. *Discriminating Data: Correlation, Neighborhoods, and the New Politics of Recognition*. The MIT Press.

Coase, Ronald. 1937. "The Nature of the Firm." *Economica* 4: 386–405.

College Board. 2005. *Enrollment Planning Services*. October. Vol. 2005. College Board. www.collegeboard.com/highered/ra/eps.html.

College Board. 2011a. *Enrollment Planning Service*. The College Board. https://collegeboardsearch.collegeboard.org/pastudentsrch/support/licensing/college-board-search-services/enrollment-planning-service.

College Board. 2011b. *Segment Analysis Service: An Educationally Relevant Geodemographic Tagging Service*. College Board. https://secure-media.collegeboard.org/mSSS/media/pdf/segment-analysis-service-overview.pdf.

College Board. 2012. *Welcome to the College Board*. College Board. https://secure-media.collegeboard.org/homeOrg/content/pdf/welcome_to_the_college_board_fullversion.pdf.

College Board. 2016. *Guide to Race and Ethnicity Reporting for AP® Online Score Reports*. https://secure-media.collegeboard.org/digitalServices/pdf/ap/ap-guide-to-race-ethnicity-reporting-schools-districts-2016.pdf.

College Board. 2023a. *Connections Geographies: Detailed Connections Geography Maps*. https://drive.google.com/file/d/1d7nVl5yL6txrG84o6l7Bg5RerqC457eS/view?usp=drive_link.

College Board. 2023b. *Detailed Connections Geography Maps: Reach More Students Where They Are*. College Board. https://cbsearch.collegeboard.org/media/pdf/connections-geographies.pdf.

College Board. n.d. *SAT Trends Dashboard Report: Interpretive Guide*. College Board. https://satsuite.collegeboard.org/media/pdf/sat-trends-dashboard-interpretive-guide.pdf.

Collins, Randall. 1979. *The Credential Society: An Historical Sociology of Education and Stratification*. Academic Press. file://C%3A%2FUsers%2Fozanj%2FDocuments%2Facademic%20papers%2Fcredentialism%2Fcredential-society-collins.doc.

Cottom, Tressie McMillan. 2017. *Lower Ed: The Troubling Rise of for-Profit Colleges in the New Economy*. The New Press.

Cottom, Tressie McMillan. 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): 441–49. https://doi.org/10.1177/2332649220949473.

Davis, Gerald F. 2016. *The Vanishing American Corporation: Navigating the Hazards of a New Economy*. First edition. Berrett-Koehler Publishers, Inc.

DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48 (2): 147–60.

Domina, Thurston, Andrew Penner, and Emily Penner. 2017. "Categorical Inequality: Schools as Sorting Machines." *Annual Review of Sociology* 43: 311–30. https://doi.org/10.1146/annurev-soc-060116-053354.

EAB. 2018. *Making Your Digital Ads Count: 15 Lessons on New and Emerging Techniques in Undergraduate Recruitment Marketing*. EAB.

EAB. 2019. *The 5 Key Stages of College Enrollment—and Which Metrics to Track During Each*. https://eab.com/insights/daily-briefing/enrollment/the-5-key-stages-of-college-enrollment-and-which-metrics-to-track-during-each/.

EAB. 2025. *About EAB*. https://eab.com/about/.

EAB. n.d. EAB: Managing School Communication During COVID-19. https://www.vistaequitypartners.com/spotlight/eab-school-communication-covid19/.

Eaton, Charlie. 2022a. *Bankers in the Ivory Tower: The Troubling Rise of Financiers in US Higher Education*. The University of Chicago Press.

Eaton, Charlie. 2022b. "Agile Predators: Private Equity and the Spread of Shareholder Value Strategies to US for-Profit Colleges." *Socio-Economic Review* 20 (2): 791–815. https://doi.org/10.1093/ser/mwaa005.

Eaton, Charlie. 2025. "What Is Private Equity? Corporate Raider, Main Street Invader, Monopoly Builder, Shadow Owner." May 12. https://www.charlieeaton.net/p/what-is-private-equity.

Espeland, Wendy Nelson, and Michael L. Sauder. 2016. *Engines of Anxiety: Academic Rankings, Reputation, and Accountability*. Russell Sage Foundation.

Espeland, Wendy Nelson, and Mitchell L. Stephens. 2008. "A Sociology of Quantification." *Archives Europeennes De Sociologie* 49 (3): 397–432. <Go to ISI>://WOS:000265147000002.

Ewing, Eve L., Samhitha Krishnan, and Brianna Lawrence. 2024. "'No Future for Black People in Chicago': Out–Migration as Slow–Motion Disaster." *Socius: Sociological Research for a Dynamic World* 10: 1–18. https://doi.org/10.1177/23780231241266511.

Feathers, Todd. 2022. "College Prep Software Naviance Is Selling Advertising Access to Millions of Students." *The Markup*, January. 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.

Fishman, Samuel H. 2020. "Educational Mobility Among the Children of Asian American Immigrants." *American Journal of Sociology* 126 (2): 260–317. https://doi.org/10.1086/711231.

Fourcade, Marion, and Kieran Joseph Healy. 2013. "Classification Situations: Life-Chances in the Neoliberal Era." Journal Article. *Accounting Organizations and Society* 38 (8): 559–72. https://doi.org/10.1016/j.aos.2013.11.002.

Fourcade, Marion, and Kieran Joseph Healy. 2024. *The Ordinal Society*. Harvard University Press.

Fuller, Sarah, Douglas Lee Lauen, and Fatih Unlu. 2023. "Leveraging Experimental and Observational Evidence to Assess the Generalizability of the Effects of Early Colleges in North Carolina." *Education Finance and Policy* 18 (4): 568–96. https://doi.org/10.1162/edfp_a_00379.

Hamilton, Laura T., Heather Daniels, Christian Michael Smith, and Charlie Eaton. 2024. "The for-Profit Side of Public u: University Contracts with Online Program Managers." *Socius* 10: 23780231231214952. https://doi.org/10.1177/23780231231214952.

Hansen, Drew. 2017. "Advisory Board Closes Sale of Its Education Division." In *Washington Business Journal*. https://www.bizjournals.com/washington/news/2017/11/20/advisory-board-completes-spinoff-sale-of-its.html.

Harcourt, Bernard E. 2007. *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*. University of Chicago Press.

Harcourt, Bernard E. 2015. "Risk as a Proxy for Race: The Dangers of Risk Assessment." *Federal Sentencing Reporter* 27 (4): 237–43. https://doi.org/10.1525/fsr.2015.27.4.237.

Harris, Chauncy D., and Edward L. Ullman. 1945. "The Nature of Cities." *The Annals of the American Academy of Political and Social Science* 242 (1): 7–17.

Hirschman, Daniel. 2021. "Rediscovering the 1." *American Journal of Sociology* 127 (3): 739–86. https://doi.org/10.1086/718451.

Hirschman, Daniel, Ellen Berrey, and Fiona Rose-Greenland. 2016. "Dequantifying Diversity: Affirmative Action and Admissions at the University of Michigan." *Theory and Society* 45 (3): 265–301. https://doi.org/10.1007/s11186-016-9270-2.

Hirschman, Daniel, and Emily Adlin Bosk. 2020. “Standardizing Biases: Selection Devices and the Quantification of Race.” *Sociology of Race and Ethnicity* 6 (3): 348–64. https://doi.org/10.1177/2332649219844797.

Holland, Megan M. 2019. *Divergent Paths to College: Race, Class, and Inequality in High Schools*. Rutgers University Press. https://doi.org/10.36019/9780813590288.

Hoover, Eric. 2023. “A Profession on the Edge: How Admissions Became the Most Reviled Office on Campus.” *The Chronicle of Higher Education*. https://www.chronicle.com/article/a-profession-on-the-edge.

Hossler, Donald, and John P. Bean. 1990. *The Strategic Management of College Enrollments*. Jossey-Bass Higher Education Series. Jossey-Bass.

Howell, Jessica, Michael H. Hurwitz, Zachary Mabel, and Jonathan Smith. 2021. *Participation in Student Search Service Is Associated with Higher College Enrollment and Completion*. College Board. https://cbsearch.collegeboard.org/pdf/college-outreach-and-student-outcomes.pdf.

Huang, Tiffany J. 2023. “Translating Authentic Selves into Authentic Applications: Private College Consulting and Selective College Admissions.” *Sociology of Education*, ahead of print. https://doi.org/10.1177/00380407231202975.

Humphery-Jenner, Mark, Zacharias Sautner, and Jo-Ann Suchard. 2017. “Cross-Border Mergers and Acquisitions: The Role of Private Equity Firms.” *Strategic Management Journal* 38 (8): 1688–700. https://doi.org/https://doi.org/10.1002/smj.2623.

Hurwitz, Michael. 2012. “The Impact of Institutional Grant Aid on College Choice.” *Educational Evaluation and Policy Analysis* 34 (3): 344–63. https://doi.org/10.3102/0162373712448957.

Jencks, Christopher, and David Riesman. 1968. *The Academic Revolution*. 1st ed. Doubleday.

Jump, Jim. 2020. Remembering Bill Royall. The Thoughtful College Search. http://www.thoughtfulcollegesearch.com/ethicalcollegeadmissions/2020/7/10/remembering-bill-royall.

Kalemli-Özcan, Sebnem, Bent E. Sorensen, Carolina Villegas-Sanchez, Vadym Volosovych, and Sevcan Yesiltas. 2024. “How to Construct Nationally Representative Firm-Level Data from the Orbis Global Database: New Facts on SMEs and Aggregate Implications for Industry Concentrations.” Journal Article. *American Economic Journal-Macroeconomics* 16 (2): 353–74. https://doi.org/10.1257/mac.20220036.

Karabel, Jerome. 1984. “Status-Group Struggle, Organizational Interests, and the Limits of Institutional Autonomy: The Transformation of Harvard, Yale, and Princeton, 1918-1940.” *Theory and Society* 13 (1): 1–40. https://doi.org/10.1007/bf00159255.

Karabel, Jerome. 2005. *The Chosen: The Hidden History of Admission and Exclusion at Harvard, Yale, and Princeton*. Houghton Mifflin Co.

Karen, David. 2002. “Changes in Access to Higher Education in the United States: 1980-1992.” *Sociology of Education* 75 (3): 191–210.

Khan, Shamus Rahman. 2010. “Getting in: How Elite Schools Play the College Game.” In *Educating Elites Class Privilege and Educational Advantage*, edited by Adam Howard and Rubén A. Gaztambide-Fernandez. Rowman & Littlefield. https://ebookcentral.proquest.com/lib/umichigan/detail.action?docID=634205.

Khan, Shamus Rahman. 2011. *Privilege: The Making of an Adolescent Elite at St. Paul's School*. Princeton Studies in Cultural Sociology. Princeton University Press.

Killgore, Leslie. 2009. "Merit and Competition in Selective College Admissions." *Review of Higher Education* 32 (4): 469–88. https://doi.org/10.1353/rhe.0.0083.

Knox, Liam. 2024. "What's Behind the ACT's for-Profit Pivot?" *Inside Higher Ed*, April 25. https://www.insidehighered.com/news/admissions/traditional-age/2024/04/25/acts-private-equity-takeover-and-future-testing.

Koppenheffer, Michael. 2021. Introducing Enroll360: Unlock the Power of a Thriving Recruitment Ecosystem. EAB. https://eab.com/insights/blogs/enrollment/introducing-enroll360-recruitment-ecosystem/.

Korver-Glenn, Elizabeth. 2022. *Race Brokers: Housing Markets and Racial Segregation in 21st century Urban America*. Oxford University Press.

Kosinski, Michal, David Stillwell, and Thore Graepel. 2013. "Private Traits and Attributes Are Predictable from Digital Records of Human Behavior." *Proc Natl Acad Sci U S A* 110 (15): 5802–5. https://doi.org/10.1073/pnas.1218772110.

Kraatz, Matthew S., Marc J. Ventresca, and Lina N. Deng. 2010. "Precarious Values and Mundane Innovations: Enrollment Management in American Liberal Arts Colleges." *Academy of Management Journal* 53 (6): 1521–45. https://doi.org/10.5465/amj.2010.57319260.

Kraus, Michael W., Ivuoma N. Onyeador, Natalie M. Daumeyer, Julian M. Rucker, and Jennifer A. Richeson. 2019. "The Misperception of Racial Economic Inequality." *Perspectives on Psychological Science* 14 (6): 899–921. https://doi.org/10.1177/1745691619863049.

Labaree, David F. 1997. *How to Succeed in School Without Really Learning: The Credentials Race in American Education*. Yale University Press.

Larson, Magali Sarfatti. 1977. *The Rise of Professionalism: A Sociological Analysis*. University of California Press.

Leeds, Daniel M., and Stephen L. DesJardins. 2015. "The Effect of Merit Aid on Enrollment: A Regression Discontinuity Analysis of Iowa's National Scholars Award." *Research in Higher Education* 56 (7): 471–95. https://doi.org/10.1007/s11162-014-9359-2.

Leyshon, Andrew, and Nigel Thrift. 1999. "Lists Come Alive: Electronic Systems of Knowledge and the Rise of Credit-Scoring in Retail Banking." *Economy and Society* 28 (3): 434–66. https://doi.org/10.1080/03085149900000013.

Litten, Larry H., Daniel J. Sullivan, and David L. Brodigan. 1983. *Applying Market Research in College Admissions*. College Entrance Examination Board.

Manson, Steven, Jonathan Schroeder, David Van Riper, et al. 2024. *IPUMS National Historical Geographic Information System: Version 19.0*. Report. IPUMS. https://doi.org/http://doi.org/10.18128/D050.V19.0.

Marcus, Jon. 2024. "Going to a 'Witch Doctor': Colleges' Reliance on Enrollment Management Firms Grows Despite Alarm over Companies' Strategies." In *Lifting the Veil on Enrollment Management: How a Powerful Industry Is Limiting Social Mobility in American Higher Education*, edited by Stephen J. Burd. Harvard Education Press.

McArthur, Daniel, and Aaron Reeves. 2022. "The Unintended Consequences of Quantifying Quality: Does Ranking School Performance Shape the Geographical Concentration of Advantage?" *American Journal of Sociology* 128 (2): 515–51. https://doi.org/10.1086/722470.

McDonough, Patricia M. 1997. *Choosing Colleges: How Social Class and Schools Structure Opportunity*. State University of New York Press.

McKelvey, Fenwick. 2022. "When the New Magic Was New: The Claritas Corporation and the Clustering of America." *Ieee Annals of the History of Computing* 44 (4): 44–56. https://doi.org/10.1109/mahc.2022.3214223.

McPherson, Miller, Lynn Smith-Lovin, and James M Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27 (1): 415–44.

Mennicken, Andrea, and Wendy N. Espeland. 2019. "What's New with Numbers? Sociological Approaches to the Study of Quantification." *Annual Review of Sociology* 45: 223–45. https://doi.org/10.1146/annurev-soc-073117-041343.

Molnar, Michele. 2020. "ACT CEO Marten Roorda Departs as Organization Announces Cost-Cutting Measures." *Edweek Market Brief*, May. https://marketbrief.edweek.org/marketplace-k-12/act-ceo-marten-roorda-departs-organization-announces-cost-cutting-measures/.

Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press.

Noel-Levitz. 1998. *National Enrollment Management Survey: Findings for Fall 1997 Four-Year Institutions*. Noel-Levitz.

Norris, Davon. 2021. "Embedding Racism: City Government Credit Ratings and the Institutionalization of Race in Markets." *Social Problems*, ahead of print. https://doi.org/10.1093/socpro/spab066.

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. First edition. Crown.

Oracle. n.d. *Loading and Assigning EPS Market Codes*. Oracle. https://docs.oracle.com/cd/E29376_01/hrcs90r5/eng/psbooks/lsad/chapter.htm?File=lsad/htm/lsad40.htm.

Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.

Patton, Michael Quinn. 2002. *Qualitative Research and Evaluation Methods*. Sage.

Poon, Martha. 2007. "Scorecards as Devices for Consumer Credit: The Case of Fair, Isaac & Company Incorporated." *Sociological Review* 55: 284–306. https://doi.org/10.1111/j.1467-954X.2007.00740.x.

Poon, OiYan, and Michael N. Bastedo. 2022. *Rethinking College Admissions: Research-Based Practice and Policy*. Harvard Education Press. https://books.google.com/books?id=IjehEAAAQBAJ.

Posselt, Julie R. 2016. *Inside Graduate Admissions: Merit, Diversity, and Faculty Gatekeeping*. Harvard University Press.

PowerSchool. 2021. Naviance by PowerSchool. https://www.powerschool.com/solutions/naviance-by-powerschool/.

Reardon, Sean F., and Kendra Bischoff. 2011. "Income Inequality and Income Segregation." *American Journal of Sociology* 116 (4): 1092–153. https://doi.org/10.1086/657114.

Ruffalo Noel Levitz. 2023. About RNL. https://www.ruffalonl.com/about-ruffalo-noel-levitz/.

Ruffalo Noel Levitz. 2025. About RNL. https://www.ruffalonl.com/about-ruffalo-noel-levitz/.

Ruffalo Noel-Levitz. 2022a. *2022 Cost of Recruiting an Undergraduate Student Report*. Ruffalo Noel-Levitz. https://www.ruffalonl.com/thank-you/download-the-2022-cost-of-recruiting-an-undergraduate-student-report/.

Ruffalo Noel-Levitz. 2022b. *2022 Marketing and Recruitment Practices for Undergraduate Students Report*. Ruffalo Noel-Levitz. https://www.ruffalonl.com/papers-research-higher-education-fundraising/marketing-and-recruitment-practices-for-undergraduate-students/.

Schmidt, Danika. 2019. Prospect Search Filters. Encoura. https://helpcenter.encoura.org/hc/en-us/articles/360035260452-Prospect-Search-Filters-.

Scott, W. Richard, and Gerald F. Davis. 2007. "The Dyadic Environment of the Organization." In *Organizations and Organizing: Rational, Natural, and Open Systems Perspectives*, edited by W. Richard Scott and Gerald F. Davis. Pearson, Prentice Hall.

Sewell, William H. 1971. "Inequality of Opportunity for Higher Education." *American Sociological Review* 36 (5): 793–93 &. <Go to ISI>://A1971K843800001.

Sewell, William H., and Vimal P. Shah. 1967. "Socioeconomic Status, Intelligence, and Attainment of Higher Education." *Sociology of Education* 40 (1): 1–23. <Go to ISI>://A1967ZD96500001.

Sewell, William H., and Vimal P. Shah. 1968a. "Parents Education and Childrens Educational Aspirations and Achievements." *American Sociological Review* 33 (2): 191–209. <Go to ISI>://A1968ZA11500001.

Sewell, William H., and Vimal P. Shah. 1968b. "Social Class, Parental Encouragement, and Educational Aspirations." *American Journal of Sociology* 73 (5): 559–72. <Go to ISI>://A1968ZA35500003.

Singer, Penny. 1988. "Behind the Flood of Mail: Name Brokers." *The New York Times*. https://www.nytimes.com/1988/01/24/nyregion/behind-the-flood-of-mail-name-brokers.html.

Smith, Jonathan, Jessica Howell, and Michael Hurwitz. 2022. "The Impact of College Outreach on High Schoolers' College Choices: Results from over One Thousand Natural Experiments." *Education Finance and Policy* 17 (1): 105–28. https://doi.org/10.1162/edfp_a_00334.

Snidal, Michael, Magda Maaoui, and Tyler Haupert. 2022. "Exodus in the American Metropolis: Predicting Black Population Decline in Chicago Neighbourhoods." *Urban Studies* 59 (15): 3114–31. https://doi.org/10.1177/00420980211070405.

Sole Source. 2022. Intersect Connection - 1 Year Subscription Renewal. University of Utah Purchasing Department - Sole Source Procurement. https://fbs.admin.utah.edu/uofubid/2022/01/07/intersect-connection-1-year-subscription-renewal/.

Stevens, Mitchell L. 2007. *Creating a Class: College Admissions and the Education of Elites*. Harvard University Press.

Stevens, Mitchell L., Elizabeth A. Armstrong, and Richard Arum. 2008. "Sieve, Incubator, Temple, Hub: Empirical and Theoretical Advances in the Sociology of Higher Education." *Annual Review of Sociology*

34: 127–51. https://doi.org/10.1146/annurev.soc.34.040507.134737.

Takamiya, Toshiro. 2005. *Mechanisms for Marketing Higher Education Information Services: The Case of the College Board*. University of Pennsylvania.

Taylor, Barrett J., Kelly Rosinger, and Karly S. Ford. 2024. "The Shape of the Sieve: Which Components of the Admissions Application Matter Most in Particular Institutional Contexts?" *Sociology of Education*, ahead of print. https://doi.org/10.1177/00380407241230007.

U. S. Census Bureau. 1994. *Geographic Areas Reference Manual*. U.S. Department of Commerce. https://www2.census.gov/geo/pdfs/reference/GARM/Ch10GARM.pdf.

U. S. Census Bureau. 2014. *Understanding Geographic Relationships: Counties, Places, Tracts and More*. https://www.census.gov/newsroom/blogs/random-samplings/2014/07/understanding-geographic-relationships-counties-places-tracts-and-more.html.

Vista Equity Partners. 2025. *EAB AI & ML: Improving the higher education enrollment process*. Https://www.vistaequitypartners.com/companies/portfolio/.

Walker, Kyle, and Matt Herman. 2024. *Tidycensus: Load US Census Boundary and Attribute Data as 'Tidyverse' and 'Sf'-Ready Data Frames*. https://walker-data.com/tidycensus/.

Weber, Max. 1948. "Bureaucracy." In *From Max Weber: Essays in Sociology*, edited by Hans Heinrich Gerth and C. Wright Mills. Routledge & K. Paul.

Yin, Robert K. 2009. *Case Study Research: Design and Methods*. 4th ed. Sage Publications.

Zemsky, Robert, and Penney Oedel. 1983. *The Structure of College Choice*. College Entrance Examination Board.

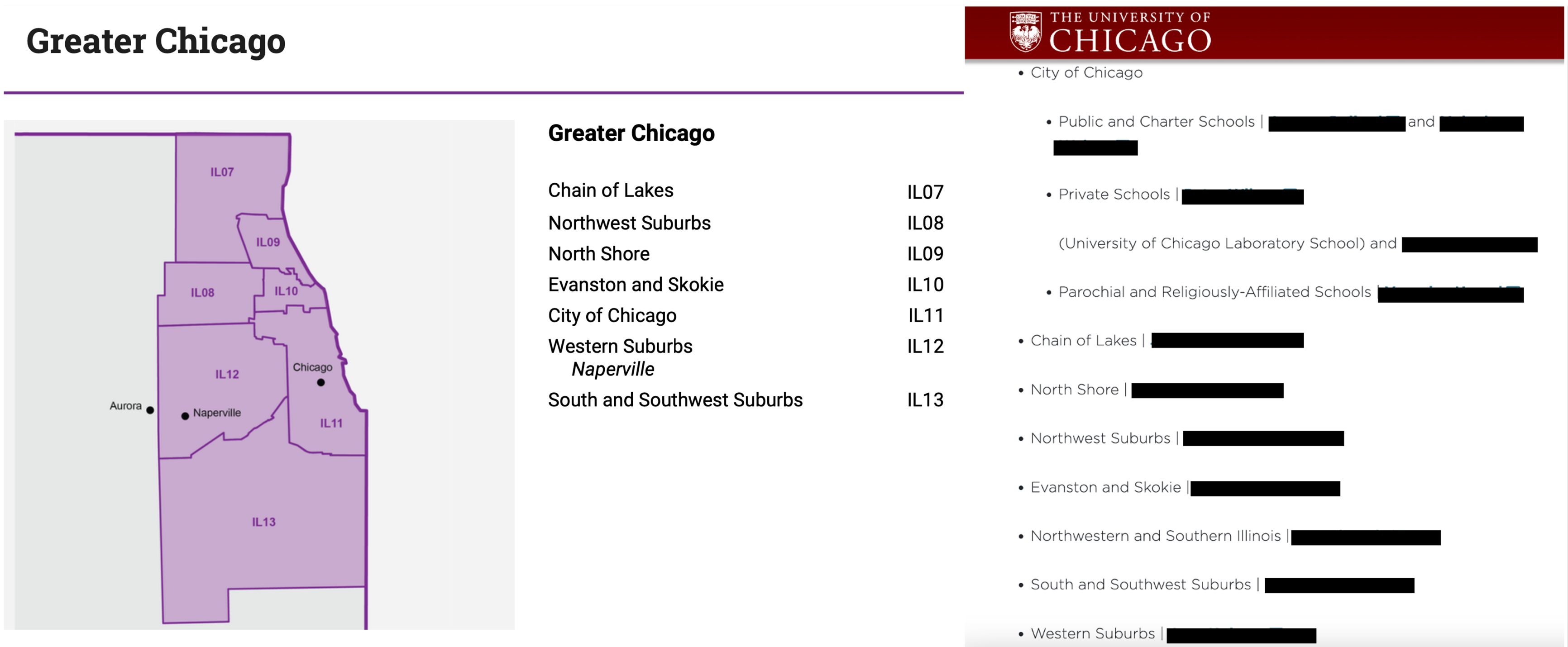


Figure 1: The College Board Enrollment Planning Service Chicago Geomarkets (left) and University of Chicago Recruiting Territories (right)

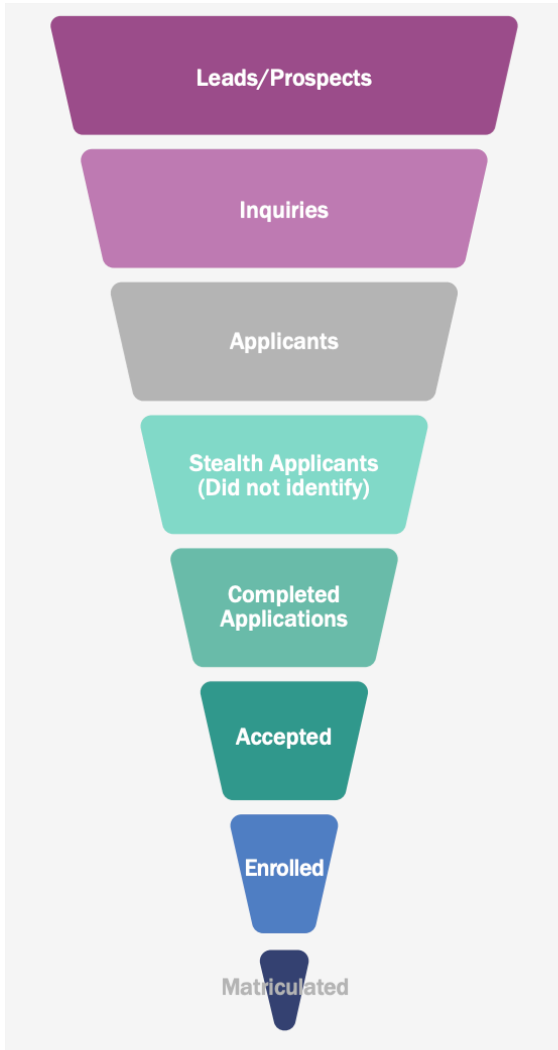


Figure 2: The Enrollment Funnel

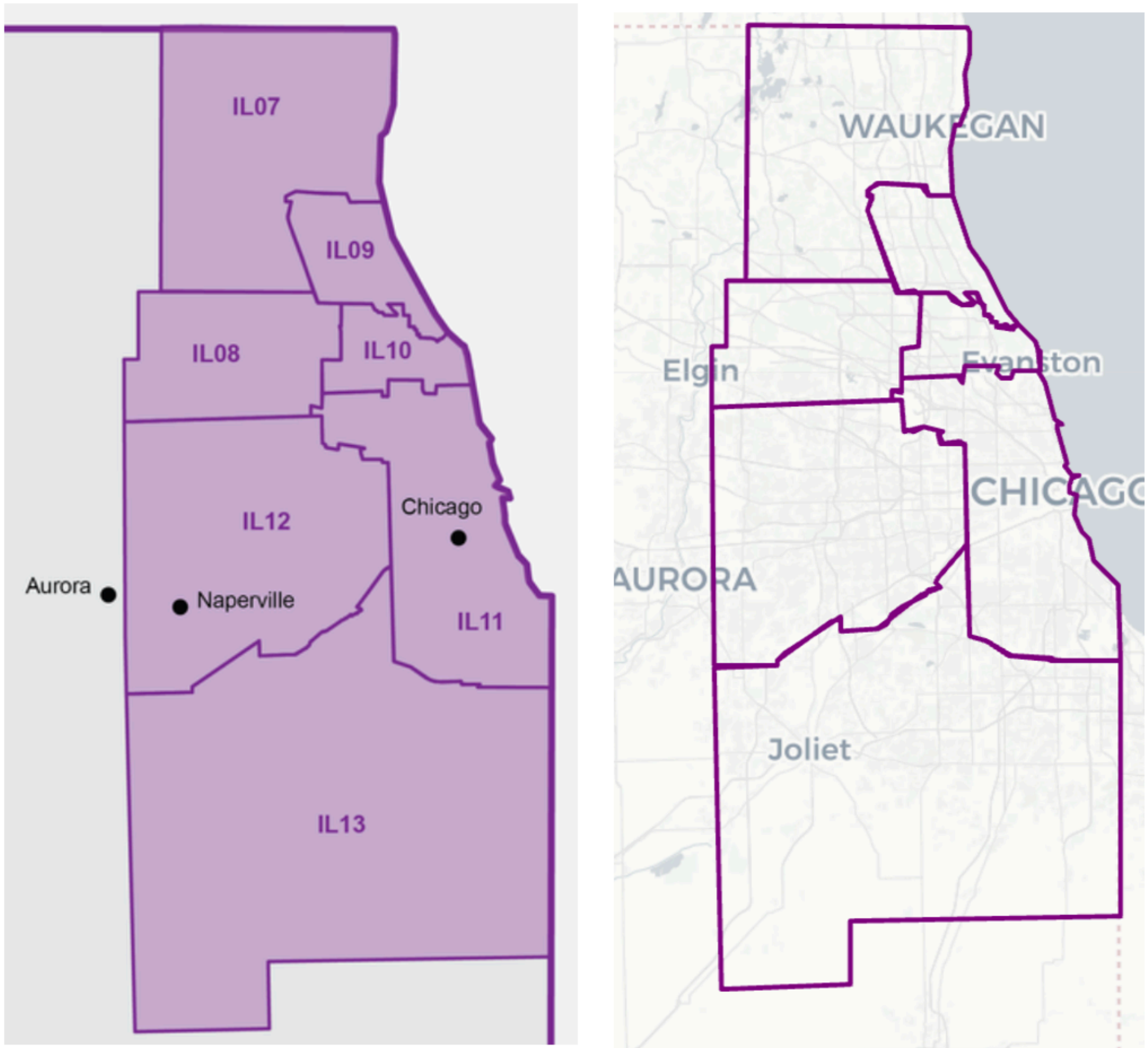


Figure 3: Chicago, College Board and Recreated Geomarkets

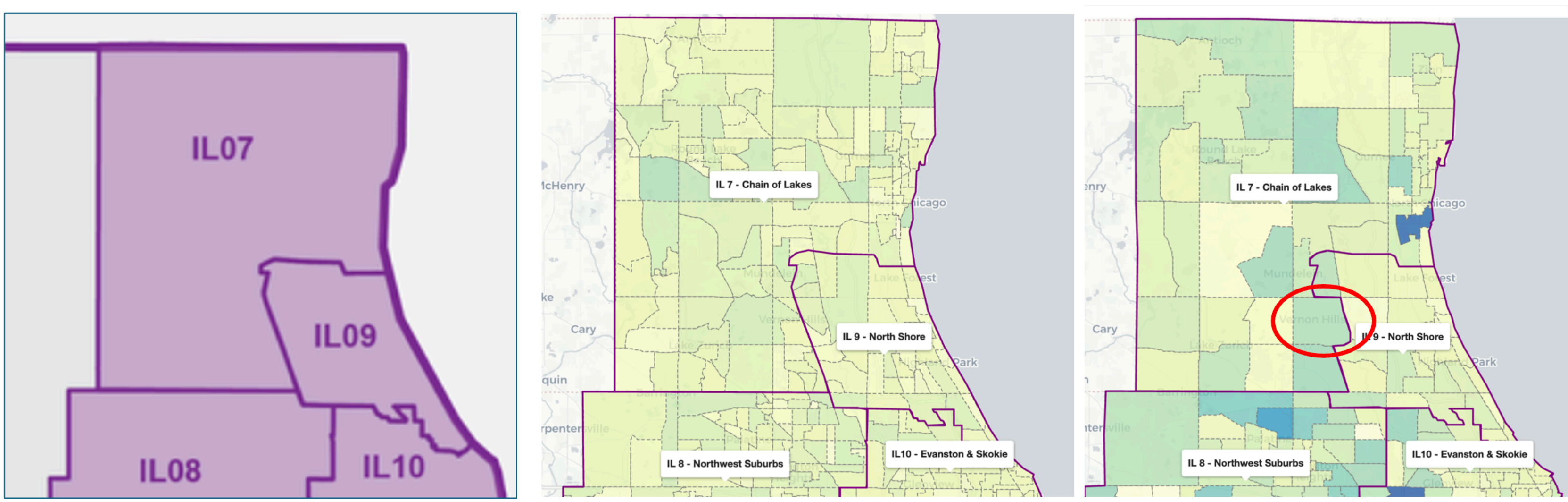


Figure 4: Chicago, College Board and Recreated Geomarket Spatial Misalignment

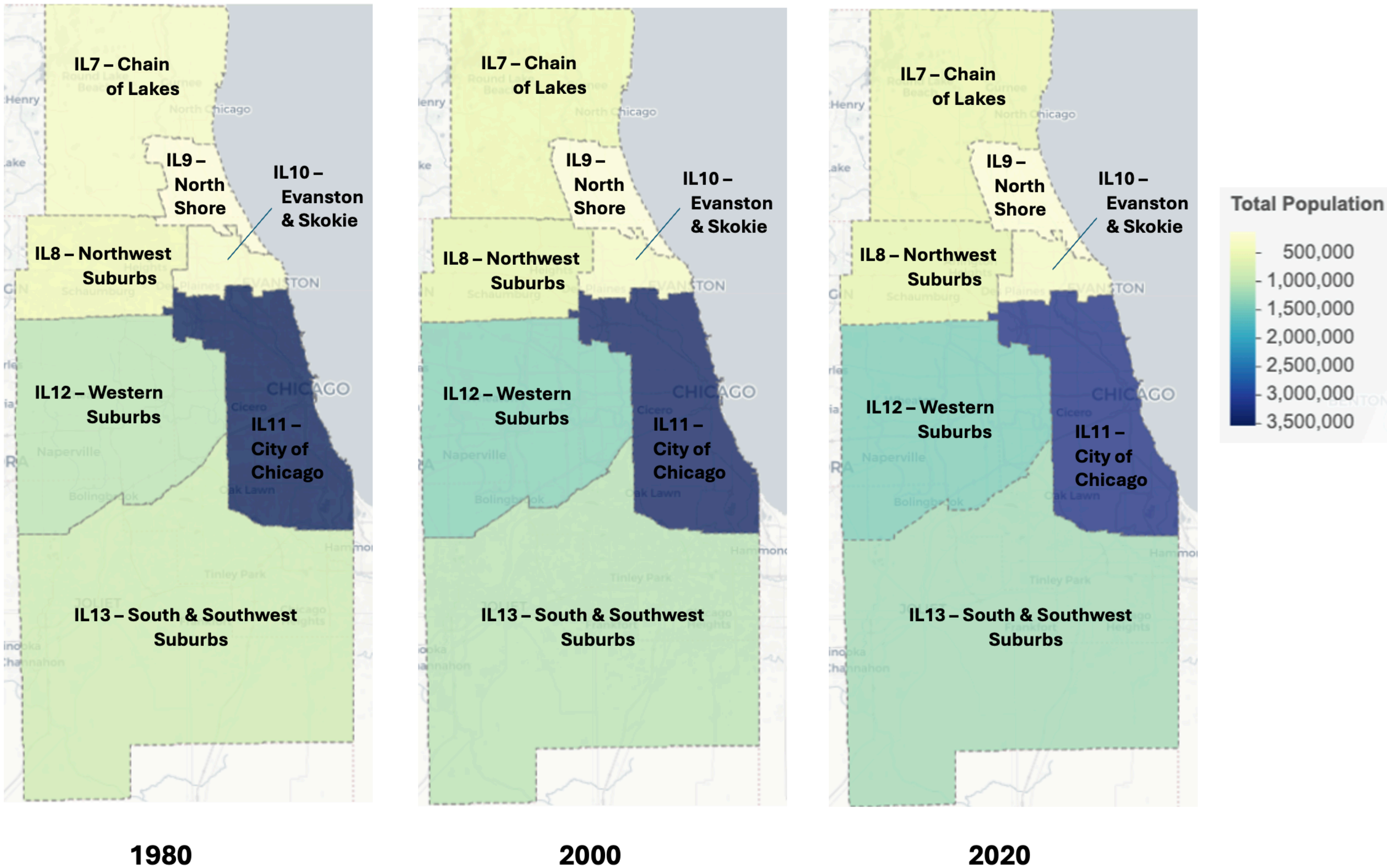


Figure 5: Chicago Geomarkets, Total Population 1980-2020


Figure 6: Racial/Ethnic Composition of Chicago Area Geomarkets, 1980-2020

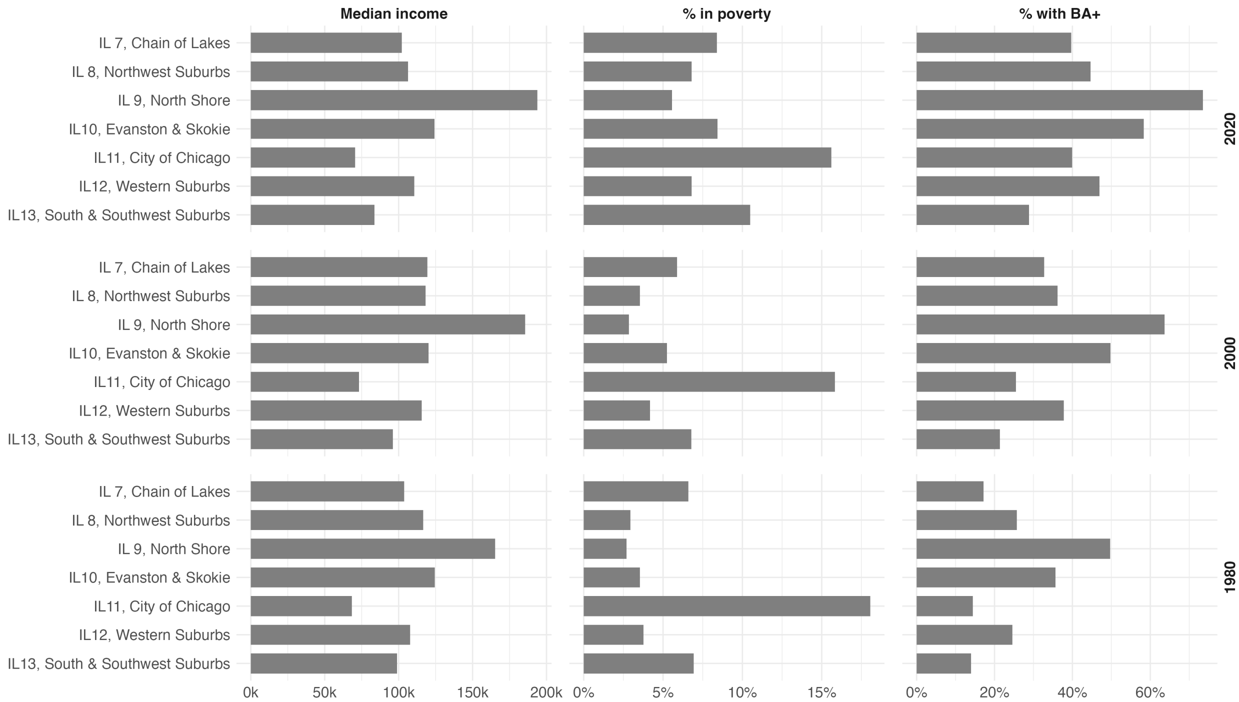


Figure 7: Socioeconomic Characteristics of Chicago Area Geomarkets, 1980-2020

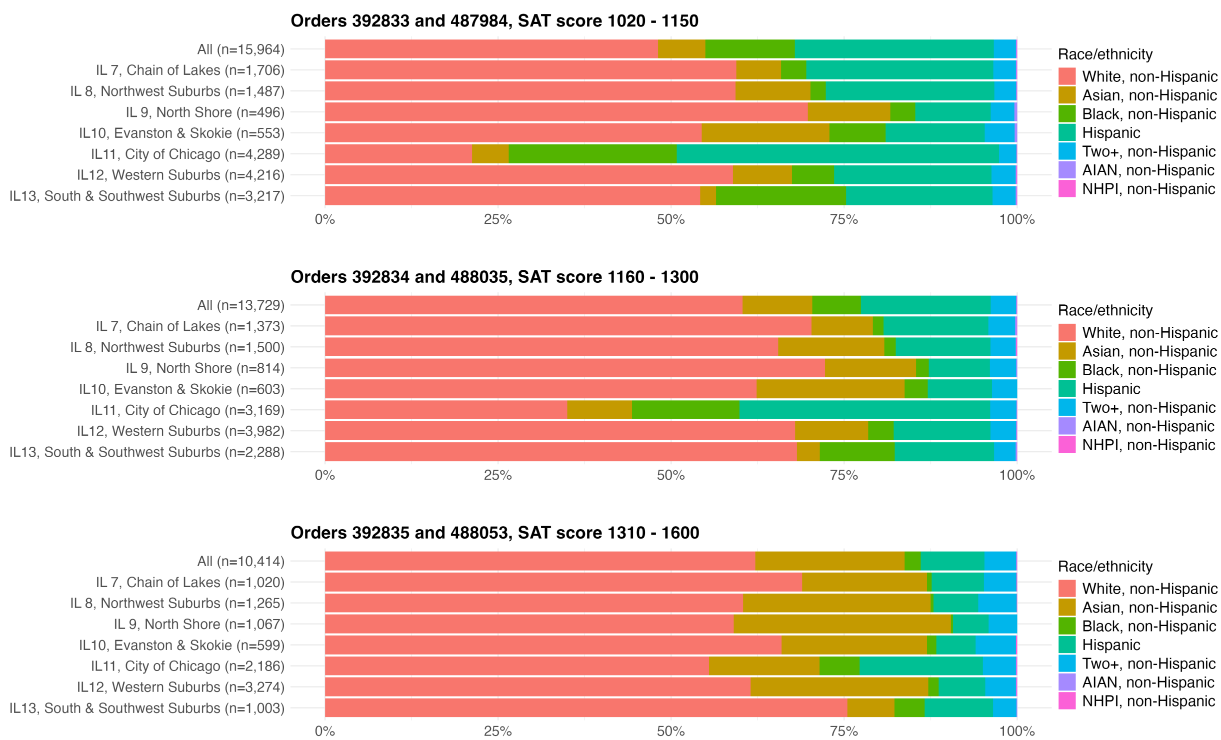


Figure 8: Racial/Ethnic Composition of Purchased Student Profiles by Geomarket, Chicago Area

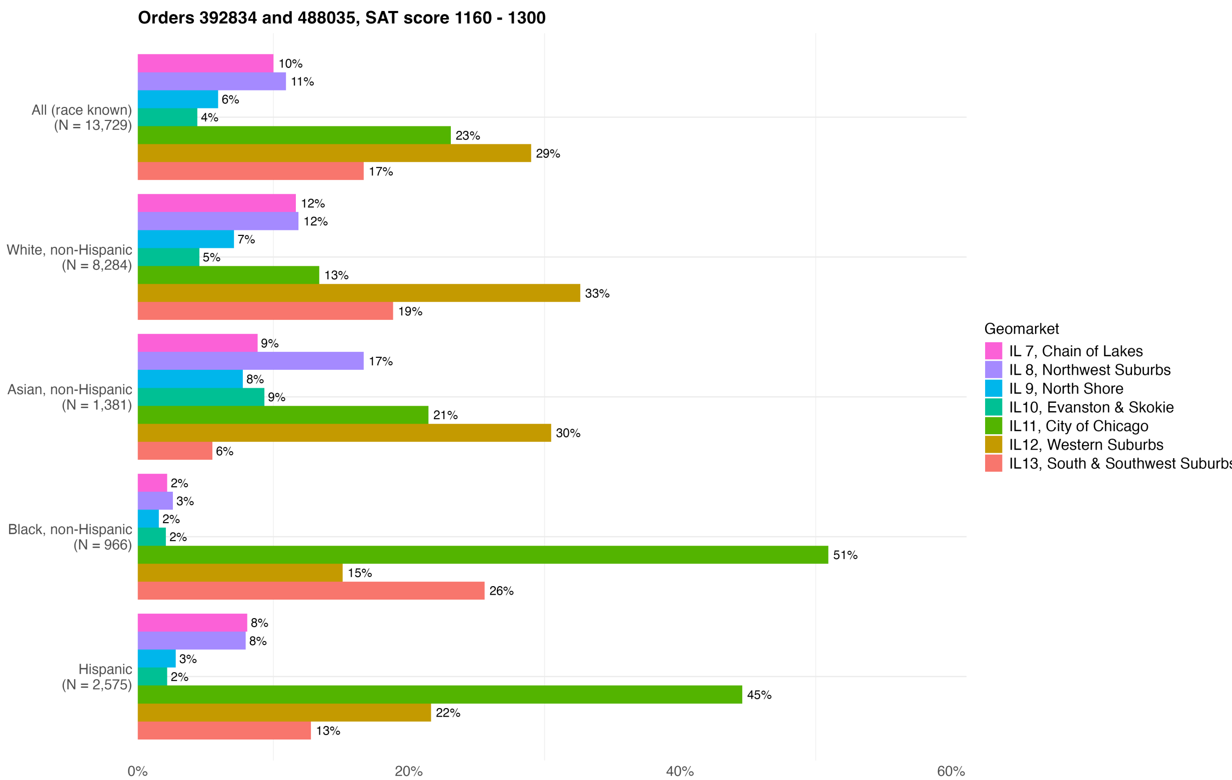


Figure 9: Chicago Geomarket Contribution to Purchased Student profiles by Racial/Ethnic Group, Middle-Range SAT orders

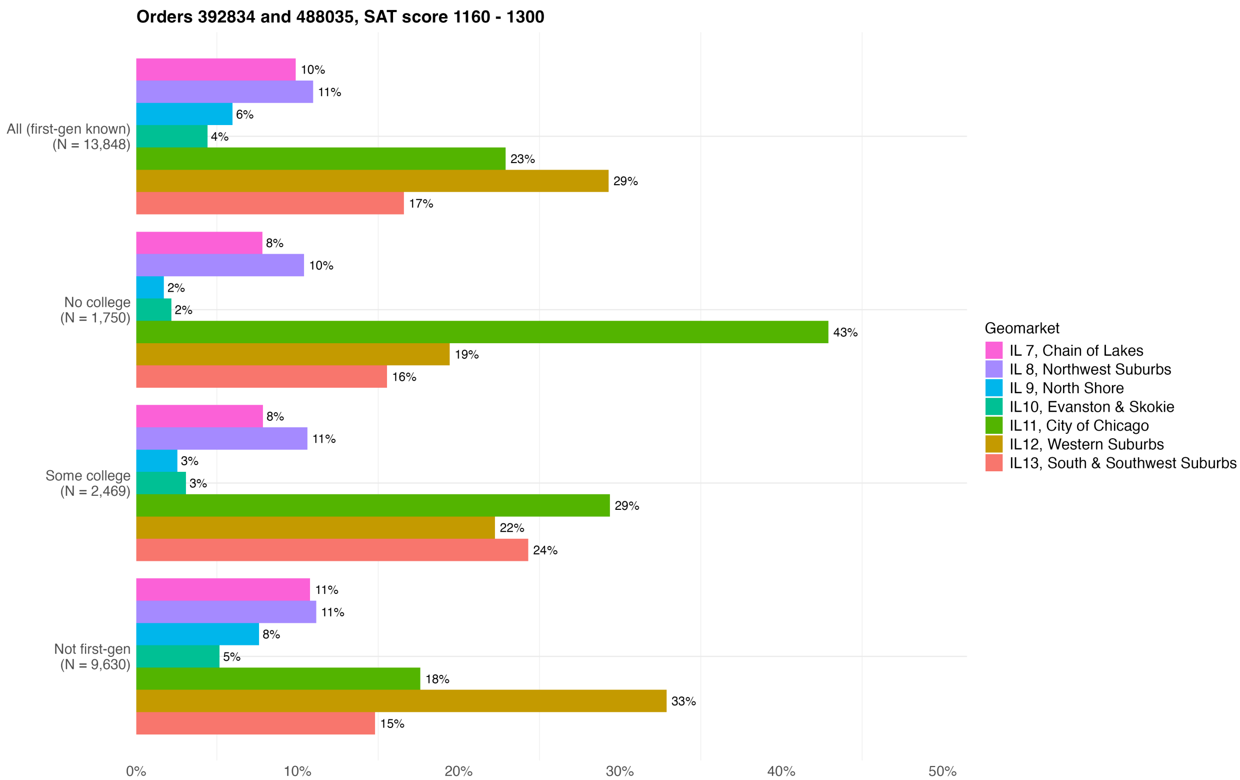


Figure 10: Chicago Geomarket Contribution to Purchased Student profiles by First-Generation Status, Middle-Range SAT orders


Figure 11: Chicago Geomarket Contribution to Purchased Student profiles by Racial/Ethnic Group and First-Generation Status, Middle-Range SAT orders

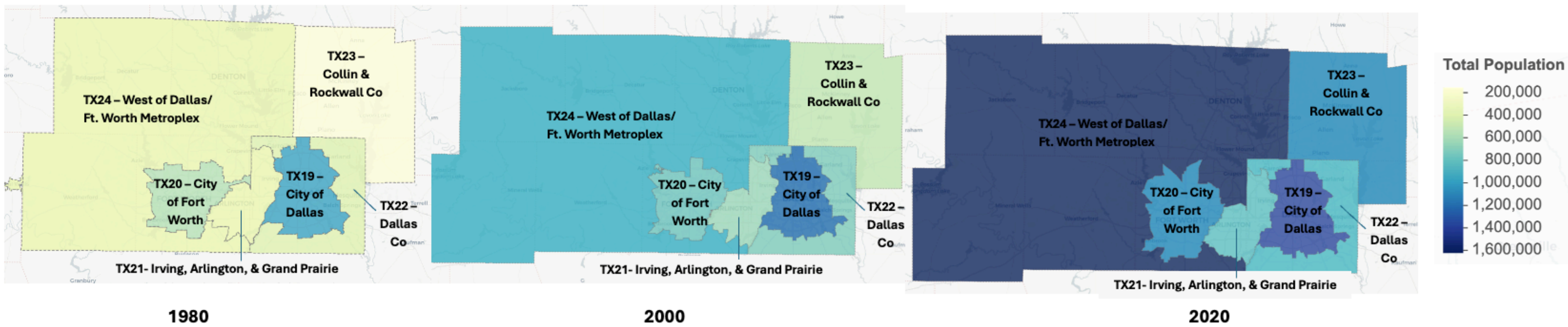


Figure 12: Dallas-Forth Worth Geomarkets, Total Population 1980-2020


Figure 13: Racial/Ethnic Composition of Dallas Area Geomarkets, 1980-2020

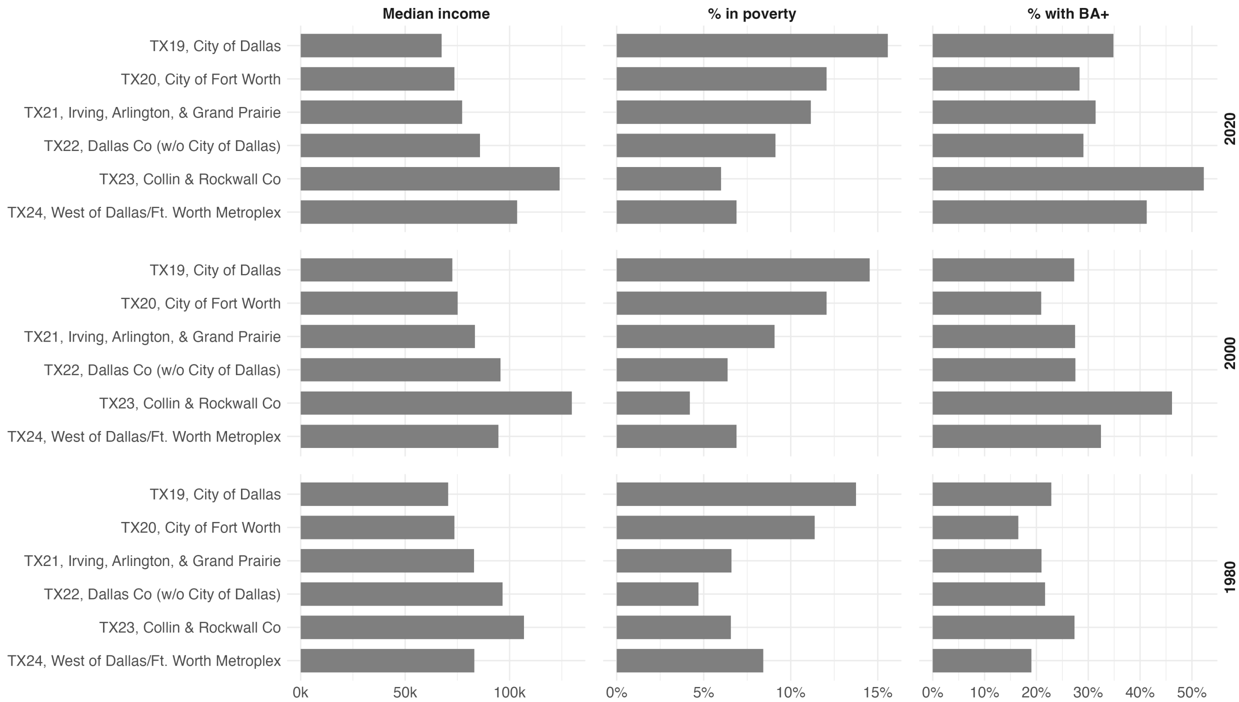


Figure 14: Socioeconomic Characteristics of Dallas Area Geomarkets, 1980-2020

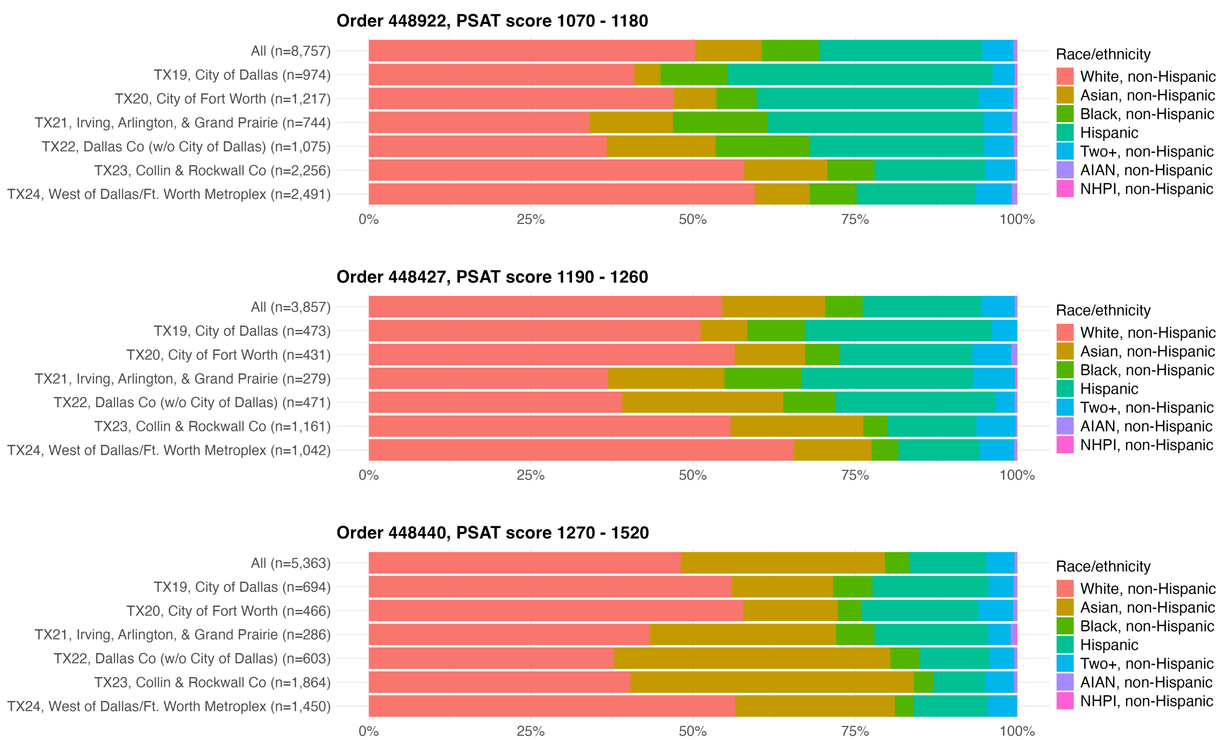


Figure 15: Racial/Ethnic Composition of Purchased Student Profiles by Geomarket, Dallas Area

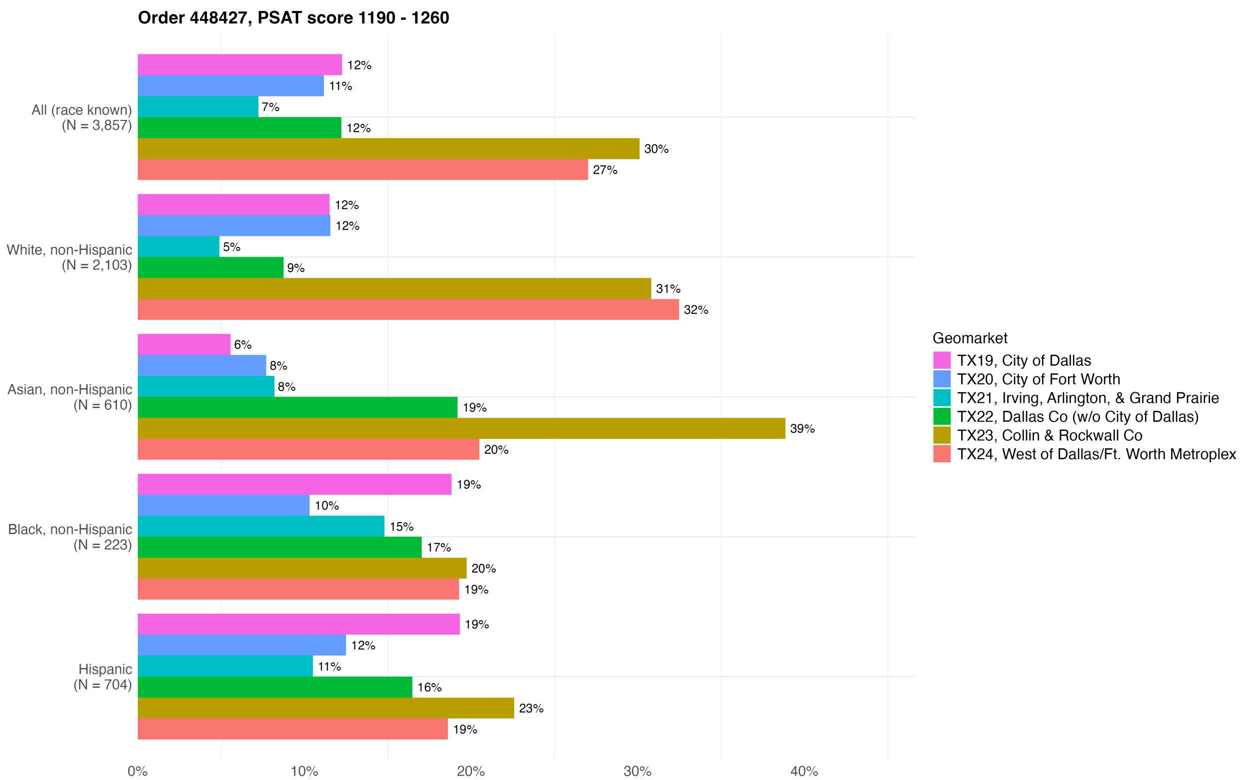


Figure 16: Dallas Geomarket Contribution to Purchased Student profiles by Racial/Ethnic Group, Middle-Range SAT orders

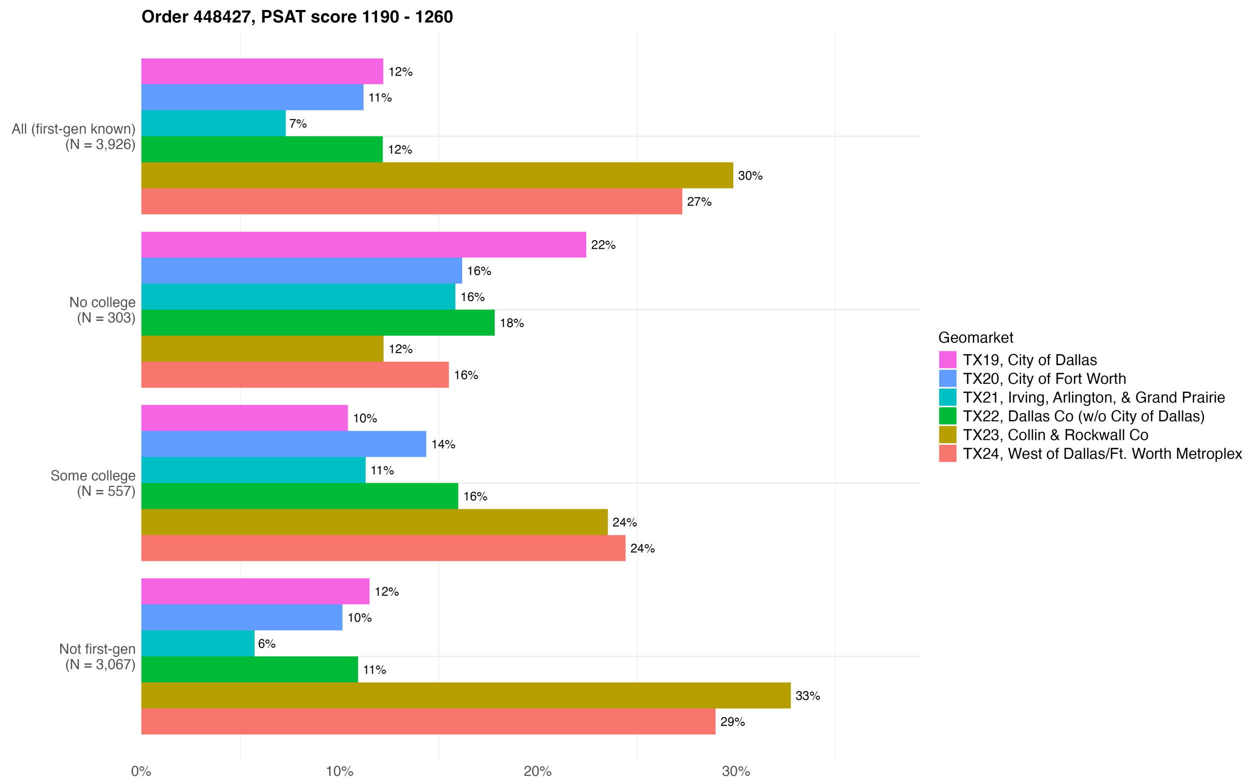


Figure 17: Dallas Geomarket Contribution to Purchased Student profiles by First-Generation Status, Middle-Range SAT orders


Figure 18: Dallas Geomarket Contribution to Purchased Student profiles by Racial/Ethnic Group and First-Generation Status, Middle-Range SAT orders

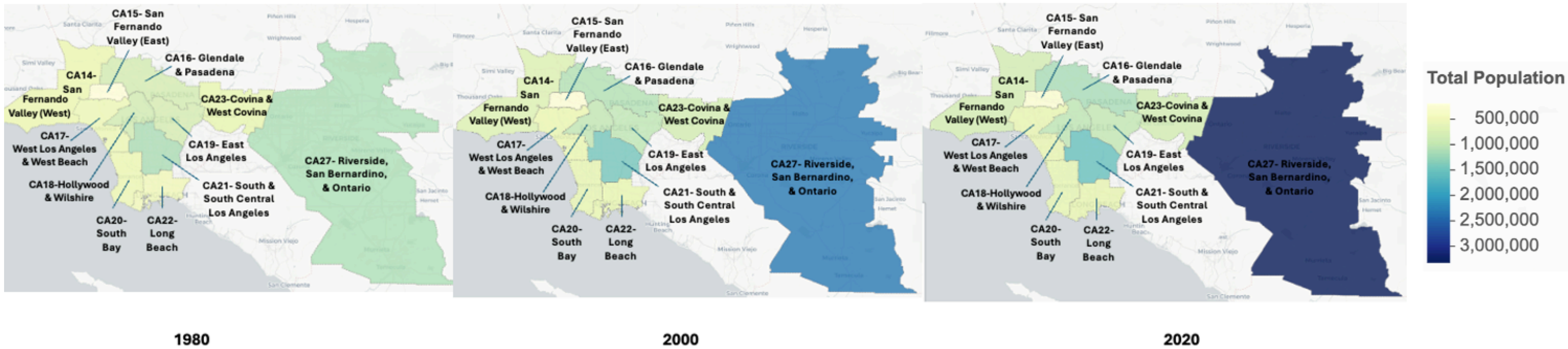


Figure 19: Los Angeles Geomarkets, Total Population 1980-2020


Figure 20: Racial/Ethnic Composition of Los Angeles Area Geomarkets, 1980-2020

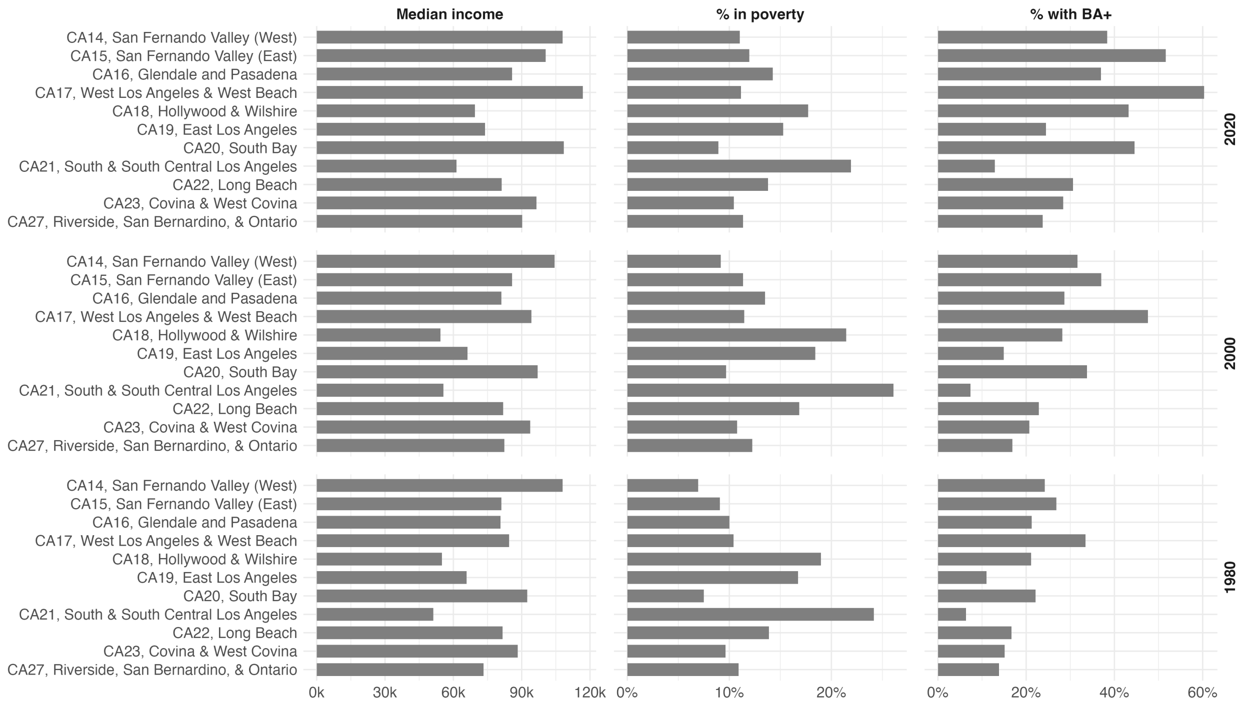


Figure 21: Socioeconomic Characteristics of Los Angeles Area Geomarkets, 1980-2020

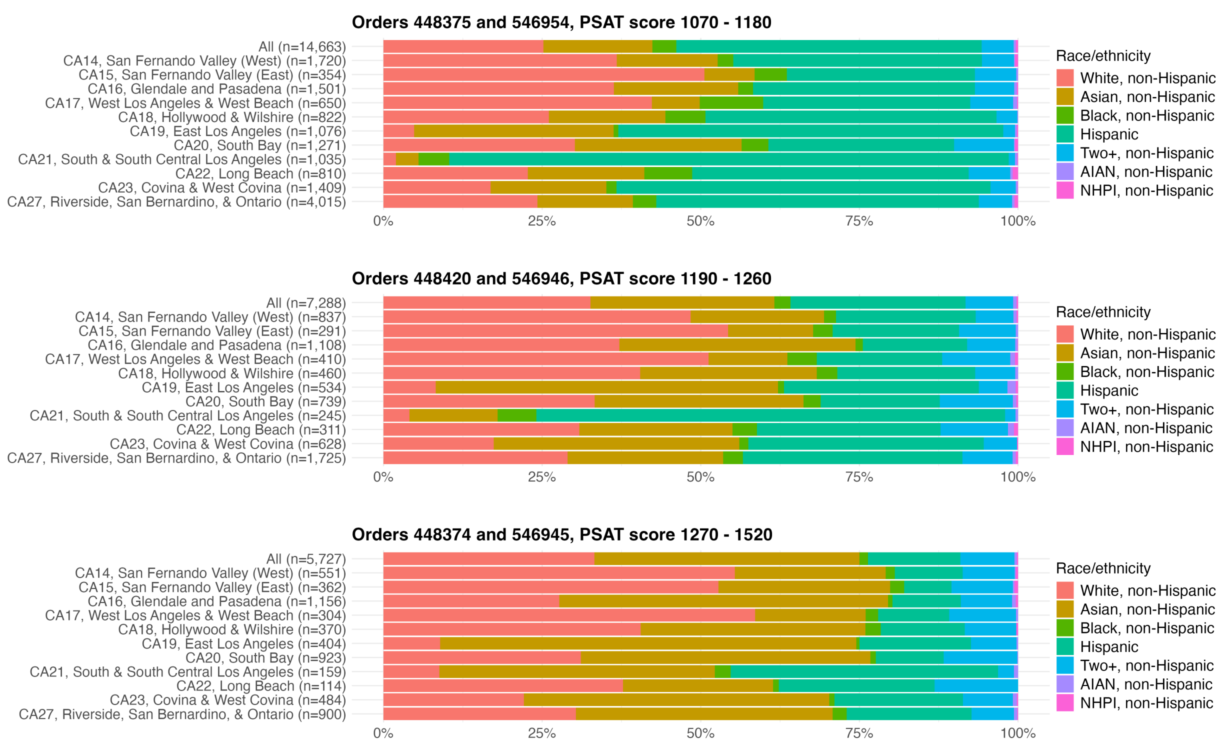


Figure 22: Racial/Ethnic Composition of Purchased Student Profiles by Geomarket, Los Angeles Area

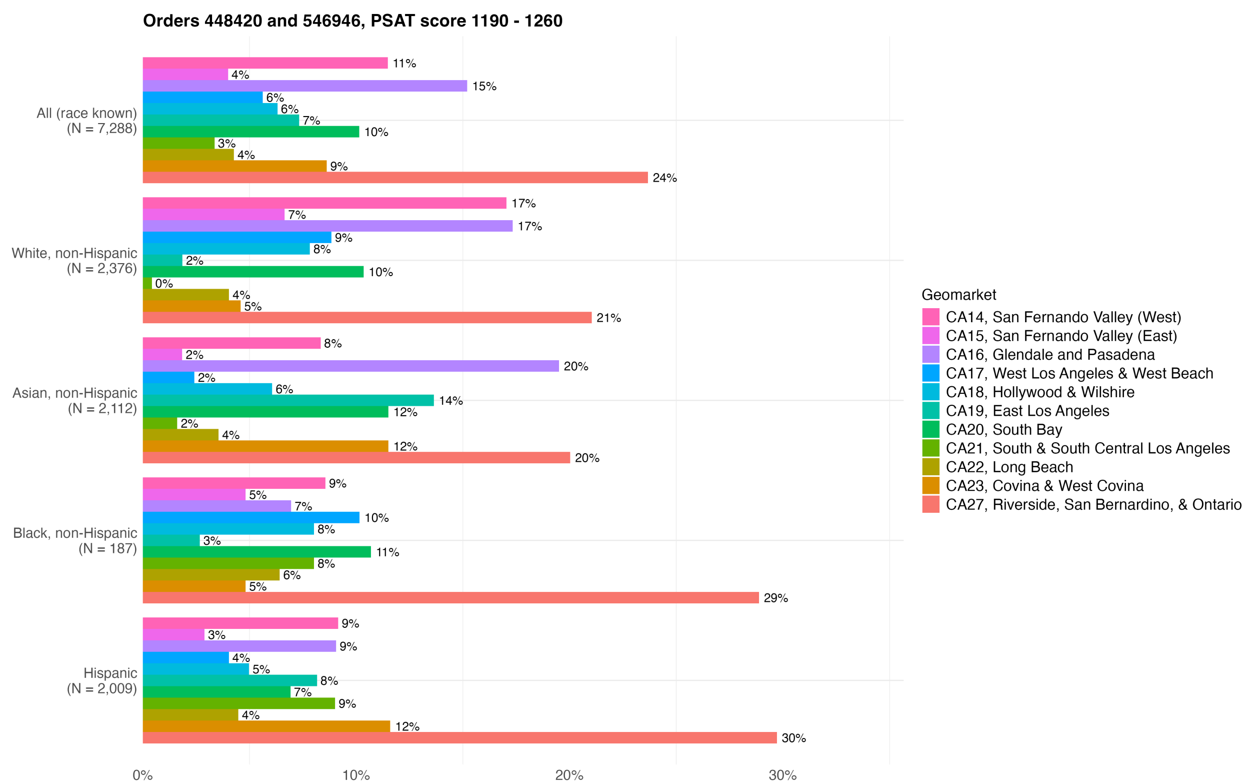


Figure 23: Los Angeles Geomarket Contribution to Purchased Student profiles by Racial/Ethnic Group, Middle-Range SAT orders

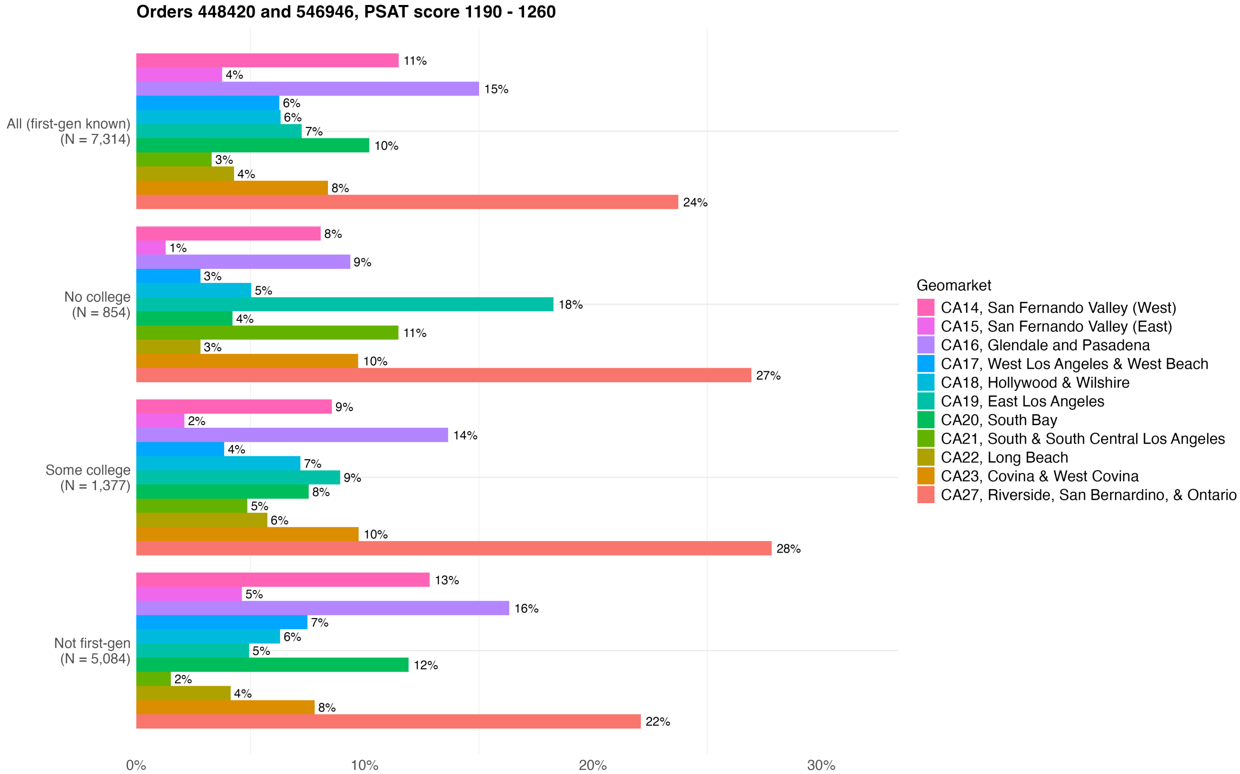


Figure 24: Los Angeles Geomarket Contribution to Purchased Student profiles by First-Generation Status, Middle-Range SAT orders

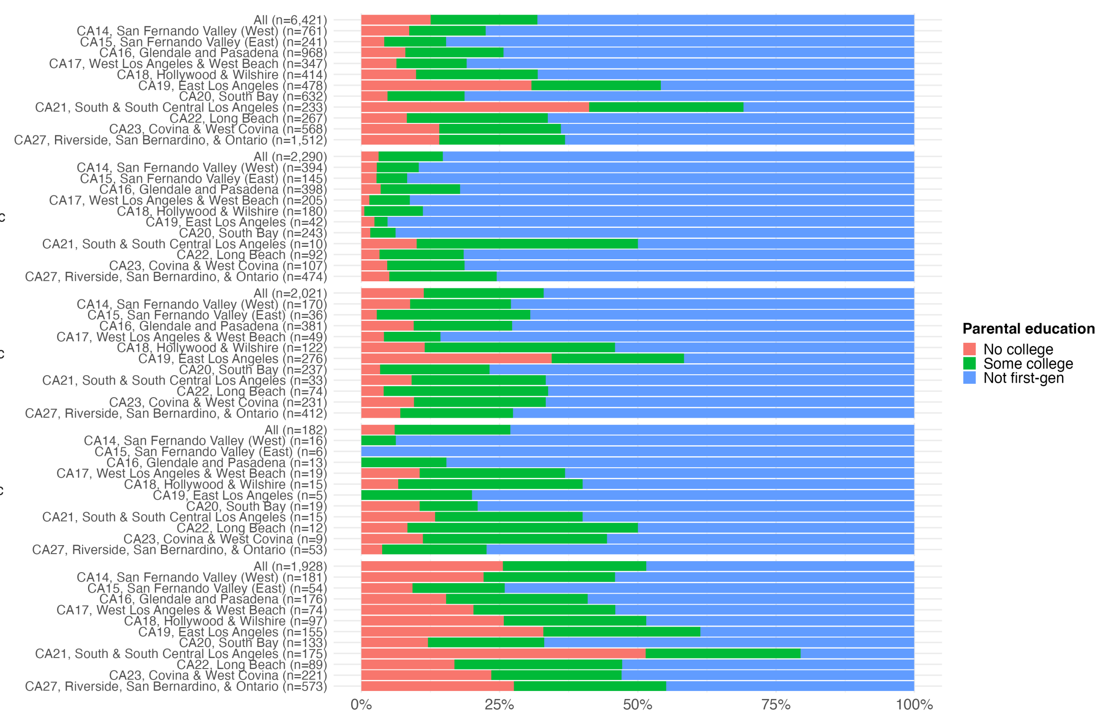


Figure 25: Los Angeles Geomarket Contribution to Purchased Student profiles by Racial/Ethnic Group and First-Generation Status, Middle-Range SAT orders