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

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

ABSTRACT

The Structure of College Choice (Zemsky & Oedel, 1983) created "Geomarkets" and the "Market Segment Model." Geomarkets carve states and metropolitan areas into smaller geographic units, meant to define local recruiting markets. The Market Segment Model predicts how student demand for a particular college varies by Geomarket, based on the socioeconomic characteristics of households. Geomarkets became an input for two College Board products that help colleges recruit students. First, the Enrollment Planning Service (EPS) software recommends specific Geomarkets and high schools from which colleges should recruit. Second, the Student Search Service sells the contact information of prospective students – referred to as "student lists" – and colleges can filter by Geomarket to determine which prospect profiles they purchase. We draw from scholarship on quantification, particularly the discussions of correlation and homophily by Chun (2021), to conceptualize how recruiting products incorporate Geomarkets. We address two research questions: What is the socioeconomic and racial variation between Geomarkets and how does this variation change over time? How does the socioeconomic and racial composition of included versus excluded prospects vary when student list purchases filter on particular Geomarkets? We answer RQ1 by analyzing Census data from 1980, 2000, and 2020. We answer RQ2 using data on student lists purchased by public universities, which we collected by issuing public records requests. We utilize a quantitative case study design. Metropolitan areas are cases. Analyses consist of descriptive statistics and interactive maps.

1 Data and Methods

1.1 Data and Variables

Our study utilizes two main data sources based on each research question. The first research question is focused on understanding the 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 tract is the smallest geographic area for which measures of both race/ethnicity and socioeconomic characteristics are available. Census tracts are statistical subdivisions of a county designed to capture relatively homogeneous demographic communities that are generally considered to be more accurate for population analyses than zip codes, which are spatial boundaries drawn to demarcate route efficiency for postal services [CITE] .

For the first research question, we use data from the 1980 and 2000 Decennial Census as well as 5-year estimates from the 2020 American Community Survey (ACS). For the 1980 Decennial Census, we specifically draw on Summary Tape File 1 (STF1) for variables about race and Hispanic origin collected via "short form" questions answered by all households and 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 households. The 1980 Decennial Census data collection period closely matches the data collection period of SAT score-sending behavior that Zemsky & Oedel (1983) utilized to create Geomarkets and the Market Segment Model. To examine how Geomarket characteristics change over time, 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 from the 2000 Decennial Census.[1] 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.[2] This data collection period mirrors the period for our primary data collection about student list purchases

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

[2]Data for the 2020 5-year ACS were retreived from the tidycensus R package, created by Walker & Herman (2024)

used to analyze RQ2, as described below.

We assigned each census tract to College Board Geomarkets by implementing a partial spatial merge. Spatial manipulations utilized functions from the *sf* R package. Data to create the shapefiles for the College Board Geomarkets were obtained from a 2012 R-bloggers post.[3] Essentially, the data file had 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. Census tract shapefiles that intersected with multiple Geomarkets (i.e., a census tract's spatial boundaries were spread across two or more Geomarkets) broken into smaller shapefiles wholly contained within a single Geomarket. The variables in these shapefiles were then replaced with the original variable value multiplied by the proportion of land area of the census tract that was contained 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 1,000 people who identify as Hispanic non-white. The partial spatial merge splits this census tract into two observations. The observation assigned to Geomarket A reports $.60 * 1000 = 600$ Hispanic non-white people and the observation assigned to Geomarket B reports $.40 * 1000 = 400$ Hispanic non-white people.

Research question 1, variables. RQ1 asks, what is the socioeconomic and racial variation between Geomarkets in metropolitan areas? And, how does this variation change over time? We created measures of race and ethnicity, specifically the number and proportion of people (all ages) in the following categories: White, non-Hispanic; Black, non-Hispanic; Asian, non-Hispanic; Pacific Islander, non-Hispanic; American Indian and Alaskan Native, non-Hispanic; and two+ races, non-Hispanic. Measures of XXX were not available for 1980 Decennial Census. We also created socioeconomic measures that approxmiate the predictors of market segment identified by Zemsky & Oedel (1983). These variables are median household income, mean household income, percent of households below the poverty line, and people age 25+ with a BA.[4] All measures were created

[3]The author of the post reported that he found the zip codes associated with Geomarkets from a Google search.
[4]STATE THAT MEAN HOUSEHOLD INCOME CREATED BY DIVIDING AGGREGATE INCOME BY NUMBER OF PEOPLE. NEED FIX THIS MEASURE THO.

at the census tract-level and the Geomarket-level by aggregating tract-level data to Geomarkets.

Research question 2, data. Geomarket is a search filter available in the College Board Student Search Service product. RQ2 asks, how does the racial and socioeconomic composition of included versus included prospects vary when student list purchases filter on particular 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 the search criteria.

The data collection sample for this project was all public universities in CA, IL, MN, and TX. Utilizing public records requests to obtain public records is a painstaking process. Initially, the majority of universities did not respond to our request or denied our request. Subsequently, we obtained pro bono representation from four law firms, which substantially increased the success of data collection. We 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 TX.

Even with firm representation, 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 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 a third-party consulting firm. Unfortunately, we were rarely able to obtain usable data from these universities. A small number of universities denied requests based on potentially questionable legal rationale, but we lacked the resources to litigate.

This article analyzes student lists purchased from College Board. We exclude student list purchases by MN public universities from this report because Minnesota is predominantly an "ACT state." 14 public universities provided PDF order summaries containing the search criteria for 830 student list purchases. We utilized the XXXX Python/R package to convert these order summaries into tabular data. About 16% of student list purchases utilized the Geomarket filter.

We received both (a) the order summary data and (b) the de-identified prospect-level student list data for 414 orders associated with 2,549,085 prospects. We draw from these data to answer RQ2. 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.[5]. 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.

In particular, we answer RQ2 by analyzing particular student list purchases that did *not* utilize the Geomarket search filter, but did purchase prospects from all Geomarkets in a particular metropolitan area. For each prospect we know their home zip code and, therefore, which Geomarket they belong to. We answer RQ2 by simulating which prospective students would have been included or excluded from the student list purchase had the purchase filtered on partiulcar Geomarkets.

Research question 2, variables. The variables of interest for RQ2 the characteristics of prospective students who would have been included/excluded by our simulations and the characteristics of their schools and neighborhoods. Prospect characteristics are derived from the pretest questionnaire administered to College Board test-takers. The de-identified student lists we received contain a standard subset of variables (e.g., HS code, HS GPA range, intended major). We create detailed measures of race and ethnicity. Information about student socioeconomic status is limited. However, we create a measure of whether the prospect is first-generation college student. We can also measure the socioeconomic characteristics of the students' home zip code and characteristics of their high school [EXPAND OR DELETE THIS SENTENCE DEPENDING ON WHETHER WE USE THESE DATA].

[5]ADD CAVEAT THAT PURCHASES EXCLUDE PROSPECT PROFILES THAT WERE PREVIOUSLY PURCHASED

1.2 Methods

Research design. We utilize a multiple, quantitative case study design in which metropolitan areas are cases.

Following XXX [CITE], we focus on a small number [HOW MANY] of metropolitan areas in order to provide sufficient analytic depth while also situating each case within the context of historic segregation and contemporary gentrification. Online Appendix XXX presents results for a larger number of metropolitan areas.

Choice of metropolitan areas is informed by several factors. Based on our read of Zemsky & Oedel (1983), we are interested in "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). Building on our conceptual framework, we are interested in racial and socioeconomic inequality between geomarkets within a metropolitan area (RQ1) and how this inequality contributes to inequality in which prospective students are targeted by student list purchases (RQ2). [DISCUSSION OF WHICH CASES TO SELECT. KARINA]. Choice of metropolitan areas is informed by data availability. Although Census data are available for all metropolitan areas (RQ1) we do not have good candidate student list purchases for all metropolitan areas (RQ2). [SAY SOMETHING ABOUT WHAT IS A GOOD CANDIDATE STUDENT LIST PURCHASE – PURCHASE WITH NOT TOO MANY CRITERIA THAT INCLUDE ALL GEOMARKETS IN A METRO].

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. For example, in the San Francisco Bay Area, how does the City of Oakland Geomarket (CA07) differ from the Alameda County Exluding Oakland Geomarket (CA08). 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 areas. 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. Interactive maps provide a more granual visualization of where prospects lived and went to high school. For example, for purchased profiles who lived in the Chester County Geomarket (PA02) [OR PHILADELPHIA PA05] and identified as Black, where in the county did these students live and which high schools did they attend?

NEXT STEPS: CREATE TABLE FOR PHILLY STUDENT LIST PURCHASES THAT FOCUSES ON FIRST GEN STATUS.

Our conceptual framework suggests that racial and socioeconomic inequality is Geomarket regions is associated with the number of geomarkets in the metropolitan area. and is associated with the patterns of racial segregation in the metropolitan area. comment more on this.

Below, we use simulations and actual student list purchases that filtered on geomarkets. In purchases that filtered on test-score and/or GPA thresholds, we can simluate who would be included and excluded had certain Geomarkets were been selected. In purchases that filter on geomarkets, we can get an initial sense of who is included in geomarkets targeted by regional state colleges versus research universities.

2 Findings

2.1 Chicago

Shifting Demographics, Persistent Segregation in Geomarket Landscape

Figure 2 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 & Oedel (1983) in the development of The Market Segment Model. Based on the color gradient scale representing total population density, Figure 2 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 occured in the Western Suburbs (1.2 Million to 1.6 Million) and the South and Southwest Suburbs (less than 1 Million to 1.3 Million) Geomarkets. The City of Chicago Geomarket, forming the metropolitan area's central core, was the only locality to experience a population decline, from 3.1 Million in 1980 to 2.8 Million in 2020, despite remaining the most densely populated Geomarket.

These population shifts have contributed to changes in the racial/ethnic composition within and across Geomarkets over time. Figure 3 illustrates the racial and ethnic composition of each Geomarkets across the three census periods. In 1980, nearly all Geomarkets were predominantly White, with proportions ranging from 83% in the South and Southwest Suburbs to approximately 93% in North Shore. In contrast, the City of Chicago had a substantially more diverse population, with White residents comprising only 46% of its total population. By 2020, several Geomarkets that were previously predominantly White shifted to a more balanced racial/ethnic distribution, including Chain of Lakes (59% White, 8% Black, 24% Hispanic; 7% Asian Pacific Islander), Northwest Suburbs (62% White, 3% Black, 18% Hispanic; 15% Asian Pacific Islander), Evanston and Skokie (64% White, 7% Black, 10% Hispanic; 16% Asian Pacific Islander), Western Suburbs (56% White, 9% Black, 23% Hispanic; 9% Asian Pacific Islander), and South and Southwest Suburbs (53% White, 24% Black, 18% Hispanic; 3% Asian Pacific Islander). 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 19 percentage point increase for Western Suburbs. 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 3 also indicates that, although the North Shore and City of Chicago Geomarkets underwent some changes in racial and ethnic composition, their overall demographic patterns remained relatively the same across the three census periods. For instance, North Shore experienced some increases in Hispanic (2% in 1980 to 8% in 2020), Asian Pacific Islander (7% in 2000 to 10% in 2020), and Black (2% in 1980 to 3% in 2020) populations yet remained a predominantly White Geomarket (74% in 2020). City of Chicago also experienced relatively modest declines in its Black population (37% in 1980 to 33% in 2020), alongside increases in its Hispanic (15% in 1980 to 26% in 2020) and Asian Pacific Islander populations (4% in 2000 to 6% in 2020), further substantiating the Geomarket as the most racially/ethnic diverse in the metropolitan area.

We also explore socioeconomic demographics - median income, percent living below poverty, and educational attainment - within and across Geomarkets over time in Figure 4. 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 $158,000 in the North Shore Geomarket to a low of $66,000 in the City of Chicago Geomarket. However, median incomes in all other Geomarkets were closer to the upper thresholds of affluence, such as $120,000 in the Northwest Suburbs, $119,000 in Evanston and Skokie, $108,000 in Chain of Lakes, $105,000 in the Western Suburbs, and $99,000 in the South and Southwest Suburbs. While overall income patterns across Geomarkets remained consistent through 2020, a median income increase for the City of Chicago Geomarket ($79,000) and decrease in the South and Southwest Suburbs Geomarket ($93,000) narrowed the income gap between the lowest and second-lowest Geomarkets to just $14,000.

Figure 4 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 had the lowest percentage of residents with a bachelor's degree or higher (13%). However, by 2020, the Geomarket experienced the largest percentage point increase in college degree attainment, reaching 38%. In comparison, all other Geomarkets experienced smaller gains in the share of residents with a bachelor's degree or higher over the same period. These trends contributed to a narrowing of the educational attainment gap across Geomarkets from 1980 to 2020.

From 1980 to 2020, these demographic shifts have transformed the Chicago metropolitan area's Geomarket landscape from largely a binary structure into a more tiered dynamic shaped by racial/ethnic and socioeconomic characteristics. 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 outmigration, as well as some narrowing of education and poverty gaps, have contributed to a more stratified Geomarket dynamic in the metropolitan area despite significant disparities remaining. The City of Chicago Geomarket remains the critical geographical hub across all Geomarkets for racial/ethnic diversity; however, other Geomarkets 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 the the Evanston and North Shore Geomarkets remain predominantly White. While some stark socioeconomic gaps within Geomarkets have narrowed since 1980, significant geographical income segregation remains across Geomarkets despite a more clearly defined middle-to-high socioeconomic Geomarket group (Chain of Lakes, Northwest Suburbs, Western Suburbs).

Composition of Students Lists by Geomarket Contributions

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 targeting prospects across the entire state of Illinois. Two of these orders applied the same filters: SAT score thresholds ranging from 1020 to 1150, GPA ranging from B- to A+, and high school class targeting students from the 2019-2020 and 2020-2021 graduating classes, respectively. The remaining four orders used similar GPA and high school class filters but differed in SAT thresholds- two targeted students with relatively middle-range scores (1160 to 1300) and two 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 5 presents the racial/ethnic composition of purchased student profiles within Geomarkets in the Chicago metropolitan area. At the lowest PSAT score thresholds (top panel, 1020-1150) the racial/ethnic distribution of prospects within Geomarkets generally reflects the demographic makeup of the metropolitan area in 2020 (see Figure 3). For example, more than 8,000 of the 16,485 total prospects in this SAT range order were concentrated in the Western Suburbs (56% White, 8% Asian, 6% Black, 26% Hispanic, 4% Multiracial) and South and Southwest Suburbs (55% White, 3% Asian, 17% Black, 22% Hispanic, and 3% Multiracial), both of which yield prospect pools comparable to the racial/ethnic makeup of their overall Geomarket populations. Similar to the larger metropolitan area, the Northshore (78% White, 8% Asian, 2% Black, 8% Hispanic) and the Northwest Suburbs (61% White, 11% Asian, 2% Black, 23% Hispanic) Geomarket pools exhibit the greatest proportions of White prospects, reflecting the Geomarket dynamics in the metropolitan area population although they yielded fewer overall students. Interestingly, while the City of Chicago is the most racially/ethnically diverse Geomarket in the larger metropolitan area, its prospect pool in Figure 5 includes nearly double the proportion of Hispanic (50%) prospects relative to the larger Geomarket resident population (26% Hispanic) in Figure 3 with a lower proportion of White prospects (14%).

However, the racial/ethnic composition of prospect pools become less similar to the overall metropolitan area as PSAT scores increase. Notably, the shares of White and Asian prospects within Geomarkets increase, while Black and Hispanic prospects decrease. For instance, Asian prospects over representation in Geomarkets relative to the overall resident population at the 1310-1600 SAT score threshold ranges from a minimum of 4 percentage points in the North Shore Geomarket (17% of prospects versus 13% of residents) to a maximum of 21 percentage points in the Chain of Lakes (28% of prospects versus 7% residents) Geomarket. The City of Chicago Geomarket exhibits the largest underrepresentation of Black prospects (7% of prospects versus 33% of residents) while Chain of Lakes exhibits the largest underrepresentation of Hispanic prospects (7% of prospects versus 24% of residents).

We also explore contributions across Geomarkets to overall prospect pools by race/ethnicity. For instance, Figure 6 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 13,643 total prospects (whose race/ethnicity is known) included in these orders by Geomarkets: 11% reside in Chain of Lakes, 13% in Northwest Suburbs, 5% in North Shore, 5% in Evanston and Skokie, 17% in City of Chicago, 28% in Western Suburbs, and 20% in South and Southwest Suburbs. These overall distributions are compared to those in the lower panels of Figure 6, which illustrate the proportional representation of each race/ethnicity group across Geomarkets.

More than half (n=8,256) of all prospects whose contact information was purchased in these middle SAT score orders identified as White. Figure 6 shows nearly all Geomarkets contributed to the pool of White prospects at rates nearly proportional to their overall contributions to all included prospects. One exception is the City of Chicago Geomarket, which contributed only 7% of White prospects, despite accounting for 17% of all included prospects. On the other hand, Western Suburbs and South and Southwest Suburbs Geomarkets 32% and 23% of White prospects while only accounting for 28% and 20% of all included prospects, respectively. Similar patterns are observed for Asian students, with proportional contributions by Geomarket largely consistent with thier overall distribution, except for Northwest Suburbs (19% versus 13% for all prospects), Evanston and Skokie (10% versus 5% for all prospects) contributing disproportionately larger shares and the South and Southwest Suburbs (8% versus 10% for all prospects) contributing smaller shares relative to their overall contribution to the larger prospect pool.

In contrast, these representational patterns are reversed for Black and Hispanic prospects in Figure 6. Among the nearly 1,000 Black students included in these orders, the City of Chicago and South and Southwest Suburbs contributed disproportionately higher shares, 48% and 30%, respectively, compared to their overall contributions of 17% and 20% of all included prospects, whereas all other Geomarkets contributed substantially smaller proportions of Black prospects. The City of Chicago Geomarket also contributed larger shares of the nearly 3,000 Hispanic students (39% versus 17% for all prospects), while all other Geomarkets contributed comparable or relatively smaller proportions of Hispanic prospects than their overall representation among all included prospects.

Appendix figures @fig-chicago-rq2-midSAT-race and **?@fig-chicago-rq2-highSAT-race** similarly depict Geomarket contributions for pools of prospects resulting from orders targeting lower-range (1020-1150) and high-range (1310-1600) SAT scores. Both figures demonstrate how representational Geomarket contributionsacross race/ethnicity, such as City of Chicago disproportionately contributing larger shares of Black and Hispanic prospects, remain consistent even at lower and higher SAT score thresholds.

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 7 presents the Geomarket distribution for the 13,749 total prospects (whose first-generations status is known) included in middle-range SAT score orders (1160-1300). Among the 1,743 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 - 38% in comparison to 17% 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 (South and Southwest Suburbs) to a nine-percentage-point difference (Western Suburbs). For prospects whose parents attended some college but did not complete their degrees ("some college"), Figure 7 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 (12%) relative to all prospects (17%) in this SAT range, whereas all other Geomarkets contributed comparable shares.

Relative to Figure 7, appendix figures @fig-chicago-rq2-midSAT-ses and **?@fig-chicago-rq2-highSAT-ses** demonstrate Geomarket contribution patterns by first-generation college student status are less pronounced in low-SAT range orders (1020-1150) and more pronounced in high-range SAT orders (1310-1600), respectively.

Finally, we examine Geomarket contributions to prospect pools by both first-generation college student status and race/ethnicity in the Chicago metropolitan area. Figure 8 presents the Geomarket distribution of the 12,981 total prospects (whose first-generation status and race/ethnicity is known) included in middle-range SAT (1160-1300) orders. For instance, among all White prospects (n=8,155) whose contact information was purchased in this SAT score range, approximately 6% were first-generation college students whose parents did not attend college, 15% were first-generation college students whose parents attended some college but did not complete their degree, and 79% were not first-generation college students. Figure 8 shows the City of Chicago and South and Southwest Suburbs Geomarkets disproportionately contribute larger shares of White first-generation college students who parents did not attend college (10% and 9%, respectively). For first-generation college students whose parents attended some college but never completed degrees, only the South and Southwest Suburbs (25%) Geomarket contributed a substantially larger share of White prospects. In contrast, Geomarkets such as Chain of Lakes (82%), North Shore (94%), Evanston and Skokie (81%), and Western Surburbs (84%) contribute a disproportionately larger share of White prospects who are not first-generation college students.

Figure 8 indicates these Geomarket patterns persist across other racial/ethnic categories. The City of Chicago Geomarket consistently contributed a disproportionately large share of first-generation college students whose parents did not attend college across all racial/ethnic categories. However, the magnitude of this disproportionality varied- from a 1-percentage-point difference for Black prospects to a 23 percentage-point difference for Asian prospects. Among Hispanic students, contributions to the first-generation prospect pool were notably balanced across Geomarkets, with the exception of North Shore and Evanston and Skokie Geomarkets, which contributed substantially lower shares of first-generation Hispanic students. Although Asian and Black prospects who are not first-generation college students make up a relatively smaller portion of the overall pool, Figure 8 suggests that such prospects are primarily contributed by the Chain of Lakes, North Shore, Evanston and Skokie, and Western Surburbs Geomarkets. Appendix figures **?@fig-chicago-rq2-lowSAT-combo** and @fig-chicago-rq2-highSAT-combo show the magnitude of these Geomarket contribution disparities become more pronounced at as SAT score thresholds increase.

The enduring segregated structure of the Chicago metropolitan area from 1980 to 2020 analyzed above plays a significant role in how Geomarkets shape the racial/ethnic and socioeconomic composition of prospect pools. Reflecting the larger metropolitan area dynamics, central 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 crucial role in fostering college access for first-generation Asian college students, contributing such prospects to the overall pool at significantly higher, disproportionate rates than any other Geomarket.

2.2 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 9 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 9, two distinct inner-core Geomarkets (City of Dallas and City of Fort Worth) are seperated by a centrally located suburban Geomarket encompassing Irving, Arlington, & 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/Ft. Worth Metroplex).

From 1980 to 2020, the population of the overall Dallas-Fort Worth metropolitan area increased substantially from approximately 2.7 Million to 6.8 Million. Figure 9 showcases all six Geomarkets experienced population growth over the three census periods. However, the most pronounced increases occured in the West of Dallas/Fort Worth Metroplex and Collin & Rockwall Counties Geomarkets, which grew from approximately 340,000 to 1.4 Million and 170,000 to 1.2 Million, respectively. These population trends, as shown in Figure 9, suggest population growth has been partly driven by urban sprawl from the metropolitan centers - City of Forth 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 10 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 62% in City of Dallas to nearly 92% 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

(58% White, 16% Black, 19% Hispanic; 6% Asian Pacific Islander), Irving, Arlington, and Grand Prairie (54% White, 12% Black, 26% Hispanic; 16% Asian Pacific Islander), and City of Fort Worth (57% White, 16% Black, 22% Hispanic; 3% Asian Pacific Islander). In contrast, Collin & Rockwall Counties and West of Dallas/Fort Worth Metroplex Geomarkets remained predominantly White in 2000 at 78% and 80%, respectively. Notably, by 2000, the City of Dallas was no longer a predominantly White Geomarket, with a racial/ethnic composition of approximately 24% Black, 30% Hispanic, and 3% Asian Pacific Islander residents.

These shifts in racial and ethnic composition within and across Geomarkets became more pronounced 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 (58% 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 (33% White, 23% Black, 32% Hispanic; 9% Asian Pacific Islander), Irving, Arlington, and Grand Prairie (30% White, 20% Black, 36% Hispanic; 12% Asian Pacific Islander), and City of Fort Worth (46% White, 15% Black, 33% Hispanic; 4% Asian Pacific Islander). Lastly, continuous increases in the shares of Black and Hispanic residents further diversified the City of Dallas, making it the most racially/ethnically diverse Geomarket in the Dallas-Fort Worth metropolitan area.

Socioeconomic demographics of median income, percent living below poverty, and educational attainment over time are presented in Figure 11 for the Dallas-Fort Worth metropolitan area. Overall, median income both within and across Geomarkets remained relatively stable across the three census periods. Between 1980 to 2000, affluence levels across Geomarkets were consistent, ranging from a low of $82,000 in City of Fort Worth to a high of $135,00 in the Collin & Rockwall Counties Geomarket. However, the substantial increase in median income for the Collin & Rockwall Counties Geomarket during this time period plateaued by 2020, reshuffling relatively affluence levels among Geomarkets. By 2020, Collin & Rockwall Counties remained the most affluent Geomarket with a median household income of $133,000, followed by West of Dallas/Fort Worth Metroplex ($116,000), and Dallas County ($90,000). In contrast, median household income for City of Dallas and City of Fort Worth remained relatively stable and continued to rank at the lower levels of affluence within the metropolitan area. [KS: NEED TO CHECK THIS; I THINK CITY OF DALLAS AND CITY OF FORT WORTH ARE MIXED UP BETWEEN GRAPH VERSUS TABLE]

Figure 11 also illustrates changes in the percentage of residents living in poverty within Geomarkets over time, revealing shifting dynamics across Geomarkets similar to those observed in median household income. In 1980, poverty rates were comparably low in both Collin & Rockwall Counties and Dallas County, moderate in the Irving, Arlington, and Grand Prarie and West of Dallas/Fort Worth Metroplex Geomarkets, and highest for City of Dallas and City of Fort Worth Geomarkets. Overtime, poverty rates declined in Collin & Rockwall Counties and West of Dallas/Fort Worth Metroplex, while increasing for all other Geomarkets. These divergent trends resulted in a widening gap in poverty levels across the Dallas-Fort Worth metropolitan area. A similar widening disparity is evident in educational attainment. Although all Geomarkets experienced increases in the percentage of residents with a bachelor's degree or higher between 1980 and 2020, the Collin & Rockwall Counties Geomarket experienced a substantially greater increase - from 23% in 1980 to 50% in 2020- maintaining the gap with other Geomarkets in terms of educational attainment.

The Dallas-Fort Worth metroplex experienced substantial demographic growth since Geomarket boundaries were dilineated by Zemsky & Oedel (1983) in the early 1980s. While all six Geomarkets experienced population increases, the most significant growth occurred in the outer suburban areas, reflecting ongoing urban sprawl. Such growth is largely attributed to increases in Hispanic, Black, and Asian populations leading to only two of the six Dallas-Fort Worth Geomarkets remaining predominantly White (Collin and Rockwall Counties and West of Dallas/Fort Worth Metroplex. On the other hand, the metropolitan area has experienced widening disparities in income, poverty, and education attainment across Geomarkets, especially between the historically affluent Collin and Rockwall Counties and the less economically advantaged inner-core Geomarkets (City of Fort Worth and City of Dallas). Together, these trends underscore the growingly racial/ethnic diverse but widening economic inequality within the evolving Geomarket landscape of the Dallas-Fort Worth metroplex.

Composition of Students Lists by Geomarkets

We also analyze student list orders collected through public requests to assess 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 2019-2020 graduating class and based on PSAT scores across three ranges: 1070-1180, 1190-1260, and 1270-1520. Similar to Chicago analyses, we focus on prospects within the middle PSAT score range (1190-1260) from the Dallas-Fort Worth metropolitan area to simulate Geomarket contributions and assess patterns of inclusion versus exclusion.

Figure 12 presents the racial/ethnic composition of purchased student profiles within Geomarkets in the Dallas-Forthworth area. At the lowest PSAT score thresholds, notable differences emerge in the racial/ethnic composition of prospects within Geomarkets when compared to the overall demographic makeup of the metropolitan area in 2020 (see Figure 10). For example, the largest number of prospects in the 1070-1180 PSAT range were from the Collin & Rockwall Counties (57% White, 7% Black, 17% Hispanic, 14% Asian) and West of Dallas/Forth Worth Metroplex Geomarkets (61% White, 7% Black, 19% Hispanic, 7% Asian), both of which yield predominantly White prospect pools. On the other hand, the Dallas County (36% White, 17% Asian, 15% Black, 28% Hispanic) and the Irving, Arlington, and Grand Prairie (34% White, 13% Asian, 15% Black, 34% Hispanic) Geomarket pools have the greatest proportions of non-White prospects (although yielding fewer overall students). Interestingly, while the City of Dallas is the most racially/ethnically diverse Geomarket in the larger metropolitan area, its prospect pool in Figure 12 includes larger proportion of Hispanic (41%) prospects but lower proportions of Asian (4%) and Black (9%) students comparatively.

These racial/ethnic composition patterns shift as PSAT scores increase. Shares of White prospects within Geomarkets remain relatively proportional to the overall 2020 metropolitan distribution in Figure 10. However, 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 13 presents the race/ethnicity of all prospects targeted in the middle PSAT score range across Geomarket contributions. Figure 13 shows nearly all Geomarkets contributed to the pool of White prospects at rates nearly proportional to their overall contributions to all included prospects, with the exception of the West of Dallas/Fort Worth Metroplex Geomarket contributing a disproportionately larger share (29% versus 23% for all prospects). Larger disparities across Geomarket contributions were evident for Asian, Black, and Hispanic students. The Dallas County (20% versus 13% for all prospects) and Colin and Rockwall Counties (47% versus 37% for all prospects) Geomarkets contributed a disproportionate 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 13. All Geomarkets besides Collin and Rockwall Counties and West of Dallas/Fort Worth Metroplex contributed equal or larger proportions of Black and Hispanic prospects relative to their contributions to the overall prospect pool in the Dallas-Fort Woth metropolitan area. However, the City of Dallas Geomarket contributed the largest disproportionate share of Hispanic prospects (19% versus 11% for all prospects), whereas the Dallas Counties Geomarket contributed the largest disproportionate share of Black prospects (21% versus 13% 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 14. Only 303 prospects of the 3,928 total 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 (14% versus 36% for all prospects) and West of Dallas/Fort Worth Metroplex (15% versus 24% for all prospects) contributed equal or larger proportions of first-generation college students (no college) relative to their contributions to the overall prospect pool. Disproportioantely larger contributions range from six percentage points by the Dallas Counties Geomarket (19% versus 13% for all prospects) to 11 percentage points by the City of Dallas Geomarket (22% versus 11% 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 3,069 prospects that are not first-generation college students, with the exception of Collin and Rockwall Counties Geomarket contributing the largest disproportinate share of not first-generation college students (40%) relative to the contributions to the overall prospect pool (36%).

Lastly, Figure 15 presents Geomarket contributions to prospect pools by both first-generation college student status and race/ethnicity in the Dallas-Fort Worth metropolitan area. Among all White prospects (n=2,094) 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 15 shows nearly all Geomarkets contribute a nearly proportionate share of prospects by first-generation college students relative to the overall pool. The only exception is the Irving, Arlington, & Grand Prairie Geomarket contributing a relatively larger share of White first-generation college students whose parents attended some college but did not complete their degree (21% versus 11% for all prospects).

Similar patterns of one or two Geomarkets disproportionately contributing larger shares of first-generation college students was evident across other racial/ethnic categories in Figure 15. For instance, City of Dallas (18%) and City of Fort Worth (24%) contributed larger proportions of first-generation college students whose parents did not attend college relative to the overall pool of all prospects (8%). On the other hand, the Irving, Arlington, & Grand Prairie (23%) and Dallas Counties (26%) Geomarkets contributed disproportionately larger proportions of Black students whose parents attended some college but did not complete their degree relative to the overall pool of all prospects (17%). For Asian prospects, City of Fort Worth and Irving, Arlington, & Grand Prairie Geomarkets contributed disproportionately larger shares of first-generation college students across both parents with no college and some college. City of Dallas (38%), City of Fort Worth (39%), Irving, Arlington, & Grand Prairie (32%), Dallas Counties (28%) all contributed larger shares of Hispanic students whose parents did not attend college relative to the overall prospect pool of these first-generation college students (23%), whereas only Dallas Counties contributed to larger shares of Hispanic students who parents attended some college but did not attain their degree (26% versus 20% for all prospects).

Analyses of student list orders across the Dallas-Fort Worth metropolitan area reveals patterns of racial/ethnic and socioeconomic stratification in Geomarket contributions. As Dallas-Fort Worth became increasingly diverse from 1980 to 2020, a more balanced racial/ethnic distribution emerged across the metropolitan area's Geomarket landscape. However, these patterns did not translate to racial/ethnic composition of student prospect pools. Suburban Geomarkets like Collin and Rockwall Counties and the West of Dallas/Fort Worth Metroplex yielded a disproportionately larger share of Asian prospects, whereas central Geomarkets such as the City of Dallas, Dallas County, and Irving, Arlington, and Grand Prairie contributed a disproportionately larger share of Hispanic and Black prospects. These racial/ethnic disparities become significantly more pronounced, particularly in the disproportionate share of Asian prospects, as PSAT score thresholds increase (see Appendix **?@fig-dallas-rq2-lowSAT-race** and **?@fig-dallas-rq2-highSAT-race**). Similar Geomarket contribution patterns are also evident by first-generation college student status, with Collin and Rockwall Counties and the West of Dallas/Fort Worth Metroplex providing larger shares of non-first generation college student prospects, and central Geomarkets (City of Dallas, City of Fort Worth, Dallas County, and Irving, Arlington, and Grand Prairie) providing disproportionately larger shares of first-generation college student prospects. Moreover, central Geomarkets are particularly important for including Black (City of Dallas and City of Fort Worth), Hispanic (City of Dallas, City of Fort Worth, Dallas County, and Irving, Arlington, and Grand Prairie), and Asian (City of Fort Worth and Irving, Arlington, and Grand Prairie) first-generation college students. While Collin and Rockwall Counties and the West of Dallas/Fort Worth Metroplex Geomarkets still contribute a notable overall number of Black, Hispanic, and Asian prospects, such students are predominantly not first-generation college students.

2.3 Los Angeles

Decentralized Diversity and Income in a Multi-Nucleo in the Los Angeles Geomarket Landscape

The 11 Geomarkets within the Los Angeles metropolitan area are illustrated in Figure 16 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 (CITE- Harris and Ullman 1945). As depicted in Figure 16, 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 of made up by 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 Bernadino, and Ontario Geomarkets are located inland.

The population growth in the Los Angeles metropolitan area increased from approximately 8 Million in 1980 to more than 12 Million in 2020. Figure 9 showcases population growth by Geomarkets over the three census periods reflect the urban sprawl forming the multi-nuclei model. The most pronounced increases occurred in South and South Central Los Angeles Geomarket, as well as the outer/inland Geomarkets of San Fernando Valley-West, Riverside and San Bernadino, and Ontario. The South and South Central Los Angeles Geomarket grew from approximately 1.2 to 1.6 Million from 1980 to 2020. In contrast, the San Fernando Valley-West and Riverside and San Bernadino, and Ontario Geomarkets grew from approximately 700,000 to 1.2 Million and from 1.1 Million to nearly 3 Million during this time, respectively.

@fig-la-rq1-race presents the racial and ethnic composition of each Geomarket within the Los Angeles metropolitan area from 1980 to 2020. Over this period, 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 62% White in Long Beach to 78% in the San Fernando Valleys (West and East) Geomarkets. By 2020, however, only the West Los Angeles and West Beach Geomarket remained predominantly White (62%). Remanining Geomarkets exhibited a more racially/ethnic diverse population by 2020, including San Fernando Valley-West- (44% White, 4% Black, 39% Hispanic; 11% Asian Pacific Islander), San Fernando Valley -East (43% White, 4% Black, 41% Hispanic; 8% Asian Pacific Islander), Glendale and Pasadena (35% White, 3% Black, 37% Hispanic; 21% Asian Pacific Islander), South Bay (29% White, 12% Black, 33% Hispanic; 21% Asian Pacific Islander), Long Beach (26% White, 11% Black, 46% Hispanic; 13% Asian Pacific Islander), Covina and West Covina (19% White, 3% Black, 52% Hispanic; 24% Asian Pacific Islander), and Riverside, San Bernardino, and Ontario (28% White, 7% Black, 54% Hispanic; 6% Asian Pacific Islander). 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 18% and 18% to 9%, respectively.

Figure 18 presents trends in socioeconomic indicators for the Los Angeles metropolitan area from 1980 to 2020. While median income increased across all of the metropolitan area over the three census periods, some shifts occurred in the relative socioeconomic levels across Geomarkets. Many Geomarkets kept their relative levels of affluence from 1980 to 2020. For instance, the Covina and West Covina, South Bay, and San Fernando Valley remained among the most affluent, whereas the San Fernando Valley East, Glendale and Pasadena, Long Beach, and Riverside, San Bernardino, and Ontario Geomarkets remained at middle ranges of affluence and the Hollywood and Wilshire, South and South Central Los Angeles, and East Los Angeles Geomarkets remained at the lowest levels of affluence across the metropolitan area during this period. The West Los Angeles and West Beach Geomarket demonstrated the most substantial increase in median household income, rising from \$112,000 in 1980 to \$132,000 in 2000, and reaching \$137,000 by 2020. This sustained growth positioned it as the most affluent Geomarket in the Los Angeles metropolitan area.

Figure 18 also illustrates shifting levels of poverty across Geomarkets in the Los Angeles metropolitan area over time. In 1980, poverty rates were relatively low (less than 12%) for all Geomarkets except for East Los Angeles (17%), South and South Central Los Angeles (24%), and Hollywood and Wilshire (18%). However, poverty rates declined in South and South Central Los Angeles (22%), remained stable in East Los Angeles and Hollywood and Wilshire, and increased in all other Geomarkets. Such shifts and West of Dallas/Ft. Worth Metroplex, while increasing for all other Geomarkets. These divergent trends resulted in more staggered poverty rates across the metropolitan area by 2020. This includes four Geomarkets with poverty rates less than 10% (San Fernando Valley- West, West Los Angeles and West Beach, South Bay, and Covina and West Covina), four Geomarkets with poverty rates ranging from 11%-15% (San Fernando Valley- East, Glendale and Pasadena, Long Beach, and Riverside, San Bernardino, and Ontario), and three Geomarkets with poverty rates ranging from 16%-22% (Hollywood and Wilshire, East Los Angeles, South and South Central Los Angeles).

Lastly, Figure 18 shows a widening disparity in educational attainment across the Los Angeles metropolitan area. Although all Geomarkets experienced increases in the percentage of residents with a bachelor's degree or higher between 1980 and 2020, the West Los Angeles and West Beach Geomarket experienced a substantially greater increase - from 35% in 1980 to 66% in 2020- widening the gap with other Geomarkets in terms of educational attainment.

The Los Angeles metropolitan area exemplifies a complex multi-nuclei urban model, characterized by decentralized population growth and increasing demographic diversity across its 11 Geomarkets. From 1980 to 2020, the region's population expanded significantly leading to overall declines in the White population and increases in Hispanic and Asian populations across Geomarkets. West Los Angeles and West Beach is the only majority White Geomarket in 2020. Socioeconomic indicators reveal enduring disparities, especially in historically lower-income Geomarkets like South and South Central Los Angeles, East Los Angeles, and Hollywood and Wilshire. Meanwhile, Geomarkets such as West Los Angeles and West Beach became increasingly affluent and highly educated, highlighting growing socioeconomic and educational divides. Given these demographic shifts within the context of a multi-nuclei urban model, the Los Angeles metropolitan area presents a distinctive Geomarket landscape case in which racial/ethnic enclaves across a range of various socioeconomic levels are distributed across multiple Geomarkets rather than concentrated in one or two.

Composition of Students 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 filtered for prospects across the entire state of California, in the 2019-2020 high school graduating class, 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.

The racial/ethnic composition of purchased student profiles within Geomarkets in the Los Angeles metropolitan area is presented in Figure 19. The more than 15,000 prospects in the resulting student lists at the low PSAT score thresholds (1070-1180) reflect very similar racial/ethnic compositions within Geomarkets than overall demographic makeup of the metropolitan area in 2020 (see Figure 17). However, the racial/ethnic composition of prospects become less proportional to overall metropolitan patterns by Geomarket as PSAT scores increase. For instance, prospect pools at middle range PSAT scores (1190-1260) within Geomarkets become disproportionately more Asian, while the representation of Hispanic and Black prospects decline substantially. This pattern becomes most pronounced at the highest PSAT score prospect pools. This disproportionate over representation for Asian prospects in comparison to the population of residents with Geomarkets ranges from 3 percentage points in the West Los Angeles and West Beach Geomarket (16% prospects versus 13% Asian residents) to 52 percentage points in the East Los Angeles Geomarket (79% prospects versus 27% residents). Figure 19 illustrates the share of Black prospects within these high PSAT orders declines to less than 3% within all Geomarkets in comparison to Black residents making up a range of 3% (Covina and West Covina) to 18% (South and South Central Los Angeles) of residents within Geomarkets. Similar underrepresentation patterns are evident for Hispanic porspects are evident across all Geomarkets execept for a nearly equal representation in South and South Central Los Angeles (73% prospects versus 71% Hispanic residents).

Figure 20 presents the race/ethnicity of all prospects targeted in the middle PSAT score range across Geomarket contributions for the Los Angeles metropolitan area. For the more than 2,500 White prospects in this pool, we see the San Fernando Valley - West (28% versus 20% for all prospects) and the West Los Angeles and West Beach (12% versus 7% for all prospects)Geomarkets contributing disproportionately larger shares. Larger disparities across Geomarket contributions were evident for Asian, Black, and Hispanic students. The Glendale and Pasadena (17% versus 14% for all prospects), East Los Angeles (11% versus 5% for all prospects) and Covina and West Covina (17% versus 11% for all prospects) Geomarkets contributed a disproportionate larger share of Asian students, whereas nearly all other Geomarkets contributed disproportionately smaller or nearly equal shares.

Representational patterns across Geomarkets are reversed for Black and Hispanic prospects in Figure 20. South and South Central Los Angeles and Riverside, San Bernadino, and Ontario Geomarkets contributed larger proportions of Black and Hispanic prospects relative to their contributions to the overall prospect pool in the Los Angeles metropolitan area. On the other hand, the San Fernando Valley- West and the Glendale and Pasadena Geomarket contributed smaller proportions of Black and Hispanic Prospects. However, the West Los Angeles and West Beach (10% versus

7% for all prospects), Hollywood Wilshire (7% versus 5% for all prospects), and the Long Beach (8% versus 5% for all prospects) Geomarkets contributed a larger share of Black prospects, whereas Covina and West Covina contributed a larger share of Hispanic students. All other Geomarkets contributed nearly equal shares of Black and Hispanic prospects.

Contributions of each Geomarket to the pool of prospects from the Los Angeles metropolitan area based on first-generation college student status are presented in Figure 21. Only 889 of the more than 8,000 total prospects (11%) in this middle PSAT score range (1190-1260) pool were first-generation college students whose parents did not attend college. Only three of the 11 Geomarkets contributed substantially larger shares of these first-generation prospects: East Los Angeles (35%), South and South Central Los Angeles (47%), and Riverside, San Bernardino, and Ontario (15%). However, disproportionate contributions become more spread across Geomarkets for the nearly 1,436 of the more than 8,000 total prospects (18%) who were first-generation college students whose parents attended but did not complete their degree. Overall, six of the 11 Geomarkets contributed larger shares of these prospects: Hollywood & Wilshire (25%), East Los Angeles (26%), South and South Central Los Angeles (28%), Long Beach (25%), Covina and West Covina (22%), and Riverside, San Bernardino, and Ontario (22%). This leads to the remaining Geomarkets - San Fernando Valley- West (80%), San Fernando Valley- East (78%), Glendale and Pasadena (77%), West Los Angeles & West Beach (87%), South Bay (81%) - contributing disproportionately larger shares of not first-generation college students relative to the overall prospect pool (71%).

Lastly, Figure 22 presents Geomarket contributions to prospect pools by both first-generation college student status and race/ethnicity in the Los Angeles metropolitan area. Among all White prospects (n=2,511) 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 22 shows all Geomarkets contribute a nearly proportionate share of prospects by first-generation college students (across both no college and some college parents) relative to the overall pool. The South and South Central Los Angeles Geomarket stands out as one exception contributing a relatively larger share of White first-generation college students whose parents did not attend college (17%) and whose parent attended some college but did not complete their degree (33%). Other Geomarket that contributed signifcantly larger shares of White first-generation prospects whose parents did not complete their degree include Long Beach (17%), and Riverside, San Bernardino, and Ontario (19%).

Geomarkets contributions to first-generation college students was mixed across other racial/ethnic categories in Figure 22. For Asian prospects (11% first-generation no college, 21% first-generation some college, and 68% not first-generation), the Hollywood & Wilshire (12%) and East Los Angeles (36%) Geomarkets contributed larger proportions of first-generation college students whose parents did not attend college. The San Fernando Valley- East (28%), Hollywood & Wilshire (35%), East Los Angeles (25%), South and South Central Los Angeles (44%), Long Beach (27%), and Covina and West Covina (25%) Geomarkets also contributed larger proportions of first-generation college students whose parents attended college but did not complete their their degree.

For Black prospects (5% first-generation no college, 20% first-generation some college, and 74% not first-generation), Geomarkets that disproportionately larger proportions of first-generation college students whose parents did not attend college include South Bay (11%), South and South Central Los Angeles (21%), Long Beach (7%), and Covina and West Covina (9%). The West Los Angeles and West Beach (26%), Hollywood & Wilshire (36%), East Los Angeles (33%), Long Beach (40%), and Covina and West Covina (27%) Geomarkets also contributed larger proportions of Black first-generation college students whose parents attended college but did not complete their their degree.

Finally, Hispanic prospects in this middle range PSAT order were on average 25% first-generation college students whose parents did not attend college, 25% first-generation college students whose parents attended some college but did not complete their degrees, and 50% were not first-generation college students. The Hollywood & Wilshire (30%), East Los Angeles (39%), South and South Central Los Angeles (54%), and Riverside, San Bernardino, and Ontario (31%) contributed disproportionate shares of Hispanic first-generation college students whose parents did not attend college. The East Los Angeles (33%), South and South Central Los Angeles (28%), Long Beach (30%), and Riverside, San Bernardino, and Ontario (27%) also contributed larger Hispanic first-generation college students whose parents attended college but did not complete their their degree.

Despite Los Angeles's multi-nuclei urban structure contributing to a more decentralized Geomarket landscape with racial/ethnic diversity across a range of socioeconomic levels more spatially distributed across the metropolitan area, Geomarkets contributions to student prospect pools reflect racial and socioeconomic stratification patterns similar to other metropolitan cases in the study. Affluent Geomarkets like San Fernando Valley–West and West Los Angeles contribute an overrepresentation of White, Asian, and not-first generation college student prospects, the magnitude of which increases substantially as score ranges increase (see Appendix **?@fig-la-rq2-lowSAT-race** and **?@fig-la-rq2-highSAT-race**). On the other hand, more diverse and historically low-income Geomarkets, like South and South Central Los Angeles and Riverside, San Bernardino, and Ontario, contribute disproportionately more Black, Hispanic, and first-generation college student prospects. Similar to Chicago and Dallas-Fort Worth, Los Angeles further solidifies the pattern of some of the lowest-income Geomarkets consistently playing crucial roles in fostering college access for first-generation Asian (East Los Angeles), Black (South and South Central Los Angeles), and Hispanic college students (South and South Central Los Angeles, East Los Angeles, Hollywood and Wilshire, and Riverside, San Berdardino, and Ontario).

3 Discussion

In their analysis of quantifying school quality in England, McArthur & Reeves (2022, p. 517) observe that "one problem with school league tables …is that the measures of school quality often merely reflect the social origins of those who attend a particular school." Similarly, considering prior research showing that SAT scores are substantially a function of social origin (Sewell & Shah, 1967), the Market Segment Model argues that student demand for higher education is mostly a function of social origin.

Institutional theory defines the organizational field 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 and products" (DiMaggio & Powell, 1983, p. 143).

4 References

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

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, *48*(2), 147–160.

Manson, S., Schroeder, J., Van Riper, D., Knowles, K., Kugler, T., Roberts, F., & Ruggles, S. (2024). *IPUMS national historical geographic information system: Version 19.0* (Report). IPUMS. https://doi.org/http://doi.org/10.18128/D050.V19.0

McArthur, D., & Reeves, A. (2022). The unintended consequences of quantifying quality: Does ranking school performance shape the geographical concentration of advantage?¹. *American Journal of Sociology*, *128*(2), 515–551. https://doi.org/10.1086/722470

Sewell, W. H., & Shah, V. P. (1967). Socioeconomic status, intelligence, and attainment of higher education. *Sociology of Education*, *40*(1), 1–23. Retrieved from <Go to ISI>://A1967ZD96500001

Walker, K., & Herman, M. (2024). *Tidycensus: Load US census boundary and attribute data as 'tidyverse' and 'sf'-ready data frames*. Retrieved from https://walker-data.com/tidycensus/

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

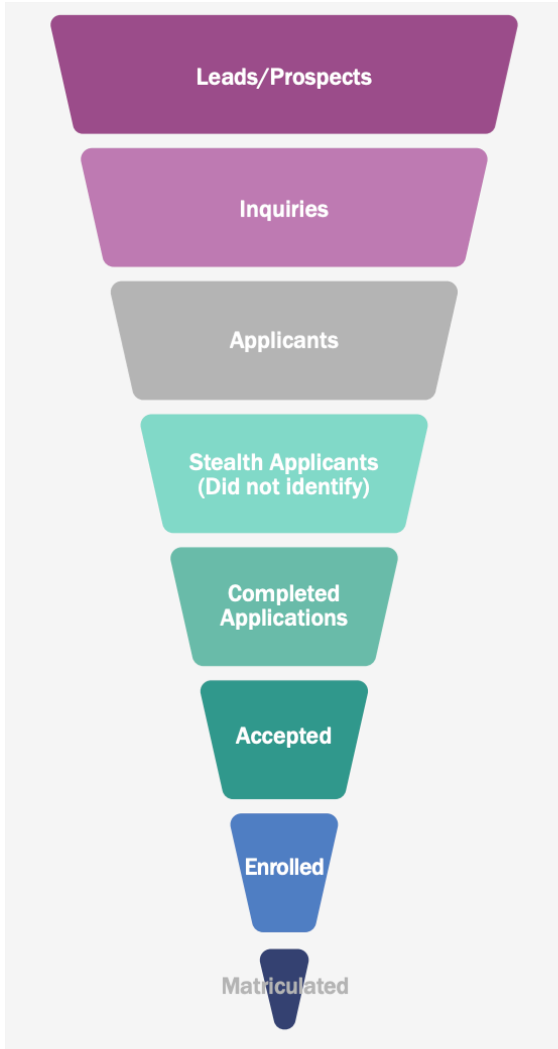

Figure 1: The Enrollment Funnel

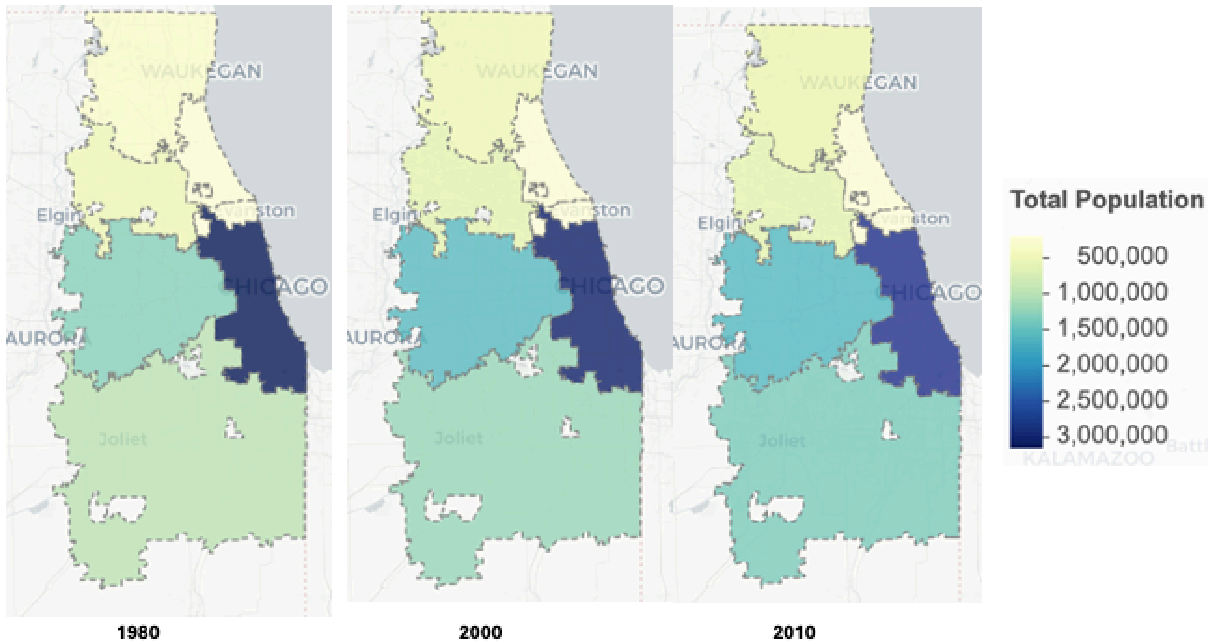

Figure 2: Chicago Geomarkets, Total Population 1980-2020

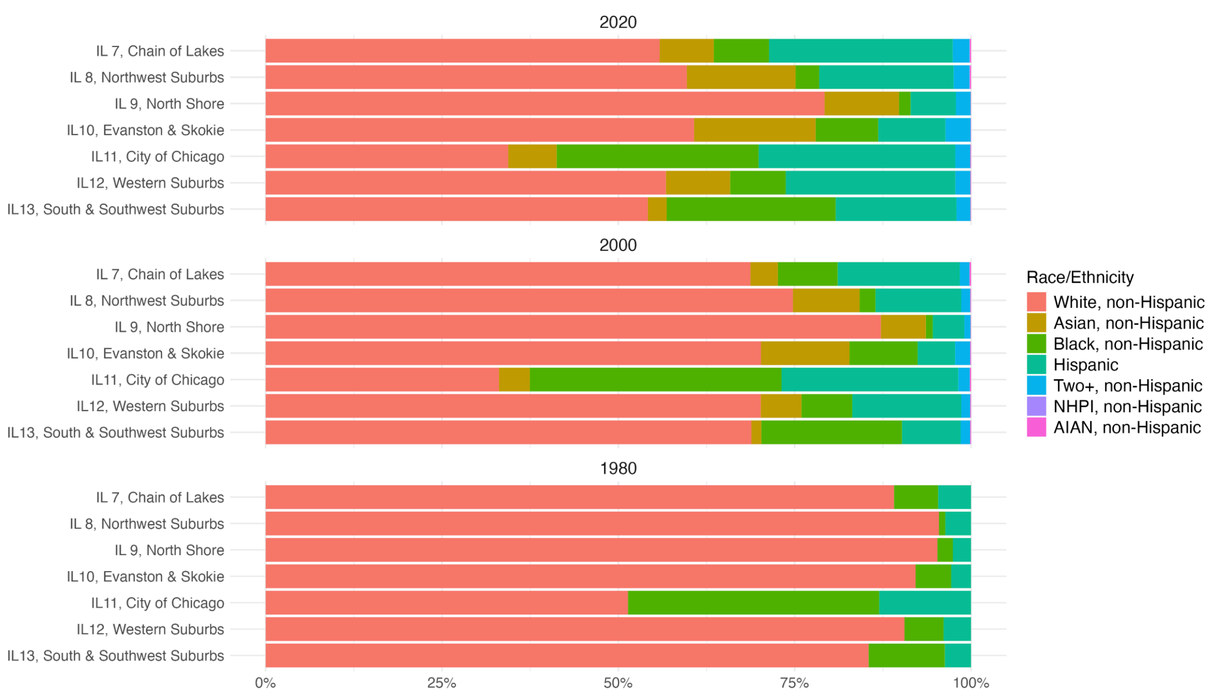

Figure 3: Racial/Ethnic Composition of Chicago Area Geomarkets, 1980-2020

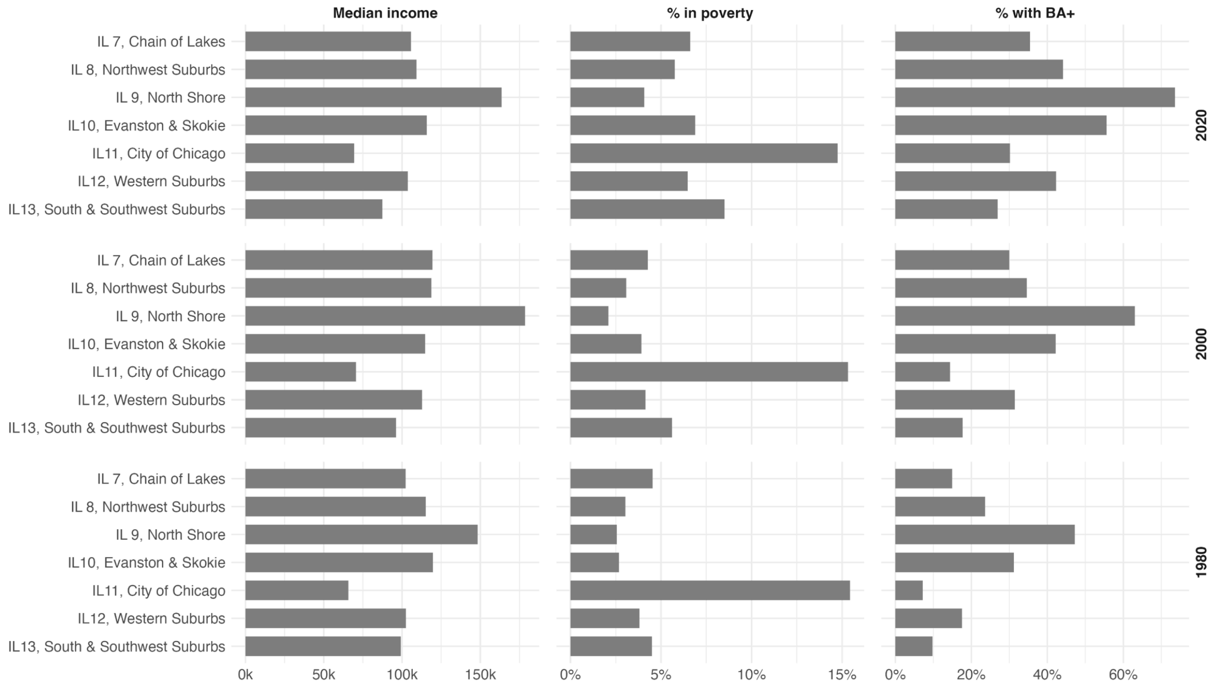

Figure 4: Socioeconomic Characteristics of Chicago Area Geomarkets, 1980-2020

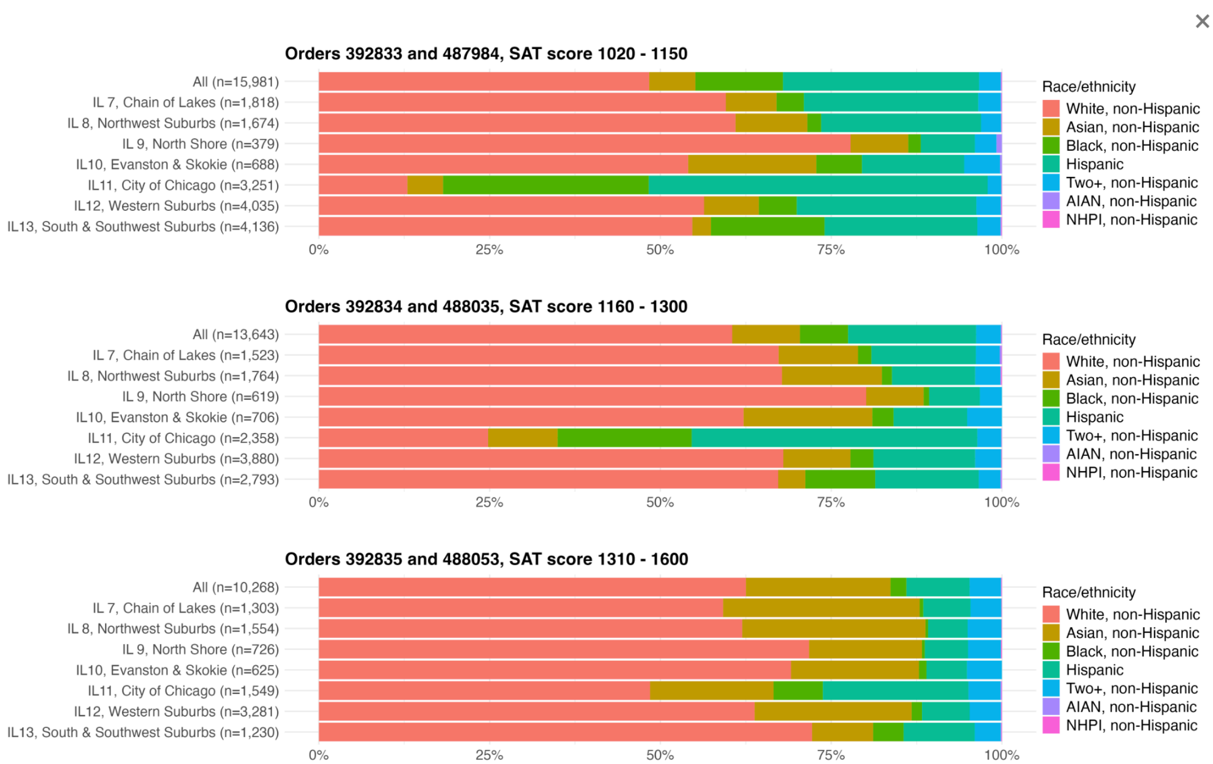

Figure 5: Racial/Ethnic Composition of Purchased Student Profiles by Geomarket, Chicago Area

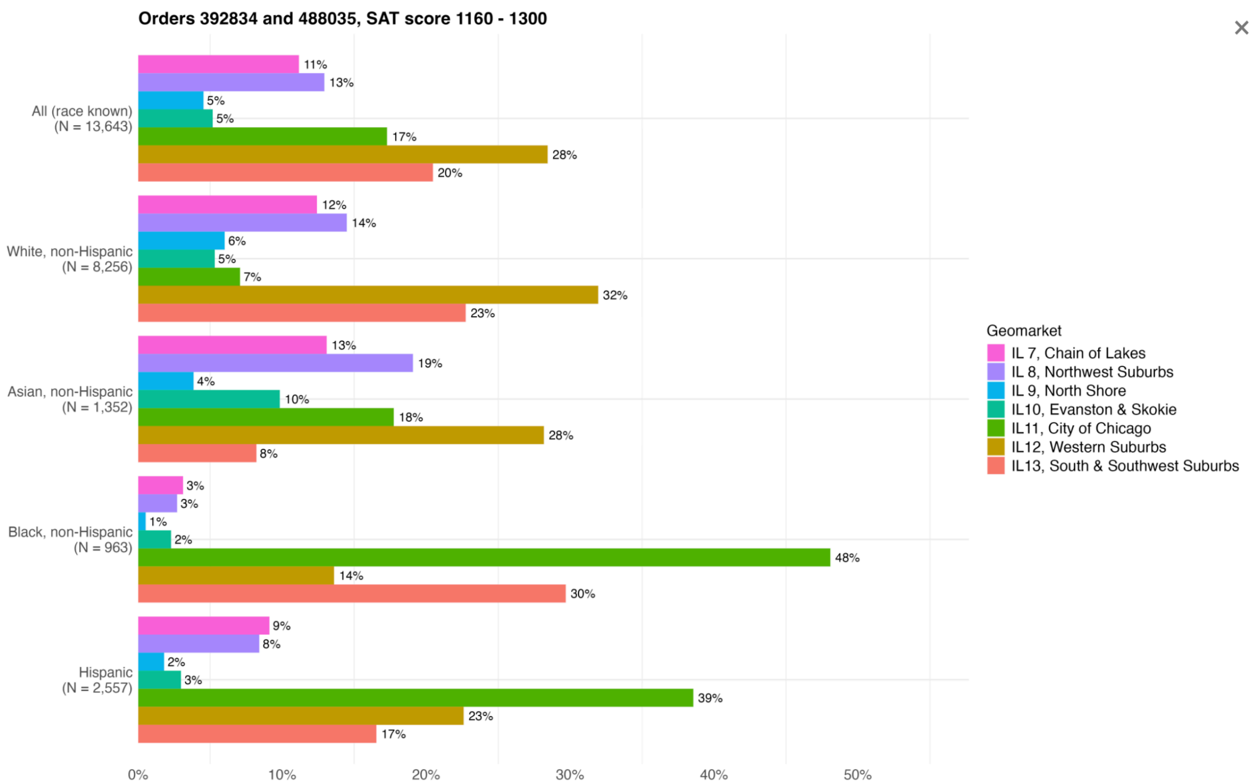

Figure 6: Chicago Geomarket Contribution to Purchased Student profiles by Racial/ethnic group, Middle-Range SAT orders

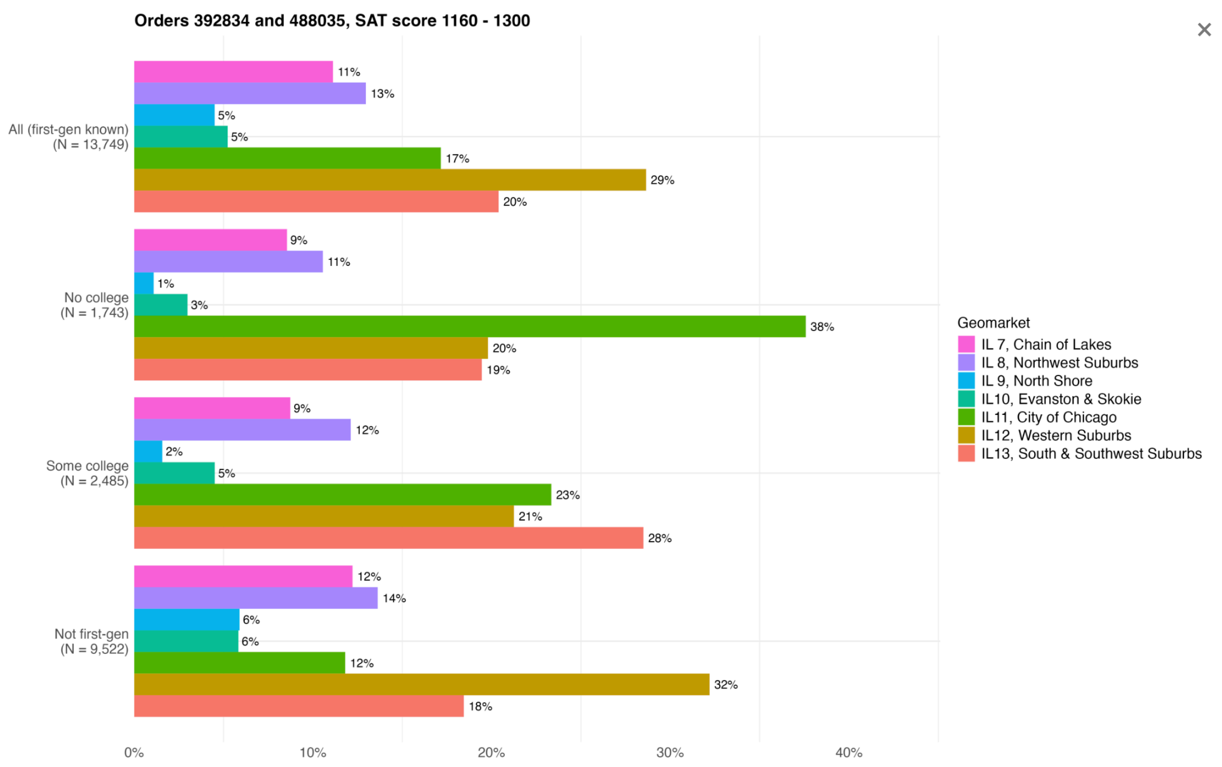

Figure 7: Chicago Geomarket Contribution to Purchased Student profiles by First-Generation Status, Middle-Range SAT orders

Figure 8: Chicago Geomarket Contribution to Purchased Student profiles by First-Generation Status, Middle-Range SAT orders

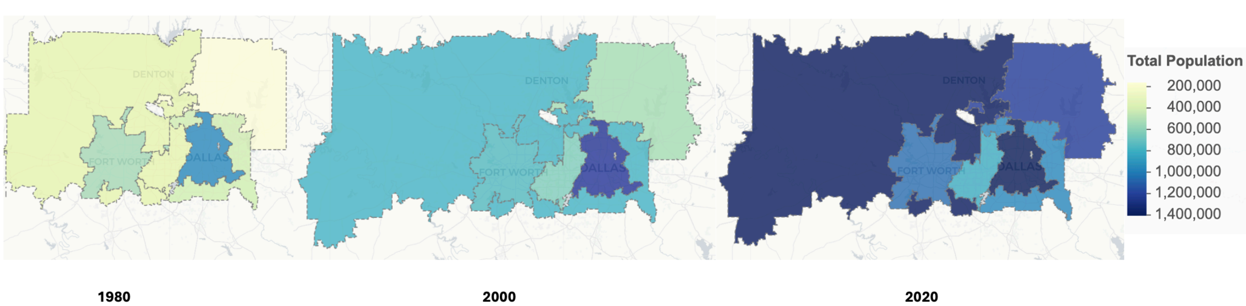

Figure 9: Dallas-Forth Worth Geomarkets, Total Population 1980-2020

Figure 10: Racial/Ethnic Composition of Dallas Area Geomarkets, 1980-2020

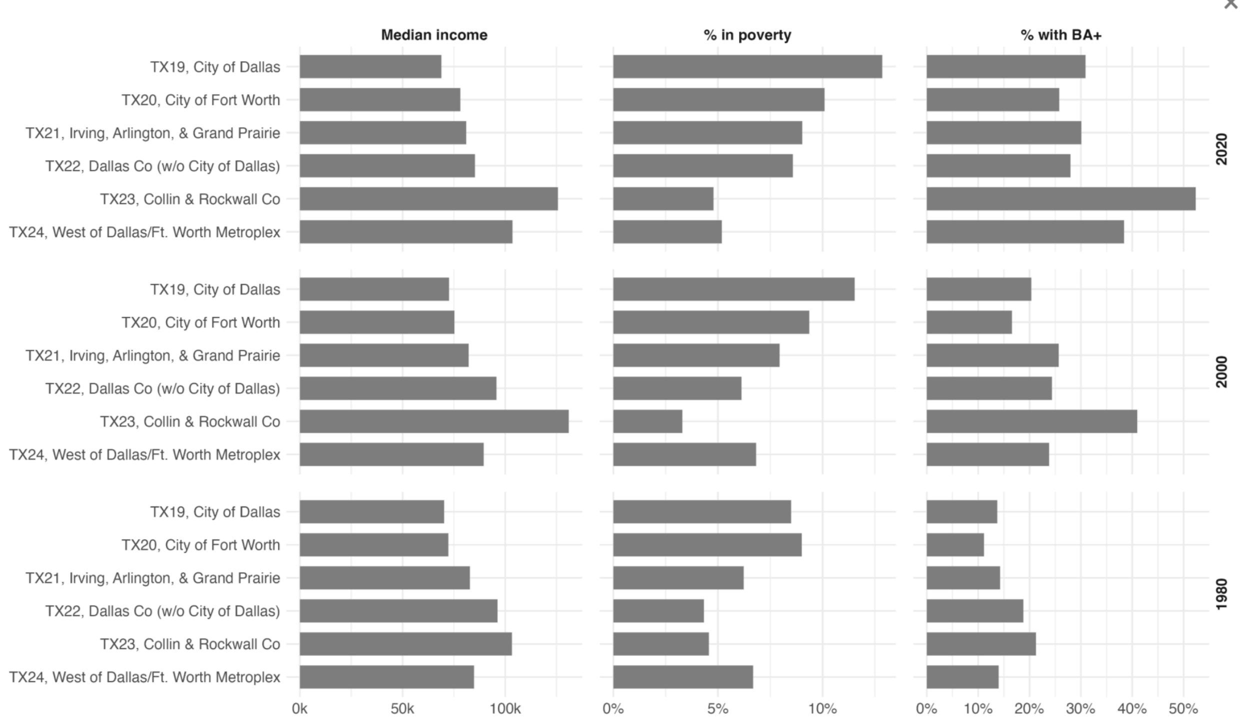

Figure 11: Socioeconomic Characteristics of Dallas Area Geomarkets, 1980-2020

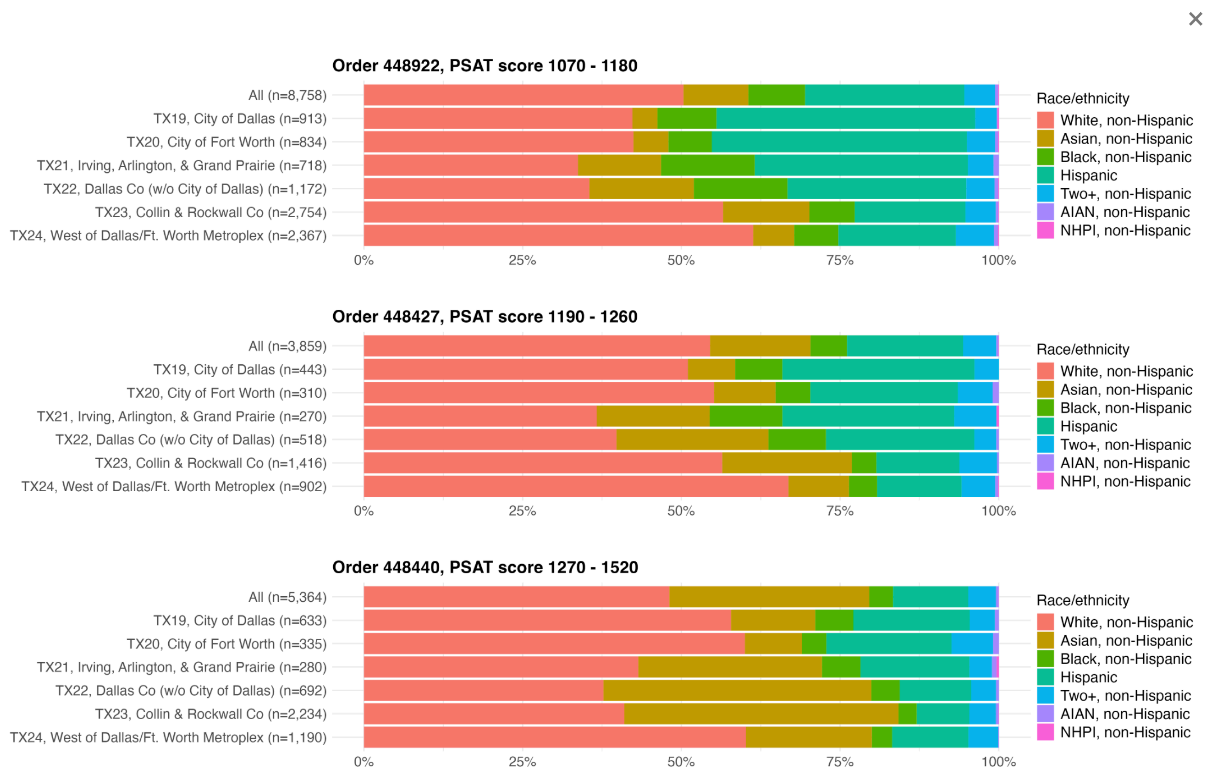

Figure 12: Racial/Ethnic Composition of Purchased Student Profiles by Geomarket, Dallas Area

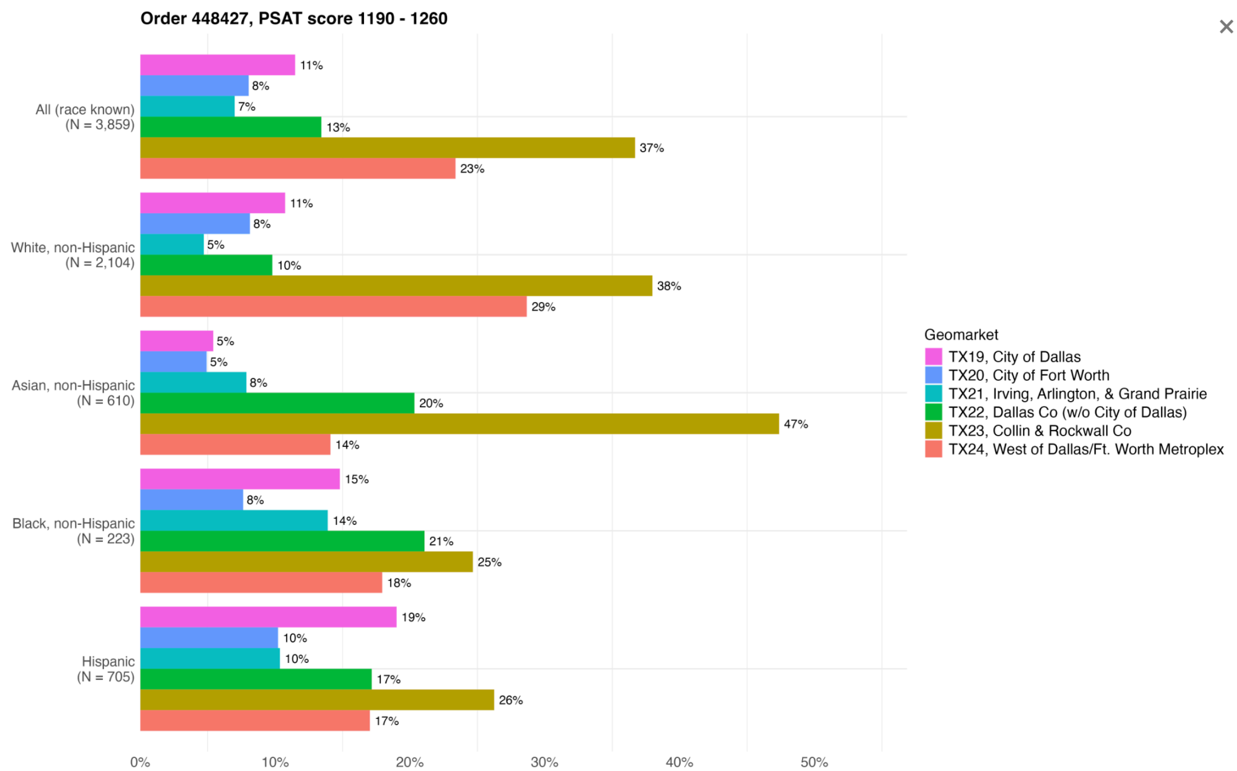

Figure 13: Dallas Geomarket Contribution to Purchased Student profiles by Racial/ethnic group, Middle-Range SAT orders

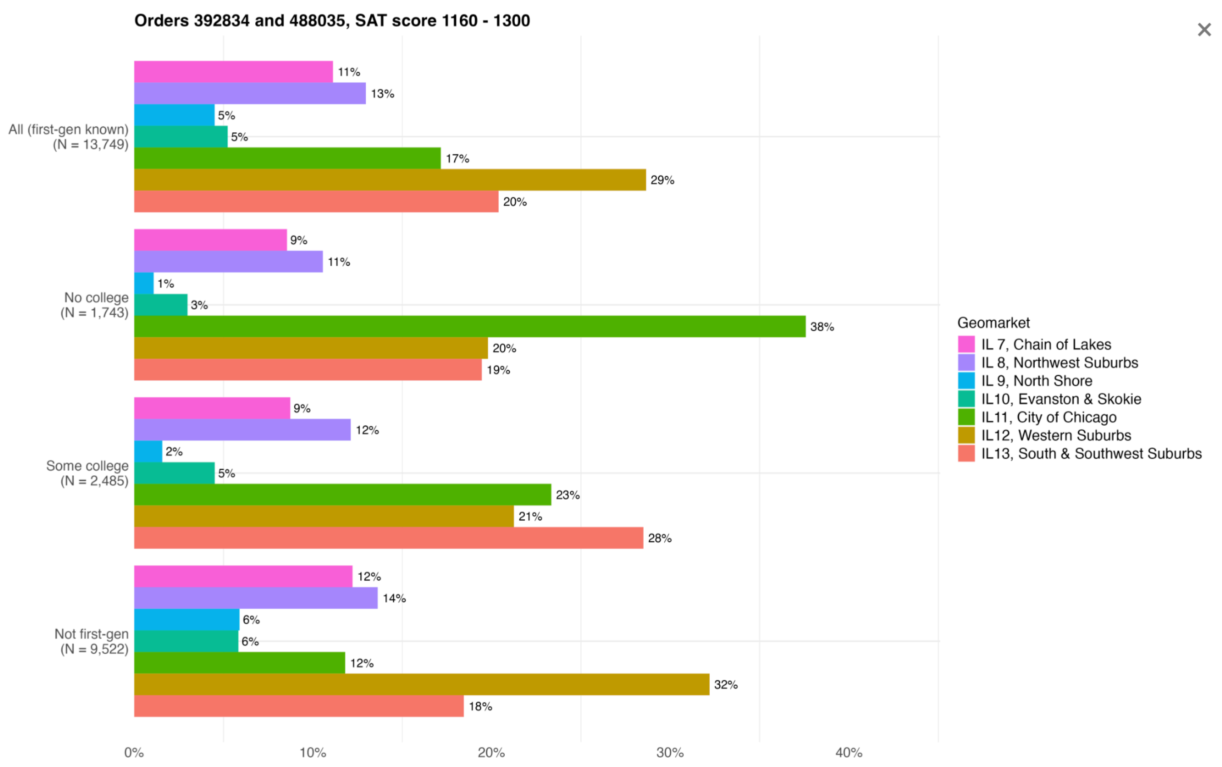

Figure 14: Dallas Geomarket Contribution to Purchased Student profiles by First-Generation Status, Middle-Range SAT orders

Figure 15: Dallas Geomarket Contribution to Purchased Student profiles by First-Generation Status, Middle-Range SAT orders

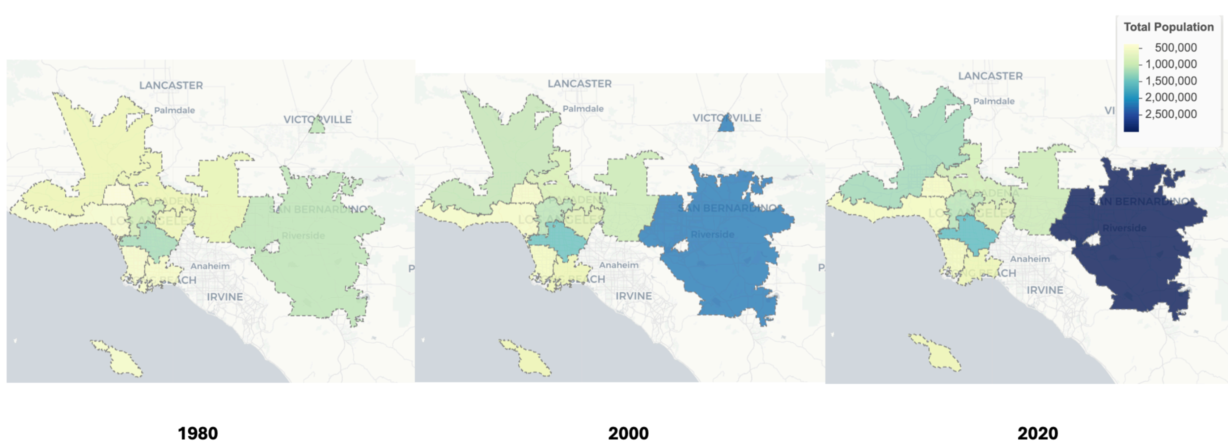

Figure 16: Los Angeles Worth Geomarkets, Total Population 1980-2020

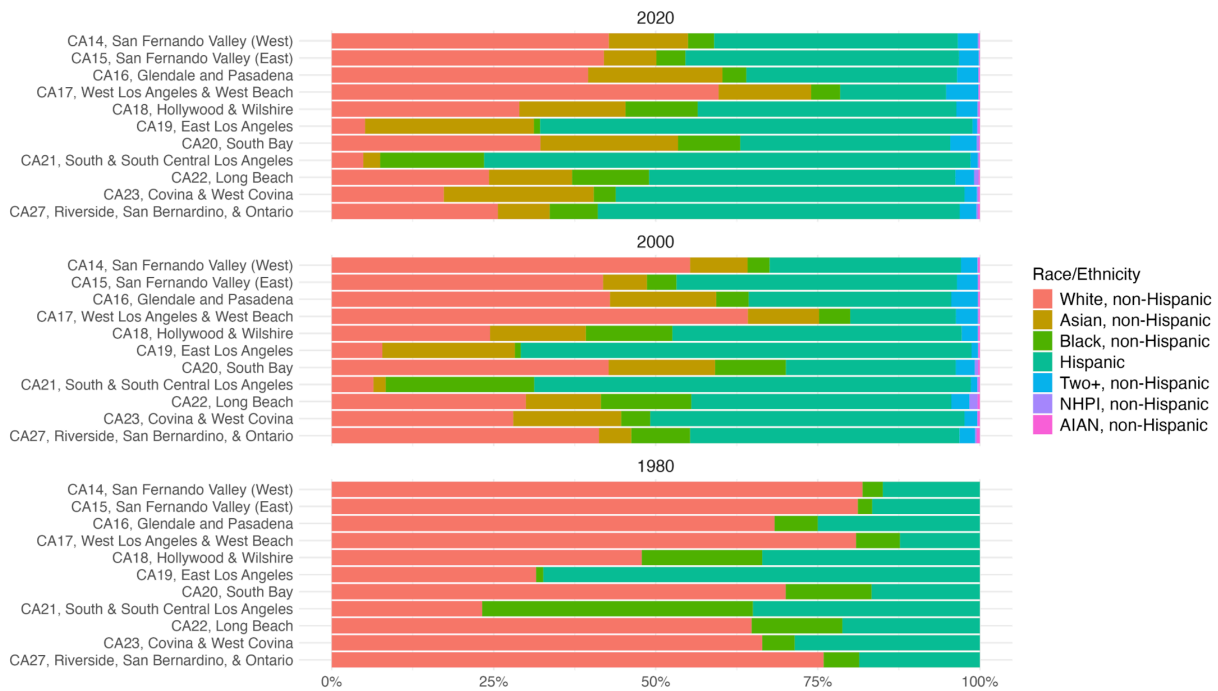

Figure 17: Racial/Ethnic Composition of Los Angeles Area Geomarkets, 1980-2020

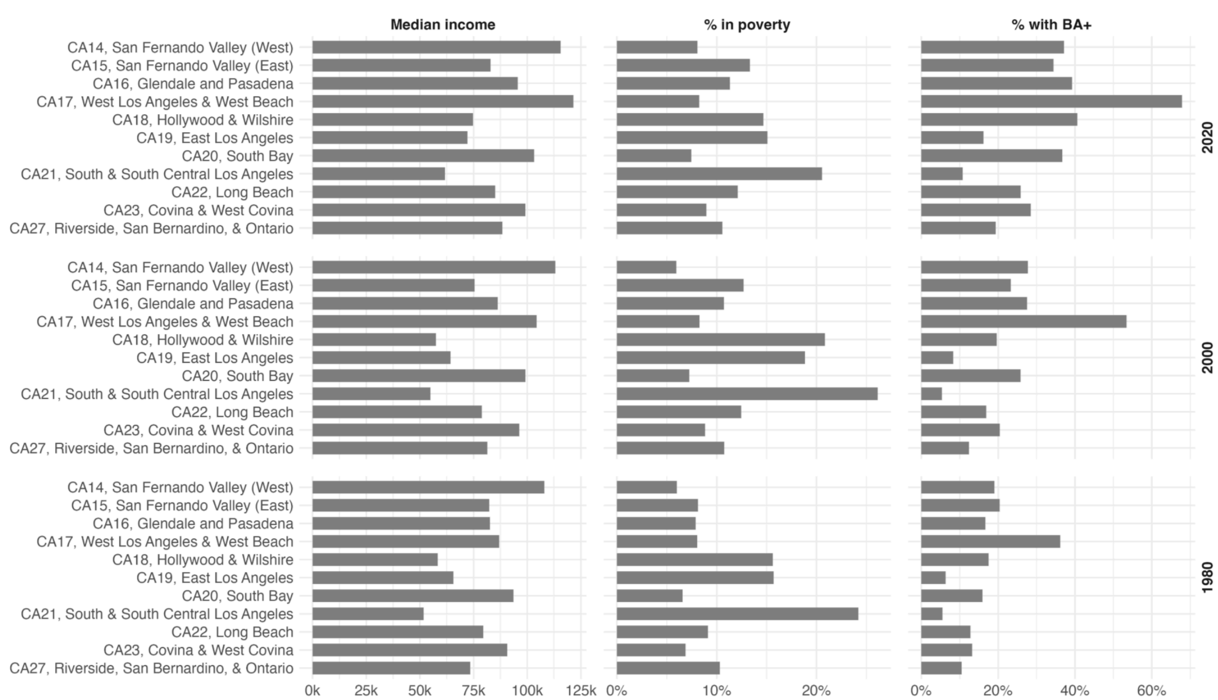

Figure 18: Socioeconomic Characteristics of Los Angeles Area Geomarkets, 1980-2020

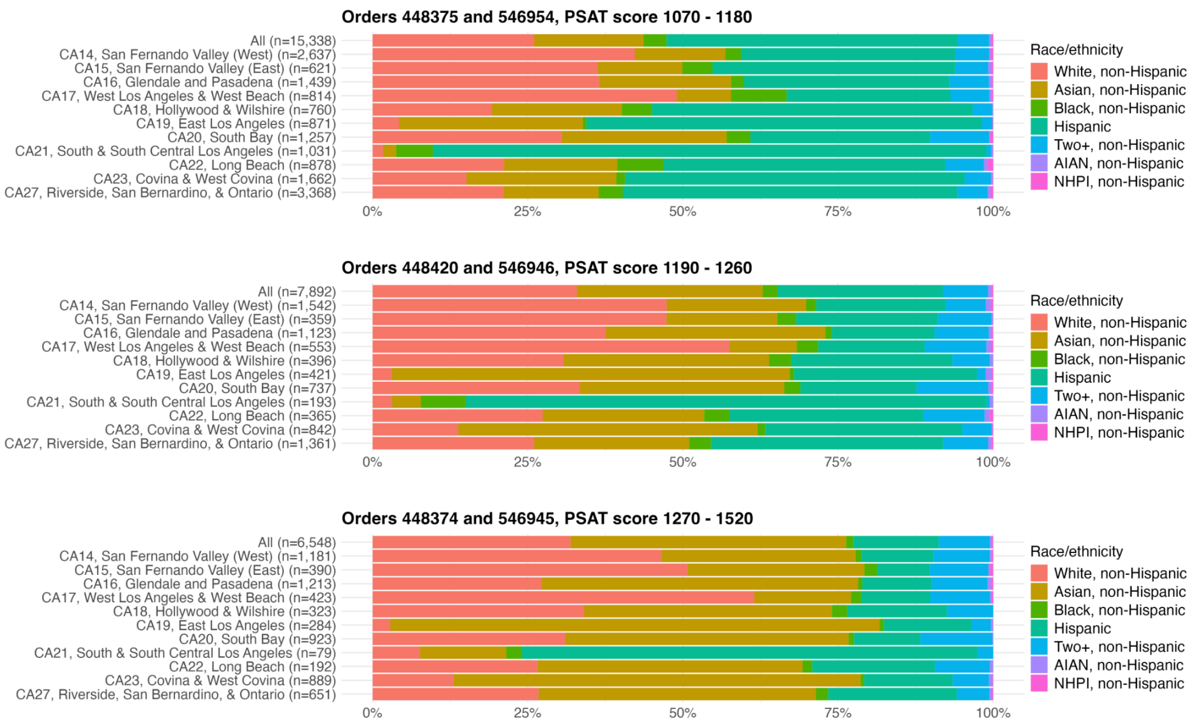

Figure 19: Racial/Ethnic Composition of Purchased Student Profiles by Geomarket, Los Angeles Area

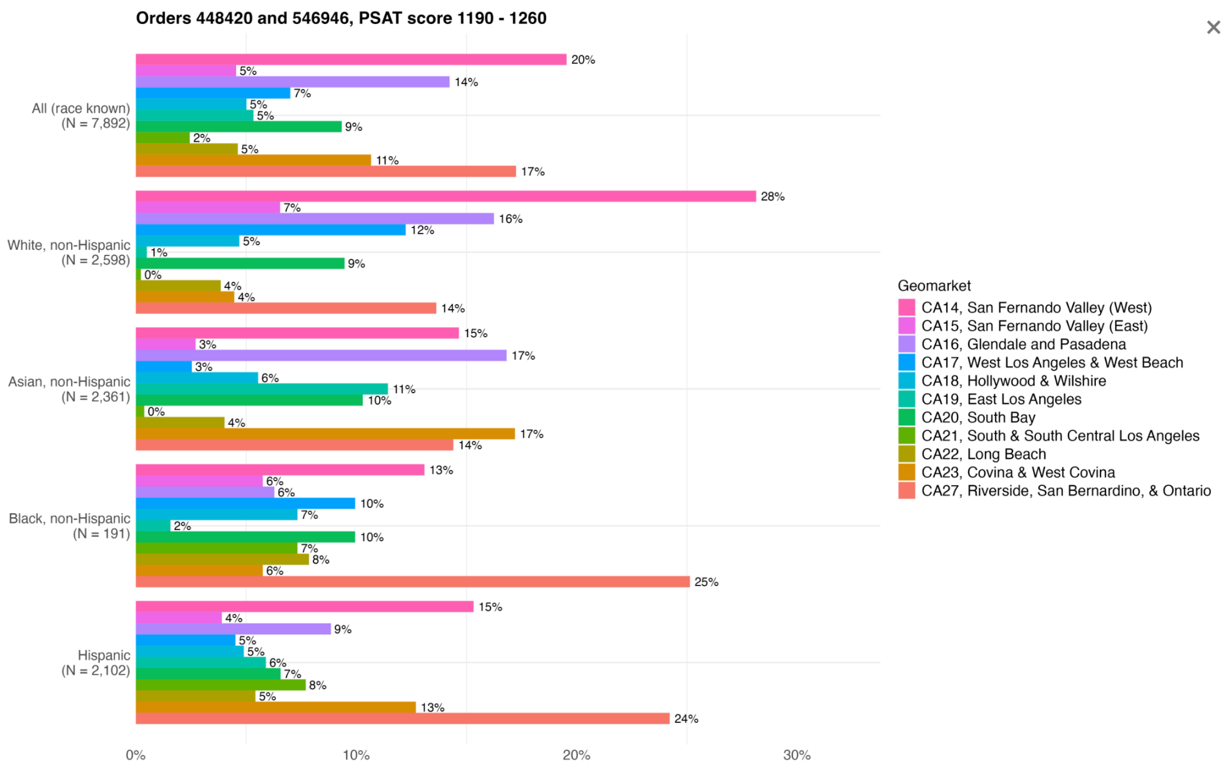

Figure 20: Los Angeles Geomarket Contribution to Purchased Student profiles by Racial/ethnic group, Middle-Range SAT orders

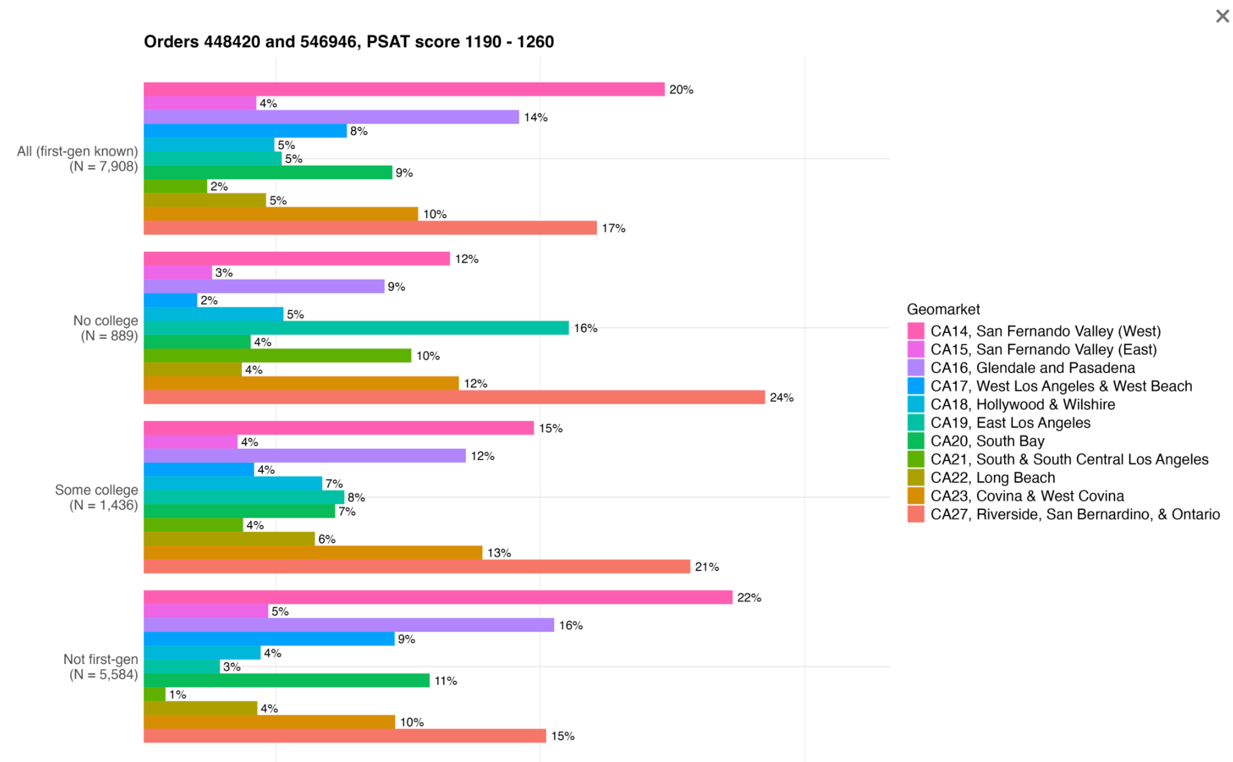

Figure 21: Los Angeles Geomarket Contribution to Purchased Student profiles by First-Generation Status, Middle-Range SAT orders

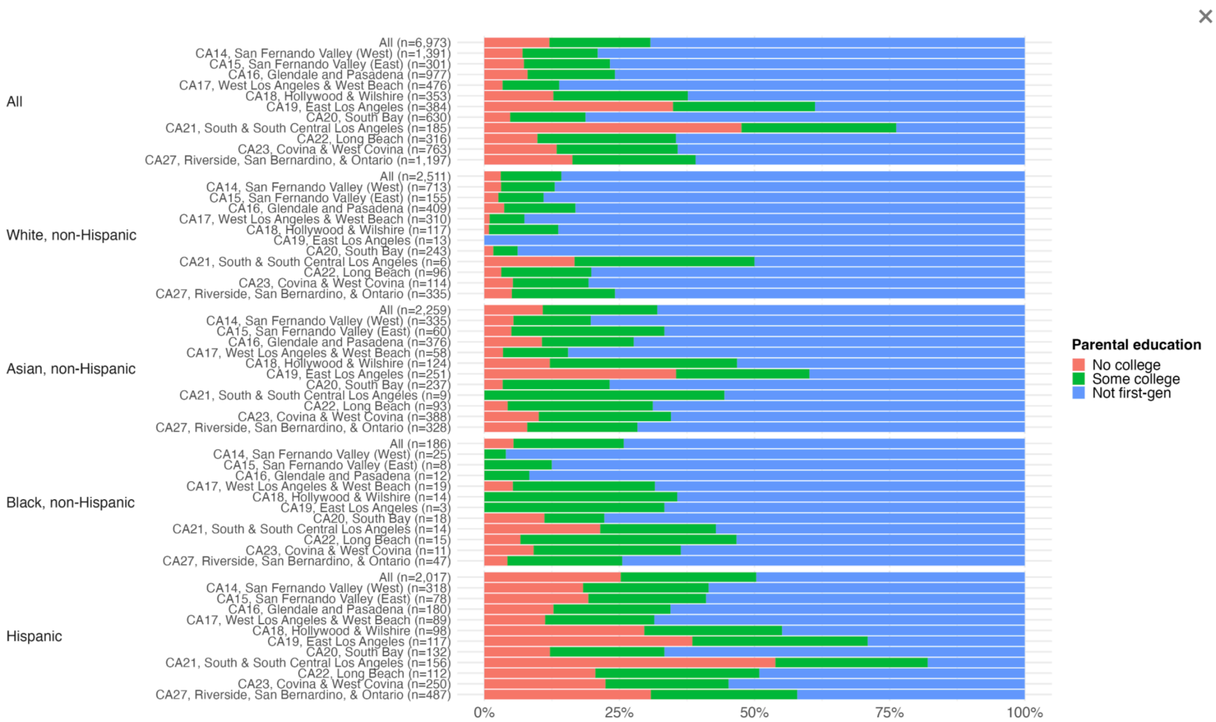

Figure 22: Los Angeles Geomarket Contribution to Purchased Student profiles by First-Generation Status, Middle-Range SAT orders

A Appendix A

Table A1 reproduces a partial, simplified version of the bottom panel of Zemsky & Oedel (1983, fig. 2.1) which represents the Market Segment Profile for "Connecticut Market 3: Fairfield County." Table A1 has separate columns for each market segment (local, in-state, regional, national) and rows show the number of test-takers and characteristics of test-takers. For example, there were 550 "local" students in Fairfield county and these students submitted SAT scores to 2.8 postsecondary institutions on average. By contrast, there were 1,664 "regional" students who submitted test scores to 4.8 institutions on average. For local students, 11.2% had family income greater than $35,000 and 9% had both parents with a BA. For regional students, 41.9% had family income greater than $35,000 and 34.0% had both parents with a BA. Each Market Segment Profile also present information about the institutions that students from each segment sent scores to.

Table A2 recreates the top panel of Zemsky & Oedel (1983, fig. 2.1) for the Fairfield County, CT local market. For example, of the 550 local students, 277 (50.4%) sent scores to institution #1, which was a private master's granting institution. Of the 1,199 in-state students, 757 (63.1) sent scores to institution #1, a public doctoral granting institution, and 515 (43.0%) sent scores to institution #2, a public master's granting institution.

The *Institutional Profile* describes students who send scores to a particular institution and which majors these students are interested in. For each institution, institutional profiles are created separately for students from a particular local market, for all students in a state, or all students in a region. Table A3 reproduces a partial, simplified version of Zemsky & Oedel (1983, fig. 2.3), the institutional profile of an anonymous institution for students from Fairfield County, CT. Table A3 shows that 58 in-state students submitted SAT scores to the institution. These 58 students represented 4.8% of the total 1,199 in-state students from Fairfield County. 69 regional students sent scores to the institution, representing 4.1% of all 1,664 regional test-takers. Of these 69 regional students, 35 expressed interest in majoring in the liberal arts. These 35 students represent 7.0% of all regional students from Fairfield County who expressed interest in the liberal arts.

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

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

Table A1: Simplified market segment profile, Connecticut Market 3: Fairfield County

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

Table A2: Top 5 institutions in terms of number of scores sent by segment, Connecticut Market 3: Fairfield County

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

Table A3: Simplified sample institutional profile for anonymous institution, students from Connecticut Market 3: Fairfield County

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