

# 1 Executive Summary

TEXT

## 2 Introduction

In the enrollment management industry, the “enrollment funnel” is a conceptual model that describes stages in the process of recruiting students. Depicted in Figure 1, the funnel begins with a large pool of “prospects” that the university would like to “convert” into customers. At each successive stage – inquiries, applicants, admits – the funnel narrows in order to convey the assumption of “melt” (e.g., a subset of “inquiries” will apply), ending with the cohort of enrolled students. Practically, the enrollment funnel informs interventions that increase the probability of “conversion” from one stage to another [CITE]. For example, emails and brochures encourage prospects and inquiries to apply. Financial aid packages – institutional grant aid and loan guidance – convert admits to enrolled students.

INSERT FIGURE 1 ABOUT HERE

At the top of the enrollment funnel, universities identify leads by buying “student lists,” which contain the contact information of prospective students. Sometimes referred to as “names,” student lists are the fundamental input for recruiting interventions that target individual prospects via mail, email, text, and on social media. The two dominant student list vendors are the College Board and ACT, which create student list products based on their database of standardized test takers. In fall 2021, the College Board Search and ACT Encoura student list products both charged \$0.50 per name CITE CB [CITE FOR ACT?]. These products enable universities to control which prospects they purchase through the use of search filters (e.g., test score range, high school GPA range, zip code).

**Student lists and student outcomes.** [CONDENSE TEXT] Research suggests that student lists substantially affect college access outcomes – and in turn degree completion

Figure 1: The enrollment funnel



outcomes – for millions of students each year (???; ???). College Board encourages test takers to opt into the College Board Student Search Service, which enables “accredited colleges, universities, nonprofit scholarship programs, and nonprofit educational organizations” [CITE](#) to “license” their contact information. (???) analyzed the college access and degree completion outcomes of SAT test-takers who graduated from high school between 2015-2018, comparing students who opted into Student Search Service to students who opted out. After controlling for covariates – including gender/sex, race/ethnicity, parental education, SAT score, and high school – 58.0% of students who participated in Search attended any college compared to 50.2% of students who opted out of Search, representing a 15.5% relative increase in the probability of college enrollment ( $((58-50.2)/50.2=15.5)$ ).

Figure X – reproduced from (???) – presents results for four-year college enrollment and degree completion. [CRYSTAL - RECREATE FIGURES 2 AND 3 FROM THIS ARTICLE; ADD FIGURE NOTES W/ LIST OF COVARIATES AND RELEVANT INFO] 41.1% of students who participated in Search attended a 4-year college compared to 32.8% of students

who opted out, representing a 25.3% relative increase in the probability of attending a 4-year college. Furthermore, change in the relative probability of attending a four-year college associated with opting in to Search out was higher for students who identified as Black (24.5%), Hispanic (34.4%), American Indian or Alaska Native (AI/AN) (23.8%), and Native Hawaiian or Pacific Islander (26.1%) than it was for students who identified as White (21.6%) or Asian (15.2%). Similarly, change in the relative probability of attending a four-year college was higher for students whose parents did not attend college (40.6%) than it was for students whose parents had a BA (18.9%).

With respect to BA degree completion, the bottom panel of Figure X shows that 20.6% of students who participated in Search obtained a BA compared within four years to 15.7% of students who opted out, representing a relative increase of 31.2%  $((20.6-15.7)/15.7=31.2)$ . Additionally, the relative increase in the probability of obtaining a BA within four years was higher for Black, Hispanic, and AI/AN students than it was for White and Asian students and higher for first-generation students than for students whose parents had a BA.

A similar study compared the college access outcomes of students who opted into versus opted out of ACT's Educational Opportunity Service (EOS) (???). EOS enables accredited postsecondary institutions and scholarship organizations to license the contact information of students who opt in. After controlling for covariates – including ACT score, number of colleges the student sent scores to, family income, parental education, degree aspirations, race/ethnicity, state – (???) found that students who opted into EOS had a 3.7 times higher likelihood (odds ratio = 3.7) of attending college than students who opted out. A second model, which restricted the sample to students who attended college, found that students who opted into EOS had 8.7% higher odds of attending a four-year college rather than a two-year college.

**Policy concerns.** Considering these findings, which suggest that student lists profoundly affect college access for groups who are underrepresented in higher education, the policy

concern is which students are being excluded from student lists. We argue that this policy concern is primarily about student list products rather than how customers (universities) use these products.

Student list products exclude students in two ways. First, universities cannot purchase the contact information of prospects who are not included in the underlying database. While students who opt out of College Board Student Search Service are making a conscious decision, students cannot opt in unless they take a College Board Assessment (PSAT, SAT, AP). College Board and ACT assessments have been criticized for racial and socioeconomic bias [CITE]. Test-taking rates differ substantially across race and class [CITE], leading to systematic racial and socioeconomic inequality in which prospective students are included in the underlying databases that College Board and ACT student list products pull from. About 1.5 million students from the high school class of 2021 took the SAT compared to about 2.2 million students from the high school class of 2020 [CITE](#) and about 1.3 million student from the high school class of 2021 took the ACT compared to about 1.7 million students from the high school class of 2020. These declines were driven by the Covid Pandemic and by the growth in test-optional and test-blind admissions policies. To the extent that student lists are an important mechanism for college access, the test-optional movement may have the unintended consequence of creating a college access crisis, in which the long-term decline in test-takers leads to fewer prospective students included in College Board/ACT student list databases which causes college access to decline.

Second, student list products exclude students by creating filters that enable universities to purchase some names but not others. While universities choose filters based on their preferences, these choices are structured by what the product allows. We are concerned that several commonly used filters systematically exclude protected classes and other populations that are underrepresented in higher education (e.g., rural students). Second, the filters on student list products enable universities to target prospects based on criteria that may

systematically exclude populations that are underrepresented in higher education. For example, College Board and ACT student list products enable universities to filter prospects by zip code, which is highly correlated with race. College Board allows universities to target prospects based on their score in some set of AP exams, but which students attend high schools with widespread access to AP classes? An emerging trend is the creation of “geodemographic” filters that enable universities to select prospects based on the past behavior of students from their high school or neighborhood (e.g., how many students from this school attended an out-of-state university?) [CITE]. Student list products enable universities to explicitly search for underrepresented groups (e.g., by race/ethnicity), but when used in conjunction with additional filters (e.g., test score range, high school characteristics) these searches tend to target more privileged members of the group.

Although research sanctioned by College Board and ACT highlights the positive relationship between opting into student list products and college access (???, ???, ???), prior research has not examined which filter criteria universities select when purchasing student lists, what are the characteristics of purchased prospects, or the relationship between filter criteria and student characteristics. we collected data about student list purchases by issuing public records requests to public universities in five states. This report focuses on student lists purchased from College Board and addresses three research questions:

1. Which filter criteria were selected in student lists purchased by universities in our sample?
2. What are the characteristics of prospects included in student lists purchased by universities in our sample?
3. What is the relationship between student list filter criteria and the characteristics of purchased prospects?

As a consequence of data collection challenges, our analysis sample cannot be considered a random sample. Nevertheless, as the first independent analysis of student list data, this

research is an important first step towards developing a body of empirical research that informs policy discourse about regulating the student list business.

### 3 Data Collection and Research Design

This section describes, first, the data collection for the broader project. Second, we describe the research methods for analyses of student list purchases from College Board, which are the focus of this report.

#### 3.1 Data Collection

**Data collection sample.** In 20XX we received funding from the Joyce Foundation and the Kresge Foundation for a research project that would utilize public records requests to collect data about recruiting behavior from all public universities in four states, California, Illinois, Minnesota, and Texas. Broadly, we sought two types of data: off-campus recruiting visits (e.g., visit from a university admissions representative to a local high school) and student list purchases.

IL and MN are focus states for the Joyce Foundation, while CA and TX are focus states for Kresge. The IL higher education system includes 3 universities in the University of Illinois system, 7 in the Illinois State University system, and 2 in the Southern Illinois University system. In MN, there are 5 universities in the University of Minnesota system and 7 in the Minnesota State University system. In CA, there are 9 universities in the University of California system and 23 in the California State University system. In TX, there are 8 universities in the University of Texas system, 4 in the Texas State University system, 11 in the Texas A&M University system, 4 in the University of Houston system, 2 in the University of North Texas system, 2 in the Texas Tech University system, and 4 independent Texas universities. We also collected data from Arizona State University and Northern Arizona University.

TABLE/FIGURE(S) describe the public universities in our data collection sample. A majority of the universities are master's or doctoral universities and located in urban areas. FIGURES/TABLES TO CREATE:

- FIGURE
  - SUB-FIGURE: NUMBER OF UNIVERSITIES BY STATE AND CARNEGIE CLASSIFICATION
  - SUB-FIGURE: NUMBER OF UNIVERSITIES BY STATE AND LEVEL OF URBANIZATION (e.g., urban, suburban, rural - you decide on the variable)
- APPENDIX TABLE
  - Same as table 2 from private school chapter, with these differences:
    - \* add column for system name and carnegie classification; delete column for USNWR rank
    - \* sort by: state, system, carnegie classification, university name

**Public records requests.** In 20XX, we signed an agreement with the Lawyers' Committee for Civil Rights Under Law (LCCR) to partner on the data collection. Although We had experience issuing public records requests from a previous data collection about off-campus recruiting visits by public research universities, the greater scale, complexity, and sensitivity of this project motivated us to find a partner with greater legal expertise. LCCR was created by the Kennedy Administration to leverage private sector legal expertise towards civil rights. LCCR typically operates by connecting projects with the pro bono efforts of corporate law firms.

LCCR recruited firms to work on our project. Several firms expressed initial interest. Unfortunately, these firms withdrew following conflict of interest checks. However, LCCR generously allocated time of an internal staff attorney towards our project. Additionally, one firm offered to help us collect data from Arizona public universities. Although Arizona was not part of the funded project, we gratefully accepted.

We began issuing public records requests in February 2020, following several months of planning and pilot requests. We issued one records request letter for each public university. An example records request letter can be found [HERE](#) [ADD LINK]. Each request letter asked for data about off-campus recruiting visits and student list purchases for the purpose of undergraduate recruiting, which were made from August 2016 to the present. For each student list purchase from a student list vendor, we requested two related pieces of data: the order summary, which specifies search criteria for the student list purchase; and the de-identified prospect-level list produced from the search criteria.

As such, we conceived of our data collection as requesting three types of information: (1) off-campus recruiting visits; (2) student list order summaries; and (3) de-identified student list data. Each request letter included examples of desired off-campus recruiting visit data [LINK], student list order summary data [LINK], and de-identified student list data [LINK] as attachments.

**Data collection challenges and successes.** [ALL/MOST OF THIS SUB-SECTION WILL GO IN AN APPENDIX IN TICAS/ACCEPT REPORT; MAYBE KEEP IN MAIN TEXT FOR REPORT TO FOUNDATIONS] TABLE(S)/FIGURE(S) summarizes the success to date of our data collection efforts and Appendix X reports the status of data collection for each university. We received usable quantitative off-campus recruiting visit data from X universities, usable student list order summary data from X universities, and usable de-identified student list data from X universities. [CRYSTAL - CREATE TABLES AND FILL IN PARAGRAPH TEXT]

Using public records requests to collect quantifiable data is difficult under the best of circumstances. Unfortunately, we started collecting data just as Covid-19 emerged. Several additional challenges – some foreseeable and tractable and others not – made data collection difficult. State public records request laws generally require public entities to redact records that contain sensitive personal information but do not require public entities to create new



records. Consider a prospect-level student list stored by a public university in a spreadsheet format. Removing personally identifiable information (e.g., fields for name, mailing address, email address) is part of the redaction process. By contrast, if information about off-campus recruiting visits stored in old emails or an antiquated calendar system, compiling these records would be considered creating new records rather than redaction. This turned out to be the most common reason we did not obtain off-campus recruiting visit data.

For student list purchases, an additional complication is that universities may purchase lists from multiple vendors. When we initiated data collection, the three largest vendors were College Board, ACT, and the National Research Center for College and University Admissions (NRCCUA), which had just been purchased by NRCCUA. In subsequent communication with universities, we narrowed our student list purchase requests to these three vendors.

Many public universities outsourced student list purchases to an enrollment management consulting firm, which created several challenges. Often, employees at these universities often lacked knowledge about student list purchases and were unfamiliar with the records we requested. Employee turnover often exacerbated this lack of institutional knowledge. Second, these universities often did not have the requested order summaries and student lists in their possession. Third, enrollment management consulting firms posed another barrier. The information firms regularly sent to universities was too aggregate and firms were unwilling to help provide the original records we sought.

In response to early data collection challenges, we created a “vendor portal document” [LINK]. This document provided detailed instructions – separately for College Board, ACT, and NRCCUA – for how to log-in to the online portal, obtain the order summary for each student list purchase, obtain student list data for each purchase, and how to de-identify these data. Unfortunately, this document was less effective for ACT and NRCCUA than it was for College Board. ACT released the Encoura platform in 20XX, following its 2018 acquisition of NRCCUA. The online portals that provided access to students lists purchased from the

legacy ACT and NRCCUA student list products were continued. Furthermore, whereas the College Board portal gave customers access to student lists purchased in the last X [CRYSTAL] years, the Encoura portal provided access to purchases within the previous 12 months.

The most common data collection challenges were universities not replying to the request or denying the request based on questionable legal grounds. We learned that obtaining these data depends on treating each university as a protracted negotiation – often several negotiations – that requires sustained effort and some degree of leverage. Even with time from a LCCR staff attorney, we lacked sufficient capacity for sustained negotiations with each university and we lacked a stick that commanded university attention.

In Spring 2021, LCCR successfully recruited three law firms to represent data collection efforts in Illinois, Minnesota, and California, respectively. However, we were unable to obtain representation for Texas. The firms produced legal research that demonstrated the legitimacy of our requests and dismantled common rationales for denial (e.g., “trade secrets”). Next, the firms systematically engaged universities, often directly engaging the university or system-level general counsel, who then directed the public records office and other offices to cooperate with our request. Firm representation substantially increased the number of successful requests. However, many universities provided legitimate reasons for not providing one or more requested data elements (e.g., records no longer exist, not required to create new records). Some universities denied request elements on grounds that firms believed lacked legal merit. However, we made a collective decision not pursue litigation.

In hindsight, we identify several changes may benefit efforts to collect public records in the enrollment management space. First, many universities were understandably overwhelmed by a request for several complicated, esoteric data sources. We could have issued several narrower requests rather than a single multi-aceted request. Related, our request for student list data should have proceeded in two stages: first, requesting contracts with a specific set of student

list vendors over a time period; and, second, for each contract received, issuing a separate records request for the student list order summaries de-identified student lists associated with that contract. Third, underestimated the extent to which university personnel were unfamiliar with student list purchases and also the number of universities that outsourced student list purchases to a consulting firm. We should have created the “vendor portal” instructions document at the beginning of the data collection process.

**Data processing.** Records received from universities were visually inspected to check whether they contained the requested fields and data structure. If not, we communicated the problems to the university and asked for revised records. Records that passed visual inspection were processed. For the current analysis, only College Board data were processed as they are the most common and easily obtained. The order summaries were generally provided as PDF files and were parsed using Python, a general-purpose programming language. The student-level lists were imported and cleaned using the statistical programming language R, and all subsequent analysis of the order summaries and lists were done in R as well.

## 3.2 Research Design

**Analysis sample.** Given that prior research has not analyzed student list purchases, this report restricts analyses to student lists purchased from College Board in order to keep results as simple as possible, with the idea that future analyses will incorporate lists purchased from ACT. Table/Figure X shows the number of public universities in our data collection sample that provided usable data about (1) student list order summaries and (2) de-identified student lists purchased from College Board. [CRYSTAL ADD TEXT SUMMARIZING TABLE]. Our analysis sample consists of X universities that provided usable data about student lists purchased from College Board, including X universities that provided both order summary and student list data, X universities that only provided order summary data, and X universities that only provided student list data[SPECIFICATIONS FOR THIS FIGURE. SHOULD SHOW:]

- WHAT TO SHOW
  - NUMBER OF UNIVERSITIES THAT DID AND DID NOT PROVIDE USABLE ORDER SUMMARY DATA
  - NUMBER OF UNIVERSITIES THAT DID AND DID NOT PROVIDE USABLE STUDENT LIST DATA
  - NUMBER OF UNIVERSITIES THAT DID AND DID NOT PROVIDE **BOTH** USABLE ORDER SUMMARY DATA AND USABLE STUDENT LIST DATA
- ORGANIZED BY [MAYBE A FIGURE WITH THREE SEPARATE PANELS? USE YOUR DISCRETION:
  - STATE
  - STATE AND CARNEGIE CLASSIFICATION
  - STATE AND LEVEL OF URBANIZATION

Table/Figure X shows the number of lists and the number of prospects purchased by universities in our analysis sample [CRYSTAL ADD TEXT] [SPECIFICATIONS FOR FIGURE (USE YOUR DISCRETION!)]

- WHAT DATA TO SHOW
  - NUMBER OF ORDERS (REGARDLESS OF WHETHER WE HAVE STUDENT LIST DATA)
  - NUMBER OF ORDERS WHERE WE HAVE STUDENT LIST DATA
  - NUMBER OF PROSPECTS PURCHASED (REGARDLESS OF WHETHER WE HAVE THE ORDER SUMMARY)
  - NUMBER OF PROSPECTS PURCHASED WHERE WE HAVE THE ORDER SUMMARY
- SHOW DATA BY:
  - STATE AND CARNEGIE CLASSIFICATION
- SUB-FIGURES?

– USE YOUR DISCRETION

To what extent are the student list purchases we analyze representative of student lists purchased by universities in our analysis sample? Universities may purchase student lists from several different vendors but College Board and ACT dominate the market. In CA and TX, the majority of test-takers take the SAT [CITE]. Historically, most IL test-takers took the ACT, but in 2016 the state signed a contract with College Board for all IL juniors to take the SAT CITE. In MN, the vast majority of test-takers take the ACT. These state-by-state differences suggest that students lists purchased from College Board will not be representative of all student lists purchases, particularly for MN public universities. Another issue is that some universities provide incomplete data about student lists purchased from College Board during the requested time period. When we received the order summary for a particular purchase but not the de-identified student list – or vice-versa – we have direct evidence of incomplete data. However, we usually cannot assess whether the data we received identifies the full set of student lists purchased from College Board.

Our inability to obtain data from all universities creates external validity concerns. Ideally, our analysis dataset includes all College Board student lists for the entire population of public universities in our data collection sample, or at least a random sample of these universities. Unfortunately, our analysis sample cannot be considered a random sample. Moreover, we suspect that non-response was systematically related to factors of substantive interest. For example, we were less successful obtaining records from universities that outsourced student list purchases to consulting firms. In turn, response bias affects external validity. Based on our results, we cannot make inferences about the population of universities in the four data collection states. Nor can we make inferences about the broader population of US public universities.

**Research questions and analyses.** Choices about research questions were informed by the limitations of our analysis sample and by substantive considerations. We cannot make

statements about university behavior that extends beyond our analysis sample. Assuming our data is representative of lists purchased by universities in our sample, we can make statements about the College Board student lists purchased by universities in our analysis sample. This reasoning suggests treating our sample as a multiple case study (Eisenhardt, 1989). The behaviors observed in our sample identify behaviors that exist in the population of public public universities, but not the prevalence of these behaviors in the population.

More substantively, during the course of data collection we realized that research should focus on the student list products themselves rather than the behavior of customers (universities) who buy the product. while universities choose which names to purchase, these choices are structured by what the product allows. Systematic inequality in purchased versus excluded names is a function of (A) which prospective students are included in the underlying data base and (B) the set of filters that universities can utilize to select prospects and finally (C) university choices about which filters to select. Prospective students who do not take College Board assessments are excluded from College Board lists. College Board filters encourage customers to purchase prospects based on their score ranges in SAT, PSAT, and AP assessments, but which students attends high schools with widespread access to AP classes? Geographic filters additionally enable to customers to filter prospects based zip code. which is highly correlated by race. More recently, College Board has created “geodemographic” filters that target prospects based on the recent college-going behaviors of nearby peers. These considerations suggest analyses that investigate the relationship between the filters chosen for a particular student list purchase and who is included in the resulting student list. The empirical analyses presented in this report are guided by three research questions, which focus on student lists purchased from College Board:

1. Which filter criteria were selected in student lists purchased by universities in our sample?
2. What are the characteristics of prospects included in student lists purchased by univer-

sities in our sample?

3. What is the relationship between student list filter criteria and the characteristics of purchased prospects?

In RQ1 the unit of analysis is the order or university-order. Analyses allow us to make statements about how orders vary – within-university and between-university variation – for universities in our sample. In RQ2 the unit of analysis is university-prospect. Analyses allow us to make statements about the characteristics of prospects targeted by universities in our sample. In RQ3 the unit of analysis is order-prospect. Analyses allow us to make statements about the relationship between filter criteria and prospect characteristics that extend to lists purchased by any university that select similar filter criteria.

Empirical analyses consist of simple descriptive statistics presented in tables, figures, and maps. For each research question, analyses are anchored by a small set of tables or figures that present results for the entire analysis sample. Next, we present analyses of selected universities, purchases and/or localities that convey commonly observed or thematically important patterns, with a focus on the nexus between race, class, and geography. For RQ2 and RQ3, we contextualize the characteristics of purchased prospects by showing the characteristics of one or more comparison groups (e.g., all high school graduates in the metropolitan area).

**Secondary data.** Analyses incorporate several secondary data sources. Integrated Postsecondary Education Data System (IPEDS) data provides characteristics of universities in the analysis sample. NCES Common Core of Data (CCD) and Private School Universe Survey (PSS), respectively, provides data about U.S. public and private high schools. The Census American Community Survey (ACS) provide data about community characteristics. We use zip-code level data from ACS 5-year estimates. [OTHER SECONDARY DATA SOURCES TO ADD?]

## 4 Results

### 4.1 Characteristics of Student List Purchases

#### 4.1.1 Total Orders and Number of Prospects

Figure X presents the 486 total orders analyzed in this report by university type and total students purchased. The 486 orders were purchased across 12 universities. The six master's universities in the study made the majority of order purchases (N=307), while research universities made 178 orders and the only baccalaureate university in the study made one order.

The number of total prospects purchased within each order varied widely. Across all 486 orders, the median number of prospects purchased per order was 3,067, whereas the mean was 950 (sd=5,619). Despite making fewer total orders than master's universities, research universities on average purchased nearly double the number of students per order (4,269 versus 2,382). The only baccalaureate college in the study purchased 5,539 prospects in their one College Board order.

#### 4.1.2 Frequency of Filters Used

Filters used to select prospect lists varied across academic criteria (e.g., GPA, PSAT, SAT, academic rank, AP Score), geographic location (e.g., zip code, state, segment, core based statistical area, geomarket, international), and demographic characteristics (e.g., high school graduation class, race/ethnicity, gender). Some geographic filters are metrics created by the College Board. For example, segment filters come from the College Board's "Segment Analysis Service" which merges demographic, geographic, and academic data on SAT test takers to create "geodemographic profiles" for college-bound students (The College Board, 2011, p. 3). These profiles are created at the neighborhood-level and at the school-level. Geomarket filters are also created by the College Board within Enrollment Management



Services, which uses information about SAT score senders from the past five admissions cycles within a specific geographic locality (e.g., counties, metropolitan areas, cities) to make projections about high school graduates in the area.

Figure X shows how often filters were used across all orders. All 486 orders filtered by high school graduation class. The most frequently used filters include GPA (94%), PSAT scores (60%), zip code (57%), and SAT scores (56%). About two in every five orders also filtered by state. Only a subset of orders filtered by race/ethnicity (15%), academic rank (10%), gender (5%), segment (5%), AP score (5%), core based statistical area (3%), geomarket (3%), and international (3%).

The three most commonly used academic filters (GPA, PSAT, SAT) were used by specifying a low and/or high threshold. Across the 464 orders using GPA, low thresholds ranged from A to C+, with the majority of orders using a low of B- (50%) or B (15%). All orders using GPA orders indicated a high threshold of A+. For orders using PSAT lowest score thresholds, 30% indicated less than 1000, 18% indicated 1000-1100, 19% indicated 1110-1200, 17% indicated 1210-1300, 7% indicated 1310-1400, and 9% indicated 1410-1500. For orders using PSAT highest score thresholds, 13% indicated less than 1000, 12% indicated 1000-1100, 26% indicated 1110-1200, 19% indicated 1210-1300, 11% indicated 1310-1400, and 19% indicated 1410-1500..<sup>1</sup> Similar thresholds and percentages were evident for orders that used SAT as a filter. Low SAT thresholds across orders were 17% for less than 1000, 30% for 1000-1100, 21% for 1110-1200, 28% for 1210-1300, 10% for indicated 1310-1400, and 6% for 1410-1500. High SAT thresholds across orders were 9% for 1000-1100, 12% for 1110-1200, 18% for 1210-1300, 5% for indicated 1310-1400, 21% for 1410-1500, and 35% indicated scores greater than 1500+.

PSAT and SAT score thresholds were consistently higher for orders by research universities in comparison to master's universities. The average low PSAT threshold score for research university orders was 1275 (about 90th percentile) versus 1034 for master's universities (about

---

<sup>1</sup>Old PSAT scores were converted to equivalent thresholds for new format: STILL NEED TO DO THIS

50th percentile). Similar patterns were evident for SAT score thresholds. The average low SAT score threshold for research university orders was 1247 versus 1110 for orders by master's universities.

Zip code, state, and segment were the most commonly used geographical filters used across all orders. While we can account for the number of orders that filtered by zipcodes (N=278), we can only analyze patterns for the 208 orders from universities that provided the list of 3-digit zipcodes filtered by.<sup>2</sup> These orders were made by two master's universities, primarily for in-state orders. Figure X shows the three-digit zip code filters used across these orders. The majority of orders (113 of 118) using zip code filters by the first master's university were to zip codes exclusively in the state where the university resides. The remaining 5 orders included in-state and neighboring state zip codes. Orders using zip code filters by the second master's university reveal similar patterns, except the majority of orders included a mix of in-state and neighboring state zip codes (70 of 108).

Figure X shows states that were selected across the 170 orders using the filter. Generally, multiple-state filters were used across out-of-state prospect orders, whereas orders that filtered by only a single state were used for purchasing in-state prospects.

The twenty two orders using segment filters were made by two public research universities in the sample. Figure X shows which neighborhood and high school segments were used across these orders and the characteristics of these segments. One university made 21 of the 22 orders using segment filters, and all 21 orders followed the same patterns of segments. These 21 orders filtered for 10 different neighborhood clusters (51, 53, 58, 60, 61, 63, 69, 70, 73, and 78). ONE SENTENCE ABOUT PATTERNS in CHARACTERISTICS OF THESE NEIGHBORHOODS. Orders also filtered by high school segment clusters 58, 63, 64, 65, 66, 68, 69, 70, 73, 75, and 79. ONE SENTENCE ABOUT PATTERNS IN CHARACTERISTICS OF THESE HIGH SCHOOLS. ONE SENTENCE ON SEGMENT

---

<sup>2</sup>The other 70 orders indicated a zip code filter; however, the actual zipcodes used to filter order lists were not provided to us after multiple requests

## ORDER BY NORTHEASTERN UNVIERSITY?

Figure X also shows that 75 (15%) and 24 (5%) of the 486 orders used a filter for race/ethnicity and gender, respectively. Most of the 75 orders using race/ethnicity filters specified multiple race/ethnicity groups. This includes 28 orders that filtered by Black, Native American, or Latinx prospects; 19 orders that filtered for Asian or White prospects; 7 orders that filtered for Native American or Latinx prospects; and 7 orders that filtered for Asian, White, or other race. The remaining orders filtered for only one race/ethnicity group, include 2 orders filtering for Native American prospects (American Indian, Alaska Native, and/or Native Hawaiian or Other Pacific Islander), 1 order filter for Black prospects, and 11 orders filtering for Latinx prospects. For the 24 orders using gender filters, 75% indicated female prospects and 25% indicate Male prospects.

### 4.1.3 Combination of Filters

Universities in the study used 39 different combinations of filters to purchase prospects. The ten most commonly used combination of filters, which account for nearly 80% of all orders, are presented in Table X. More than half of all orders analyzed used a combination of high school graduation class, zip code, GPA, and PSAT and/or SAT scores to filter prospect lists. This includes 143 orders (29%) that used PSAT scores, 107 orders that used SAT scores (22%), and 28 orders that used both PSAT and SAT (5%) in addition to graduation class, zip code, and GPA. Other orders used a similar pattern of filters except for using state geographic filters instead of zip code, including 18 orders using PSAT scores and another 18 orders using SAT scores.

About 8% of orders (39 of 486) used high school graduation class, state, sat, psat, gpa, and class rank filters while specifying for the race/ethnicity of prospects. The second most commonly used filter combination that specified the race/ethnicity of prospects includes 16 orders that used it in combination with graduation class, state, PSAT, and GPA filters.

The remaining combinations followed similar general patterns in targeting both the academic and geographical characteristics of prospects as the top combinations described above; however, they used other types of filters (e.g., gender, AP scores, segment, geomarket). Thirteen orders used graduation class, state, SAT, PSAT, and GPA filters in combination with segment and gender. Orders that used the segment filter used both neighborhood and high school clusters. For example, one specific order included a filter for census tracts assigned to neighborhood cluster 51, but further filtered by only including prospects from schools assigned to high school clusters 65, 68, 70, and 79 were included.

Eleven orders used graduation class, state, GPA in combination with AP scores. AP score filters tended to be grouped into “fields” and varied across score thresholds. Most orders targeted prospects scoring from 3-5 on AP exams in either STEM fields (e.g., Physics, Calculus, Biology, Chemistry, Computer Science) or Humanities and Fine Arts (e.g., Spanish, French, Art History, Music Theory), although some STEM orders filtered for prospects scoring a 4 or 5 on their AP exams.

Lastly, ten orders used graduation class, SAT, and geomarket. These orders filtered for geomarkets within Texas and Louisiana. Louisiana geomarket orders targeted prospects in the Baton Rouge and Shreveport areas. Texas geomarkets filtered for include but are not limited to the Dallas/Ft. Worth areas, Central Gulf Coast, Wharton County, Victoria County, City of San Antonio, Southwest Houston Metropolitan Area, Brazos and Trinity Valley, Austin and Central Texas, Galveston area, East Texas.

## **4.2 Characteristics of Prospects**

Our analysis on the characteristics of prospects purchased by universities includes 388 orders resulting in 1,967,352 prospects. Figure X shows the total number of prospects by domestic versus international status. Of the nearly 2 Million prospects purchased, 95% of them were domestic students.

Figure X also shows the number of domestic prospects purchased by in-state versus out-of-state and by institutional type. Overall, the majority of prospects purchased across all orders by all institutions in the study were in-state students (58%). However, the percent of in-state versus out-of-state prospects varied across institutional type. Research universities purchased more students overall and a greater proportion of out-of-state students than master's universities. For example, research universities in the study purchased more than 1.2 Million prospects of which 62% were out-of-state. In comparison, master's universities purchased less than 650,000 prospects of which only 4% were out-of-state students.

Below we also describe the racial, economic, and schooling characteristics of domestic prospect lists across institutional type and in-state versus out-of-state. The last sub-section then describes the characteristics of international prospects purchased.

#### **4.2.1 Racial Characteristics**

Figure X presents the racial characteristics of all domestic prospects resulting from the 388 purchased orders by in-state versus out-of-state for research and master's university purchases. Race/ethnicity of prospects is collected from the College Board's voluntary demographic questionnaire completed by students when taking the SAT; therefore, we report racial characteristics as self-identified and include the percentage of students that did not report their race/ethnicity. About 38% of all domestic prospects self-identified as White, 23% as Latinx, 16% as Asian, 6% as Black, 5% as Multiracial, and 13% did not report their race/ethnicity.

Out-of-state prospects purchased tended to be more White and Asian than in-state prospects purchased, although a larger percentage of in-state prospects did not report their race/ethnicity. For example, when universities in the study made order purchases for prospects residing in different states than where their campus is located, these lists resulted in prospect lists made up of 43% White students, 22% Asian students, 19% Latinx students, 6% Black, 6% multiracial, and 4% no response. In comparison, purchases for prospects residing in the

same state as the institution's campus results in lists made up of 34% White students, 11% Asian, 25% Latinx, 5% Black, 4% multiracial, and 19% of students that did not report their race/ethnicity. While out-of-state prospect lists included a larger proportion of White and Asian students, these orders also include a critical mass of Native American students (N=7,885 DOUBLE CHECK THIS) that can be glossed over when only looking at overall proportions.

The differences in the racial characteristics of in-state versus out-of-state prospects are likely also a function of purchases made by research versus master's universities. While research universities made more out-of-state than in-state prospect purchases, the differences in the racial characteristics of both groups were relatively small in comparison to purchases by master's universities. For example, White students made up 43% and 40% of out-of-state and in-state prospect lists purchased by research universities, respectively. In comparison, White students made up 45% of out-of-state prospect lists and only 30% of in-state prospect lists purchased by master's universities. However, purchases by master's universities also included a much larger proportion of students that did not report their race/ethnicity.

#### **4.2.2 Economic Characteristics**

Figure X presents the average median income of the zip code where prospects live by in-state versus out-of-state status for research and master's university purchases. Purchased prospects, across all orders by the 12 universities in the study, live in areas with an average median household income of \$92,000.

Overall, Figure X shows out-of-state prospects tended to live in more affluent areas than in-state prospects. Across all institution types, when universities in the study made order purchases for prospects residing in different states than where their campus is located, these lists resulted in prospects that live in areas where the average median household income is \$100,000. In comparison, purchases for prospects residing in the same state as the institution's campus resulted in prospects that live in areas where the average median household income

is \$87,000.

This disparity is also likely driven by several differences across purchases by research versus master's universities. For example, out-of-state prospects purchased by research universities live in areas where the average median household income is \$101,000, whereas in-state prospects purchased live in areas with a \$91,000. However, the opposite pattern is evident for purchases by master's universities. Out-of-state prospects purchased by master's universities on average live in less affluent areas (\$69,000 median household income) than in-state prospects (\$84,000 median household income).

#### **4.2.3 High Schools Attended**

Given some of the College Board products link individual prospects to high schools for services like high school segment clusters, we are able to analyze some of the characteristics of high schools purchased prospects attend.

Figure X presents school type for purchased prospects by in-state versus out-of-state status for research and master's university purchases. Overall, 82% of prospects purchased attend public high schools, 9% attend private schools, and 9% did not report their high school. While these overall proportions are comparable to national averages, there are several differences across in-state versus out-of-state purchases by research and master's universities. For research universities, students attending private high schools made up a greater share of out-of-state prospect lists (12%) than in-state lists purchases (7%). Student list purchases by master's universities also exhibited this pattern, where private high school students made up 10% of out-of-state prospects versus 6% of in-state prospects purchased.

#### **4.2.4 International Prospects**

Nearly 100,000 of the 2 Million prospects purchased were international students. Figure X presents the countries from which these prospects were purchased, with XYZ representing the total number of prospects purchased by country. The top ten countries, which account

for more than 60% of international prospects purchased across the 388 orders, include India (18%), China (10%), Singapore (6%), South Korea (6%), Canada (5%), United Arab Emirates (5%), Pakistan (4%), Taiwan (3%), Saudi Arabia (3%), and Thailand (3%).

### 4.3 Filter Criteria and Characteristics of Prospects

We analyze the relationship between filter criteria and the characteristics of purchased prospects in two different ways. First, we analyze prospect characteristics (e.g., race/ethnicity, income, in-state versus out-of-state) across individual filters to understand broad patterns. Second, we analyze prospect characteristics across common combinations of filters. Here we use selected universities, purchases and/or localities that convey commonly observed or thematically important patterns across combinations of filters. We also contextualize the characteristics of purchased prospects by showing the characteristics of one or more comparison groups based on the selected examples.

#### 4.3.1 Prospect Characteristics Across Individual Filters

Table X presents the characteristics of prospects by individual filters. For each column, averages are reported across all prospects that were purchased via orders using the specified column filter, which includes orders that used the specified filter in combination with other filters..<sup>3</sup>

Focusing on the racial characteristics of prospects, student lists with the largest percentages of White and Asian prospects result when orders use PSAT, gender, segment, or CBSA filters. For example, orders that specify a gender filter result in prospect lists that are less than 10% Black, Latinx, and/or Native American. This pattern is consistent in prospect lists that use segment or CBSA filters, although the disparity is not as large for orders using a PSAT filter (30% Black, Latinx, Native America). On the other hand, orders that filter by specifying

---

<sup>3</sup>Given we present all prospects across individual filters that are used in combination with others, total number of prospects summed across columns will exceed our grand total of 1,967,352 prospects



particular race/ethnicity groups result in lists that have fewer White and Asian prospects and greater proportions of Black, Latinx, Native American, and multiracial prospects. This coincides with descriptive findings above that suggest more than half of all orders using a race/ethnicity filter specified Black, Native American, and/or Latinx prospects.

Similar disparities are evident across the economic characteristics of prospect lists by filters used. Orders using PSAT, gender, segment, or CBSA filters result in prospect lists with the largest average median household incomes. Orders using a CBSA filter showcase the upper extreme of this pattern, resulting in lists where the average prospect lived in a zipcode where the median household income \$113,000. Similarly, orders using race/ethnicity filters showcased the lower extreme. When universities purchased orders that filtered for specific race/ethnicity groups, the resulting lists included prospects that lived in zip codes where the average median household income was less than \$85,000.

Not surprisingly, orders using geographic filters result in specific patterns of in-state versus out-of-state prospects. However, analyzing the residency status of prospect lists across filters can help us develop insights into how specific filters are used to target prospects geographically. For example, orders using segment and CBSA filters are likely used for targeting out-of-state students, as the use of these filters result in prospect lists made up of 85% and 96% out-of-state prospects, respectively. However, orders filtering for prospects within specific state(s) result in list that are nearly equal proportions of out-of-state and in-state students. Coinciding with descriptive statistics detailed above and data limitations (i.e., we only received zip codes used to filter order lists by two master's universities in our sample), nearly 98% of prospects resulting from orders using a zip code filter were in-state students. Similar to disparities in racial and economic characteristics of prospects, orders using a gender filter also resulted in geographical disparities (94% out-of-state versus 6% in-state).

Lastly, table X shows the difference in proportions of prospects that attend public versus private schools does not change significantly across filters used. For example, orders that

specify a CBSA result in student lists where on average 14% of prospects attend private schools, which is the maximum proportion across all filters. In comparison, orders that use race/ethnicity or state filters result in students lists with the minimum proportion of prospects attending private schools (7%).

### **4.3.2 Prospect Characteristics Across Combinations of Filters**

#### **4.3.3 In-State Order**

We begin analyzing the characteristics of prospects across some of the most common combination of filters. Many of the orders in our analysis filtered by high school graduation class, PSAT scores, GPA, and zip code. Analyses above suggest orders using this combination of filters were made by master’s universities to identify in-state prospects. We “zoom” into orders by one of these master’s universities in order analyze in-depth patterns in the racial and economic characteristics of prospects that result from this combination of filters.

A master’s university located in Texas made 65 orders using graduation class, PSAT scores, GPA, and zip code filters. These orders targeted 2019-2022 high school graduating classes with minimum PSAT scores ranging from 920-1300 and maximum SAT scores ranging from 970-1450. The university also filtered for prospects with GPAs ranging from a low of C+ to a high of A+. Lastly, prospects were also filtered using a series of 3-digit zip codes that included both in-state and neighboring state zip codes.

For each three-digit zip code used as an order filter, we compare the average racial and economic characteristics of the resulting purchased prospects to the zip code’s population of 15-19 year-olds. We select the zip code population as the comparison group for purchase prospects filtered by graduation class, PSAT scores, GPA, and zip code filters for various reasons. First, the university included a three-digit zip code filter which encapsulates all students living within in a specified area (usually a City or County), which should hypothetically cast a wider, more equitable net than filtering explicitly by 5-digit zip codes. However, we analyze

whether this wider net is compromised when it is used along with test scores and GPA. Secondly, master's universities are like to serve regional student markets, rather than the entire state where they are located. Therefore, we use a filtered zip code's population as a comparison group for prospects purchased within that zip code rather than all zip codes in the state from which prospects were not purchased from.

For example, the three-digit zip code 752 was included in 27 of the 65 orders that filtered across graduation class, PSAT scores, GPA and zip code. The 49 five-digit zip codes nested within the three-digit 752 zip code includes communities across the Dallas metropolitan area (see Figure X). The student lists for these orders resulted in 8,707 purchased prospects living in one of the 47 five-digit zip codes. About 40% of purchased prospects identified as White, 39% as Latinx, 8% Black, 5% Asian, and 3% multiracial. In comparison, the population of 15-19 year olds in these 49 five-digit zip codes are 24% White, 43% Hispanic, 20% Black, 3% Asian, and 3% multiracial.

Figure X presents the average percentage point difference between the racial/ethnic composition of students lists and of the 15-19 year old population across all zip codes filtered by the master's university in this analysis. For example, Zip Code 752 in Figure X shows a 16 percentage point difference between White prospects purchased living in the zip code and the overall percentage of White 15-19 year olds living in the zip code ( $40\% - 24\% = 16$  percentage points), suggesting White prospects are overrepresented in purchased lists relative to the population of their home zip code. Similarly, the figure shows a -12 percentage point difference between the Black prospects purchased from the zip code and the percentage of Black 15-19 year olds living in the zip code ( $8\% - 20\% = -12$  percentage points), suggesting Black prospects are underrepresented in purchased lists relative to the population of their home zip code.

Overall, Figure X suggests prospect lists from orders that filtered using graduation class, PSAT scores, GPA, and zip code tend to be more White and Asian and less Black and Latinx relative

to the population of 15-19 year olds from zip codes filtered by. The overrepresentation of White prospects in comparison to zip code populations ranged from 2 to 36 percentage points for 11 of the 23 zip codes filtered by, whereas three zip codes had proportional representation (0 percentage point difference). The remaining 9 zip codes had an underrepresentation of White prospects, where eight orders ranged from 1 to 6 percentage points and one order had -21 percentage points. The overrepresentation of Asian prospects purchased across the 23 zip codes ranged from 1 to 17 percentage points.

Black prospects were underrepresented relative to zip code populations across all zip codes used and in the greatest magnitude across all racial/ethnic groups. The underrepresentation of Black prospects purchased across the 23 zip codes ranged from 2 to 47 percentage points. Latinx students were also underrepresented across 12 of the 23 zip codes, although the magnitude was not as large as (ranging from 1 to 10 percentage points). Three of the zip codes filtered by had proportional representation of Latinx students (0 percentage point difference), whereas the remaining zip codes had an overrepresentation ranging from 1 to 4 percentage points.

Economic disparities are also evident when we compare the average median income of prospects relative to the average median household of the zip code population. Figure X presents the difference in average median household income between prospects (measured at the 5-digit zip code level) and the average median household income across all nested 5-digit zip codes within the 3-digit filters used. Because both prospects' and population income is measured at the zip code level across 5- versus 3-digits, a disparity indicates more prospects living in affluent 5-digit zip codes were purchased than the average median household income across all 5-digit codes within the 3-digit filter.

Figure X shows student lists resulting from 22 of the 23 zip code filters included more affluent prospects than the average population of those zip codes. This disparity was largest for prospects purchased from 774 zip codes (communities within Richmond, Texas), which

had an average median income nearly \$40,000 greater than the average median income of the population across all nested 774 zip codes. Prospects purchased from 758 zip codes (communities within Palestine, Texas) were the only across all filters to have an average median income less than the population across all nested zip codes (\$43,000 versus \$50,000).

#### **4.3.4 Out-of-State Order**

## **5 References**

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, 14(4), 532–550. <https://doi.org/10.5465/amr.1989.4308385>