College Choice and Intended Major in the U.S.

PRELIMINARY AND INCOMPLETE.

Emily E. Cook*

January 15, 2019

Abstract

Universities compete for new cohorts of high school graduates each year, both through prices and through other characteristics—including the number and type of majors offered. As financial and competitive pressures have increased in the past decade, many universities have made or proposed changes to the number and type of undergraduate majors that they offer, seeking to gain a competitive edge by aligning their academic offerings with the changing preferences of high school graduates. Through descriptive analysis I demonstrate that universities vary substantially in the number and type of majors that they offer, and that students' preferences over majors have evolved in the last decade. I model the application, admissions, and enrollment process and estimate the parameters using a panel of state- and institution-level data. Counterfactuals explore applications, admissions selectivity, and enrollment outcomes by student intended major and institution type, under three types of policy counterfactuals: 1) university-level reorganization of majors, 2) tuition-free policies, and 3) expanded mandatory admissions testing policies.

*Department of Economics, University of Virginia, emilycook@virginia.edu. I thank Federico Ciliberto, Sarah Turner, Gaurab Aryal, Simon P. Anderson, Ekaterina Khmelnitskaya, Dan Savelle, Cailin Slattery, and many other students and faculty at the University of Virginia for helpful advice and comments. Victoria Roloff provided excellent research assistance. I also wish to acknowledge the Bankard Fund for Political Economy and the National Science Foundation for financial support. The results contained herein were derived in part from data provided by the College Board. All errors are mine.

1 Introduction

Citing "declining financial resources, demographic changes...and rising competition among public and private universities," the University of Wisconsin at Stevens Point (UWSP) in 2018 proposed to cut 6 majors, including French, German, history, geology, geography and art, while adding new programs in environmental science and other high-demand majors (Strauss, 2017; Herzog, 2018). The leadership at UWSP are not alone in their efforts to cut costs and align majors offered with changing student preferences. Among others, the University of Southern Maine, the University of Akron, Goucher College, and universities across the University of North Carolina system have all cut undergraduate major programs within the last five years (Stancill, 2015; University of Southern Maine, 2015; Strauss, 2017; Tkacik, 2018; McCafferty, 2018). While the cuts draw public attention, the news is not all about cuts; many universities have introduced new majors, particularly in "career-focused" fields such as computer science, environmental science, graphic design, and cybersecurity. On average, the number of majors offered at four-year degree-granting universities increased by 11.7% among public universities and 16% among private universities between the 2006-07 and 2015-16 academic years.[1]

As suggested by the leadership at UWSP, the U.S. higher education market has recently been marked by changing financial conditions, increasing competition for students and, in some geographic areas, declining populations of high school graduates. Appropriations per student to public four-year institutions declined 27% between the 2007-08 and 2012-13 academic years, and despite some recent increases in appropriations, the decade 2006-07 to 2015-16 saw a net 18.9% decrease. Students are sending more applications than ever before, with the average number of applications per first-time first-year enrollee in four-year degree-granting public and private non-profit colleges and universities increasing from 5 to 7.2.[2] The increase in applications per matriculant is perhaps the most salient way to see the changing competition in this industry, as it suggests that students consider and compare a broader array of options than they did decades ago. Finally, demographic shifts impact competition for students: the total number of U.S. high school graduates is increasing over time, but the rise has slowed substantially. The number of public and private high school graduates rose 19.5% between 1996-07 and 2005-06 and 10.6% between

[1]This is among majors that existed in 2006. In addition, universities have created new programs that were unclassified by the National Center for Education Statistics as of 2006.

[2]Applications per student has likely increased due to technological improvements that have increased high school graduates' access to information and reduced the cost of distance to home, continuing trends that were noted by (Hoxby, 1997).

2006-07 and 2015-16, but the increase is only projected to be 3.7% between 2016-17 and 2025-26 (U.S. Department of Education, 2016). While national levels of graduates were rising, 11 individual states (largely in the Midwest) experienced declines in the number of high school graduates of 2% or more between 2006-07 and 2015-16.

In this climate, university leaders may wonder whether the selection of majors offered at their university will affect applications and enrollment. Is reorganization of majors offered an effective method to boost demand among high school graduates? To what extent can reorganization of majors offered offset price reductions, while hold applications and enrollment steady?[3] The answers depend upon whether high school students have preferences over majors that affect their college choice. To the extent students do consider their major options at the application and enrollment stages, a strategy that keeps the number and type of majors offered in line with student preferences may be effective at boosting demand while minimizing expenditures on unpopular majors.

For the state and federal governments, typical policy interventions in higher education involve either direct price regulation for public universities, or administration of scholarship and aid programs. One of the most politically visible proposals is the tuition-free proposal that has been championed by politicians including President Obama, Hillary Clinton, and Bernie Sanders. A central question for the policy debate surrounding this proposal is how a reduction in tuition will affect college choice outcomes. My model allows me to study questions such as: Which students would apply to college if tuition were free? How are the potential enrollment effects distributed across students with varying scores and intended majors? If free tuition were implemented at public universities only, how would it affect applications, selectivity and enrollment at private universities, holding capacity (and private universities' tuition) constant?

The final policy application of my model is to mandatory college admissions testing. State-level mandatory admissions testing policies provide free registration for college admissions tests and require all students to sit for a college admissions test prior to high school graduation. Clearly these policies reduce the pecuniary cost of application, but they also reduce the non-pecuniary cost of college application for high school graduates, as the non-pecuniary testing cost becomes a component of the cost of graduating high school. Mandatory testing policies have been shown to have small positive impacts on college enrollment (see Klasik (2013); Hurwitz et al. (2015); Goodman (2016); Hyman (2017)). Between 2001 and

[3]It is clear that the optimal university policy depends not only on demand elasticities but also on cost variables and competition from other universities. A supply-side model is required to study the optimal university policy, and is left to future work.

2015, more than half of U.S. states have adopted a mandatory testing policy. I study the counterfactual applications, admissions selectivity, and enrollment by major and college type under national mandatory college admissions testing.

To study the university, state and federal policies described in the last three paragraphs, I build and estimate a discrete-choice model that incorporates the application, admission, and enrollment stages of the college admissions process. In the model, a student's enrollment utility depends on university characteristics (including majors, net price, and admissions standards), the student's admissions test score, and their intended major. Universities choose an admissions threshold subject to a capacity constraint and a minimum acceptable student match value. Students' application decisions are based on the expected value of their applications, and the application cost.

The model and estimation strategy build upon previous work on demand estimation in general and the college admissions problem in particular. Gentzkow (2007) and Fan (2013) provide frameworks for modeling and estimating multiple-discrete choice problems that are similar to my college application context. Within the college choice literature, the work by Arcidiacono (2005), Howell (2010), Fu (2014), Chade et al. (2014), and Kapor (2016) provide the foundation for my paper.

The remaining sections of this paper are as follows: Section 2 describes the contribution of this work to the related literature. Section 3 outlines the college choice model, which includes the applications, admissions, and enrollment stages of the college admissions process. Section 4 describes the data sources. Section 5 provides a descriptive analysis of majors offered, university characteristics and students' intended majors, which provides evidence that intended major may affect college choice. Section 6 details the empirical strategy for estimating the model. In Sections 7 and 8 I present the results and conclusions.

2 Related Literature

Besides adjustments to prices and majors offered, university leaders have multiple other tools to deal with changing market conditions, some of which have been explored in prior research. These tools include reducing cost by substituting tenure-track faculty for non-tenured teaching staff (Bettinger and Long, 2010), and enrollment of more international or (particularly for public universities) out-of-state students (Bound et al., 2016). The question addressed in this paper—whether adjustments to majors offered is another effective tool by which universities can affect student demand—has been unexplored to date.

My work is complementary to other multiple discrete choice approaches to demand estimation, such as those used by Gentzkow (2007) and Fan (2013). In both of these papers on the newspaper industry, estimating equations are derived from utility functions for each individual newspaper. Similarly, I derive choice probabilities for application portfolios (analogous to a "bundle" in the newspaper market) from the utility of enrollment at individual universities, following much of the existing structural work on college choice (Arcidiacono, 2005; Howell, 2010; Fu, 2014; Kapor, 2016). At the application stage, students then make a single discrete choice among application portfolios. The approach is distinct from Gentzkow (2007) and Fan (2013) in that high school graduates may only attend one university in the final stage, whereas newspaper consumers may actually read more than one newspaper. Thus in my model, the utility from an application portfolio is an expected value of enrolling in the best option after admissions decisions are revealed.

Unlike the existing structural work on college choice, I study college choice with data aggregated at the institution, state and year level, supplemented by information from individual student survey data. These data combined with the method of moments estimation approach provide some key advantages. First, I can model a large set of individual universities (instead of grouping universities into "types" or using a limited set of universities, as often done due to data constraints, research setting, or computational limitations—see Epple et al. (2006), Fu (2014), and Kapor (2016)). Second, the data provide information on the high school graduates' intended major at the state and year level (data that is unavailable in any publicly available individual survey).[4] Third, the data are available for the decade from 2006-07 to 2015-16, and for all universities that accept the SAT and ACT in admissions. Individual surveys are generally available for a single cohort. Because my model is estimated off of such a large panel of individual universities across state/year markets I am able utilize geographic, institution and time variation in estimation of the value of university and student characteristics. Finally, the estimation framework allows for a methodologically straightforward approach to correct for endogenous university characteristics, in the spirit of Berry et al. (1995).

[4]Some recent NCES surveys such as the Education Longitudinal Study of 2002 (ELS2002) and the High School Longitudinal Study of 2009 (HLS2009) ask about intended major, but only after matriculation at a specific university. Thus the answers are limited by the majors available at the university the student attends.

3 College Choice Model

There are two types of agents in the model: high school graduates (students) and universities. The timing is as follows: first, students observe net price, majors offered, and exogenous characteristics of institutions, and choose an application portfolio (a set of universities to which they apply—this can include no university).[5] Second, universities notify students of their admissions decisions. Finally, students choose which university to attend from the set of universities that accepted them, or they may choose not to attend. Details on each stage follow below, starting with the enrollment utility, then the application utility and the model of admissions. The last three subsections derive aggregate moments, discuss the market definition, and outline equilibrium in admissions.

3.1 Student Utility from Enrollment

Students pick an application portfolio that maximizes their expected utility of enrollment. Thus, the first step toward a fully specified model of college choice is to define the utility of enrolling in a college. The specification assumes that the student is free to choose any major upon enrollment.[6]

Let $l = 1...L$ denote the high school graduate's state of residence, and $t = 1...T$ denote the academic year. There are $J + 1$ universities $j = 0, 1, ..., J$, where $j = 0$ represents the outside option of two-year colleges, other educational certificate programs, or the labor market. Each university may offer majors within K possible major groups, $k = 1, ..., K$. Let $\mathbf{M_{jt}}$ be a $K \times 1$ vector where each element, M_{jkt} is the number of majors university j offers in major group k in year t. Universities also have exogenous characteristics, $\mathbf{X_{jlt}}$, including private/public control, distance to each state, instructional expenditures per student, and selectivity ranking. Universities also have a net price, p_j, which is tuition minus average grants and aid.[7]

The utility individual i receives from enrolling at university j is:

$$u_{ijlt} = \alpha + \beta_0 p_{jlt} + \boldsymbol{\beta_1} \mathbf{M_{jt}} + \boldsymbol{\beta_2} \mathbf{X_{jlt}} + \xi_{jlt} + \sum_{k=1}^{K} \rho_k M_{jkt} m_{ik} + \eta_{ijlt},$$

[5]Note that this is a bit of a simplification, because net price for the upcoming year is typically announced after application deadlines. Essentially, I assume that students are able to make an accurate prediction of net price for the upcoming year based on available information, such as last year's net price.

[6]That is, I do not model the admissions process for "limited enrollment" majors that have admissions criteria beyond admission to the university.

[7]In future work I will incorporate financial aid, which depends on student income level.

where m_{ik} are indicators for whether student i has a preferred major k. These indicators have known population probabilities.[8] The student-specific error, η_{ijlt}, is distributed Type I Extreme Value. This error represents information that becomes known to the student after their acceptance and that may have different value to different students (e.g., information on food options gained during a campus tour). The university-specific error ξ_{jlt} represents the mean value of university characteristics unobserved to the researcher but known to students before application. For example, this could be something like a university sports team having a successful season.

The term $\sum_{k=1}^{K} \rho_k M_{jkt} m_{ik}$ represents heterogeneous preferences for a university that offers the vector $\mathbf{M_j}$. If a student prefers to major in engineering, the indicator m_{ik} is equal to 1 for the k that includes engineering, and 0 otherwise. The additional utility the student who prefers the engineering major receives from enrolling at a university that offers M_{jkt} majors in the engineering category is captured by ρ_k.[9]

Let $\mathbf{D_i}$ be a $(J+1) \times 1$ vector where the element $D_{ij} = 1$ if the student i is offered admission to school j and $D_{ij} = 0$ otherwise. Students may always choose not to attend any college ($D_{i0} = 1$), in which case they receive mean utility of 0 for the outside option.

From here, I drop the state and time subscripts for notational ease. Letting $\nu_{ij} = u_{ij} - \eta_{ij}$ Given the assumptions outlined so far, the individual choice probability at the enrollment stage is given by:

$$\mathbb{P}(Enroll_{ij}|\mathbf{D_i}, \mathbf{m_i}) = \frac{e^{\nu_{ij}}}{\sum_{j\ s.t. D_{ij}=1} e^{\nu_{ij}}} \tag{1}$$

3.2 Student Utility from Application

Each student chooses an application portfolio to maximize his expected utility of enrollment, given a probability of admission. To derive this, we can first find the utility i receives from being accepted to $\mathbf{D_i}$. This is the expected value of the maximum utility among the universities represented by $\mathbf{D_i}$, which can be expressed as follows:

$$v(\mathbf{D_i}|\mathbf{m_i}) = ln(\sum_{j\ s.t.\ D_{ij}=1} e^{\nu_{ij}}) + \gamma, \tag{2}$$

[8] In the data, the population probabilities for the entire high school graduating cohort in each state and year are not observed. Only the distribution among test-takers is observed. See Appendix C.2 for details on the estimation of the population probabilities.

[9] This model does not incorporate students' decisions to switch between majors (or between universities) after matriculation; instead, I model the utility of applicants for enrollment, as perceived at the time of application. Thus the results are to be interpreted only as effects on first-time first-year enrollment, not as effects on total enrollment.

where γ is the Euler–Mascheroni constant.

The utility of an application portfolio is the expected utility of enrollment given the equilibrium probability of acceptance at each university. Let $\mathbf{Y_i}$ be a vector indicating the student's application portfolio. $\mathbf{Y_i}$ is a $(J+1) \times 1$ vector where the element $Y_{ij} = 1$ if the student applies to school j and 0 otherwise. The student always "applies" to the outside option $j = 0$. I use s_i to denote student i's academic skill, as measured by a normalized SAT or ACT score. Let the probability of being admitted to $\mathbf{D_i}$, conditional on the student's skill level, s_i, and application portfolio, $\mathbf{Y_i}$ be denoted by $\mathcal{P}(\mathbf{D_i}|s_i, \mathbf{Y_i})$. Students correctly anticipate this probability prior to application. The specification of this probability is discussed in Section 3.3.

Then the expected value of an application portfolio Y_i for individual i is:

$$V(\mathbf{Y_i}|s_i, \mathbf{m_i}) = \sum_{\mathbf{D_i} \subseteq \mathbf{Y_i}} (\mathcal{P}(\mathbf{D_i}|s_i, \mathbf{Y_i}) v(\mathbf{D_i}|\mathbf{m_i})) - c(\mathbf{W}|\mathbf{Y_i}) - \epsilon_{\mathbf{Y_i}}, \tag{3}$$

where $c(\mathbf{W}|\mathbf{Y_i}) + \epsilon_{\mathbf{Y_i}}$ represents the cost of application to $\mathbf{Y_i}$. $\mathbf{W}$ represents variables affecting cost, and $\epsilon_{\mathbf{Y_i}}$ is a Type I Extreme Value error that captures unobserved elements of the cost of application. This unobserved component could include application completion time associated with different application procedures across schools (everything from different log-in requirements for the online application system to different application essay topics or lengths).[10]

The application cost depends upon the number of applications and whether or not the state has a mandatory college admissions testing policy, as discussed in the introduction:

$$c(\mathbf{W}|\mathbf{Y_i}) = \omega_1 ln(\sum_j Y_{ij}) + \omega_2 mtest, \tag{4}$$

where $mtest$ is an indicator for whether the state has implemented mandatory admissions testing.

3.3 Admission Probabilities

Derivation of the expected value of an application portfolio requires a model of admissions, which I outline here. This model is in the spirit of previous work by Kapor (2016).

[10] I assume this cost error is uncorrelated with the individual preferences for specific universities. I also assume it is independent across application sets. The latter is a weakness that will be addressed in future work through a Generalized Extreme Value assumption on the error term, as in (Bresnahan et al., 1997) and (Arcidiacono, 2005).

Each student has an academic skill level, s_i, which is known to the student and the university. The university also sees a university-specific match value, ζ_{ij}, which is distributed $N(0, \sigma^2)$. Students know only its distribution. Universities see only their own match value, not other universities' match values. The match value captures idiosyncratic components of the match between student and university. For example, it might be positive if an applicant plays the trumpet and the university needs more trumpet players in its marching band.

The admission decision is is based on the sum of the university-specific match and the observed skill: $s_{ij} = s_i + \zeta_{ij}$. Universities maximize student quality, which is the sum of matriculants' s_{ij}, constrained by capacity and a minimum match value, $\underline{s}_{jt}^{low}$. Kapor (2016) shows that this model yields cutoff values as optimal admissions rules. Thus, for the rest of this paper an admissions rule is a threshold, which is the lower bound on s_{ij} that j will admit.

Each private university has exogenous capacity for first-year students, κ_{jt}, while public universities have exogenous in-state and out-of-state (κ_{jt}^{is}, κ_{jt}^{os}) capacities. Thus public universities can be thought of as two separate universities (in-state and out-of-state) for the purpose of admissions—they have separate admissions thresholds and capacities for in-state and out-of-state students. Universities choose admissions thresholds $\underline{s}_{jt}$ (or $\underline{s}_{jt}^{is}$, $\underline{s}_{jt}^{os}$ for public universities) that set first-year enrollment equal to capacity in expectation, unless the admissions threshold required to fill capacity would be below the lower bound, $\underline{s}_{jt}^{low}$. This lower bound allows that some (non-selective) schools have excess capacity that cannot be filled with qualified candidates.

Let N_{jlt} denote the number of students who applied to j from state l at time t. As defined earlier, $\mathbf{Y_i}$ is a vector of indicators representing i's application decisions. Then the total enrollment in time t at a private university j is the probability that a student enrolls in university j, times the total number of students in state l, summed across all states in which the university competes. From the private university's perspective at the admissions stage, this is:

$$\kappa_{jt} = \sum_{l \in L} N_{jlt} \int_{\underline{s}_{jt}}^{\infty} \mathbb{P}(Enroll_{ij} | D_{ij} = 1, \mathbf{m_i}) dF_{lt}(s_{ij}, \mathbf{m_i} | Y_{ij} = 1). \tag{5}$$

The term inside the integral $\mathbb{P}(Enroll_{ij} | D_{ij} = 1, \mathbf{m_i})$ is the probability of enrolling in j after having been admitted to j, conditional on intended major. This is equation 1, weighted by the probability of combinations of other acceptances conditional on academic skill and summed across those combinations.

The expression $F_{lt}(s_{ij}, \mathbf{m_i}|Y_{ij} = 1)$ is the distribution of student match value and intended major conditional on application.

Public universities each have two similar equations that determine in-state and out-of-state thresholds:

$$\kappa_{jt}^{is} = N_{jt}^{is} \int_{\underline{s}_{jt}^{is},}^{\infty} \mathbb{P}(Enroll_{ij}|D_{ij} = 1, \mathbf{m_i}) dF_t^{is}(s_{ij}, \mathbf{m_i}|Y_{ij} = 1), \tag{6}$$

and:

$$\kappa_{jt}^{os} = \sum_{l \in L^{os}} N_{jlt} \int_{\underline{s}_{jt}^{os},}^{\infty} \mathbb{P}(Enroll_{ij}|D_{ij} = 1, \mathbf{m_i}) dF_{lt}(s_{ij}, \mathbf{m_i}|Y_{ij} = 1). \tag{7}$$

Students rationally anticipate their admissions probability, which is $\mathcal{P}_j(s_i) = Prob(s_{ij} > \underline{s}_j) = 1 - \Phi(\underline{s}_j - s_i)$ at private universities, where $\Phi(\cdot)$ is the normal CDF with mean 0 and variance σ^2.[11] At public universities, the probabilities depend upon the relevant in-state or out-of-state threshold. Given these admissions probabilities and the assumption of independence in admissions conditional on test scores, students have the following probability of being admitted to a set $\mathbf{D_i}$ conditional on their academic skill and application decisions:

$$\mathcal{P}(\mathbf{D_i}|s_i, \mathbf{Y_i}) = \prod_{j \ s.t. Y_{ij}=1,\ D_{ij}=1} \mathcal{P}_j(s_i) \prod_{j \ s.t. Y_{ij}=1,\ D_{ij}=0} [1 - \mathcal{P}_j(s_i)]. \tag{8}$$

3.4 Aggregate Applications, Enrollment and Academic Skill of Enrollees

I derive three expressions that are matched to the data by state and year. The first is the market penetration of each university, defined as the proportion of high school graduates in each state and year who apply to the university. The second expression is the expected first-year enrollment of the university. The third is the interquartile range of academic skill (as measured by test scores) among first-year students at each university. In future work, these three will be supplemented by the proportion of students who submit a given count of applications within each state and year.

The probability that an individual applies to a given application portfolio $\mathbf{Y_i}$ is:

$$\mathbb{P}(Apply_{i\mathbf{Y}}|s_i, \mathbf{m_i}) = \frac{e^{\tilde{V}(\mathbf{Y}|s_i, \mathbf{m_i})}}{\sum_{\mathbf{Y} \in \Upsilon} e^{\tilde{V}(\mathbf{Y}|s_i, \mathbf{m_i})}}. \tag{9}$$

[11] Equilibrium in admissions is discussed in Section 3.6.

Summing across all $\mathbf{Y}$ that include j gives the probability that i applies to university j:

$$\mathbb{P}(Apply_{ij}|s_i,\mathbf{m_i}) = \frac{\sum_{\mathbf{Y}\ s.t.\ Y_j=1} e^{\tilde{V}(\mathbf{Y}|s_i,\mathbf{m_i})}}{\sum_{\mathbf{Y}\in\Upsilon} e^{\tilde{V}(\mathbf{Y}|s_i,\mathbf{m_i})}}. \tag{10}$$

Finally, the market penetration function for j is:

$$\mathbb{P}_j = \int \mathbb{P}(Apply_{ij}|s_i,\mathbf{m_i})dF(s_i,\mathbf{m_i}). \tag{11}$$

Equation 11 becomes the first estimating equation. The second is derived from first-year enrollment. Equation 5 describes how a (private) university chooses the admissions threshold conditional on applications. As the researcher I do not observe the distribution of demographics for each university's set of applicants in each year, so I do not utilize this equation directly. Instead, I take the expectation of enrollment conditional on admissions thresholds, market demographics, and other demand parameters. Note that this function describes expected enrollment, which may be different from capacity for less-selective institutions, as described in Section 3.3.

The unconditional probability of enrolling in university j is:

$$\mathbb{P}(Enroll_{ij}|s_i,\mathbf{m_i}) = \sum_{\mathbf{Y_i}}\sum_{\mathbf{D_i}\subseteq\mathbf{Y_i}} \mathbb{P}(Apply_{i\mathbf{Y}}|s_i,\mathbf{m_i})\mathcal{P}(\mathbf{D_i}|s_i,\mathbf{Y_i})\mathbb{P}_i(Enroll_j|\mathbf{D_i},\mathbf{m_i}). \tag{12}$$

Note that each component inside the integral has been described before. They are, respectively, the probability of application to a specific application portfolio conditional on academic skill and intended major (equation 9), the probability of acceptance to a set of universities conditional on applications and observed skill, (equation 8), and the probability of enrollment conditional on the admissions decisions and intended majors (equation 1).

And the enrollment share of the university is:

$$\mathbb{P}(Enroll_j) = \int\sum_{\mathbf{Y_i}}\sum_{\mathbf{D_i}\subseteq\mathbf{Y_i}} \mathbb{P}_i(\mathbf{Y_i})\mathcal{P}(\mathbf{D_i}|s_i,\mathbf{Y_i})\mathbb{P}_i(Enroll_j|\mathbf{D_i},\mathbf{m_i})dF(s_i,\mathbf{m_i}) \tag{13}$$

Enrollment in university j is this probability times the total number of high school graduates in each state, summed over all states.[12]. For public universities there are separate equations for in-state and

[12]The probability is 0 for states in which the university does not compete according to the definition in Section 3.5

out-of-state enrollment. For this expression I add back in the market and year subscripts, denoting the result of equation 13 by $\mathbb{P}_{lt}(Enroll_j)$. The total (domestic) first-year enrollment is then given by:

$$Enroll_{jt} = \sum_{l \in L} N_{jlt} \mathbb{P}_{lt}(Enroll_j) \tag{14}$$

From the model, we can compute the density function of s_i among enrollees:

$$P[s_i|Enroll_j] = \frac{P[s_i]P[Enroll_{ij}|s_i]}{P[Enroll_j]} \tag{15}$$

The three components of the last expression are either known or computed from the model as described above. Once the distribution of scores is obtained, the IQR can be matched to data.

3.5 Market definition

Students may apply to any university in their consideration set, which is made up of universities that compete in their state. Universities may compete in multiple states, but most universities in the U.S. are competitive in only a few states. This is important, because it alters the set of competitors that each institution faces. To determine which universities compete in each state, I rank the states (within university) by the proportion of the state's high school graduates that attend that university. Then, starting from the top of the list, I select the states that together account for at least 85% of the university's total enrollment. The university competes in this list of states. In these computations I used all ten years of data to smooth out any fluctuations by year.

Figure B1 demonstrates the outcome of this method for the College of William and Mary, Duke University, Old Dominion University, and the University of Virginia. Duke is well-known nationwide and highly prestigious, so its enrollment pulls from a large number of states across the nation. The University of Virginia, while a public institution, is one of the top-ranked public institutions in the nation, so it also draws a large number of students from other states, particularly in the Northeast. The College of William and Mary, a small but highly-ranked college, draws most of its students from the Northeast as well, while Old Dominion University (a less selective public university located in Virginia) draws nearly all of its enrollment from Virginia. All four institutions are competitors in Virginia, while of these four, only Duke competes for students from Texas.

3.6 Equilibrium in Admissions

I use a rational expectations equilibrium concept for the admissions stage (as in Chade et al. (2014) and Kapor (2016)). A rational expectations equilibrium is a set of enrollment decisions, $Enroll_{ij}\ \forall\, i$, admissions thresholds, $\underline{s}_{jt}$ (or $\underline{s}^{is}_{jt}$, $\underline{s}^{os}_{jt}$ in the case of public universities) $\forall\, j, t$, and application portfolios, $\mathbf{Y_i}\ \forall\, i$, such that:

1. $Enroll_{ij}$ maximizes u_{ij} given offers $\mathbf{D_i}\ \forall\, i$.
2. $\underline{s}_{jt}$ sets κ_{jt} (or κ^{is}_{jt}, κ^{os}_{jt} in the case of public universities) equal to expected enrollment at j in time t, conditional on applications.
3. Application portfolios $\mathbf{Y_i}$ maximize $V(\mathbf{Y_i}|s_i, \mathbf{m_i})\ \forall\, i$.

Proposition 1. *A rational expectations equilibrium exists.*

To show existence of equilibrium, I need to show that there is a vector of admissions thresholds $\underline{s}$ that solves the enrollment equation for every university, and in-state and out-of-state status in the case of public universities. Recall that this equation is the following, in the case of private universities:

$$\kappa_{jt} = \sum_{l \in L} N_{jlt} \int_{\underline{s}_{jt}}^{\infty} \mathbb{P}(Enroll_{ij}|D_{ij} = 1, \mathbf{m_i}) dF_{lt}(s_{ij}, \mathbf{m_i}|Y_{ij} = 1).$$

This equation, and the corresponding equations for public universities, define $\underline{s}_{jt}$ implicitly as a function of $\underline{s}_{-jt}$, as well as application and enrollment decisions.

Define the implied best-response function as $h(\underline{s}) \equiv \underline{s}^*(\tilde{Y}(\underline{s})) : R^J \to R^J$, where $\tilde{Y}$ is the distribution of applications induced by $\underline{s}$. Because the share of applications and the density of student-college match change continuously in $\underline{s}$, the best-response function is continuous.

We have assumed a lower bound $\underline{s}^{low}_j$ on the admissions thresholds (see Section 3.3). An upper bound on the admissions threshold $\underline{s}^{high}_j$ is the threshold that would be chosen if all students applied only to j. Thus Brouwer's fixed point theorem applies, and $h(\underline{s})$ has a fixed point.

Proposition 2. *There is a unique equilibrium cutoff vector $\underline{s}$, conditional on applications.*

Following Kapor (2016), I can also show a unique cutoff vector $\underline{s}$ conditional on applications. Consider two admissions sets $D \subseteq D'$. The following holds in the logit model described here:

$$\mathbb{P}(Enroll_{ij}|D') \leq \mathbb{P}(Enroll_{ij}|D). \tag{16}$$

Define $\underline{s}^*(\underline{s}) : R^{|J|} \to R^{|J|}$ by $\underline{s}^*(\underline{s}) = \underline{s}_j^*(\underline{s}_{-j})$. The function $\underline{s}^*$ is monotonic in $R^{|J|}$, because of the enrollment condition above. Therefore, by Tarski's Fixed Point Theorem, its fixed points form a complete lattice and there is an inf ($\underline{s}_L^*$) and sup ($\underline{s}_H^*$) set of equilibrium cutoffs.

Assume $\underline{s}_L^* \neq \underline{s}_H^*$. For some student i, this means that $D_i(\underline{s}_H^*) < D_i(\underline{s}_L^*)$, i.e. there is some college j that does not admit i under $\underline{s}_H^*$ that would have under $\underline{s}_L^*$. The share of students attending college j must weakly decrease (by equation 16) when college j does not admit i. But j is at full capacity under both cutoffs:

$$\mathbb{P}(Enroll_{ij}|\underline{s}_H^*) = \mathbb{P}(Enroll_{ij}|\underline{s}_L^*), \tag{17}$$

which is a contradiction.

4 Data

The data elements required include university characteristics, applications to each university by state and year, enrollment in each university from each state in each year, and student demographic information. To obtain all of these elements I use data from several sources including the ACT, the College Board, and the National Center for Education Statistics' (NCES) Integrated Postsecondary Education Data System (IPEDS).

The university characteristics come from the IPEDS database from 2006-07 through 2015-16.[13] I include every active 4-year public or private non-profit institution located in the 50 states and D.C. that requires the SAT or ACT for admission and that is not listed as a "Special Focus Institution" or "Tribal College" based on the Carnegie Classification of Institutions' Basic Classification. The population of universities under study, and their average tuition, grants and aid, and majors offered are summarized in Table A3. There are approximately 500 public universities and 900 private universities included in this study in each year. Average tuition and average grants and aid rose markedly over the decade for both public and private universities, and the number of majors has increased by nearly 6 majors on average

[13]The IPEDS data and SAT/ACT data are available earlier than 2006-07, but differences in the way variables are recorded over time cause inconsistencies in the required data if I go back further.

among both public and private non-profit universities.

I use degrees awarded by major to infer majors offered at each institution in each year. Using basic internet searches, I confirmed that the year degrees are first awarded in a major typically aligns with the year the major is first offered. A college major is measured and defined using the six-digit major code in NCES' 2010 Classification of Instructional Programs (CIP) system.[14] The six-digit codes are highly disaggregated, as illustrated in Table A1, which lists several History codes as an example. Each six-digit CIP code is associated with a two-digit CIP Code that classifies broad areas of study. Because it is infeasible to model hundreds of these individual six-digit majors separately, I combine two-digit CIP codes to form these groups: 1) Arts and Architecture, 2) Business, 3) Education, 4) Engineering, Math, and Science, 5) Health, 6) Social Science and Humanities, and 7) Vocational and Other.[15] I then measure majors offered within each group using the count of six-digit majors offered within each major group. Table A2 shows the major groups and the maximum number of majors that are possible within each group.

The IPEDS data provide enrollment by state by year, but not applications to each institution by high school resident state and year. Data from the SAT and the ACT admissions tests help fill this gap. When students take the SAT or ACT, they have the option to send their score reports to institutions to supplement their college applications. This provides a proxy for applications.[16] I obtain, from both the College Board and the ACT, a table showing the total score reports sent to each institution in the country, by student state of residence and graduation year. These provide the basis for computing applications by university, student state of residence, and year. To infer applications from score reports, I first combine the information from the two tests, accounting for the fact that some students may send reports from both tests. Then I scale the score-sending data to match total applications by university, which is available in IPEDS. This step accounts for the students who send score reports but subsequently do not apply. The process for inferring applications from score-sending data is described in full in Section C.1.

From the data on SAT and ACT test-takers, I also have the distribution of preferred college major among test-takers, and their average admissions test scores. I use this information to estimate the full distribution among all high school graduates. Details on this process are included in Section C.2.

[14] The classification system was updated in 2010. I use the crosswalk available from the NCES to align data post-2010 with the pre-2010 data, and I excluded majors first introduced in the 2010 classification.

[15] The groups were chosen first to align with the categories of major preference in the College Board and ACT data, and second, to combine areas of study that are similar in content.

[16] Beginning with Card and Krueger (2005), who show a high correlation between test-score sending and applications, the use of SAT and ACT test score-sending as a close approximation for applications has been generally accepted in the higher education literature.

Finally, a sample of high school students from the Educational Longitudinal Study of 2002 (ELS 2002) provides information on the application portfolios, acceptance outcomes, enrollment decisions, and student characteristics for a sample of students who were sophomores in high school in 2002. I use this sample to inform restrictions on the choice sets faced by students in the model (see Section 6.3). In future work, I will also derive some moments from this individual data and use them to supplement estimation.

5 Descriptive Analysis

Three empirical observations motivate and inform the economic model. First, there is variation in majors offered across universities. Second, the total number of majors offered is positively correlated with university net price, size, selectivity, and non-tuition revenue sources, although patterns of association differ by major category. Third, the distribution of students' intended majors varies over time, student state of residence, and by performance on college admissions tests.

Figure B2 shows the number of majors at public and private non-profit universities and the universities' full-time equivalent (FTE) undergraduate enrollment.[17] The graph shows a strong correlation between enrollment and majors offered, but also shows variation among institutions of the same size. The variation in majors offered is fundamental to my research question: students may observe these differences and consider them in their college application and enrollment decisions.

Table 1 shows how the variation in majors offered is correlated with university characteristics, using a regression of the number of majors offered within selected major categories on university characteristics. Fixed effects for the college state, college type (Carnegie Classification – see definition in Section 4), and private control are included but not shown. The findings support some generalizations—institutions that offer more majors rely less on tuition and more on appropriations, charge students more, and are larger and more selective (i.e. have higher average admissions test scores). However, there is some nuance depending on major type.

The extent to which a university relies on tuition as a revenue source is negatively correlated with the total number of majors, although the association is small: for every one percentage point increase in tuition

[17]Throughout the analysis, I will study only public and private non-profit four-year universities that require college admissions tests. For-profit universities typically do not require admission tests and are non-selective, so they typically target a different type of student than four-year private non-profit and public universities do. In the 2015-16 IPEDS data, 53% of four-year for-profit universities reported that they had an open admissions policy. Of the remaining 47%, only 2% reported an admissions test requirement for application, and 21% reported that they recommended an admissions test for application.

Table 1: Majors Offered and University Characteristics

	(1) All Majors	(2) Arts	(3) Bus.	(4) Educ.	(5) Eng., Math & Sci.	(6) Health	(7) Soc. Sci. & Human.	(8) Vocational, Other
Tuition Reliance	-1.943**	0.335**	0.816***	0.297	-2.469***	0.353***	-1.012***	-0.263
	(0.778)	(0.163)	(0.183)	(0.224)	(0.202)	(0.106)	(0.277)	(0.165)
Approp./ Student	0.372***	0.006	0.029*	0.032	0.122***	0.038***	-0.032	0.178***
	(0.069)	(0.015)	(0.016)	(0.020)	(0.018)	(0.009)	(0.025)	(0.015)
Endow./ Student	-8.910	-5.300**	-23.120***	-2.695	-2.443	-5.142***	35.552***	-0.001***
	(10.292)	(2.156)	(2.421)	(2.958)	(2.671)	(1.399)	(3.669)	(0.000)
Net Price	0.469***	0.096***	0.051***	-0.099***	0.128***	0.003	0.311***	-0.021***
	(0.035)	(0.007)	(0.008)	(0.010)	(0.009)	(0.005)	(0.012)	(0.007)
Average SAT	0.060***	0.013***	-0.012***	-0.006*	0.040***	-0.013***	0.061***	-0.023***
	(0.012)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)
UG Enroll. (1,000s)	2.346***	0.312***	0.342***	0.194***	0.427***	0.180***	0.582***	0.308***
	(0.037)	(0.008)	(0.009)	(0.011)	(0.010)	(0.005)	(0.013)	(0.008)
Constant	8.735**	-2.261***	0.447	-0.599	4.791***	0.276	2.546*	3.536***
	(3.757)	(0.787)	(0.884)	(1.080)	(0.975)	(0.511)	(1.339)	(0.798)
N	12,071	12,071	12,071	12,071	12,071	12,071	12,071	12,071
R^2	0.626	0.363	0.429	0.329	0.661	0.375	0.580	0.396

Note: This table presents a university-level linear regression of the number of majors offered within selected major categories on university characteristics using source data from IPEDS from 2006-07 to 2015-16. All dollar amounts are in thousands of 2015 dollars. Fixed effects for private control, Basic Carnegie Classification (see definitions in Section 4) and the state in which the college is located are included but not shown. Standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

as a percent of revenue, the number of majors decreases by -0.01943. The coefficients within major category vary substantially: universities that rely more on tuition offer more Arts, Business, and Health programs, and fewer programs in Engineering, Math and Science, and Social Science and Humanities. State appropriations per student is positively related to the total number of majors: a $1,000 increase in state appropriations per student on average is associated with 0.37 additional majors offered. Endowments per student have no significant effect on the total number of majors, but the effect varies significantly by type of major. More Social Sciences and Humanities majors are offered at institutions with high endowments, while other majors are typically negatively associated with endowments.

Net price (tuition minus average grants and aid) is positively related to majors offered overall and within most fields, perhaps suggesting that students are willing to pay more for schools that offer more majors. A $1,000 increase in net price is associated with 0.47 additional majors. Interestingly, net price is negatively correlated with Education and Vocational majors. This may reflect the fact that careers in Education and Vocational fields are not typically well-compensated, so students may have a lower willingness to pay for that degree.

More selective and larger institutions offer more majors, but while size is positively related to majors offered in every category, selectivity is not. Institutions offering more Business, Education, Health, and Vocational programs are typically less selective, while institutions offering more majors in the Arts, Engineering, Math and Science, and Social Science and Humanities are typically more selective.

The institutional correlations between selectivity and majors offered is consistent with variation in academic ability across students with different major interests—variation that is demonstrated in admissions testing data. In Section 4 and Appendix Section C.2 I detail the data and the process by which I estimate the distribution of intended major for the population of high school graduates in each state and year. The results, shown in Figures B3 and B4, show that the proportion of students intending to major in Education, the Arts, Business, and Social Sciences and Humanities has declined over the decade 2006-2015 in most states, while majors in Engineering, Math and Sciences, Vocational Fields, and Health have increased in popularity over time. Students intending to major in Vocational fields typically have lower admissions test scores than those in any other intended major. Education and Health majors are typically the next lowest-scoring fields, followed by the Arts and Business. Students with intended majors in Engineering, Math and Sciences and the Social Sciences and Humanities are typically the highest performers on admissions tests.

Overall, the facts explored here are consistent with my hypothesis that intended major may drive college choice. That selectivity and net price are positively correlated with the total number of majors offered is consistent with a positive effect of majors offered on student demand. When there are differences by major type, these patterns are consistent with differences in the student population by major. While these correlations are suggestive, estimation of the full model will allow me to determine to the effect of intended major on college choice.

6 Estimation

6.1 Identification

The parameters of the model are the parameters governing the utility of enrollment (α, β_0, $\boldsymbol{\beta_1}$, $\boldsymbol{\beta_2}$, and ρ_k for each k), the variance of the match value between students and universities (σ^2), the admissions thresholds $\underline{s}_{jlt}$ for private universities and $\underline{s}_{jlt}^{is}$ and $\underline{s}_{jlt}^{os}$ for public universities, and the application cost parameters (ω_1 and ω_2).

In section 3 I presented three equations that I match to the data. They are the market penetration function, or, applicants as a proportion of high school graduates (equation 11), total freshman enrollment (14) and the interquartile range of test scores among matriculants (equation 15).

The parameters that govern the mean utility of enrollment (α, β_0, $\boldsymbol{\beta_1}$, $\boldsymbol{\beta_2}$) are identified by the covariation of the market penetration and the university and student characteristics. All else equal, universities with a higher mean utility will receive applications from a greater proportion of high school graduates. Note that all of these parameters are interacted both with the corresponding component of the mean utility of enrollment and the probability of admission, since they are contained in the expected value term. This means that, for example, α does not simply govern the average utility of college attendance. Instead, it also reflects the extent to which students value the probability of admission. As α increases, students will be more likely to apply to every college, but the effect of α varies with admissions probability: high α implies that students will apply more often to "safe" options.

Across-market covariation between applicant shares and demographics identifies the parameters on student characteristics interacted with university characteristics. For example, as the proportion of high school students who intend to major in Business varies across markets, the direction and extent of the variation in market penetration among universities with varying numbers of Business majors identifies the ρ_k for Business.

Given the demand parameters and match value distribution, the admissions thresholds ($\underline{s}_{jlt}$ for private universities and $\underline{s}_{jlt}^{is}$ and $\underline{s}_{jlt}^{os}$ for public universities) are the values that set expected enrollment exactly equal to observed enrollment — in other words, they are the error in the enrollment equation (equation 14). σ^2 is identified in by the interquartile range of observed skill at each university. If the observed skill distribution at each university has a low variance (low interquartile range), this means that the observed skill is an informative component of the admissions criteria. As the interquartile range of skill within university widens, the match value must have a higher variance.

6.2 Instruments

IN PROGRESS.

The endogenous variables that affect demand include net price and majors offered within each category. These may be correlated with the ξ_j, which are unobserved to the econometrician but known by the universities. Thus, I need appropriate instruments for tuition and majors offered. Instruments come from three sources: supply-side factors influencing cost and budget, competitor's exogenous characteristics, and the overlapping structure of the regional markets defined in Section 3.5.

Certain conditions arguably are exogenous and determine the available resources at a university and the cost of offering additional majors in certain categories. These include state appropriations per student, endowment per student, and the number (and size) of graduate programs in each major category. Universities that are land-grant institutions (established by one of the Morril Acts of the 1800's) had a large initial endowment and a foundational mission to provide a wide-ranging array of majors, particularly in Agriculture and Engineering.

The approach used by Berry et al. (1995), who use exogenous characteristics of competitor's products as instruments, also applies naturally to my model. Following this approach, I can use competitor's $\mathbf{X_j}$ as institution-year specific instruments.

In a context with overlapping exogenous regional markets, Fan (2013) uses demographics in competitor's markets as instruments. To illustrate her approach applied to my model, assume that universities A and B are both competing in Virginia, and assume university A competes in New Jersey, Maryland and Virginia, while university B competes only in Virginia. Then instruments for B's tuition and majors offered are demographic characteristics of potential applicants in New Jersey and Maryland. Using this method, I can obtain instruments drawn from the major preferences and the skill distribution of potential applicants.

6.3 Computation and Choice Sets

The computation of the market penetration function is challenging because it involves computing the expected value of application for all possible application sets in each market, which in turn requires computation of the probability of admission and expected value of enrollment for all possible admissions outcomes given each application portfolio. If I was to assume any combination of the approximately 1,400 universities in the country could be considered by each student, the potential number of application portfolios, 2^{1400}, would render the problem impossible to solve.

I have already described one limitation on the problem that makes it substantially more tractable: I limit the options of students depending on where they live. If a student lives in Colorado, empirically, he is highly unlikely to attend James Madison University—a large public university in Virginia. Thus, I rule out application portfolios that include James Madison University for students in Colorado. Section 3.5 describes the empirical method by which I limit the regional market presence of universities. The number of universities plausibly considered by students in each state varies, but the market with the largest number of universities has about 300 universities competing for these students. This is still intractable, at 2^{300} possible portfolios.

Substantial computational savings are also generated by limiting the number of applications that an individual student may send. In the ELS 2002 data, I observe that 82% of students apply to 4 or fewer four-year universities that use the SAT or ACT in their admissions process, so I number of applications a student can send to four. This means that in the market with 300 universities, there are now $\binom{300}{4} + \binom{300}{3} + \binom{300}{2} + 301$ possible combinations: approximately 335 million.

I further limit the problem by making restrictions on the types of universities that are combined in a single application portfolio. Evidence from the individual data suggest that it is highly unlikely that students apply to, for example, three highly selective institutions (like Harvard, Yale and Stanford) and a non-selective university that accepts nearly all applicants. To establish these rules, I examine patterns of selectivity in the ELS 2002 data using Barron's rankings from the 2004 *Profiles of American Colleges*. I characterize application portfolios by the number of universities in each of six selectivity categories. I find that 95% of students' application portfolios fall in the top 74 patterns of selectivity categories, so I limit the possible portfolios to follow these patterns. The resulting set of portfolios is still very large, so I utilize parallel computing in Matlab to evaluate the market penetration function at each iteration of the optimization procedure.

6.4 Estimation Method

IN PROGRESS.

The parameters of the model are estimated using the Generalized Method of Moments. As in Berry et al. (1995) and Fan (2013), a key challenge is that the error in the market penetration function enters non-linearly in a way that makes analytic inversion of the market penetration function infeasible. I explore two ways of dealing with this problem. The first is the nested fixed point algorithm proposed by Berry et al. (1995), which is useable with my model despite the differences between their market share and my market penetration function. A second method, proposed by Lee and Seo (2015), speeds up the computation using a Taylor approximation to the market penetration function. In the author's simulations, this method achieved results very similar to those found using the nested fixed point algorithm, provided the number of markets was sufficiently large.

7 Results

IN PROGRESS.

8 Conclusion

IN PROGRESS.

References

Arcidiacono, P. (2005). Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings? *Econometrica 73*(5), 1477–1524.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica 63*(4), 841–890.

Bettinger, E. P. and B. T. Long (2010). Does Cheaper Mean Better? The Impact of Using Adjunct Instructors on Student Outcomes. *The Review of Economics and Statistics 92*(August), 598–613.

Bound, J., B. Braga, G. Khanna, and S. Turner (2016). A Passage to America: University Funding and International Students. *NBER Working Paper*.

Bresnahan, T. F. ., S. Stern, and M. Trajtenberg (1997). Market Segmentation and the Sources of Rents from Innovation : Personal Computers in the Late 1980s. *The RAND Journal of Economics 28*(0).

Card, D. and A. B. Krueger (2005). Would the elimination of affirmative action affect highly qualified minority applicants? Evidence from California and Texas. *Industrial and Labor Relations Review 58*(3), 416–434.

Chade, H., G. Lewis, and L. Smith (2014). Student portfolios and the college admissions problem. *Review of Economic Studies 81*(3), 971–1002.

Cook, E. E. (2019). Mandatory Testing and the Pool of College Applicants.

Epple, D., R. Romano, and H. Sieg (2006). Admission, Tuition, and Financial Aid Policies in the Market for Higher Education. *Econometrica 74*(4), 885–928.

Fan, Y. (2013). Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market. *American Economic Review 103*(5), 1598–1628.

Fu, C. (2014). Equilibrium Tuition, Applications, Admissions, and Enrollment in the College Market. *Journal of Political Economy 122*(2), 225–281.

Gentzkow, M. (2007). Valuing New Goods in a Model with Complementarity: Online Newspapers. *American Economic Review 97*(3), 713–744.

Goodman, S. (2016). Learning from the Test: Raising Selective College Enrollment by Providing Information. *The Review of Economics and Statistics 98*(October), 671–684.

Herzog, K. (2018, nov). UW-Stevens Point rolls out transformation that would cut 6 liberal arts degrees , focus on careers.

Howell, J. S. (2010). Assessing the Impact of Eliminating Affirmative Action in Higher Education. *Journal of Labor Economics 28*(1), 113–166.

Hoxby, C. M. (1997). How the Changing Market Structure of U.S. Higher Education Explains College Tuition.

Hurwitz, M., J. Smith, S. Niu, and J. Howell (2015). The Maine Question: How Is 4-Year College Enrollment Affected by Mandatory College Entrance Exams ? *Educational Evaluation and Policy Analysis 37*(1), 138–159.

Hyman, J. (2017). ACT for All: The Effect of Mandatory College Entrance Exams on Postsecondary Attainment and Choice. *Education Finance and Policy 12*(3), 281–311.

Kapor, A. (2016). Distributional Effects of Race-Blind Affirmative Action. *Working Paper*.

Klasik, D. (2013). The ACT of Enrollment: The College Enrollment Effects of State-Required College Entrance Exam Testing. *Educational Researcher 42*(April), 151–160.

Lee, J. and K. Seo (2015). A computationally fast estimator for random coefficients logit demand models using aggregate data. *RAND Journal of Economics 46*(1), 86–102.

McCafferty, R. A. (2018, aug). University of Akron to cut almost 20% of degrees, degree tracks.

Pallais, A. (2015). Small Differences That Matter: Mistakes In Applying To College. *Journal of Labor Economics 33*(2), 493–520.

Stancill, J. (2015, may). Dozens of Degree Programs Phased Out in UNC System.

Strauss, V. (2017, mar). A University of Wisconsin campus pushes plan to drop 13 majors – including English, history and philosophy.

Tkacik, C. (2018, aug). Maryland's Goucher College is eliminating several majors, including math.

University of Southern Maine (2015). Discontinued Programs.

U.S. Department of Education, N. (2016). Table 219.10. High school graduates, by sex and control of school: Selected years, 1869-70 through 2026-27.

Appendix A Tables

Table A1: CIP Major Examples

Major Category	CIP Code and Description
Social Sci. and Humanities	54.0101 History, General.
Social Sci. and Humanities	54.0102 American History (United States).
Social Sci. and Humanities	54.0103 European History.
Social Sci. and Humanities	54.0104 History and Philosophy of Science and Technology.
Social Sci. and Humanities	54.0105 Public/Applied History.
Social Sci. and Humanities	54.0106 Asian History.
Social Sci. and Humanities	54.0107 Canadian History.
Social Sci. and Humanities	54.0199 History, Other.

Note: The table provides examples of majors that can be offered within social sciences and history, to demonstrate the level of detail in the major classification. This is not an exhaustive list of majors.

Table A2: Majors

Major Category	Number of Majors
Arts and Architecture	61
Business and Communications	119
Education	94
Eng., Math, and Sciences	261
Health	210
Social Science and Humanities	258
Vocational, Other	294

Note: The table lists the major categories used in analysis, and the total number of majors that fall into these categories. A university may offer any number of majors up to the total number of majors available in each category.

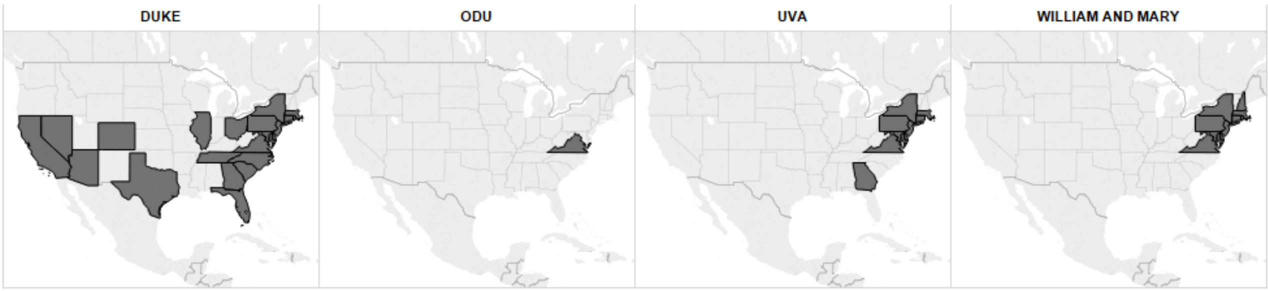
Table A3: Summary of Universities

Year	Public Universities					Private Universities				
	Count	Tuition (In-State)	Avg Grants & Aid	Avg # Majors	St. Dev. Majors	Count	Tuition	Avg Grants & Aid	Avg # Majors	St. Dev. Majors
2006-07	525	4,948	3,664	50.01	28.46	918	23,153	11,480	36.55	20.08
2007-08	526	5,061	4,036	51.07	28.95	918	23,904	55,240	37.47	20.45
2008-09	526	5,170	4,352	52.06	29.29	916	24,340	21,982	38.50	21.01
2009-10	526	5,593	5,073	52.10	29.01	913	25,549	16,589	38.83	20.67
2010-11	536	5,849	5,438	51.51	29.01	947	26,181	15,405	37.98	20.89
2011-12	537	6,153	5,311	51.85	28.97	941	26,559	15,246	38.76	21.03
2012-13	537	6,330	5,380	52.57	29.45	935	27,158	15,781	39.47	21.37
2013-14	533	6,473	5,588	53.36	29.96	935	27,752	16,392	40.01	21.61
2014-15	490	6,685	5,939	55.48	30.26	857	29,486	17,757	42.28	21.39
2015-16	489	6,904	6,078	55.86	30.21	853	30,451	18,529	42.40	21.57

Note: The table summarizes several key characteristics of 4-year degree-granting private non-profit or public institutions of higher education in the 50 United States and D.C. that require the SAT or ACT for admissions and that are not listed as a "Special Focus Institution" or "Tribal College" based on the Carnegie Classification of Institutions' Basic Classification. Average grants and aid is calculated as the average grants and aid received by all students who were enrolled at the university, including those who received no grants and aid. Grants and aid includes all federal, state, local, and institutional funds that are not required to be repaid by students. In the model and estimation, tuition for public universities will be state-residence specific, although only in-state tuition is summarized here. All monetary values are in 2015 $.

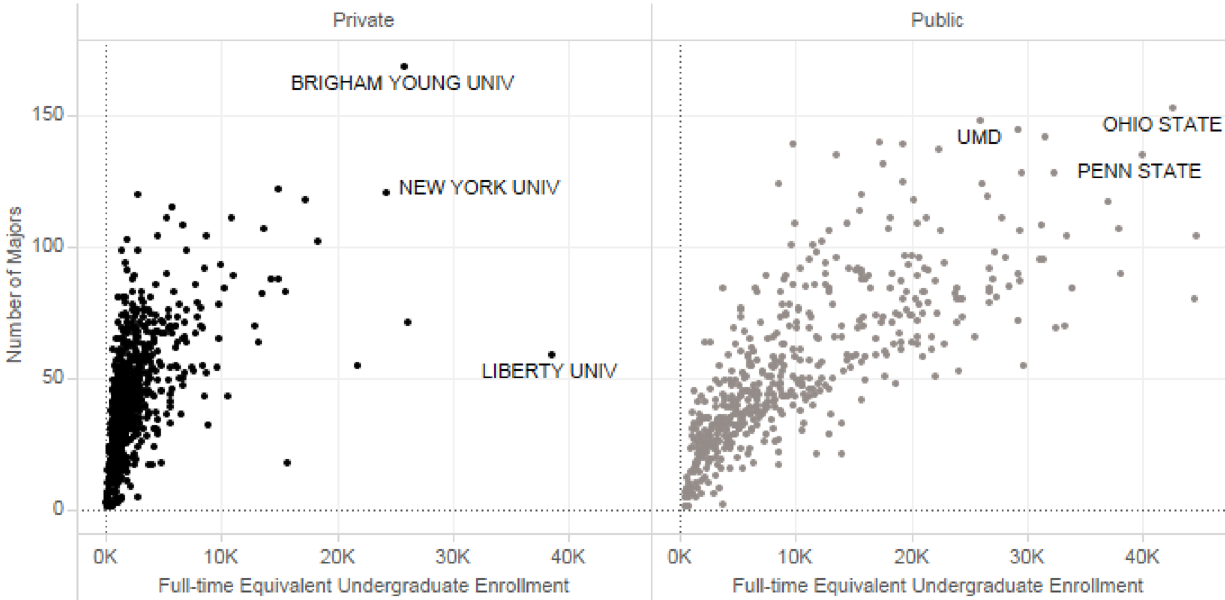
Appendix B Figures

Figure B1: University Market Presence



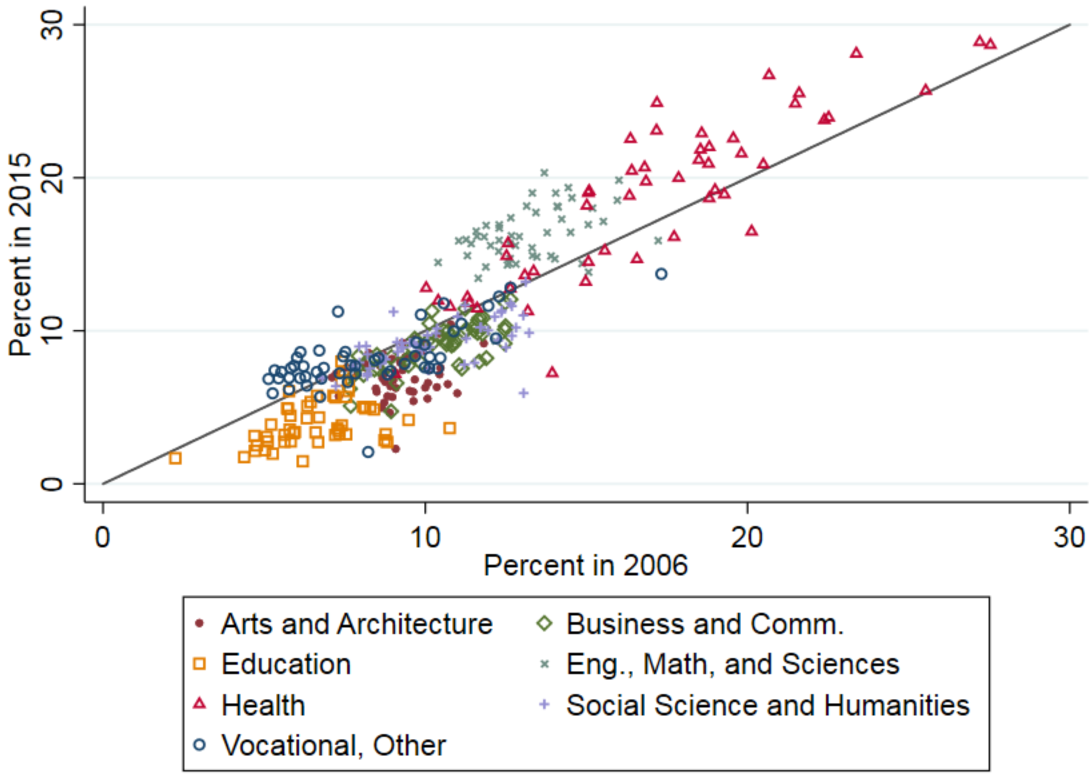
Note: This figure shows the states that each institution competes in (market presence) for applications and enrollment. To define market presence, states are ranked (within university) by the proportion of the state's high school graduates that attend that university. Then starting from the top of the list, states are included in the university's market until the included states make up at least 85% of the university's enrollment.

Figure B2: Number of Majors versus Enrollment



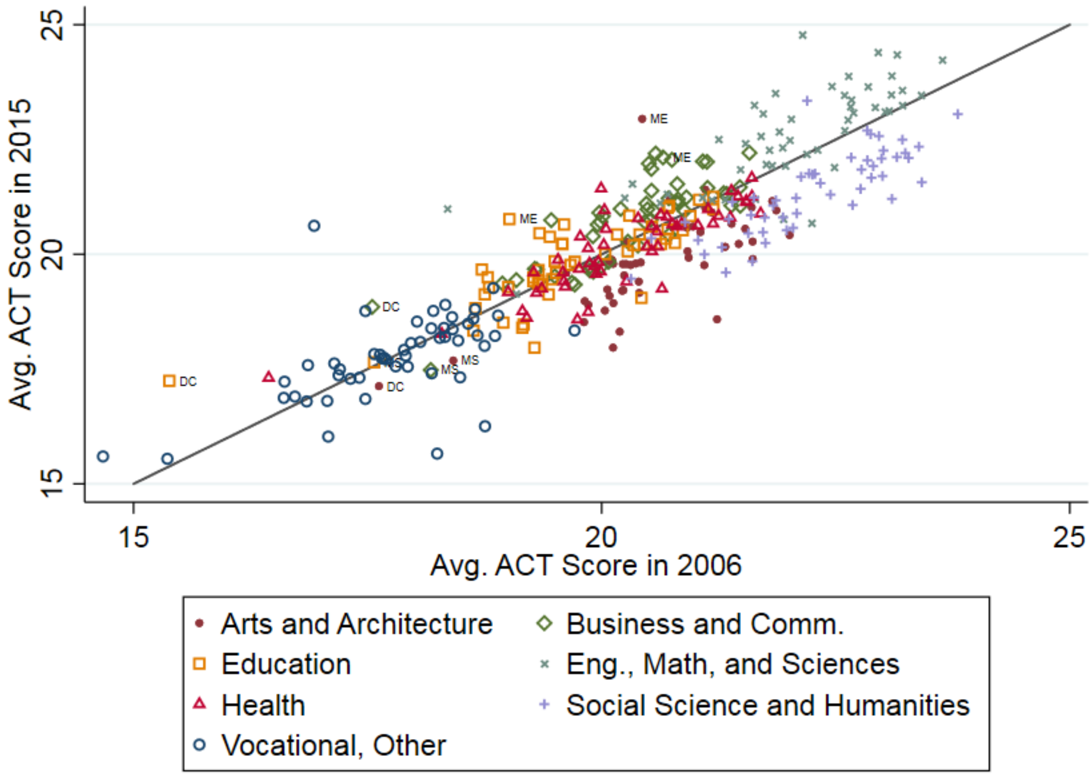
Note: This graph shows the availability of majors among private and public universities on the vertical axis and the total undergraduate enrollment on the horizontal axis. Public universities are in gray and private universities in black. This was generated using the 2015 IPEDS data.

Figure B3: Shares by State and Major over Time



Note: This graph shows the percent of students reporting each intended major by state in 2006 and 2015.

Figure B4: Scores by State and Major over Time



Note: This graph shows average admissions test scores (on the ACT score scale) by state and intended major in 2006 and 2015.

Appendix C Data Processing Details

C.1 Inferring Applications from Score Reports and IPEDS Data

The NCES' IPEDS database discussed in Section 4 provides applications, admissions, and enrollment to each institution in each year. IPEDS does provide a breakdown of enrollment by student state of residence, but applications and admissions numbers are not available by student state of residence. While many of the parameters in my model are identified using variation in application shares, institution characteristics and student demographics at the level of institution and/or year, incorporating the cross-state variation allows for stronger identification of parameters on student demographics. Further, modeling the problem at the state-level allows me to include state-level differences in pricing, admissions testing policies, and admissions thresholds set by universities.

To split the applications reported in IPEDS by state, I use admissions testing score report data from the ACT and the College Board (the creators of the SAT). This data provides the number of score reports sent to each university from high school residents in each state. Score reports are required for admission at the universities that I study, so the score report data helps me to approximate applications by state and year. However, not every score report sent is necessarily attached to an application, because students can choose not to follow-up with an application after having sent a score report.

When students take the SAT or ACT, they are entitled to four free score reports if they list the recipient institutions prior to a deadline that occurs several days after the testing date. To send additional reports and/or send reports after the deadline, students must pay a fee (around $12) or obtain a fee waiver available to low-income students.[18] If a student chooses to apply to a university, any score reports sent by the ACT or College Board on behalf of the student are matched with the student's application by the admissions office.

Sending a score report to a university is an indication of serious interest in applying to the university. Free score reports are not unlimited, and the cost of sending a report encourages students to pick a small number of schools to receive reports. As a result, students typically do apply to the schools to which they send score reports; at the institution/year level, the correlation coefficient between applications and score reports is 0.88.

Despite the fact that score-sending is a strong signal of interest in a school, there is a non-negligible

[18] See https://collegereadiness.collegeboard.org/sat/register/fees and https://www.act.org/content/act/en/products-and-services/the-act/scores/sending-your-scores.html.

amount of score-sending that is not accompanied by an application Pallais (2015). Thus, score report data cannot be used directly as applications without some adjustment. This section describes the process by which I approximate applications from the SAT and ACT score-report data. The process has three steps: first, I combine the ACT and SAT data; second, I compute a weight on the score reports from universal testing states to account for the impact of universal testing policies; and third, I adjust total score reports by institution to match the total applications reported in IPEDS.

To combine the data from the ACT and the SAT, I merged the two sources by institution, student state of residence, and year. I then merged the IPEDS data at the institution level to the testing data. The data do not contain identifiers that allow direct merges between sources, so I utilized a mapping between ACT numbers, College Board Codes, and the IPEDS unitid, which I obtained from Wintergreen Orchard House, a Division of Carnegie Dartlet. In addition, some matching on institution name was required. Once I merged these data, I added the ACT and SAT reports to create the total number of reports sent by student state and year to each institution. As this stage, I made no adjustment for the possibility that a single student could send score reports from both the SAT and the ACT, as this is accounted for later.

When a mandatory testing policy is implemented, total score-sending increases on average by approximately 30% Cook (2019). Further, it is plausible that a higher proportion of the score reports induced by the policy are unlikely to be accompanied by an application, as these score reports are being sent by marginal applicants—students who would not apply to a four-year college that requires an admissions test in absence of the testing policy. I account for this by running the following regression at the institution level, where $reports\_nmand$ is the total number of score reports from states without mandatory testing at year t and $reports\_mand$ is the total number from states with mandatory testing:

$$applications_{jt} = a\ reports\_nmand_{jt} + b\ reports\_mand_{jt} \tag{18}$$

This yields coefficient estimates of 0.852 and 0.793 for a and b, respectively. This indicates that score reports do convert to applications at a lower rate in mandatory testing states, as expected. To account for the differences in score-sending and application behavior across states with different testing policies, I multiply score reports in mandatory testing states by b/a, or 0.93.

The last step is to scale the score reports in each state to match the total number of applications to j in time t. By university, I estimate applications from each state and year by scaling the number of score reports to the university from each state and year by $applications_{jt}/total\_reports_{jt}$, where $total\_reports_{jt}$ is the sum of score reports sent to the institution after the adjustment for mandatory testing. The result gives the estimate of applications by state and year. Dividing the estimated applications by the number of high school graduations in each state gives the market penetration, which I match to equation 11.

C.2 Estimating the Joint Distribution of Scores and Major Preference

Both the ACT and the College Board provide publicly available data on the number of students and their average score by planned major, by state and year. The planned major is collected from students when they register for the ACT and SAT tests, so this raw data is for test-takers only. In my estimation I require the underlying joint distribution of major preference and score for *all* high school graduates. This section describes how I estimate the population joint distribution of scores and major preference. The process has three steps: first, I combine the SAT and ACT data; second, I estimate the population proportion of students with each intended major and their average scores; and third, I generate a normal distribution of scores within each intended major category, assuming the standard deviation of scores does not differ across major categories.

The first step is to combine the data from the two tests (ACT and SAT). While the majority of students only take one of the two tests, some students take both. This complicates the averaging across the data from each test, as I need to account for the proportion of the student population that is counted in both data sources. Out of necessity, I assume that the percent of students who take both tests ("dual-testing rate") is independent of major preference.[19] I compute the dual-testing rate by state from the Education Longitudinal Study of 2002 (ELS 2002), and assume that the rate within state is constant across my time period.[20] Once these rates r_l have been computed, I compute the number of students reporting each major k as the sum of the reported count for each test (n_{klt}^{SAT}, n_{klt}^{ACT}), weighted to reflect dual-testing:

$$n_{klt} = n_{klt}^{SAT} + n_{klt}^{ACT}(1 - r_l). \tag{19}$$

When computing average scores by major I account for the fact that scores may be related to the decision to take both tests. From ELS 2002, I compute the average scores of single-test-takers, $\bar{ACT}_{lt}^{single}$

[19]The individual data do not include a major preference reported prior to matriculation.

[20]Students may also take either the ACT or SAT multiple times, but this is not a concern because the ACT and SAT data report each student only once.

and take the ratio of this average to the overall average $\bar{ACT}_{lt}^{overall}$. This, combined with the assumption that dual-testing is independent of planned major, allows me to find the single-test-takers score average within major. To summarize, I use the following formula for the average score by major across the two tests, where ACT and SAT scores have been aligned on the same scale:

$$\bar{score}_{klt} = \frac{1}{n_{klt}} n_{klt}^{SAT} \bar{SAT}_{klt} + n_{klt}^{ACT}(1 - r_l) \frac{\bar{ACT}_{lt}^{single}}{\bar{ACT}_{lt}^{overall}} \bar{ACT}_{klt}. \tag{20}$$

Table C1 shows the percent of students who indicate each major preference by year (averaged across states) and the standard deviation of these percentages across states. From this table it is immediately apparent that the percent of students who take the test but do not report a preferred major is significantly higher among the 2006-2009 graduating classes than among the graduating classes in years following. This is likely a result of a shift from paper registrations to online registrations during this period. In order to have comparable distributions of preference over time, I must account for the changes in the reporting pattern. In addition, the distribution in Table C1 is still conditional on taking one of the two admissions tests, so there is a group of non-tested students whose major preferences and potential scores are not included.

To estimate the population distribution of intended major and mean scores by intended major, I take advantage of two sources of exogenous variation. First, while is unlikely that non-response is uncorrelated with underlying major preference and score, the shift in non-response between years 2006-2009 and 2010 onward is a result of exogenous changes in the survey format. Second, the existence of universal admissions testing policies that vary over state and year provide a source of exogenous variation in the population of testers — the total joint distribution of major preference and score is observed in 80 state/year combinations. The desired counterfactual distribution is the one that would be observed if reporting followed the pattern in years 2010 and later, and universal testing policies were implemented in all states and years.

To compute the counterfactual distribution across majors, I first estimate a multinomial logit model where the proportion of students in each major group within each state and year is a function of state fixed effects, the existence of a universal testing policy, a linear state-specific time trend, year fixed effects (where I constrain years 2010 and onward to have identical fixed effects), and a state/year/major error. I include "No Response" as a separate intended major group, and use "Undecided" as the base category.

Thus the proportion of students in a given major category k is modeled as:

$$p_{klt} = \frac{e^{(a_{k1}+b_{k1}mand_{lt}+b_{kl2}t+\tau_{1kt}+\varsigma_{1kl}+e_{1klt})}}{1+\sum_{k\neq 0} e^{(a_{k1}+b_{k1}mand_{lt}+b_{kl2}t+\tau_{1kt}+\varsigma_{1kl}+e_{1klt})}}, \tag{21}$$

where $\delta_{2010} = \delta_{2011} = \delta_{2012} = \delta_{2013} = \delta_{2014} = \delta_{2015} = 0$, so that these years form the excluded category.

Taking the log of the ratio of the proportion in major k, p_{klt}, to the proportion in the base category ("Undecided"), p_{0lt}, leads to the following OLS model:

$$log(p_{klt}) - log(p_{0lt}) = a_{k1} + b_{k1}mand_{lt} + b_{kl2}t + \tau_{1kt} + \varsigma_{1kl} + e_{1klt}. \tag{22}$$

I run this regression on each major category separately using all 510 states and years to estimate the coefficients. I then predict the shares for the entire population by setting $policy_{lt} = 1$ for all states and setting $\delta_{2006} = \delta_{2007} = \delta_{2008} = \delta_{2009} = 0$. The resulting prediction yields some non-response, but the patterns of non-response are now consistent over time. I combine the remaining non-responders with the undecided students for the purpose of estimating the college choice model.

To estimate average scores by major for the entire population, I run the following OLS regression using all 510 state and year observations and all major categories, including non-response:

$$sc\bar{o}re_{klt} = a_{k2} + b_{k4}mand_{lt} + b_{kl5}t + \tau_{2kt} + \varsigma_{2kl} + e_{2klt}. \tag{23}$$

The regression results for the proportion in each major and average scores are included in Tables C3 and C4, and the predicted shares and average scores are summarized in Tables C5 and C6. The predictions are obtained by setting setting $mand_{lt} = 1$ for all states and setting $\delta_{2006} = \delta_{2007} = \delta_{2008} = \delta_{2009} = 0$.

Variation in the share of intended major across states and time is critical for identification of the effect of major preference on college choice, while variation in mean scores by intended major aids in identification of the effect of admissions probability on application portfolio choice. The correlation between mean scores and intended major provides a demand-side explanation for some of the patterns I provided in Section 5, for example, that the number of majors offered in certain major categories (such as Engineering, Math and Sciences) is positively correlated with the average scores of the student body, while the number of Education majors is negatively correlated with average scores.

Table C1: Average Percentage of Students by Intended Major

Year	Statistic	Arts and Architecture	Business and Comm.	Education	Eng., Math & Sciences	Health	Social Sci. & Humanities	Vocational, Other	Undecided	Unreported
2006	Mean	7.45	11.30	5.76	13.13	14.64	8.07	5.31	6.26	28.09
	SD	0.71	1.43	1.30	1.36	4.28	0.76	1.62	3.46	7.02
2007	Mean	6.88	10.45	5.28	12.28	14.11	7.86	4.73	8.16	30.25
	SD	0.82	1.83	1.28	1.61	3.81	1.21	1.48	4.05	7.28
2008	Mean	7.27	12.43	5.97	13.93	15.99	9.23	5.41	9.50	20.28
	SD	1.04	1.70	1.44	1.92	4.52	1.78	1.91	4.78	10.21
2009	Mean	7.88	12.72	6.16	15.55	17.48	9.78	6.08	9.70	14.65
	SD	1.10	1.83	1.49	1.93	4.70	1.96	2.22	4.19	9.80
2010	Mean	8.09	11.94	6.33	16.78	18.73	10.89	6.53	11.46	9.25
	SD	0.97	1.98	1.53	1.93	4.51	2.97	2.14	3.90	6.55
2011	Mean	7.89	10.86	5.93	17.08	19.49	11.14	6.51	11.80	9.30
	SD	1.11	1.84	1.59	1.93	4.77	2.77	2.16	3.58	7.53
2012	Mean	8.00	10.69	5.42	17.06	20.29	11.64	6.48	11.79	8.65
	SD	1.26	2.11	1.46	2.10	4.88	2.37	1.80	3.82	7.73
2013	Mean	7.68	10.53	5.04	17.37	19.90	10.96	6.77	11.54	10.23
	SD	1.33	2.25	1.43	2.31	4.78	2.22	2.02	3.83	7.82
2014	Mean	7.21	10.65	4.87	18.24	20.02	10.64	6.85	11.46	10.05
	SD	1.14	2.31	1.44	2.64	4.64	2.03	2.13	3.99	7.94
2015	Mean	7.05	10.71	4.67	18.61	19.40	9.98	6.77	11.36	11.45
	SD	1.21	2.42	1.43	2.88	4.47	2.03	2.13	3.95	8.85

Note: This table shows, by year, the state-level average and standard deviation of the percentage of students reporting each intended major upon registration for the SAT and ACT. The data is taken from annual state reports generated by the College Board and the ACT, then combined as described in the text.

Table C2: Average Scores by Preferred Major

Year	Statistic	Arts and Architecture	Business and Comm.	Education	Eng., Math & Sciences	Health	Social Sci. & Humanities	Vocational, Other	Undecided	Unreported
2006	Mean	21.35	20.86	20.32	23.20	20.65	22.94	19.08	21.52	21.48
	SD	0.82	0.76	0.99	0.93	0.90	0.75	0.91	0.89	0.72
2007	Mean	21.43	20.86	20.31	23.24	20.67	22.67	19.05	21.28	21.50
	SD	0.83	0.88	1.05	1.02	1.03	0.75	0.92	0.90	0.88
2008	Mean	21.38	20.83	20.22	23.18	20.79	22.51	18.99	21.83	21.43
	SD	0.85	0.82	0.93	1.10	0.94	0.84	0.92	1.00	1.15
2009	Mean	21.35	20.95	20.30	23.24	20.90	22.49	19.00	21.96	20.47
	SD	0.94	0.85	0.94	1.12	0.98	0.88	1.01	1.24	1.28
2010	Mean	21.34	21.04	20.38	23.32	20.88	22.50	18.94	22.06	19.11
	SD	0.91	0.91	0.87	1.13	0.91	0.85	0.98	1.19	1.23
2011	Mean	21.14	21.05	20.44	23.26	20.81	22.19	18.92	22.05	19.11
	SD	0.95	0.92	0.92	1.16	0.91	0.87	0.90	1.13	1.31
2012	Mean	21.11	21.02	20.43	23.32	20.68	21.95	18.92	22.11	18.78
	SD	0.93	0.96	0.85	1.11	0.97	0.86	0.96	1.16	1.35
2013	Mean	20.76	20.99	20.32	23.28	20.63	21.89	18.76	22.00	18.68
	SD	1.01	0.91	0.85	1.09	0.88	0.85	1.01	1.23	1.44
2014	Mean	20.79	21.10	20.33	23.37	20.68	21.86	18.71	22.02	18.28
	SD	1.06	0.97	0.90	1.16	0.91	0.85	1.04	1.25	1.35
2015	Mean	20.77	21.17	20.37	23.41	20.68	21.88	18.75	22.06	18.17
	SD	1.13	0.96	0.90	1.15	0.92	0.82	1.03	1.21	1.34

Note: This table shows, by year, the state-level average and standard deviation of the mean scores among students reporting each intended major upon registration for the SAT and ACT. The data is taken from annual state-level reports generated by the College Board and the ACT, then combined as described in the text.

Table C3: Multinomial Logit Regression: Distribution across Majors

	(1) Arts & Architecture	(2) Business & Comm.	(3) Education	(4) Eng., Math & Sciences	(5) Health	(6) Social Sci. & Humanities	(7) Vocational, Other	(8) No Response
Universal Testing	0.406***	0.239***	0.219***	0.256***	0.311***	0.266***	0.701***	1.325***
	(0.036)	(0.037)	(0.039)	(0.034)	(0.032)	(0.028)	(0.043)	(0.087)
Year==2006	0.537***	0.677***	0.426***	0.601***	0.497***	0.384***	0.587***	2.250***
	(0.029)	(0.030)	(0.032)	(0.028)	(0.026)	(0.023)	(0.035)	(0.071)
Year==2007	0.187***	0.309***	0.097***	0.210***	0.159***	0.064***	0.164***	2.001***
	(0.026)	(0.027)	(0.028)	(0.025)	(0.023)	(0.020)	(0.032)	(0.064)
Year==2008	0.109***	0.347***	0.121***	0.153***	0.110***	0.070***	0.115***	1.319***
	(0.024)	(0.025)	(0.025)	(0.022)	(0.021)	(0.018)	(0.028)	(0.057)
Year==2009	0.156***	0.326***	0.156***	0.183***	0.131***	0.079***	0.154***	0.739***
	(0.021)	(0.022)	(0.023)	(0.020)	(0.019)	(0.016)	(0.025)	(0.051)
Observations	510	510	510	510	510	510	510	510
R-squared	0.96025	0.96629	0.95726	0.95784	0.95311	0.98154	0.92477	0.96212

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: This table shows the results of the multinomial logit regressions described in the text, where the dependent variable is the proportion of students reporting each intended major.

Table C4: OLS Regression: Scores by Major

	(1) Arts & Architecture	(2) Business & Comm.	(3) Education	(4) Eng., Math & Sciences	(5) Health	(6) Social Sci. & Humanities	(7) Vocational, Other	(8) No Response
Universal Testing	-1.190***	-0.598***	-0.792***	-1.166***	-0.758***	-0.566***	-1.342***	-2.526***
	(0.089)	(0.076)	(0.074)	(0.085)	(0.073)	(0.092)	(0.092)	(0.177)
Year==2006	-0.361***	-0.023	-0.103*	0.063	-0.306***	0.154**	-0.011	1.563***
	(0.073)	(0.062)	(0.061)	(0.069)	(0.059)	(0.075)	(0.075)	(0.145)
Year==2007	-0.178***	-0.052	-0.111**	0.071	-0.257***	-0.000	-0.007	1.750***
	(0.065)	(0.056)	(0.055)	(0.062)	(0.053)	(0.068)	(0.067)	(0.130)
Year==2008	-0.123**	-0.117**	-0.202***	-0.028	-0.109**	-0.051	-0.041	1.851***
	(0.059)	(0.050)	(0.049)	(0.056)	(0.048)	(0.061)	(0.060)	(0.116)
Year==2009	-0.028	-0.015	-0.107**	0.014	0.051	0.049	0.021	1.113***
	(0.052)	(0.045)	(0.044)	(0.050)	(0.043)	(0.054)	(0.054)	(0.104)
Observations	510	510	510	510	510	510	510	510
R-squared	0.93753	0.94539	0.95058	0.95478	0.95440	0.92219	0.93272	0.92552

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: This table shows the results of the OLS regressions described in the text, where the dependent variable is the average score among students reporting each intended major. State-specific fixed effects and time trends are included but not shown.

Table C5: Average Predicted Share of Students by Preferred Major

Year	Statistic	Arts and Architecture	Business and Comm.	Education	Eng., Math & Sciences	Health	Social Sci. & Humanities	Vocational, Other	Undecided	Unreported
2006	Mean	9.22	10.37	6.65	13.20	17.03	10.29	8.29	8.70	16.23
	SD	0.99	1.45	1.49	1.45	4.31	1.83	2.48	4.40	7.20
2007	Mean	8.97	10.25	6.29	13.56	17.26	10.21	8.30	8.66	16.49
	SD	0.93	1.43	1.37	1.41	4.36	1.75	2.36	4.17	7.32
2008	Mean	8.72	10.12	5.94	13.92	17.48	10.12	8.30	8.62	16.77
	SD	0.92	1.42	1.30	1.37	4.44	1.68	2.26	3.94	7.53
2009	Mean	8.47	9.99	5.62	14.27	17.69	10.01	8.30	8.59	17.07
	SD	0.94	1.42	1.27	1.36	4.53	1.61	2.18	3.72	7.83
2010	Mean	8.22	9.84	5.31	14.62	17.88	9.90	8.29	8.56	17.39
	SD	1.00	1.42	1.28	1.37	4.64	1.56	2.11	3.52	8.21
2011	Mean	7.98	9.68	5.02	14.95	18.06	9.77	8.28	8.52	17.72
	SD	1.07	1.44	1.31	1.41	4.77	1.52	2.05	3.32	8.68
2012	Mean	7.74	9.52	4.75	15.29	18.23	9.64	8.26	8.50	18.07
	SD	1.16	1.47	1.36	1.49	4.92	1.49	2.02	3.14	9.22
2013	Mean	7.50	9.35	4.50	15.61	18.39	9.51	8.25	8.47	18.43
	SD	1.25	1.51	1.41	1.60	5.08	1.49	2.00	2.99	9.81
2014	Mean	7.27	9.18	4.26	15.92	18.53	9.36	8.23	8.45	18.80
	SD	1.34	1.55	1.47	1.75	5.26	1.50	1.99	2.86	10.46
2015	Mean	7.04	9.00	4.03	16.23	18.66	9.22	8.20	8.44	19.18
	SD	1.43	1.60	1.53	1.93	5.46	1.52	2.00	2.76	11.13

Note: This table summarizes the predictions from the multinomial logit regressions in Table C3. The predictions are summarized across states, and the means and standard deviations are reported. State-specific fixed effects and time trends are included but not shown.

Table C6: Average Predicted Scores by Preferred Major

Year	Statistic	Arts and Architecture	Business and Comm.	Education	Eng., Math & Sciences	Health	Social Sci. & Humanities	Vocational, Other	Undecided	Unreported
2006	Mean	20.57	20.30	19.67	22.01	20.23	22.24	17.80	21.08	17.49
	SD	0.81	0.82	1.01	1.04	0.97	0.80	0.90	0.93	0.98
2007	Mean	20.49	20.35	19.68	22.07	20.21	22.14	17.80	21.11	17.37
	SD	0.81	0.82	0.99	1.03	0.95	0.80	0.86	0.93	0.92
2008	Mean	20.41	20.39	19.69	22.14	20.20	22.04	17.80	21.14	17.25
	SD	0.81	0.83	0.96	1.02	0.93	0.80	0.84	0.94	0.89
2009	Mean	20.33	20.44	19.71	22.20	20.18	21.94	17.79	21.17	17.13
	SD	0.82	0.84	0.94	1.02	0.92	0.80	0.82	0.97	0.88
2010	Mean	20.25	20.48	19.72	22.26	20.17	21.85	17.79	21.20	17.00
	SD	0.84	0.85	0.93	1.02	0.92	0.80	0.82	1.00	0.91
2011	Mean	20.17	20.52	19.73	22.32	20.15	21.75	17.78	21.23	16.88
	SD	0.86	0.88	0.92	1.03	0.91	0.82	0.82	1.04	0.96
2012	Mean	20.09	20.57	19.75	22.38	20.14	21.65	17.78	21.26	16.76
	SD	0.90	0.90	0.91	1.05	0.91	0.83	0.84	1.09	1.03
2013	Mean	20.01	20.61	19.76	22.44	20.12	21.55	17.78	21.29	16.64
	SD	0.94	0.93	0.91	1.08	0.91	0.85	0.87	1.14	1.12
2014	Mean	19.93	20.66	19.78	22.50	20.11	21.45	17.77	21.32	16.52
	SD	0.98	0.96	0.92	1.11	0.92	0.87	0.91	1.20	1.22
2015	Mean	19.85	20.70	19.79	22.56	20.09	21.35	17.77	21.34	16.40
	SD	1.03	1.00	0.93	1.15	0.93	0.90	0.96	1.26	1.34

Note: This table summarizes the predictions from the OLS regressions in Table C4. The predictions are summarized across states, and the means and standard deviations are reported.