Cook, Emily E.

Working Paper
Market Structure and College Access in the US

CESifo Working Paper, No. 11733

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

CESifo WORKING PAPERS

11733
2025

March 2025

Market Structure and College Access in the US

Emily E. Cook

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 11733*

Market Structure and College Access in the US

Abstract

State governments provide grants to students to subsidize college attendance. In response, colleges can adjust their tuition, aid policies, and admission standards, affecting equilibrium enrollment and pass-through of aid to students. To quantify demand- and supply-side responses, I develop a model that incorporates the geographic nature of the market and strategic competition between individual colleges. Simulations demonstrate that college responses are meaningful: when students receive $1,000 to attend in-state public colleges, these colleges absorb over 40% of the subsidy on average and raise admissions standards, reducing the enrollment effect of the policy. Close competitors see enrollment declines of 2-3%.

JEL-Codes: I230, I280, L130, L310.

Keywords: college choice, college admission, college enrollment, financial aid, mixed oligopoly, non-profit firms, demand estimation.

Emily E. Cook
Department of Economics
Texas A&M University, College Station / TX / USA
ecook4@tamu.edu

February 28, 2025
I thank Sarah Turner, Federico Ciliberto, and Gaurab Aryal for their invaluable advice on my dissertation, which laid the foundation for this paper. I also thank Flavio Cunha, Michael Dinerstein, Ian Fillmore, Jeremy Fox, Chao Fu, Doug Harris, Jean-Francois Houde, Adam Kapor, Ekaterina Khmelnitskaya, Lance Lochner, Chris Neilson, Jeff Smith, Kevin Stange, Todd Stinebrickner, Chris Taber, and Basit Zafar for helpful conversations. This work benefited from input from participants at many conferences and seminars. Ed Hall and the Advanced Research Computing Services at the University of Virginia, Brian Streicher and Technology Services at Tulane University, and Texas A&M High Performance Research Computing provided computational resources and technical support. Victoria Roloff provided excellent research assistance. The Bankard Fund for Political Economy at UVa and the National Science Foundation provided financial support. The results contained herein were derived in part from data provided by ACT Research, and in part from data provided by the College Board. All errors are mine.

1 Introduction

Inflation-adjusted tuition and fees have risen 125% at public four-year colleges in the last thirty years, and 80% for private four-year colleges during the same period (Ma and Pender 2022), leading to concerns about college access and affordability.[1] Because college completion is associated with higher earnings and increased economic mobility,[2] policymakers seek to improve college affordability through policies such as government grant aid to students (Page and Scott-Clayton 2016; Herbaut and Geven 2020). In equilibrium, these policies affect student decisions *and* the tuition, aid policies, and admissions standards set by colleges. As famously suggested by Bennett (1987), colleges may respond to the availability of financial aid by increasing tuition, diminishing the effect of the policy on prices paid by students.

To understand how public and private colleges' tuition, financial aid policies, admission standards, and student enrollment respond to subsidies, I develop an equilibrium model in which public and private colleges compete strategically for students. I build on previous work on student application and enrollment choices (Arcidiacono 2005; Howell 2010; Kapor 2024), as well as previous work on equilibrium in higher education (Epple, Romano, and Sieg 2006; Fu 2014; Epple et al. 2017; Epple et al. 2019).[3] The most significant differentiator between this work and other equilibrium models is that I model the incentives facing *individual* colleges, accounting for the geographic nature of the market.[4] This approach allows me to simulate responses to regional policies, such as the types of policies typically implemented at the state level. This framework also allows me to examine policy responses across individual colleges, to investigate how public colleges' actions affect private colleges, and to understand to what extent state-level policies affect colleges in neighboring states.

My equilibrium model features four stages: first, colleges choose tuition, aid policies, and an admission threshold; second, high-school seniors observe colleges' choices and exogenous characteristics and choose a set of colleges to apply to; third, colleges offer admission and (potentially) financial aid to a subset of applicants; and finally, students make their enrollment decisions. Students base their application decisions on the expected value of an application portfolio, taking into account their probability of admission and likelihood of receiving financial aid. A multiple-discrete choice problem arises in the application stage because a student can choose more than one school to apply to at a time. I model the application stage as a choice between application portfolios, which converts the multiple-discrete choice problem to a single-discrete choice problem.[5] At the enrollment stage, students choose the single college that provides the highest utility from among the colleges that admitted the student. On the supply-side, colleges have heterogeneous preferences

[1] Increases in aid have partially offset these tuition increases (Cook and Turner 2022).

[2] See Heckman, Lochner, and Todd (2006), Goldin and Katz (2008), and Chetty et al. (2020).

[3] Other related work includes Gordon and Hedlund (2019) and Gordon and Hedlund (2022), which imbed monopolistically competitive colleges into an incomplete markets, life-cycle model to study the causes of the increase in tuition since 1987. A different approach to modeling price discrimination is presented in Fillmore (2023). He models aid as the result of an auction in which colleges bid for students.

[4] The other equilibrium models cited in this paragraph group colleges into a few types and assume all colleges within type are identical. The types are defined without regard to physical location. While suited to some research questions, that approach does not allow for the simulation of policies that affect specific colleges or groups of colleges within a region.

[5] This solution has been used in other contexts. For example, Gentzkow (2007) estimates demand for newspapers, where consumers can choose to read more than one newspaper. In his model, consumers choose between "bundles" of the available newspapers. See his paper for a more thorough review of multiple-discrete choice models.

over profit, their admission standards and academic characteristics of students, and the mean income of the student body.

The empirical method incorporates the regional nature of competition in this market. In general, students must travel away from home to attend a four-year college, making distance a key determinant of college choice. As a result, competition between colleges is regional. Some colleges, such as Harvard, draw students from all over the nation, while some colleges (particularly less-selective public colleges) draw the vast majority of their students from in-state. I capture this feature of the market empirically using data on enrollment from each student state of residence to each college from the Residence and Migration Survey of the Integrated Postsecondary Education Data System (IPEDS).[6] For this empirical market definition, I use *historic* data relative to my period of study in order to avoid defining markets based on the model outcome. I estimate the model using data on applications and enrollment by college and student state of residence, where ACT and College Board score reports are used as a proxy for applications.[7] To estimate features of the admissions process, college financial aid functions, and some aspects of heterogeneity across students, I supplement the state-by-college aggregate data sources with moments computed from student survey data, specifically, the National Center for Education Statistics (NCES)' High School Longitudinal Study of 2009 (HSLS:09).

I estimate the model in several steps. First I obtain the federal and state grant aid functions and admissions parameters from regressions estimated on the student survey data. Then I estimate demand and supply separately, extending traditional approaches to estimation of differentiated-products models of oligopoly competition, as in Berry (1994) and Berry, Levinsohn, and Pakes (1995), to this setting. Notable differences between the higher education market and a typical consumer product market motivate substantial deviations from the standard model. The institutional details that drive these differences include: 1) colleges are not profit maximizing, 2) consumers (students) must apply and be accepted in order to "purchase" the product, and 3) colleges engage in third-degree price discrimination—they allow prices to vary by student admission test scores and family income. Thus, I supplement the traditional market share measures with aggregate and individual moments that help identify parameters of the application and admissions process.

With estimates in hand, I interpret the demand-side results using several counterfactual simulations that hold the college choices fixed while allowing student demand to adjust. Under these assumptions, dropping tuition at all colleges across the nation by 1% would increase enrollment by 0.7%, all else equal. The enrollment effects are strongest at lower income levels; students with family incomes between \$0 and \$30k would see a 1.01% change in enrollment from a 1% decrease in all college tuition, while students with family incomes greater than \$110k would only increase enrollment by 0.46%. Decreasing admissions standards across all colleges by 1% is estimated to increase total enrollment by 0.52%, with the strongest effects also at the lowest income levels.

Turning the focus to state-level policies, I simulate pricing, admissions, aid policies, and enrollment

[6] IPEDS is a set of surveys of the population of colleges in the United States and is maintained by the National Center for Education Statistics.

[7] Students send score reports to the colleges that they apply to. I have the number of score reports sent by high school graduating seniors in each state to each college, which I use to proxy applications. To the extent that students may sent some score reports without following through with a complete application, I may underestimate application costs.

outcomes after changes in the generosity of a statewide financial aid program. I implement a $1,000 state grant to in-state students to attend in-state public colleges, using two states as examples: Louisiana and Virginia. These states provide a contrast in that Louisiana has a public system largely dominated by a single college, Louisiana State University (LSU), while Virginia's public system features multiple large and high-quality public institutions, with public in-state enrollment more evenly dispersed across these colleges.

For each example state, I first simulate student responses to grant aid, holding college choices fixed. In Louisiana, student responses to this grant would increase first-year enrollment by 8.5%. For Virginia, the corresponding number is 6.6%. In these demand-side simulations, private colleges can lose up to 6% of their enrollment, with the largest effects among private colleges that draw a substantial portion of their students from in-state and that are similar to regional public colleges in terms of selectivity. I then simulate the equilibrium effects when the supply-side is allowed to adjust. The net enrollment increase is 4.1% in Louisiana and 3.4% in Virginia, with the most affected private colleges losing 2-3% of their enrollment. The difference between the demand-only and equilibrium enrollment changes is largely due to the public colleges' response, which is to raise tuition, increase price discrimination through institutional aid, and raise admissions standards. Ultimately, public in-state colleges absorb over 40% of the subsidy.

The finding that colleges absorb a substantial share of the grant aid intended for students supports the Bennett hypothesis: that is, that market responses to financial aid will result in higher tuition.[8] It has proven difficult for economists to find credible empirical evidence on the Bennett hypothesis, in part because the programs in question (Pell grants, federal loans, or even statewide grants) are implemented on a large scale, making it difficult to cleanly apply quasi-experimental methods to obtain policy responses at the college level. Despite these challenges, there are a few studies that have examined the link between aid programs and tuition using these methods (Singell and Stone 2007; Turner 2012; Turner 2017; Cellini and Goldin 2014), with varying conclusions depending on the specific aid program and the institution type. Most relevant to my study, Long (2004) studies the impact of the Georgia HOPE scholarship on college tuition and aid, among other measures. The Georgia HOPE program is a state grant aid program similar to the one that I model, with two major exceptions: it is a merit-based award that requires certain high school achievement to qualify, and it funds attendance at in-state private schools. Long (2004) uses a difference-in-differences approach, comparing outcomes at Georgia colleges to outcomes in colleges in neighboring southern states, before and after the introduction of the HOPE scholarship in 1993. The findings are generally supportive of the Bennett hypothesis, though in this particular setting, it is the private colleges that responded the most, in some cases recouping 30% of the grant aid.

My study complements existing literature on the Bennett hypothesis by using a structural approach to assess supply-side responses to subsidies. The method has several advantages. For example, it allows me to isolate the effect of an aid policy from other potential concurrent policies and does not rely on assumptions about a control group. Alongside this contribution, the method also provides insight into the functioning of the higher education market more generally; I examine own- and cross-price elasticities for individual

[8] William Bennett's 1987 article on the topic referred specifically to federal loans. He stated that "increases in financial aid in recent years have enabled colleges and universities blithely to raise their tuitions, confident that Federal loan subsidies would help cushion the increase" (Bennett 1987). But a similar argument applies to grant aid, which functions as a subsidy to students and boosts demand.

colleges, document the elasticity of enrollment to tuition at different income levels, estimate college-level objectives that are fully heterogeneous by college, account for strategic competition across colleges, and show that college responses to government grant aid can occur *both* through pricing policies and through admissions standards.

2 Data

Estimation of the model relies on data on college characteristics and student choices at both the aggregate and individual levels. These data sources are described in Sections 2.1 and 2.2. The model is informed by an empirical description of market structure, which is the subject of Section 2.3.

2.1 College Data

Table 1 summarizes the college-level information, which comes from IPEDS for the academic year 2013-14.[9] The main source of price variation is between the in-state public tuition at a mean of \$8,370, and the out-of-state public and private tuition levels, which are \$19,850 and \$32,470 on average, respectively. Student grant aid from all sources (federal, state, and college) can make up a large share of tuition — or even exceed tuition at some colleges. For example, an average student with a family income between \$30,000 and \$48,000 would receive \$10,070 in total grant aid at a public college, generating a net tuition of \$8,370-\$10,070 = -\$2,330 at an in-state public college. The excess grant aid can typically be used to help pay for room and board.

The measure for selectivity is the 25th percentile ACT score within the freshman entering class.[10] ACT scores are measured on a scale from 1 to 36, and the mean among test-takers was approximately 21 in 2013. Public colleges tend to be less selective than private colleges; the average of the 25th percentile ACT score is 20.1 at public colleges compared to 22.3 at private colleges.

I include several other college characteristics as controls: instructional spending per student, the degree completion rate (measured as completion within 6 years of enrollment), whether the college has a graduate program, whether the college has a religious affiliation, whether the college is a historically black college or university (HBCU), discrete measures of the campus size, and the urbanicity of the location of the campus. I also include the Carnegie Classification, which is largely a function of the number and type of degree programs offered at the college. The last panel of Table 1 summarizes these characteristics.

[9] IPEDS is a collection of college surveys on a range of topics such as admissions and enrollment, finances, financial aid, and degrees granted. These surveys include the entire population of post-secondary colleges in the US that participate in federal student aid programs under Title IV (such as Pell Grants and federal student loans). I focus on four-year public or private non-profit colleges located in the 48 contiguous states or the District of Columbia. I include two-year colleges and for-profit colleges in the outside option along with the labor market, because the substitution between the for-profit four-year and traditional public and private non-profit four-year sectors is limited. For-profit colleges comprised approximately 3 percent of first-time four-year enrollment among recent high school graduates in the 2013-14 academic year. After imposing some restrictions on size and selectivity, the population for analysis includes 421 public colleges and 454 private colleges. I restrict analysis to colleges that are classified in the Carnegie Classification system as Baccalaureate, Masters, or Doctoral/Research level colleges. Finally, I restrict the sample to colleges that are somewhat large or selective. Colleges included have either: total undergraduate enrollment greater than 2,000, a Barron's selectivity rank of "Competitive" or higher, or receive applications from more than 2% of students in at least one state.

[10] I convert SAT scores to the ACT scale where required.

2.2 Student Data

To estimate the college choice model, I use aggregate data on applications and enrollment by student state of residence, as well as individual-level data on student decisions from the NCES. The IPEDS Residence and Migration Survey provides enrollment in each college by student state. I use these data to generate enrollment shares by state, defined as a college's enrollment from a state as a proportion of the state's high school graduating class. The second column of Table B2 shows the distribution of enrollment shares across all states and colleges. This distribution is heavily right-skewed, with the enrollment share reaching only 1% of high school graduates at the 90th percentile, and then 17% at the maximum.

I use data from the SAT and the ACT admissions tests as a measure of applications by student state of residence. When students take the SAT or ACT, they have the option to send their score reports to colleges to supplement their college applications. Because the colleges in my study require an SAT or ACT score report to accompany each application, these reports provide a proxy for applications.[11] The College Board and the ACT each provided a table showing the total score reports sent to each college in the country, by student state of residence and graduation year. I link the SAT and ACT data by college name and then sum the total score reports across the two tests.[12] An example of these data for Connecticut is presented in Table B1. I construct an "application penetration" measure using the score-send data by dividing the total number of score reports sent to each college from each state by the number of high school graduates from that state. The application penetration measure is summarized in the first column of Table B2. The maximum value is 0.38, which means that at most, a university receives applications from 38% of high school graduates in a single state. The average application penetration is 0.02, meaning that on average across college-state pairs in my sample, 2% of students from the state send an application to the college.

In addition to these sources of aggregate data, I have access to a sample of high school students from the National Center for Education Statistics' High School Longitudinal Study of 2009 (HSLS:09). This survey provides information on the application portfolios, admission outcomes, enrollment decisions, and student characteristics for a sample of students who were freshmen in high school in 2009.

2.3 Empirical Features of the Market Structure

College enrollment in the US is highly regional (Hoxby 1997; Smith, Pender, and Howell 2018; Hoxby and Turner 2019; Fu et al. 2022). In fact, 74% of four-year college enrollees attended a college in their state of residence as their first primary college, based on student records data from the HSLS:09. Yet, competition between colleges is linked nationally. A regional college that serves only residents of Virginia may compete for Virginia students with another less regional college, such as Duke, which in turn competes

[11] Card and Krueger (2005) show a high correlation between test-score sending and applications, so the SAT and ACT report-sending data may be a close approximation for applications. However, to the extent that sending a score report is easier than sending an actual application, my estimates may overstate the number of applications sent and understate admissions rates. This should affect the interpretation of the application numbers in my results, but it does not affect the enrollment numbers and characteristics of enrollees, which are of primary importance for estimating the supply-side. Further, since I measure selectivity using characteristics of enrollees, this should not affect the interpretation of colleges' admission standards.

[12] This does imply some double-counting in the case that a student sends *both* an ACT and an SAT score with their application, which is likely an unusual situation. To the extent this double-counting exists, it will be reflected in a lower application cost in my estimates.

with a regional college in Georgia for Georgia residents. In this way, the regional colleges in Virginia and in Georgia compete indirectly. This is similar to the chain-linked competition among local newspapers studied by Fan (2013).

A careful analysis of strategic competition in this market incorporates the regional nature of enrollment while recognizing that the market is integrated through chain-linked competition. To do this, I develop a method based on Fan (2013) to generate sets of competitor colleges in each state. Mechanically, the method is as follows:

1. Generate enrollment shares for each college, j, and state of residence, l: $e_{jl} = \frac{n_{jl}}{N_l}$. The numerator is the number of high school graduates from state l who attended college j, and the denominator is the number of high school graduates from state l.[13]

2. Sort the data by college and descending enrollment share.

3. Starting from the top of the list for each college, select the states that together account for at least 85% of the college's total enrollment.[14][15]

In these computations, I used the eight years of data *prior* to the year of this analysis (2005-06 through 2012-13) to smooth out year-to-year fluctuations and to avoid using current-year enrollment patterns to define the sets of competitors in each market.

As an example, Figure 2 demonstrates the outcome of this method for Tulane University, Louisiana State University (LSU) and Louisiana Tech University. Tulane University is a well-known private university, and its enrollment comes from a large number of states. Still, Tulane draws more enrollment from southern states than a similar college further north would (e.g. Georgetown University). LSU is the flagship public university, so it enrolls students primarily from Louisiana, but also attracts a non-negligible share of students from Texas and Mississippi. Finally, Louisiana Tech draws the overwhelming majority of students from Louisiana, as is the case for many regional public colleges. The example also illustrates that this method—which is based on enrollment *shares* by state—does not automatically choose the largest states in terms of population. This is why, for example, Tulane is marked as competing in Vermont but not in North Carolina.

A drawback of this method is that counterfactual conclusions are conditional on this definition of regional competition: there is no role for expanding the regional draw of enrollment, even if a college were to become more attractive by reducing prices or admission standards. This is a reasonable assumption if the regional draw of enrollment is primarily a function of the knowledge students have about their options — e.g. the reason a student in New York would not choose a regional college in Louisiana is in part because the student is not aware of the college in Louisiana. Further, even if there is a margin for counterfactual pricing and aid policies to broaden the regional draw of enrollment, the changes would have to be very large

[13]The total number of high school graduates by state is obtained from NCES data as collected in the *Knocking at the College Door* report produced by the Western Interstate Commission for Higher Education (WICHE).

[14]A larger enrollment threshold would capture more of a college's demand, while a smaller threshold limits the computational challenges that arise from larger choice sets.

[15]I also require that at least 0.15% of students in the state apply to the college.

to impact the conclusions substantially. The average enrollment share in the state-university pairs dropped by my regional market definition is 0.0001.[16]

3 Model

There are two types of agents in the model: students, i, and colleges, j.[17] There are $J+1$ colleges $j = 0, 1, ..., J$, where $j = 0$ represents the outside option, including for-profit colleges, open-admission or non-selective four-year colleges, two-year colleges, and the labor market. The timing of the model is illustrated in Figure 1. First, each college simultaneously chooses tuition, a financial aid policy, and an admission policy. Each student then observes the colleges' choices and exogenous characteristics and chooses a set of colleges to apply to—an *application portfolio*. After applications are received, colleges notify students of their admission outcomes. Each student then chooses which college, if any, to enroll in from among the colleges that have offered admission.

The following sections cover details about functional forms and the information available to each agent at each stage. In the first two subsections below, I describe the financial aid and admissions policies. Next, I describe the student's problem at the enrollment and application stages and then move to the supply-side.

3.1 Pricing

Before making an enrollment decision, student i will observe the *net tuition* that they would pay at each college, defined as tuition minus grants received from the federal and state governments and the college. The net tuition student i pays to attend college j is given by:

$$p_j(inc_i, s_i, l_i, \chi_{ij}) = t_j(l_i) - a^f(inc_i) - a^s(inc_i, s_i, l_i) - a^c_j(inc_i, s_i, \chi_{ij}), \tag{1}$$

where $t_j(l_i)$ is tuition at college j, which varies across states only by in-state status for public colleges.[18] $a^f(inc_i)$ is the level of the Federal Pell Grant as a function of income inc_i. State grant aid $a^s(inc_i, s_i, l_i)$ is a function of the student income level and observed college admission test score, s_i, as well as features of the state grant aid program available in the student's state, l_i. Finally, $a^c_j(inc_i, s_i, \chi_{ij})$ is college j's grant aid to student i as a function of the student's income, test score, and an idiosyncratic determinant of aid, χ_{ij}, which follows the standard normal distribution. Students do not observe χ_{ij}, but they observe the aid offer upon acceptance to a college. Thus to notate the price from the student's perspective upon admission, I use $p_j(inc_i, s_i, l_i, a^c_{ij})$.

The government aid functions are exogenous functions of observable characteristics and thus can be estimated separately from the rest of the model, as discussed in Section 5.1. The college aid function merits further discussion because parameters of this aid function are endogenous choices of the college.

[16] This is after filtering to four-year colleges located in the US and with Baccalaureate, Masters, or Doctoral Carnegie Classification codes, but without imposing further restrictions on size as described in Section 2.1.

[17] Students are high school graduating seniors, and colleges include most four-year colleges in the US.

[18] Reciprocity agreements give some out-of-state students in-state status for the purpose of pricing. I take these agreements as given in the analysis.

Colleges choose aid levels as a function of student income and test scores as well as some information about the student that is idiosyncratic (χ_{ij}). The specification allows for the existence of merit-based and need-based aid depending on the degree to which the college considers family income and test scores in the aid function. In general, students are quite uncertain about their financial aid outcomes before applying, an empirical reality that is captured by the idiosyncratic component of the aid function. I parameterize aid as a proportion of tuition. It is infeasible to model aid as continuous, so I discretize college aid into four levels. The latent value that determines the aid level is the aid index $\tilde{a}_{ij}^c = a_{j1}^c + a_{j2}^c exp(s_i) + a_{j3}^c inc_i + \chi_{ij}$, where the coefficients vary by college j. Colleges can thus choose to what degree they consider merit (ACT scores) or income in the aid policy by adjusting the coefficients on those two factors in the aid index. The aid index maps to realized aid as a proportion of tuition according to an ordered probit.[19]

3.2 Admissions

Admissions outcomes for each student are based on an admission index, $s_{ij} = s_i + \zeta_{ij}$, where s_i is the student's college admission test score, and $\zeta_{ij} \sim N(0, \sigma_\zeta^2)$ is an idiosyncratic measure of student fit with the particular college. Prior to the admission stage (Stage 3 in Figure 1), each college j only knows the distributions of test scores s_i and the student fit ζ_{ij} among all high school graduates. At the admission stage, the values of s_i and ζ_{ij} for all of college j's applicants are revealed to college j. College j can then construct the admission index for each applicant and compare the index to the pre-determined admission threshold, $\underline{s}_j$, to determine the admission outcome.

Student i knows the value of s_i but knows only the distribution of ζ_{ij}. The student also knows the admission threshold for each college j, and can thus construct the conditional probability of admission. Some additional notation is useful for representing the admissions probability. Let $\mathbf{D}_i$ be a $(J+1) \times 1$ vector, where $D_{ij} = 1$ if i was offered admission at college j and $D_{ij} = 0$ otherwise, whether this is because the student did not apply or because the college rejected the student. Further, let $\mathbf{Y}_i$ be a vector indicating a 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. Then the probability that a student is accepted to j conditional on application is:

$$\mathscr{P}(D_{ij} = 1 | s_i, \underline{s}_j, Y_{ij} = 1) = \mathscr{P}(s_i + \zeta_{ij} > \underline{s}_j) = \Phi\left(\frac{s_i - \underline{s}_j}{\sigma_\zeta}\right), \tag{3}$$

where $\Phi(\cdot)$ is the standard normal CDF.

[19] Specifically, college aid is the following:

$$a_j^c(inc_i, s_i, \chi_{ij}) = \begin{cases} 0 & \text{if} \quad \tilde{a}_{ij}^c \leq k_1 \\ 0.3 \times t_j(l_i) & \text{if} \quad k_1 < \tilde{a}_{ij}^c \leq k_2 \\ 0.6 \times t_j(l_i) & \text{if} \quad k_2 < \tilde{a}_{ij}^c \leq k_3 \\ 1.1 \times t_j(l_i) & \text{if} \quad \tilde{a}_{ij}^c \geq k_3 \end{cases}. \tag{2}$$

3.3 Student Utility from Enrollment

With the net tuition and admission probability as specified above, the utility student i receives from enrolling at college j is:

$$u_{ijl}(inc_i, s_i, l_i, X_{ij}, a^c_{ij}) = \beta_{0i} + \beta_{1i} p_j(inc_i, s_i, l_i, a^c_{ij}) + \boldsymbol{\beta_{2i}} \mathbf{X_{ijl}} + \xi_{jl} + \eta_{ijl}.$$

Students may always choose not to attend any college ($D_0 = 1$), in which case they receive a mean utility that is normalized to zero.

This specification allows the value of attending any college, β_{0i}, to vary by student. In particular, $\beta_{0i} = \beta_{00} + \beta_{01} ln(inc_i)$. Similarly, the disutility of net tuition varies by income level of the student, so that $\beta_{1i} = \beta_{10} + \beta_{11} ln(inc_i)$. Students may also value exogenous characteristics of the college, $\mathbf{X_{ijl}}$, including (but not limited to) the distance from the student's residence to the college, whether or not the college is in-state, the public/private status of the college, and controls for quality such as a measure of instructional spending per student.[20] The value placed on instructional quality (measured by instructional spending per student) is allowed to differ across students with different ACT scores and income levels; in this way, $\boldsymbol{\beta_{2i}}$ varies across students. The utility function also includes a measure of college quality that is observed by the student but not by me, captured in ξ_{jl}. Finally, the student-college idiosyncratic component of utility, η_{ijl}, is distributed Type I Extreme Value. Students learn this value after application (e.g., this could represent information about campus dining options presented during a campus tour).

Due to the Type I Extreme Value assumption on η_{ij}, the probability that a student enrolls in a particular college among colleges that admitted the student is a conditional logit probability. If the college does not admit the student, then the probability of enrollment is zero. Let $\mathbf{a^c_i}$ be the set of college aid offers that the student faces. Then, letting $\nu_j(inc_i, s_i, l_i, X_{ij}, a^c_{ij}) = u_{ij}(inc_i, s_i, l_i, X_{ij}, a^c_{ij}) - \eta_{ij}$, the probability that i enrolls in j given a set of admission and aid offers can be written:

$$\mathbb{P}(Enroll_{ij} = 1 | \mathbf{D}_i, inc_i, s_i, l_i, \mathbf{X_i}, \mathbf{a^c_i}) = \begin{cases} \frac{e^{\nu_j(inc_i, s_i, l_i, X_{ij}, a^c_{ij})}}{\sum_{j' \ni D_{ij'}=1} e^{\nu_{j'}(inc_i, s_i, l_i, X_{ij}, a^c_{ij})}}, & \text{if } \mathbf{D}_{ij} = 1 \\ 0, & \text{if } \mathbf{D}_{ij} = 0. \end{cases} \tag{4}$$

3.4 Application Value Function

Each student chooses an application portfolio to maximize their expected utility of enrollment net of application costs. To derive this expected value, 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 colleges with $D_{ij} = 1$:[21]

$$v(\mathbf{D}_i | inc_i, s_i, l_i, X_i, \mathbf{a^c_i}) = ln \left(\sum_{j \ s.t. \ D_{ij}=1} e^{\nu_j(inc_i, s_i, l_i, X_{ij}, a^c_{ij})} \right) + \gamma, \tag{5}$$

[20] Just as with the other college characteristics, this instructional spending per student is constant and exogeneous. In other words, the college has pre-committed to a certain level of instructional spending per student.

[21] The log-sum formula follows from the assumption that η_{ij} are Type-I Extreme Value with location and scale parameters normalized to 0 and 1, respectively.

where γ is the Euler–Mascheroni constant. This constant cancels out of the choice probabilities as it does not change the relative utility of any option, and will be dropped from here on.

Recall that $\mathbf{Y_i}$ is a vector representing the application set. If the student applies to j, $Y_{ij} = 1$, otherwise $Y_{ij} = 0$. The probability of admission to j conditional on application, $\mathscr{P}(D_{ij} = 1|s_i, \underline{s}_j, Y_{ij} = 1)$, is described in Section 3.2. Because the unobserved component of the student application index is independent across colleges, the student has the following probability of being admitted to a set $\mathbf{D}_i$:

$$\begin{aligned}&\mathscr{P}(\mathbf{D}_i|s_i, \mathbf{Y}_i)\\ &= \prod_{j\ \ni Y_j=1,\ D_j=1} \mathscr{P}(D_{ij} = 1|s_i, \underline{s}_j, Y_{ij} = 1) \prod_{j\ \ni Y_j=1,\ D_j=0} [1 - \mathscr{P}(D_{ij} = 1|s_i, \underline{s}_j, Y_{ij} = 1)].\end{aligned} \tag{6}$$

Let $\mathbf{A}$ represent the set of possible aid offers from colleges that admitted the student, and $\mathscr{P}(\mathbf{a^c}_i|inc_i, s_i, \mathbf{D}_i)$ represent the probability of receiving a specific combination of aid offers.[22] The value of applying to Y_i is the expected utility of application net of application costs:

$$\begin{aligned}&V(\mathbf{Y}_i|inc_i, s_i, l_i, \mathbf{X_i}) =\\ &\left[\sum_{\mathbf{D}_i \subseteq \mathbf{Y}_i} \mathscr{P}(\mathbf{D}_i|s_i, \mathbf{Y}_i) \sum_{\mathbf{a^c}_i \subseteq \mathbf{A}} (\mathscr{P}(\mathbf{a^c}_i|inc_i, s_i, \mathbf{D}_i) v(\mathbf{D}_i|inc_i, s_i, l_i, \mathbf{X_i}, \mathbf{a^c}_i))\right] - c_i(\mathbf{Y}_i),\end{aligned} \tag{7}$$

where $c_i(\mathbf{Y})$ represents i's cost of application to $\mathbf{Y}$. The cost depends on the application cost associated with each college-market pair, plus an idiosyncratic component that varies across application sets:

$$c(\mathbf{Y}_i, l_i) = \left(\sum_{j\ \ni Y_{ij}=1} c_{jl}\right) + \epsilon_{i\mathbf{Y}_i}, \tag{8}$$

where $c_{jl} = \beta^{\mathbf{c}} X^c_{jl} + \xi^c_{jl}$. The idiosyncratic cost, $\epsilon_{\mathbf{Y_i}}$, is distributed Type I Extreme Value. Students may have different costs to apply to different sets of colleges for many reasons, including school and parental guidance.[23]

3.5 Aggregation to Enrollment Shares and Application Penetration

Based on the model so far, I can construct enrollment shares and application penetration measures, which are useful for estimation. First, note that each student faces a unique set of application portfolios from which to choose (a choice set), which I denote Υ_i. These choice sets are in part determined by the student's geographic location. The choice sets are discussed in further detail in Section 5.3.

Conditional on the choice set Υ_i and student characteristics, the probability that a student applies to an

[22]This probability is given by the ordered probit aid function as described above.

[23]Because idiosyncratic cost shocks are drawn at the application portfolio level, there is some probability that a student may find it less costly to apply to a large application set rather than a smaller subset of the same colleges. The likelihood of this situation is small because c_{jl} is relatively large compared to the variance of a standard Type-I EV distribution. See Figure B2 for the distribution of estimated costs across all state-college pairs. The assumption greatly facilitates estimation of the model.

application portfolio $\mathbf{Y}$ is:

$$\mathbb{P}(Apply_{\mathbf{Y}}|inc_i, s_i, l_i, \mathbf{X_i}, \Upsilon_i) = \frac{e^{\tilde{V}(\mathbf{Y}|inc_i, s_i, l_i, \mathbf{X_i})}}{\sum_{\mathbf{Y}' \in \Upsilon_i} e^{\tilde{V}(\mathbf{Y}|inc_i, s_i, l_i, \mathbf{X_i})}}, \tag{9}$$

where $\tilde{V}(\mathbf{Y}|inc_i, s_i, l_i, \mathbf{X_i}) = V(\mathbf{Y}|inc_i, s_i, l_i, \mathbf{X_i}) - \epsilon_{i\mathbf{Y}}$.

Aggregating across portfolios that contain j for each student and across all students yields the application penetration for college j for a given state:

$$\mathbb{P}(Apply_j|l_i = \ell) = \frac{1}{N_\ell} \sum_{i \ni l_i = \ell} \sum_{\mathbf{Y} \ni Y_j = 1} \mathbb{P}(Apply_{\mathbf{Y}}|inc_i, s_i, l_i, \mathbf{X_i}, \Upsilon_i). \tag{10}$$

In estimation, this model-derived application penetration will be matched to the observed proportion of students who apply to each college in each state. Note that the sum of the market penetration of applications yields a number greater than one, because each application portfolio will be counted as many times as colleges it contains. In other words, if a student applies to three colleges, that student is counted in the market penetration of applications for each of the three colleges.

Another equation used in estimation is the enrollment share by student state. For each student, the probability of enrolling in college j, unconditional on admission decisions, is:

$$\begin{aligned} \mathbb{P}(Enroll_j|s_i, inc_i, l_i, \mathbf{X_i}, \Upsilon_i) = &\sum_{\mathbf{Y} \in \Upsilon_i \ni Y_j = 1} \underbrace{\mathbb{P}(Apply_{\mathbf{Y}}|s_i, inc_i, l_i, \mathbf{X_i}, \Upsilon_i)} \times \\ &\sum_{\mathbf{D} \in \mathbf{Y}} \underbrace{\mathscr{P}(\mathbf{D}|s_i, \mathbf{Y}_i)} \underbrace{\mathbb{P}(Enroll_j|inc_i, l_i, \mathbf{X_i}, \mathbf{D})}. \end{aligned} \tag{11}$$

Each component in this expression (as indicated by the brackets) has been described before. They are, from left to right, the probability of application to a specific application portfolio conditional on the student's characteristics and choice set (equation 9), the probability of acceptance to a set of colleges conditional on the student test score and colleges applied to (equation 6), and the probability of enrollment conditional on the student's admission and aid outcomes (after integrating out the aid outcomes in equation 4). Aggregating to the market-level yields the following expression for the enrollment share:

$$\mathbb{P}(Enroll_j|l_i = \ell) = \frac{1}{N_\ell} \sum_{i \ni l_i = \ell} \mathbb{P}(Enroll_j|s_i, inc_i, l_i, \mathbf{X_i}, \Upsilon_i). \tag{12}$$

Section 5.4 demonstrates how I use the enrollment shares and application penetration in estimation.

3.6 The College Objective Function

In most markets, firms are assumed to maximize economic profit and have no direct preference over consumer characteristics. But in the context of the US higher education market, the firm objective function must be adapted to account for colleges' preferences over student characteristics. Previous research on the supply-side incentives in US higher education agrees on this point, though specific functional forms vary

(Rothschild and White 1995; Epple, Romano, and Sieg 2006; Fu 2014; Epple et al. 2017; Epple et al. 2019; Blair and Smetters 2021).

I take a flexible, empirically-driven approach to the objective function, allowing colleges to vary heterogeneously in the degree to which they value profit (π_j), their admissions standard ($\underline{s}_j$), the mean academic fit of the student body ($\bar{s}_j$), and a function of mean income ($f(\mu_j^{inc})$).[24] Each of these factors enter a Cobb-Douglas objective function with parameters that are fully heterogeneous across colleges. I assume constant returns to scale. Ultimately, estimation shows that for most colleges the four terms all have positive (non-zero) exponents, but the specification does not presuppose that all of the terms enter positively. Thus, it is possible that some colleges may not place any weight on some of these factors, while other colleges do.

All of the terms in the college objective are functions of all college choice variables (tuition levels, aid coefficients, and admission thresholds), but to keep the notation concise I drop the dependence on the endogenous variables in the notation. On another technical point: in the objective function, I convert all admissions-related measures (the admissions standard and the mean academic fit) to percentiles relative to the population of high school graduates. The conversion to percentiles ensures that no student is infinitely better than others from the college's point of view, although the support of the unobserved component of the admission index, ζ_{ij}, is all real numbers. In terms of notation, I let $P(\bar{s}_j)$ represent the percentile rank of the mean student's admission index in the population of high school graduates, and $P(\underline{s}_j)$ represent the percentile rank of the admission threshold, which is equivalent to the percentile rank of the student at the minimum standard for admission.

Each college maximizes:

$$W_j(t_j(l_i), \mathbf{a_j^c}, \underline{s}_j) = \pi_j^{(1-\lambda_{1j}-\lambda_{2j}-\lambda_{3j})} P(\underline{s}_j)^{\lambda_{1j}} (P(\bar{s}_j) - P(\underline{s}_j))^{\lambda_{2j}} \left(f(\mu_j^{inc})\right)^{\lambda_{3j}}, \tag{13}$$

where

$$\pi_j = \underbrace{\sum_{i \ni Enroll_{ij}=1} t_j(l_i) - a^c(s_i, inc_i, \chi_{ij}) + \psi_j}_{\text{Revenue}} - \sum_{i \ni Enroll_{ij}=1} \underbrace{(\alpha_1 + \alpha_2 q_j + \alpha_3 X_j^s + \omega_j - I[j \in J^{pub} \;\&\; l_i = l_j](\alpha_4 + \omega_j^{is}))}_{\text{Cost}}. \tag{14}$$

The first bracket is total revenue. This is tuition, $t_j(l_i)$— which varies by in-state status for public colleges, minus the college aid amount, plus ψ_j, which represents revenue from private gifts and federal, state, and local appropriations. The second bracket represents the student's cost to the college. This varies by college

[24]Specifically, to ensure that the function is decreasing in average incomes and bounded below at zero, I use $f(\mu_j^{inc}) = \frac{1}{\mu_j^{inc}} - min\left(\frac{1}{\mu_j^{inc}}\right)$, where $min\left(\frac{1}{\mu_j^{inc}}\right)$ is the lowest observed value of the fraction $\frac{1}{\mu_j^{inc}}$. Without the adjustment for the $min\left(\frac{1}{\mu_j^{inc}}\right)$ term, there is a potential corner solution to the college objection function for certain parameter values, which involves a very high average income.

characteristics and quality, but importantly also by the number of students enrolled, q_j. The number of enrolled students is derived from the demand system described above. Allowing the per-student cost to depend upon current enrollment implies a (potentially) increasing marginal cost. Finally, public colleges may have an incentive to enroll in-state students, which reduces the effective marginal cost. This is captured by the term: $I[j \in J^{pub} \ \& \ l_i = l_j](\alpha_4 + \omega_j^{is})$, where $I[j \in J^{pub} \ \& \ l_i = l_j]$ is an indicator function for in-state public colleges. The parameter α_4 is the average difference in revenue between in-state and out-of-state students.[25]

Colleges choose tuition ($t_j(l_i)$) the aid policy parameters on income and ACT scores (a^c_{j1} and a^c_{j2}) and the admission threshold ($\underline{s}_j$) to maximize Equation 13. As mentioned above, tuition levels are constrained to be the same across student locations except in the case of public colleges, which may charge different tuition levels for in-state and out-of-state students. Tuition, aid policy parameters, and admissions standards are solutions to the first-order conditions of the college problem. To find the college-specific parameters, this system of first-order conditions can be solved for λ_{1j}, λ_{2j}, λ_{3j}, the marginal cost, and in the case of public colleges, the differential cost for in-state students, ($\alpha_4 + \omega_j^{is}$).

4 Identification

In any market equilibrium model, the main identification challenge is to separate the effect of endogenous variables on supply and demand, knowing that observed equilibrium values of endogenous variables are generated by the interaction of both sides of the market. This affects the current cross-sectional study in two ways. First, differences across colleges in tuition, financial aid policies, and admission thresholds may be correlated with unobserved determinants of demand, in particular, the ξ_{jl} in utility. This follows because colleges make their decisions about tuition, aid policies, and admissions standards taking into account expected demand, which is a function of ξ_{jl}. So, for example, I cannot assume that the price is independent of the unobserved determinants of demand when estimating the price coefficient. Second, when estimating the effect of enrollment quantity on marginal cost, α_2, I cannot assume that the college's enrollment is independent of unobserved determinants of cost, because enrollment is a function of endogenous choices of the college. Thus, I must identify the effect of price and admissions standards on utility and the effect of enrollment on marginal cost without using these conditions.

4.1 Demand-Side

The instrument for net tuition in demand estimation is an indicator for in-state public colleges. I include indicators for whether a college is in-state and the private/public status in the utility function, and the instrument is the interaction between these two variables. The relevance of the instrument comes from the public college's incentive to enroll in-state students: this incentive generates a reduction in the price paid for in-state students at public colleges only. The exogeneity assumption is that the difference in demand

[25] Note that the data are unclear as to what portion of state appropriations and other revenue sources are linked to in-state students, so I must estimate the revenue difference between these student types. This is identified by pricing variation conditional on other differences in in-state and out-of-state student characteristics.

between public and private colleges would be the same among in- and out-of-state colleges, aside from the price differential at in-state public colleges and the other factors included in my controls, such as distance to college.[26]

4.2 Supply-Side

On the supply-side, the college-specific parameters and the marginal cost are found as solutions to the first-order conditions of the college's problem given in Equation 13. After finding the value of the marginal cost for each college, I set it equal to the model expression for marginal cost ($\alpha_1 + 2\alpha_2 q_j + \alpha_3 X^s_j + \omega_j$). The identification challenge is to isolate the effect of enrollment (q_j) on marginal cost. Enrollment is a function of college choices, which are in turn a function of unobservable cost shifters (ω_j). I need an instrument that is uncorrelated with the cost-shifters but is correlated with enrollment. The instrument I use is enrollment predicted as a function of distance and the regional market definitions, using the estimated effect of distance from the demand estimation process. In constructing the instrument, I also set aside the role of the admission threshold, as it is a function of cost-shifters. Thus the instrument is the predicted enrollment as a function of distance alone, where students have 100% probability of admission at every college. In essence, I set aside every determinant of demand except for distance when I generate the instrument.

5 Estimation and Results

Estimation proceeds in several steps. First, I estimate the federal and state aid functions and the admissions probability using individual student data. These steps are described in Sections 5.1 and 5.2. Then I simulate a representative sample of 1,000 students in each state, with students varying by location, ACT score, and family income. This step is discussed in Appendix A. After I have a simulated sample of students, I simulate each student's choice set, as described in Section 5.3. I estimate demand and college financial aid policies simultaneously using the simulated method of moments. Finally, taking the demand estimates as given, I estimate supply based on the colleges' first-order conditions. Sections 5.4 through 5.6 describe these final two steps.

5.1 Federal and State Aid Functions

The price that students pay for college is determined in part by federal and state grant aid, as shown in Equation 1. The Pell Grant is by far the federal government's largest college grant aid program, and awards aid largely on the basis of need. State grant aid amounts are generally determined by student family income or some combination of income and merit. I estimate Federal Pell and state grant aid policies in the HSLS:09 survey using enrolled students. Throughout the rest of the estimation procedure, the parameters of these policies are fixed.

For federal aid, I run a Tobit regression of Pell grant aid amounts on a quadratic in family income. Pell aid is bounded below at $0 and above at $5,650, which is the maximum Pell grant set by the government

[26] One might argue that loyalty to the in-state public college sports team might result in a violation of this assumption. However, the inclusion of available controls for college sports activity (such as NCAA membership) does not significantly alter the estimates.

in the applicable year. I exclude students with incomes above \$75k, assuming that they will not receive any Pell grant.[27] The results for the federal aid equation are shown in Table B3. The predicted Pell grant at a family income of \$30k is \$3,795 and at a family income of \$60k is approximately \$823.

State grant aid is more complicated, varying across states both in total generosity and in the degree to which the state program rewards merit or need. I distinguish between three groups of states: states with large merit aid programs, California (which is large enough to have enough observations in HSLS:09 to run a state-specific regression), and all other states, which have relatively small programs that typically tend to put more emphasis on need. The list of states with large merit aid programs is derived from Fitzpatrick and Jones (2012).[28] For each type of state, I run a Tobit regression of student state aid on student characteristics, whether the college is public or private, and the generosity of the state's grant aid program as measured by the total state grant aid budget per high school graduate. Aid amounts are bounded below by 0. A student's state grant eligibility is determined by the student's state of residence. I assume that the aid is only applicable at in-state colleges, which is typically the case with these grant aid programs.

The estimates from these regressions are shown in Table B4 are consistent with expected patterns: for example, in states with large merit aid programs, a student with a family income of \$30k is expected to be offered \$467 more in state grant aid than an identical student with a \$100k family income. In states that focus less on merit aid and more on need-based aid ("All Other States" in the table), the income gradient is much steeper: a student with a family income of \$30k is expected to be offered \$3,230 more than an identical student with a family income of \$100k. A regression for California that includes the ACT score shows that the ACT does not have an economically or statistically significant effect on aid received in that state, so I leave out the ACT score measures for California.

5.2 Admission Probability

The admission probability as described in Section 3.2 depends on only the college's threshold, the student's ACT score, and an idiosyncratic measure of academic fit. The assumption that the idiosyncratic component is normally distributed leads to a simple expression for the probability of admission, given in Equation 6.

In the HSLS:09 data, I observe students' application sets and admission outcomes, and could therefore estimate the parameters of Equation 6 using a Probit regression on these data, provided I observe the students' test scores and the college admission thresholds. A complication arises here because college thresholds are unobserved, although I do see student admission test scores. While I don't observe admission thresholds, I do have information on selectivity; in particular, I have the 25th percentile of test scores among enrolled students at each college. I parametrize the threshold as a quadratic function of the 25th percentile test scores and estimate the parameters from the Probit regression on admission outcomes. The estimating

[27] Approximately 8% of students in the HSLS:09 who have family incomes above \$75k, attend a college in my sample, and have National Student Loan Database (NSLDS) records received a non-zero Pell grant. However, the overall rate of Pell receipt is even lower for this income range because students who did not receive any federal aid are not in the NSLDS system. Imputing zero Pell grant for students without NSLDS records yields a rate below 5%.

[28] States with large merit aid programs at this time were: AR, GA, FL, KY, LA, MD, MI, MS, NM, NV, SC, SD, TN, WV, and WY.

equation is a Probit where the probability of admission is given by:

$$\mathscr{P}(D_{ij} = 1 | s_i, \underline{s}_j, Y_{ij} = 1) = \mathscr{P}(s_i + \zeta_{ij} > \underline{s}_j) = \Phi\left(\frac{s_i - \underline{s}_j}{\sigma_\zeta}\right),$$

where $\Phi(\cdot)$ is the standard normal CDF. Defining $ACT25_j$ as the observed 25th percentile of the ACT score among admitted students at college j, I assume $\underline{s}_j = \beta_0^{adm} + \beta_1^{adm} ACT25_j + \beta_2^{adm} ACT25_j^2$. Note that since the coefficient on s_i is constrained to be one, I can identify the standard deviation σ_ζ.

The estimates of σ_ζ, β_0^{adm}, β_1^{adm}, and β_2^{adm} are provided in Table 2. Figure 3 shows the admission threshold function, marking the admission threshold for certain example colleges. The admission threshold can also be interpreted as the ACT score that gives a student a 50% chance of admission. The graph shows that the admission threshold for Vanderbilt was approximately 35, while the threshold for Tulane was approximately 25. A regional public college like Nicholls State in Louisiana was much less likely to reject students, with an admission threshold of approximately 12. Another way to think of this is that admission to Vanderbilt was far from guaranteed even at the highest possible ACT score of 36, while an average ACT score near 20 would nearly ensure admission to Nicholls State.

5.3 Choice Sets

As discussed in Section 3, each student may apply to any school within their individual-specific choice set. In my setting, choice sets are both exogenous and unobserved, so I must make an assumption about the distribution of choice sets. There is a wide range of assumptions that would rationalize observed data. One potential assumption is that all students have full information and consider every college. At the other extreme, I could assume every student always applies to every college in their choice set, so choice sets are indistinguishable from application portfolios (in this case, there would be effectively no choice at the application stage).

All research on the application portfolio choice problem must make some assumption about the distribution of choice sets, but the assumptions vary. Howell (2010) utilizes an approach that allows estimation of a full-information model, and Arcidiacono (2005) uses a sampling approach to select choice sets from an exogenously determined distribution. I use an assumption that students' choice sets are limited by their state and by random information that generates differences in choice sets across students in the same state. This assumption is most similar to the approach used by Arcidiacono (2005). I use this assumption both because limited information is most consistent with the institutional setting, and because it helps resolve the computational difficulty of estimating the demand-side.[29]

Substantial computational savings are also generated by limiting each student to send a maximum of four applications. In the ELS:2002 survey data from the NCES, which includes primarily students who would apply to college in 2004, 86% of college applicants apply to four or fewer of the four-year colleges that are included in my analysis, so the assumption is not terribly restrictive.[30]

[29]The computation is challenging because students have a portfolio choice problem: the number of possible application sets gets very large with even a small number of colleges in the choice set. In addition, for every application portfolio there are many different admission outcomes and aid offers, complicating the computation of the expected value of a given admission portfolio.

[30]The HSLS:09 data do not record all applications sent by each student, so it is not possible to get a corresponding number for

In Section 2.3, I described the geographic limitations on a college's market. This corresponds to a geographic limitation on each student's choice set, as students must choose from only the colleges that compete in the student's state. After applying this limitation, the market with the largest number of colleges has just over 150 colleges competing for these students (see Figure B1). Combining this assumption with the limitation that students apply to no more than four colleges still yields nearly 21 million possible portfolios in a state in which 150 colleges compete. I further limit the problem by making restrictions on the types of colleges that are combined in a single application portfolio. Evidence from the individual data suggests that it is highly unlikely that students apply to, for example, three highly selective colleges (like Harvard, Yale, and Stanford) and a non-selective college. To establish these results empirically, I examine patterns of selectivity in the ELS data using Barron's rankings from the 2004 *Profiles of American Colleges*.[31] I characterize application portfolios by the number of colleges in each of six selectivity categories. I find that 95% of students' application portfolios fall in the top 74 combinations of selectivity categories, so I limit the possible portfolios to follow these patterns.

After applying these restrictions, I simulate choice sets of 20 colleges randomly from the colleges that compete in the student's state. These choice sets correspond to the Υ_i in the model. There is no data to inform the likelihood of specific choice sets, but each option must be observed at least enough times that the applicant share data can be rationalized (e.g. a college that receives applications from 25% of all high school graduates must be in the consideration set for at least 25% of students). To achieve this property when sampling, I weight each college by the application share in the corresponding state.

5.4 Estimation Strategy for College Choice and Financial Aid Policy Parameters

I use the method of moments to estimate simultaneously the enrollment utility and application cost parameters as well as parameters of the college financial aid policy, as in Equation 2. College-level financial aid policy parameters must be estimated because they are not observed in the data. This is similar to how admissions thresholds are unobserved and are therefore estimated as discussed in Section 5.2. In this section, I first describe the rationale for estimating the college financial aid policies simultaneously with the demand-side of the model. Then I discuss the moments I use in estimation and further details of the estimation process.

Federal and state aid policies can be estimated separately from the rest of the model, as described in Section 5.1, because there is no selection on unobservables.[32] College-level financial aid policies, on the other hand, cannot be estimated separately from the rest of the model due to a selection problem. Colleges use information other than test scores and income levels to set aid amounts. Students will see their aid amount only upon admission to the college, and then make their enrollment decisions based on this aid amount. Thus, students with higher aid amounts will select into college attendance. This implies that in

students applying to college in 2013. The best information I have from that survey says that 78% of all applicants apply to four or fewer colleges among the entire population of post-secondary colleges (including 2-year colleges and other schools not included in my analysis). Of course, the share of applicants applying to four or fewer colleges in my subset of four-year colleges studied is larger than 78%.

[31] I use ELS here instead of the HSLS:09 because the HSLS:09 only records details on up to three applications per student.

[32] This assumption reflects published policies, where aid amounts are generally determined by observed income levels and/or test scores.

a sample of *enrolled* students such as my HSLS:09 sample, the aid amounts will generally be higher than expected for the average high school graduate with the same ACT score and income. By estimating the college's aid policies together with the demand system, I account for the selection process through the demand system.

While I do not observe college aid policies directly, I do have data that is informative about these policies. From the college-level IPEDS data, I observe the net tuition paid by enrolled students from five different income levels after all sources of aid (federal, state, and college) have been applied. I also see the average college aid to all students. Table 1 shows the average net tuition by income group and public/private college type from these data. In the HSLS:09 student survey, I see the identity of each enrolled college and the amount of aid each student received from the college, as well as the student's family income and ACT score.

Recall that the college aid a student receives depends upon the aid index $\tilde{a}^c_{ij} = a^c_{j1} + a^c_{j2} s_i + a^c_{j3} inc_i + \chi_{ij}$, where the a^c_{j1}, a^c_{j2}, and a^c_{j3} are the college policy parameters. These are estimated as functions of observed data:

$$\begin{aligned} a^c_{j1} &= \beta^{aidc}_1 X^{aidc}_{1j} \\ a^c_{j2} &= \beta^{aidc}_2 X^{aidc}_{2j} \\ a^c_{j3} &= \beta^{aidc}_3 X^{aidc}_{3j}, \end{aligned} \tag{15}$$

where the X variables are constructed from the data described in the previous paragraph as detailed below.

With the specification for the college financial aid policy in hand, I estimate the enrollment utility, application cost, and financial aid policy parameters using the method of moments. The first two moment equations are based on the enrollment and application shares. The instruments for mean utility, $\mathbf{Z_{jl}}$, contain all exogenous college characteristics in the utility function as well as the instrument for price as discussed in Section 4. I assume that these instruments are uncorrelated with the unobserved component of mean utility, ξ_{jl}, and thus can use the moment condition $E[\mathbf{Z_{ijl}} \xi_{jl}] = 0$. Similarly, I assume that $E[\mathbf{X^c_{jl}} \xi^c_{jl}] = 0$ in the application cost function.

To estimate the coefficients on interactions between college and student characteristics, I match moments from the student-level data on enrolled students that capture how student and college characteristics co-vary. Specifically, I match the average log distance to college among enrollees; coefficients from a regression of student income on a public in-state indicator, the log of the college's instructional spending, the student ACT score, and a constant; and coefficients from another regression of enrolled student ACT scores on the log of the college's instructional spending. The values of these moments are shown in Panels A through C of Table 3.[33]

To estimate the aid policy parameters, I match regression coefficients from another auxiliary model estimated in the HSLS:09 data. In particular, I regress the student's college aid as reported in HSLS:09 on X^{aidc}_{1j}, X^{aidc}_{2j} interacted with student test scores (s_i), and X^{aidc}_{3j} interacted with family income (inc_i). The

[33]I target all of these regression coefficients with the exception of the coefficient on the ACT score in the student income regression, which is included in the auxiliary model as a control.

college-level variables in X include several measures constructed from the net tuition and aid data from IPEDS. In particular, X_{1j}^{aidc} includes a constant and the average college aid as a proportion of tuition. X_{2j}^{aidc} and X_{3j}^{aidc} include a constant and the average aid (as a proportion of tuition) in the income groups above and below \$75k.[34] I also match the proportion of students at each of the four discretized college aid levels, in order to estimate the cutpoints in the college aid function. The estimates for the auxiliary college aid regression and the proportion of students in each aid category are shown in Table 3.

To summarize, I form the GMM objective function using the following moments:

$$E[\mathbf{Z_{ijl}}\xi_{jl}] = 0$$
$$E[\mathbf{X^c_{jl}}\xi^c_{jl}] = 0$$
$$E[\mathbf{b^{model}} - \mathbf{b^{aux}}] = 0,$$

where $\mathbf{b^{aux}}$ are the moments from the HSLS:09 data, and $\mathbf{b^{model}}$ are the model counterparts. The model is exactly identified; the number of target moments is equal to the number of parameters.

To use the first two moments above requires having the values of ξ_{jl} and ξ^c_{jl} as a function of the model parameters. This is a common situation when estimating models of demand from market shares. Analytic inversion of the market shares for the mean utility is often impossible or prohibitively difficult, as it is in my case (Berry, Levinsohn, and Pakes 1995). In my setting, there is an additional twist relative to the traditional framework, as there are two state-by-college level data shares to match: enrollment shares and application penetration. Recall that ν_{jl}, the mean value of utility at the college-market level, is linear in certain parameters ($\theta^{l,\nu}$), data, and the unobserved college quality ξ_{jl}. The mean application cost, c_{jl} is a function of another set of linear parameters ($\theta^{l,c}$) and the college-state level unobservable in application cost. Both the mean utility and the mean application cost are functions of a common set of non-linear parameters, θ^{nl}.

Equations 10 and 12 cannot be inverted analytically for $\boldsymbol{\nu}$ and $\mathbf{c}$, but a fixed-point algorithm similar to those used in Rust (1987) and Berry, Levinsohn, and Pakes (1995) allows numeric inversion, conditional on values for θ^{nl}. I stack the following two equations and iterate until convergence:

$$\boldsymbol{\nu}^{h+1} = \boldsymbol{\nu}^h + \log(\mathbf{E}) - \log(e(\boldsymbol{\nu}^h, \mathbf{c}^h, \theta^{nl}))$$
$$\mathbf{c}^{h+1} = \mathbf{c}^h - (\log(\mathbf{App}) - \log(app(\boldsymbol{\nu}^h, \mathbf{c}^h, \theta^{nl})),$$

where E and App are the enrollment shares and application penetration in the data, respectively, and $e(\boldsymbol{\nu}^h, \mathbf{c}^h)$ and $app(\boldsymbol{\nu}^h, \mathbf{c}^h)$ are the model-derived counterparts. Once the $\boldsymbol{\nu}$ and $\mathbf{c}$ are obtained, the linear utility and application cost parameters are found using IV-GMM, where the $\boldsymbol{\nu}$ and $\mathbf{c}$ are the dependent variables. The unobserved quality and unobserved component of application cost are constructed as the residuals.

Once I obtain the residuals, I construct the GMM objective function from the moments listed above.

[34] Before taking this average aid as a proportion of tuition, I reduce the aid amount in each income category by the predicted state grant aid in order to better approximate college aid.

Using the strategy of Berry, Levinsohn, and Pakes (1995), I search for the parameters θ^{nl} that set this GMM objective function to zero. In every iteration of the search for θ^{nl}, I compute the residuals ξ_{jl} and ξ^{c}_{jl} using the fixed-point algorithm described above.

5.5 College Choice and Financial Aid Policy Parameters: Results

Table 4 shows the estimates for the college aid function. From these estimates, we see that colleges observed to have higher aid (as a proportion of tuition) for students with lower income levels have aid policies with steeper slopes in both income and ACT scores. Table 5 shows the enrollment utility estimates. These are most easily interpreted through a few simulations for the demand-side alone. Table 6 shows the total applications and enrollment under two counterfactual simulations assuming there are no supply-side adjustments. In the first simulation, I decrease tuition by 1% at all colleges, while maintaining all of the same financial aid policies, so that financial aid adjusts proportionately to tuition. In Table 6, we see that a 1% decrease in tuition at all colleges is predicted to increase the number of applications sent by 5,817 (0.11%) and the number of students enrolled by 7,019, or 0.70% of baseline enrollment. Thus market demand is modestly inelastic with respect to tuition. Decreasing the admission thresholds by 1% at all colleges increases applications by 4,501 (0.09%), and enrollment by 5,180 (0.52%). It is reasonable that college attendance overall is not very sensitive to market-wide changes in admissions selectivity; many colleges already have low admissions standards relative to the distribution of student test scores, implying that admissions standards aren't a significant barrier to attendance at *any* four-year college in the sample. However, enrollment at specific high-selectivity colleges may be much more sensitive to admissions thresholds.

The bottom five rows in Table 6 show the effects of the two simulated changes on enrollment by income level. The highest-income students (family incomes above $110k) are not very price-sensitive, with a price elasticity of enrollment equal to 0.46, while students from the lowest-income households have a price elasticity of enrollment equal to 1.01. There is also a strong relationship between income and the enrollment effect of an admissions threshold decrease. A 1% decrease in admissions thresholds across all colleges increases enrollment for the lowest-income group by 0.77%, but only by 0.35% for the highest-income group. This pattern is sensible because students from high-income families are more likely to go to college in the baseline, both because of differences in the utility of attending college and because higher-income students have higher test scores on average are therefore more likely to meet admission standards.

The two simulations shown so far are market-level changes, but I can also use the model to examine own- and cross-price elasticities at the college level. To illustrate typical patterns, I calculate these elasticities for four example colleges in Louisiana and Virginia, including the flagship public college and a notable private university in each state: Louisiana State University (LSU) and Loyola in Louisiana, and University of Virginia (UVA) and Liberty University in Virginia. These calculations adjust tuition downwards by 1%, assuming that financial aid policies remain the same (in other words, aid will adjust proportionally with tuition). Based on these calculations, I find that at LSU, the own-price elasticity is 0.89, while at Loyola it is 1.79, reflecting differences in selectivity at the colleges (LSU is more selective than Loyola), in baseline tuition (LSU had a low in-state public tuition of $7,873 while Loyola charged a much steeper private tuition of $36,610), and in the number of states in which the university competes. I find a similar pattern for

Virginia, where the flagship public institution has a lower elasticity than a well-known private institution: UVA has an own-price elasticity of 1.28 while Liberty has an own-price elasticity of 2.26.

Figure 4 shows cross-price elasticities for the two Louisiana colleges, presented as the percent change in enrollment due to a 1% decrease in the respective colleges' tuition. Panel A shows the top 25 colleges in terms of cross-price elasticity with respect to LSU. The largest cross-price elasticities (around -0.12) are for regional public colleges in Louisiana (e.g., University of Louisiana in Lafayette, Nicholls State University, and University of New Orleans) and relatively non-selective private colleges in the New Orleans area, such as Dillard University and Xavier University. Panel B shows the top 25 colleges in terms of cross-price elasticities with respect to Loyola (New Orleans). Note the difference in the y-axis between LSU and Loyola: since Loyola's enrollment is relatively low and is spread out across the country, the cross-price elasticity between Loyola and any specific competitor university is smaller. However, there is still a strong regional component to Loyola's enrollment, making moderately or less-selective colleges in Louisiana its strongest competitors, as measured by cross-price elasticity. Other colleges that show up in the top 25 include colleges in Florida, Georgia, and Mississippi. Figure 5 shows a similar graph for Virginia. The patterns across college types are similar to what we saw for Louisiana: changes in tuition at the flagship public university primarily affect other public colleges as well as private colleges located in-state. Effects on out-of-state public colleges and selective private colleges are smaller because their enrollment is dispersed across many states. However, one thing we can see in the Virginia example that is not as clear in the Louisiana example is that the cross-price elasticities do depend on how similar the two institutions are in terms of selectivity. UVA has a 25th percentile ACT score of 28, while Liberty's is 19. The colleges with the strongest cross-price elasticities are those most similar in terms of selectivity and location.

In the final two demand-side exercises, I simulate the effects of a $1,000 increase in financial aid for in-state students to use at public colleges. I do this separately for public colleges in Louisiana and in Virginia, holding the supply-side fixed. Although I could do this type of simulation for any state in the US, I limit the counterfactuals to two example states to keep the computation and presentation of results tractable, and I will maintain the same example states in the equilibrium counterfactuals for comparison. I chose Louisiana and Virginia as the examples because these states are both mid-sized states with very different public systems in terms of market concentration measures. The fact that these are mid-sized states is helpful in a practical sense, as the equilibrium counterfactuals take longer to run with a larger set of colleges. These two states are interesting to compare because the public systems have different structures. In Louisiana, LSU dominates the market, enrolling upwards of 9% of all high school graduates. The other public colleges in Louisiana are distinctly smaller, less selective, and regional. In Virginia, the public system is quite robust with enrollment spread across multiple high-quality public colleges, some of which rank among top public colleges nationally— in particular, William and Mary, University of Virginia, and Virginia Tech. In addition to these colleges that stand out on a national level, there is a second group of public colleges in Virginia that are relatively large and still moderately selective, including George Mason University, James Madison University, and Virginia Commonwealth University. With these two example states chosen, I move forward with demand-side simulations in each state, where the state government offers all in-state students $1,000 to attend an in-state public college.

Figure 6 shows the effect on enrollment in each state. The additional grant aid increases enrollment between 7 and 12 percent at Louisiana in-state public colleges, and at Virginia public colleges the range is a bit wider, between nearly 6 and 12 percent. The demand-side subsidy has little effect on highly-selective private colleges that compete in these two states. This is both because these colleges have very strict admissions standards–limiting the number of potential students who could be admitted and who are induced to apply through this subsidy—and because they draw students from all over the country, limiting the potential for a single state to significantly impact total enrollment. Similarly, out-of-state public colleges don't see a large direct demand-side effect. However, less-selective private colleges lose enrollment to the public colleges, up to 6% of current enrollment. The net effect with demand-side responses alone is to increase enrollment in Louisiana by 8.5% and enrollment in Virginia by 6.6%. However, supply-side responses will dictate how much of the state grant aid passes through to students, and whether admissions standards and institutional aid will adjust. These supply-side responses may have substantial implications for equilibrium enrollment. Thus, in the next two sections, I turn to the supply-side estimation and equilibrium simulations, comparing the results to those with demand-side responses alone.

5.6 Supply-Side Estimation and Results

On the supply-side, the model features college-specific marginal cost and preferences over profit, admissions thresholds and academic fit, and average family income of the student body, as shown in Equation 13. The estimation process for the supply-side parameters involves solving the first-order conditions of the college problem. For private colleges and public colleges that compete only in-state, the system of first-order conditions has four equations corresponding to the choice of tuition, admission threshold, and the financial aid policy coefficients on income and admission test scores. For public colleges that compete both in-state and out-of-state, the first-order conditions are a system of five equations, with the additional equation coming from the choice of the out-of-state tuition level. Letting $\boldsymbol{\theta_j^{sup}}$ denote the vector of choice variables for the college, the system of first-order conditions can be written as:

$$\begin{aligned}(1-\lambda_{1j}-\lambda_{2j}-\lambda_{3j})\frac{1}{\pi_j}\frac{\boldsymbol{\partial \pi_j}}{\boldsymbol{\partial \theta_j^{sup}}}+\lambda_{1j}\frac{1}{P(\underline{s}_j)}\frac{\boldsymbol{\partial P(\underline{s}_j)}}{\boldsymbol{\partial \theta_j^{sup}}}\\+\lambda_{2j}\frac{1}{P(\bar{s}_j)-P(\underline{s}_j)}\frac{\boldsymbol{\partial P(\bar{s}_j)-P(\underline{s}_j)}}{\boldsymbol{\partial \theta_j^{sup}}}+\lambda_{3j}\frac{1}{f(\mu_j^{inc})}\frac{\boldsymbol{\partial f(\mu_j^{inc})}}{\boldsymbol{\partial \theta_j^{sup}}}=\mathbf{0}\end{aligned} \tag{16}$$

The last three terms are functions of college-specific preference parameters (the λ's) and objects that can be calculated directly from the demand system at current values of the college choices. In other words, I can compute derivatives in the last three terms numerically, and then treat these derivatives as data.

The first term is not quite as simple because it includes an expression for marginal cost, including the cost parameters that must be estimated. For public colleges, this term is also a function of the heterogeneous cost differential for in-state students ($\alpha_4+\omega_j^{is}$), but for clarity, I will focus on the case of private colleges. In this case, the cost is: $q_j(\alpha_1+\alpha_2 q_j+\alpha_3 X_j^s+\omega_j)$. We can divide the cost into the linear part, $(q_j * \nu_j^s)$, where $\nu_j^s=\alpha_1+\alpha_3 X_j^s+\omega_j$, and the part that is non-linear in enrollment $(\alpha_2 q_j^2)$. Given a value for the non-linear parameter α_2, I can write $\frac{1}{\pi_j}\frac{\partial \boldsymbol{\pi_j}}{\partial \boldsymbol{\theta_j^{sup}}}$ as a function of demand-side derivatives and the linear component of cost, ν_j^s.

Given a value of α_2, I can solve the system of first-order conditions for each college to get the values of λ_{1j}, λ_{2j}, and λ_{3j}, along with the linear components of marginal cost ν_j^s, and the cost differential for in-state students $(\alpha_4 + \omega_j^{is})$.[35] Once these solutions are obtained, I estimate α_1 and $\boldsymbol{\alpha_3}$ using OLS, where the dependent variable is $\nu_j^s - 2\alpha_2 q_j$. Then, I can generate the marginal cost residuals ω_j.

I estimate the value of α_2 using GMM. Letting Z_j^{sup} be the instrument for enrollment in the marginal cost function, the moment used to estimate α_2 is $E[Z_j^{sup}\omega_j] = 0$. The process of estimating α_2 is iterative: for each value of α_2, I find the solutions to the first-order conditions using the process described above and generate the marginal cost residuals. Using those residuals, I find the value of the GMM objective function. I search over values of α_2 until the GMM function is minimized.

The results of this process are summarized in Table 7, which shows the mean value of the preference parameters, the cost differential for in-state students, and the coefficients in the marginal cost function. The first three rows show the mean preference parameters (exponents λ_1, λ_2, and λ_3 in the college objective function) across colleges for the measure of academic fit, admission rate, and a function of mean income. The estimates suggest that on average, colleges seek to enroll a student body that is relatively lower-income and higher academic fit than what profit maximization alone would imply. Figure 7 shows the distribution of the preference parameters, demonstrating that there is substantial heterogeneity across colleges.

The next section of Table 7 shows the marginal cost estimates. Private colleges have a much higher marginal cost on average than public colleges, by approximately \$20,565 per student. Colleges with a higher completion rate also have much higher marginal cost: a one percentage point increase in the completion rate is associated with a marginal cost increase of approximately \$194. It is likely that private colleges and colleges with higher completion rates offer more support services and amenities to students, raising the cost per student. Marginal cost is increasing the number of enrollees ($\alpha_2 = 0.003$), although the estimate is not statistically significant. The last row of the table shows that the average benefit to public colleges for enrolling an in-state student, relative to an out-of-state student, is approximately \$10,145.

6 Counterfactual Financial Aid

In the last several decades, many state governments have created and expanded financial aid programs that subsidize students who attend in-state colleges. These programs typically favor public colleges, although some programs do offer aid for attending in-state private colleges. Such policies affect student decisions, but they also affect college decisions. To illustrate the economic importance of the supply-side responses, I simulate the equilibrium responses to additional financial aid offered by two example states. These simulations assume colleges respond to each other's choices in a Nash setting. I use iterated best-response to solve for the equilibrium.[36] As discussed in Section 5.5, I chose Louisiana and Virginia as the example states. Mirroring the demand-side simulations, I simulate the effects of an additional \$1,000 grant

[35] I solve the first-order conditions using numerical methods. I also confirm numerically that the second-order conditions are met.

[36] The results are robust to starting values that assume public universities either fully absorb the state grant aid with tuition increases or allow full pass-through to students (no change in tuition).

to in-state students to attend in-state public colleges, separately in each state. As shown in Section 5.5, a comparable demand-side simulation (holding all college responses fixed) generated an 8.5% increase in enrollment in Louisiana and a 6.6% increase in Virginia.

When colleges competing for in-state students respond to the counterfactual policy, I find that total Louisiana enrollment increases by 4.1% and total Virginia enrollment increases by 3.4%, a large reduction relative to the demand-side effects alone. This result underscores the importance of the supply-side response when estimating the effect of subsidies in this market. Public colleges benefit from the additional revenue generated by the demand-side subsidy, choosing to raise tuition and increase admissions standards in response.

Table 8 shows the levels of the college choices before and after the policy, for each college that competes in Louisiana per the definitions discussed in Section 2.3. Focusing first on in-state public colleges, we see that these colleges consistently increase the posted tuition levels and increase admissions standards. Among the two colleges that compete for out-of-state students, the out-of-state tuition also increases. Of the \$1,000 that is initially provided as a subsidy to students, between \$300 and \$800 passes through to increased tuition. This does not necessarily translate to equivalent increases in the average price paid by students, however, because grant aid adjusts as a proportion of tuition, as in the model. While some public colleges (LSU, for example) increase the degree to which merit and need are considered in the aid function, most colleges increase the admissions threshold while simultaneously reducing the ACT and income gradients in the aid function. Patterns for Virginia are similar, as shown in Table 9, except that the responses are more similar across the public universities, reflecting the fact that no single university dominates in Virginia. Ultimately, the net tuition received by public colleges in each state increases by over \$400 per student due to the introduction of the grant.

The in-state private colleges that compete most closely with public colleges —for example, Dillard University and Centenary College in Louisiana—generally respond to the negative demand shock by reducing tuition. Most out-of-state public and private colleges do not change their choices substantively in the counterfactual, but this is not surprising as the initial demand-side effects were small for many of the nationally competitive private and out-of-state public colleges.

Figure 8 shows the equilibrium enrollment effects in a way that is directly comparable to the demand-side results shown in Figure 6. After the colleges respond to the initial demand shock, the combination of increased tuition and increased admission thresholds at the public colleges reduces the enrollment effects substantially. For example, LSU's enrollment is predicted to increase by 8% under the demand-only simulation, but increases by only approximately 2% in equilibrium. Comparing the two figures shows that responses by affected private colleges lessen the initial enrollment impact of their negative demand shock.

7 Conclusions

The effects of state-level financial aid programs are complex, involving responses by students and colleges. Standard market equilibrium models are not easily applied to this mixed market of public and nonprofit private firms that care about the characteristics of the students they enroll. This paper develops an equilibrium model that is appropriate for the unique aspects of this particular market. The model is distinct from previous models of higher education, as it takes into account competition between colleges with substantial heterogeneity in characteristics along with the geographic nature of the market.

The college choice framework involves three stages: application, admission, and enrollment. While the expressions derived from this model are not the same as those found in typical logit demand applications, I demonstrate that a combination of aggregate (state-level) data on enrollment and applications and a modified version of the fixed-point algorithm used in Berry, Levinsohn, and Pakes (1995) can be used to estimate the parameters of this model. I supplement the aggregate data with moments from an individual student survey to estimate college financial aid policies and allow for several dimensions of student-level heterogeneity in utility. I also use student-level survey data to estimate parameters governing the admission process.

On the supply-side, individual colleges set prices, aid policies, and admissions selectivity in a strategic environment. Colleges (potentially) value profit, their admission standards and academic characteristics of students, and the mean income of the study body. I find heterogeneous preferences for each of these measures as solutions to a system of first-order conditions for each university's problem.

The estimated demand-side effects of financial aid are large: as an example, a $1,000 scholarship to students in Louisiana to attend an in-state public college increases enrollment by 8.5%. The same experiment in Virginia increases enrollment by 6.6%. When public colleges are allowed to respond to this policy by adjusting tuition, aid policies and admissions standards, they raise tuition and admissions standards. Adjustments to the aid function vary among colleges with different preferences. Increases in posted tuition range from approximately $300 to nearly $800 at the high end. However, the model also suggests that admissions is another important dimension along which colleges may respond to demand shocks. Colleges value student ability, so when demand increases, colleges respond by raising admissions standards. The private colleges most affected by the policy generally respond by decreasing tuition. Taking all of these market responses into account, the final estimated enrollment effect of the policy is a 4.1% increase in overall enrollment in Louisiana and a 3.4% increase in Virginia. The private colleges that compete most closely with public colleges lose as much as 3% of their enrollment.

This model has limitations that provide direction for additional research. For example, the model does not allow for exit. Thus, when small private colleges are adversely affected by counterfactual policies that provide subsidies to public colleges, the model predicts that they would stay open but adjust their prices, aid policies, and admission thresholds. In some cases, private colleges may in fact close with relatively large changes that favor public colleges. Another potential direction for future research is to incorporate endogenous quality differentiation, as this model holds instructional spending and other non-price, non-admissions characteristics fixed in equilibrium. Finally, this model has assumed that there is no collusion or joint setting of endogenous variables. This may not hold in all cases, particularly for public colleges run by governing boards or for colleges that have strong incentives for collusion over tuition, admissions standards, and aid policies. There is a need for research examining the prevalence and possible effects of joint decision-making among colleges.

References

Arcidiacono, Peter (2005). “Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?” *Econometrica* 73.5, pp. 1477–1524.

Bennett, William J. (1987). *Our Greedy Colleges*.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995). “Automobile Prices in Market Equilibrium”. *Econometrica* 63.4, pp. 841–890.

Berry, Steven T (1994). “Estimating Discrete-Choice Models of Product Differentiation”. *The RAND Journal of Economics* 25.2, pp. 242–262.

Blair, Peter Q. and Kent Smetters (2021). “Why Don’t Elite Colleges Expand Supply?”

Card, David and Alan 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, pp. 416–434.

Cellini, Stephanie Riegg and Claudia Goldin (2014). “Does federal student aid raise tuition? New evidence on for-profit colleges”. *American Economic Journal: Economic Policy* 6 (4), pp. 174–206.

Chetty, Raj, John N. Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan (2020). “Income Segregation and Intergenerational Mobility across Colleges in the United States”. *Quarterly Journal of Economics* 135.3, pp. 1567–1633.

Cook, Emily E. and Sarah Turner (2022). “Progressivity of pricing at US public universities”. *Economics of Education Review* 88.

Epple, Dennis, Richard Romano, Sinan Sarpça, and Holger Sieg (2017). “A general equilibrium analysis of state and private colleges and access to higher education in the U.S.” *Journal of Public Economics* 155.September 2013, pp. 164–178.

Epple, Dennis, Richard Romano, Sinan Sarpça, Holger Sieg, and Melanie Zaber (2019). “Market power and price discrimination in the US market for higher education”. *RAND Journal of Economics* 50.1, pp. 201–225.

Epple, Dennis, Richard Romano, and Holger Sieg (2006). “Admission, Tuition, and Financial Aid Policies in the Market for Higher Education”. *Econometrica* 74.4, pp. 885–928.

Fan, Ying (2013). “Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market”. *American Economic Review* 103.5, pp. 1598–1628.

Fillmore, Ian (May 2023). “Price Discrimination and Public Policy in the US College Market”. *The Review of Economic Studies* 90 (3), pp. 1228–1264.

Fitzpatrick, Maria D and Damon Jones (2012). “High education, merit-based scholarships and post-baccalaureat migration”.

Fu, C (2014). "Equilibrium Tuition, Applications, Admissions, and Enrollment in the College Market". *Journal of Political Economy* 122.2, pp. 225–281.

Fu, Chao, Junjie Guo, Adam J. Smith, and Alan Sorensen (2022). "Students' heterogeneous preferences and the uneven spatial distribution of colleges". *Journal of Monetary Economics* 129, pp. 49–64.

Gentzkow, Matthew (2007). "Valuing New Goods in a Model with Complementarity: Online Newspapers". *American Economic Review* 97.3, pp. 713–744.

Goldin, Claudia and Lawrence Katz (2008). *The Race between Education and Technology*. Belknap Press for Harvard University Press.

Gordon, Grey and Aaron Hedlund (2019). "Accounting for the Rise in College Tuition". In: *Education, Skills, and Technical Change: Implications for Future U.S. GDP Growth*. Ed. by Charles Hultan and Valerie Ramey. University of Chicago Press, pp. 357–394.

— (2022). "Accounting for Tuition Increases across U.S. Colleges".

Heckman, James J, Lance J Lochner, and Petra E Todd (2006). "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond". In: *Handbook of the Economics of Education*. Vol. 1.

Herbaut, Estelle and Koen Geven (2020). "What works to reduce inequalities in higher education? A systematic review of the (quasi-)experimental literature on outreach and financial aid". *Research in Social Stratification and Mobility* 65.November 2019, p. 100442.

Howell, Jessica S. (2010). "Assessing the Impact of Eliminating Affirmative Action in Higher Education". *Journal of Labor Economics* 28.1, pp. 113–166.

Hoxby, Caroline and Sarah Turner (2019). "The Right Way to Capture College "Opportunity"". *Education Next* 19.2.

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

Kapor, Adam (2024). "Transparency and Percent Plans".

Long, Bridget Terry (2004). "How Do Financial Aid Policies Affect Colleges? The Institutional Impact of the Georgia HOPE Scholarship". *Journal of Human Resources* 39.4, pp. 1045–1066.

Ma, Jennifer and Matea Pender (2022). *Trends in College Pricing and Student Aid 2022*. Tech. rep.

Page, Lindsay C. and Judith Scott-Clayton (2016). "Improving college access in the United States: Barriers and policy responses". *Economics of Education Review* 51, pp. 4–22.

Rothschild, Michael and Lawrence J White (1995). "The Analytics of the Pricing of Higher Education and Other Services in Which the Customers Are Inputs". *Journal of Polical Economy* 103.3, pp. 573–586.

Rust, John (1987). "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher". *Econometrica* 55.5, pp. 999–1033.

Singell, Larry D. and Joe A. Stone (June 2007). "For whom the Pell tolls: The response of university tuition to federal grants-in-aid". *Economics of Education Review* 26.3, pp. 285–295.

Smith, Jonathan, Matea Pender, and Jessica Howell (2018). "Competition among Colleges for Students across the Nation". *Southern Economic Journal* 84.3, pp. 849–878.

Turner, Lesley J (2017). "The Economic Incidence of Federal Student Grant Aid".

Turner, Nicholas (Aug. 2012). "Who benefits from student aid? The economic incidence of tax-based federal student aid". *Economics of Education Review* 31.4, pp. 463–481.

Table 1: College Characteristics

	Public		*Private*	
	Mean	SD	Mean	SD
Panel A: Price, Aid, and Selectivity				
Tuition	8.37	2.45	32.47	8.78
Out-of-State Tuition	19.85	7.31	.	.
Room and Board	9.78	2.07	11.42	2.45
Avg. Grant All Sources, Income 0-30k and Fed Aid Recipient	11.48	3.93	31.01	13.03
Avg. Grant All Sources, Income 30-48k and Fed Aid Recipient	10.07	3.58	30.03	12.84
Avg. Grant All Sources, Income 48-75k and Fed Aid Recipient	7.03	3.14	27.24	12.02
Avg. Grant All Sources, Income 75-110k and Fed Aid Recipient	4.76	2.35	23.50	9.96
Avg. Grant All Sources, Income 110k+ and Fed Aid Recipient	4.24	2.16	18.69	6.17
25th Percentile ACT	20.11	2.68	22.31	3.70
Panel B: Other Measures				
Instructional Exp. per FTE Student	9.41	4.53	14.36	13.76
Completion Rate within 150% of normal time	0.56	0.15	0.68	0.15
Any Grad Program	0.97	0.17	0.82	0.38
Religious	0.00	0.00	0.57	0.49
Historically Black College or University	0.07	0.26	0.03	0.17
Carnegie Classification: R1	0.17	0.37	0.07	0.26
Carnegie Classification: R2	0.16	0.37	0.04	0.21
Doctoral/Research Universities	0.06	0.25	0.06	0.24
Carnegie Classification: Masters (larger programs)	0.33	0.47	0.26	0.44
Carnegie Classification: Masters (medium programs)	0.10	0.31	0.11	0.31
Carnegie Classification: Masters (smaller programs)	0.06	0.24	0.04	0.19
Carnegie Classification: BA–Arts & Sciences	0.03	0.17	0.30	0.46
Carnegie Classification: BA–Diverse Fields	0.06	0.25	0.12	0.32
NCES Locale: City—Large	0.15	0.36	0.24	0.42
NCES Locale: City—Midsize	0.14	0.35	0.13	0.33
NCES Locale: City—Small	0.19	0.40	0.14	0.35
NCES Locale: Suburb	0.20	0.40	0.28	0.45
NCES Locale: Town or Rural	0.31	0.46	0.21	0.41
Observations	421		454	

NOTE: The table summarizes the characteristics of 4-year degree-granting private non-profit or public colleges in this analysis. All monetary values are in thousands of 2019 $.

Table 2: Admissions Parameters

	estimate	se
Variance of academic fit, σ_{ζ}^{2}	7.015	0.234
Admission Threshold: Constant	22.14	4.191
Admission Threshold: 25th Percentile ACT	-2.107	0.358
Admission Threshold: 25th Percentile ACT, Sq	0.078	0.008
Obs	9,790	

NOTE: Results obtained from a Probit regression of student admission outcomes on the student's ACT score and a quadratic in the 25th percentile of ACT scores at the college. The coefficient on the student ACT score is constrained to be 0, consistent with the empirical model. This allows for the identification of the variance in the Probit regression.

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09), High School Transcript and Second Follow-up. U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System 2013, Institutional Characteristics and Admissions Surveys.

Table 3: Auxiliary Regressions and Other Target Moments from HSLS:09

	estimate	se
Panel A: Means		
log(Distance)	-0.685	0.030
Panel B: log(Income) Regression Coefficients		
Public In-State	-0.092	0.040
log(Instructional Spending)	0.108	0.040
ACT	0.037	0.004
Constant	-1.297	0.116
Panel C: ACT Regression Coefficients		
log(Instructional Spending)	4.040	0.157
Constant	14.500	0.388
Panel D: College Aid Regression Coefficients		
Average College Aid	0.555	0.061
Income	0.084	0.031
Income * Aid for Income $<$ \$75k	-0.131	0.037
Income * Aid for Income $\geq$ \$75k	0.025	0.054
exp(ACT/100)	1.667	0.174
exp(ACT/100) * Aid for Income $<$ \$75k	-0.062	0.040
exp(ACT/100) * Aid for Income $\geq$ \$75k	0.157	0.065
Constant	-2.118	0.216
Panel E: Share of Students by College Aid as Proportion of Tuition		
College Aid = 0	0.439	0.013
College Aid = 0.3	0.219	0.011
College Aid = 0.6	0.222	0.011
College Aid = 1.1	0.119	0.008

NOTE: The table shows the moments that I match from student survey data. Panel A shows the mean of the log distance to college for all enrolled students, where distance is measured in hundreds of miles. Panel B shows the coefficients from a regression of student income on college and student characteristics, among all enrolled students. Panel C shows the coefficients from a regression of student ACT scores on the log of instructional expenditures and a constant. The final two panels show the target moments for estimating the college aid function: Panel D shows the coefficients from an OLS regression of college aid (as a proportion of tuition) on listed variables, including student characteristics (ACT and income level), and college-level measures of aid generosity, including average college aid as a proportion of tuition (labeled as ”Average College Aid”), as well as average aid as a proportion of tuition after all sources of aid, by income group. Panel E shows the proportion of students by discrete aid category.

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09), Base Year, First Follow-Up, and Student Records.

Table 4: Institutional Aid Policy Estimates

	estimate	se
Panel A: Coefficients		
Average Institutional Aid as Share of Tuition	1.313	0.703
Income	0.692	0.274
Income * Institution Avg Net Tuition for Income $<$ \$75k	-1.453	0.687
Income * Institution Avg Net Tuition for Income $\geq$ \$75k	0.616	0.326
exp(ACT/100)	5.475	0.507
exp(ACT/100) * Institution Avg Net Tuition for Income $<$ \$75k	0.771	0.546
exp(ACT/100) * Institution Avg Net Tuition for Income $\geq$ \$75k	-0.398	0.352
Panel B: Cut Points		
k1	7.758	0.265
k2	8.530	0.189
k3	9.538	0.111

NOTE: The table shows estimates of the college aid policy parameters. Panel A shows the coefficients and Panel B shows the cut points for the ordered probit college aid function (see Equation 2).

Table 5: Utility Parameter Estimates

	coef	se
Panel A: Mean Utility Estimates		
Constant	9.414	0.733
Religious	0.119	0.067
Room and Board	-0.080	0.015
Public	-1.091	0.088
In-State	-1.839	0.056
Any Graduate Program	-0.245	0.091
Completion Rate (150 % Normal Time)	1.187	0.281
log(Instructional Exp per Student)	-1.450	0.085
HBCU	-0.635	0.117
NCES Locale: City—Midsize	0.492	0.079
NCES Locale: City—Small	0.638	0.080
NCES Locale: Suburb	0.341	0.077
NCES Locale: Town or Rural	0.609	0.080
Panel B: Student-College Interactions		
Net Tuition	-0.180	0.032
ACT Score	-0.288	0.029
ACT × log(Instructional Exp per Student)	0.257	0.046
log(Distance)	-0.859	0.054
log(income)	1.089	0.290
log(income) × Net Tuition	0.05	0.033
log(income) × log(Instructional Exp per Student)	1.075	0.353

NOTE: The table shows the estimated enrollment utility parameters. State fixed effects and indicators for Carnegie Basic Classification categories are included in the regression but not displayed.

Table 6: Demand-Side Simulations

Outcome Measure	Baseline Values	1% Decrease in Tuition: Change	1% Decrease in Tuition: % Change	1% Decrease in Admissions Thresholds: Change	1% Decrease in Admissions Thresholds: % Change
Apps	5,153,747	5,817	0.11%	4,501	0.09%
Enroll (Total)	1,001,774	7,019	0.70%	5,180	0.52%
Enroll: $0-$30k	57,141	576	1.01%	442	0.77%
Enroll: $30-$48k	144,668	1,439	0.99%	1,047	0.72%
Enroll: $48-$75k	150,833	1,412	0.94%	1,020	0.68%
Enroll: $75-$110k	165,731	1,376	0.83%	975	0.59%
Enroll: $110k+	483,402	2,216	0.46%	1,695	0.35%

NOTE: The table shows application and enrollment changes with two counterfactual simulations on the demand-side, holding the college choices fixed.

Table 7: Supply-Side Estimates

parameter	estimate	se
Objective Function Parameters (Means)		
λ_1 (Mean of Admission Threshold Exponent)	0.192	0.006
λ_2 (Mean of Exponent on Academic Fit)	0.576	0.005
λ_3 (Mean of Exponent on Income Measure)	0.031	0.001
Marginal Cost Parameters:		
α_1: Constant	-18.43	7.074
α_2: Enrollment	0.003	0.003
α_3: Private	20.565	1.211
α_3: Completion Rate (150% Normal Time)	19.382	3.482
α_3: Religious	-5.594	1.111
α_3: Room and Board	0.153	0.212
α_3: Any Graduate Program	-0.091	1.595
α_3: HBCU	9.549	1.861
α_3: NCES Locale: City—Midsize	3.397	1.348
α_3: NCES Locale: City—Small	1.707	1.306
α_3: NCES Locale: Suburb	2.768	1.182
α_3: NCES Locale: Town or Rural	5.569	1.268
α_3: Carnegie Classification: R1	-2.854	6.760
α_3: Carnegie Classification: R2	-5.994	6.736
α_3: Doctoral/Research Universities	-8.386	6.805
α_3: Carnegie Classification: Masters (larger programs)	-7.422	6.668
α_3: Carnegie Classification: Masters (medium programs)	-5.915	6.743
α_3: Carnegie Classification: Masters (smaller programs)	-2.628	6.833
α_3: Carnegie Classification: BA–Arts & Sciences	0.605	6.588
α_3: Carnegie Classification: BA–Diverse Fields	-3.461	6.718
α_4 (In-state Student Cost Differential)	10.145	0.339

NOTE: The table shows estimates for the supply-side cost and objective function preference parameters, as described in Section 5.6.

Table 8: Counterfactual Responses: Louisiana

Type	Name	In-State Tuition		OOS/Private Tuition		Admission Threshold		Aid Coef on ACT		Aid Coef on Income		Enrollment	
		Before	After	Before	After	Before	After	Before	After	Before	After	Before	After
In-State Public	LA STATE U	7.873	8.653	25.790	26.785	14.930	14.943	6.158	6.191	-0.633	-0.681	4,906.8	5,004.4
	LA TECH U	7.302	7.730			12.369	12.399	5.950	5.871	-0.251	-0.230	1,228.9	1,302.3
	UNIV OF NEW ORLEANS	6.578	6.952			12.417	12.441	5.930	5.874	-0.214	-0.191	695.0	753.2
	NICHOLLS STATE	6.468	6.920			11.314	11.351	5.822	5.761	0.006	0.033	1,100.0	1,170.4
	U OF LA MONROE	6.318	6.711			11.314	11.354	5.824	5.743	-0.028	-0.011	1,159.9	1,230.0
	NORTHWESTERN STATE U	6.246	6.660			10.313	10.350	5.898	5.824	-0.173	-0.148	996.0	1,064.8
	U OF LA LAFAYETTE	6.192	6.698			12.417	12.449	5.869	5.741	-0.082	-0.038	2,469.9	2,561.7
	GRAMBLING STATE	5.950	6.323	15.118	15.645	8.404	8.424	5.886	5.935	-0.242	-0.258	534.0	579.6
	SOUTHEASTERN LA	5.715	6.216			11.314	11.352	5.846	5.712	-0.048	-0.005	2,288.9	2,380.7
	MCNEESE STATE	5.701	6.088			11.314	11.359	5.925	5.826	-0.194	-0.162	1,134.0	1,202.9
Out-of-State Public	UNIV OF AL	9.450	9.450	23.950	23.950	13.370	13.370	5.916	5.917	-0.228	-0.228	4,296.6	4,296.1
	UNIV OF MS	6.760	6.758	17.728	17.727	12.417	12.417	6.297	6.299	-0.947	-0.948	2,722.8	2,721.9
	UNIV SOUTHERN MS	6.744	6.738	15.024	15.036	11.273	11.273	6.053	6.062	-0.505	-0.510	1,368.9	1,366.5
	JACKSON STATE UNIV	6.348	6.343	15.552	15.555	8.944	8.944	6.026	6.028	-0.348	-0.349	803.4	802.7
	ALCORN STATE UNIV	6.108	6.107	6.108	6.117	8.488	8.487	5.965	5.965	-0.317	-0.317	419.0	418.8
	TX SOUTHERN UNIV	7.492	7.502	18.566	18.596	8.339	8.339	5.634	5.638	0.391	0.391	888.9	888.8
Private	SPRING HILL COLL			30.924	30.920	12.209	12.205	5.928	5.925	-0.264	-0.262	199.5	197.4
	HOWARD UNIV			22.783	22.783	12.417	12.417	5.675	5.675	0.169	0.169	1,328.5	1,327.6
	SPELMAN COLL			24.634	24.634	11.314	11.314	5.644	5.644	0.317	0.317	326.0	325.9
	EMORY UNIV			44.008	44.008	25.382	25.382	6.020	6.020	-0.422	-0.422	1,217.8	1,217.6
	CLARK ATLANTA UNIV			21.100	21.100	8.665	8.665	5.619	5.620	0.416	0.416	477.0	476.7
	XAVIER UNIV			20.560	20.456	10.813	10.815	5.679	5.673	0.270	0.275	402.5	395.4
	TULANE UNIV OF LA			46.930	46.930	24.619	24.619	5.821	5.821	-0.045	-0.045	1,410.5	1,407.7
	LOYOLA UNIV			36.610	36.447	13.677	13.677	5.843	5.839	-0.089	-0.087	346.9	341.5
	DILLARD UNIV			15.778	15.740	8.833	8.836	5.767	5.764	0.055	0.058	119.0	115.6
	CENTENARY COLL OF LA			30.740	30.559	13.384	13.386	5.865	5.859	-0.119	-0.116	88.0	86.1
	MS COLL			14.868	14.844	12.265	12.265	5.883	5.882	-0.172	-0.172	444.0	442.7
	MILLSAPS COLL			32.520	32.519	14.806	14.806	5.948	5.943	-0.306	-0.303	122.5	120.9
	VANDERBILT UNIV			42.978	42.978	34.888	34.888	6.154	6.154	-0.723	-0.723	1,165.6	1,165.3
	RHODES COLL			39.794	39.794	21.506	21.506	5.874	5.872	-0.153	-0.152	245.0	243.8
	TX CHRISTIAN UNIV			36.590	36.590	17.643	17.643	5.915	5.915	-0.197	-0.197	1,478.9	1,478.5
	SOUTHERN METHODIST UNIV			43.800	43.800	22.325	22.325	5.936	5.936	-0.269	-0.269	1,018.8	1,018.3
	RICE UNIV			38.941	38.941	32.062	32.062	6.152	6.152	-0.690	-0.690	694.9	694.7
	BAYLOR UNIV			35.972	35.972	17.494	17.494	5.801	5.801	0.000	0.000	2,466.5	2,466.4

NOTE: The table shows counterfactual choices of colleges before and after Louisiana offers $1,000 to Louisiana students to attend in-state four-year public colleges.

Table 9: Counterfactual Responses: Virginia

		In-State Tuition		OOS/Private Tuition		Admission Threshold		Aid Coef on ACT		Aid Coef on Income		Enrollment	
Type	Name	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After
In-State Public	WILLIAM AND MARY	14.267	14.673	38.440	39.205	24.524	24.528	6.382	6.412	-1.105	-1.149	1,131.9	1,173.1
	UVA	12.668	13.104	40.054	40.192	25.200	25.207	6.189	6.215	-0.740	-0.795	2,650.7	2,729.0
	VA COMMONWEALTH UNIV	12.002	12.547			12.417	12.436	5.833	5.774	-0.042	-0.030	2,923.8	3,019.7
	VA TECH	11.455	12.052	27.211	27.930	18.396	18.408	6.030	6.034	-0.413	-0.424	4,309.7	4,400.9
	LONGWOOD UNIV	11.340	11.774			11.111	11.136	5.796	5.772	0.034	0.039	996.9	1,065.0
	CHRISTOPHER NEWPORT UNIV	11.092	11.540			15.093	15.116	5.856	5.825	-0.055	-0.048	1,139.0	1,210.6
	GEORGE MASON UNIV	9.908	10.532	28.592	29.165	15.316	15.328	5.813	5.799	0.008	0.012	2,150.8	2,232.2
	UNIV OF MARY WA	9.720	10.142	22.590	23.009	13.677	13.704	6.130	6.138	-0.573	-0.582	851.0	910.3
	JAMES MADISON UNIV	9.176	9.731	23.654	24.388	15.093	15.110	6.149	6.154	-0.640	-0.660	3,877.2	3,963.1
	RADFORD UNIV	8.976	9.488			10.157	10.192	5.824	5.762	-0.037	-0.026	1,809.8	1,898.9
	OLD DOMINION UNIV	8.550	9.038			10.848	10.879	5.958	5.885	-0.272	-0.254	1,680.9	1,766.8
	VA STATE UNIV	7.784	8.206	17.192	17.679	8.147	8.165	5.893	5.910	-0.174	-0.172	832.9	888.7
	NORFOLK STATE UNIV	7.126	7.680	20.596	21.334	8.866	8.879	5.766	5.784	0.023	0.030	724.3	781.0
Out-of-State Public	CONCORD UNIV	6.160	6.162	13.490	13.532	9.578	9.576	5.904	5.906	-0.206	-0.202	395.0	394.5
	SHEPHERD UNIV	6.256	6.253	15.840	15.839	10.368	10.367	6.005	6.014	-0.377	-0.387	599.0	598.3
	COASTAL CAROLINA UNIV	9.760	9.745	22.770	22.770	10.368	10.368	5.863	5.864	-0.077	-0.077	1,761.9	1,761.0
	WV UNIV	6.456	6.444	19.632	19.619	11.837	11.836	6.149	6.151	-0.678	-0.680	3,933.6	3,931.6
	GA TECH	10.650	10.650	29.954	29.954	24.524	24.524	6.045	6.046	-0.453	-0.453	1,864.8	1,864.0
Private	RANDOLPH COLL			32.850	32.744	12.417	12.418	5.842	5.835	-0.108	-0.106	105.4	102.8
	RANDOLPH MACON COLL			34.850	34.849	12.417	12.418	5.771	5.763	0.048	0.052	323.8	315.7
	HAMPDEN SYDNEY COLL			37.352	37.351	13.677	13.678	5.796	5.791	-0.003	-0.002	276.8	270.3
	HOLLINS UNIV			33.320	33.216	12.417	12.418	5.828	5.821	-0.083	-0.081	69.4	67.8
	UNIV OF LYNCHBURG			33.565	33.564	10.368	10.368	5.789	5.787	0.017	0.018	404.8	397.0
	BRIDGEWATER COLL			29.090	29.045	11.314	11.316	5.836	5.835	-0.096	-0.095	424.8	419.0
	MARYMOUNT UNIV			26.430	26.383	11.314	11.316	5.770	5.767	0.055	0.056	244.9	241.7
	EMBRY RIDDLE			31.334	31.334	12.417	12.417	5.706	5.698	0.201	0.206	78.0	77.5
	ALDERSON BROADDUS UNIV			22.740	22.720	10.017	10.016	5.911	5.911	-0.210	-0.211	190.0	188.9
	WA AND LEE UNIV			44.507	44.507	27.995	27.995	6.189	6.189	-0.767	-0.767	251.9	250.6
	UNIV OF THE SOUTH			34.793	34.793	19.329	19.329	6.076	6.074	-0.527	-0.524	245.0	243.8
	HAMPTON UNIV			20.724	20.701	10.485	10.486	5.780	5.780	0.043	0.044	586.9	584.7
	UNIV OF RICHMOND			45.320	45.320	23.033	23.033	6.048	6.046	-0.479	-0.477	534.4	532.7
	VA UNION UNIV			14.930	14.946	7.968	7.933	5.612	5.717	0.374	0.305	256.9	256.2
	DAVIDSON COLL			42.849	42.849	24.058	24.058	6.125	6.123	-0.647	-0.645	247.7	247.1

NOTE: The table shows counterfactual choices of colleges before and after Virginia offers $1,000 to Virginia students to attend in-state four-year public colleges. There are 92 total colleges competing for students in Virginia. To keep the table length reasonable I show all in-state public colleges, the top 5 public out-of-state colleges and the top 15 private colleges, as ranked by percentage change in enrollment.

Figure 1: Model Timing

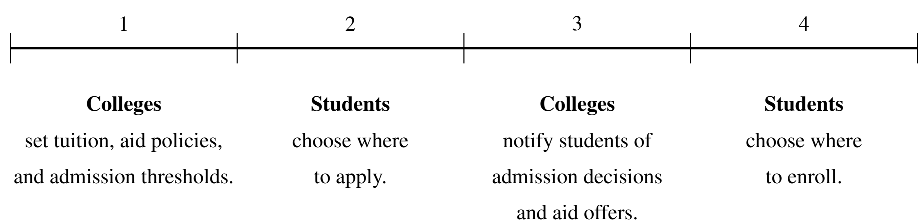
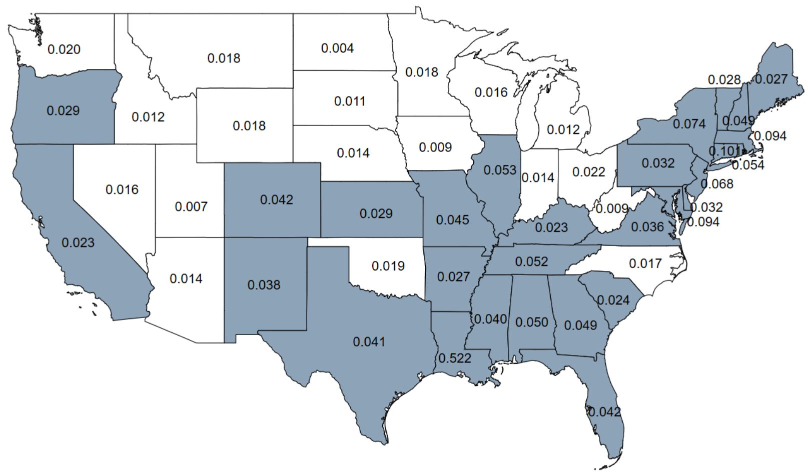
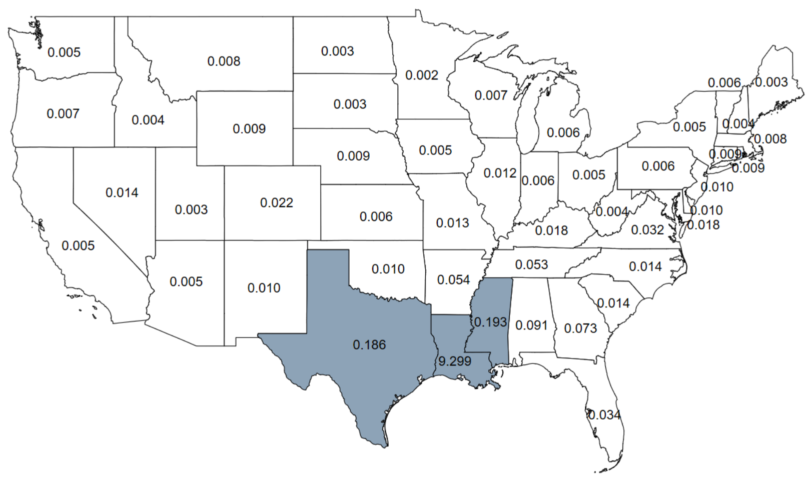


Figure 2: Market Definition

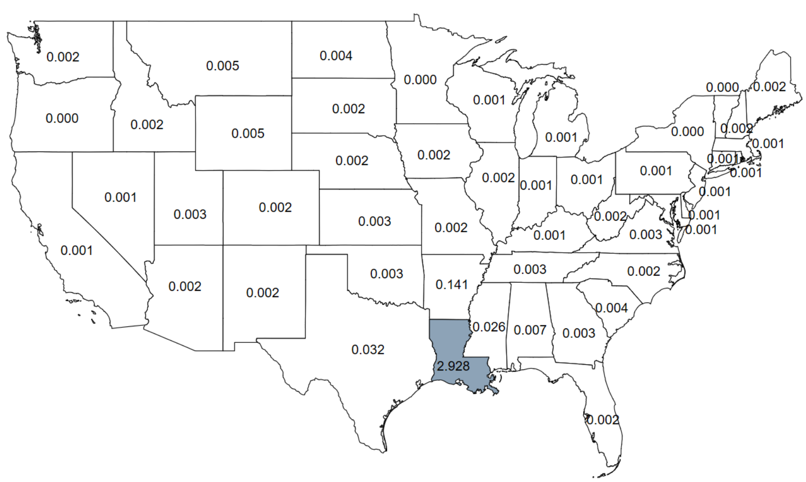
(a) Tulane



(b) LSU

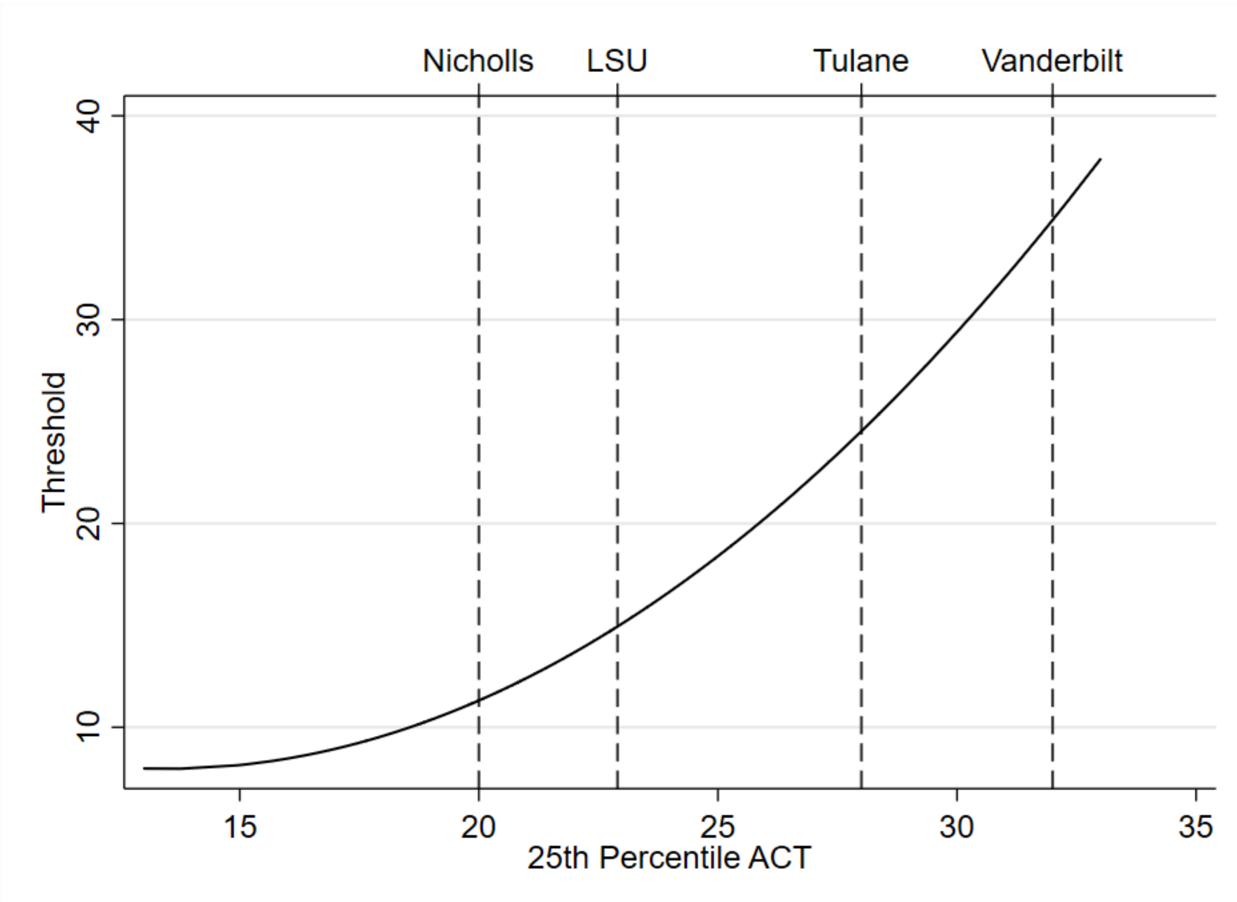


(c) Louisiana Tech



NOTE: The figure shows states by the percentage of high school graduates enrolled in the respective college on average between 2006 and 2012. For example, LSU enrolls 9.3% of Louisiana's high school graduating class on average during this period. Blue-shaded states are those in which the college competes, based on the method described in Section 2.3.

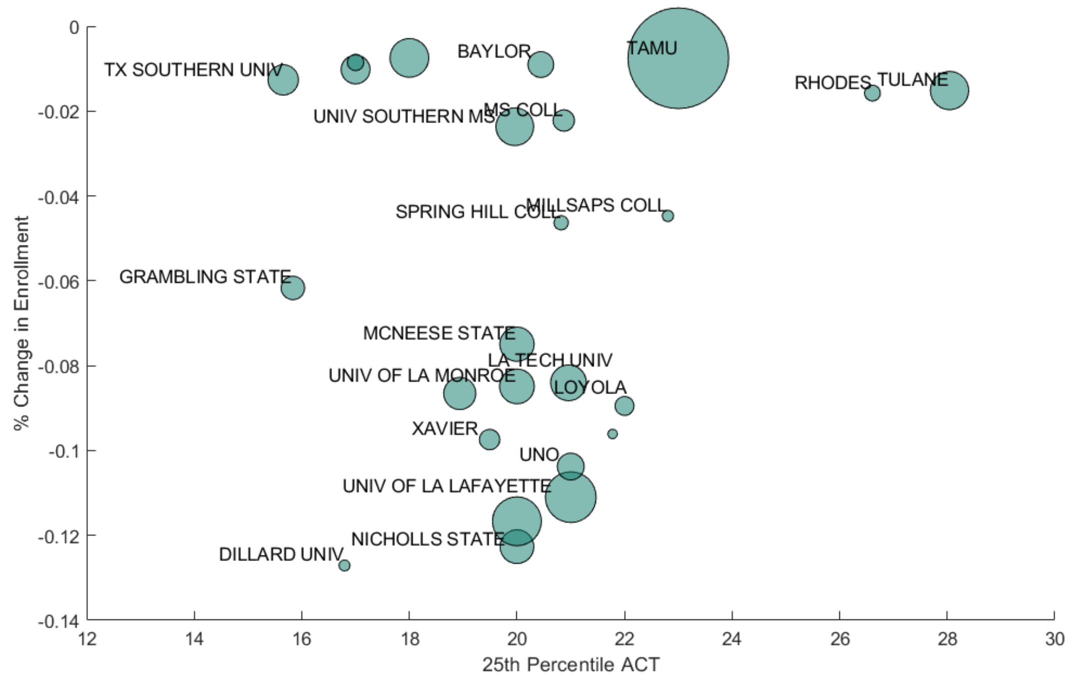
Figure 3: Admission Threshold Function and Example Colleges



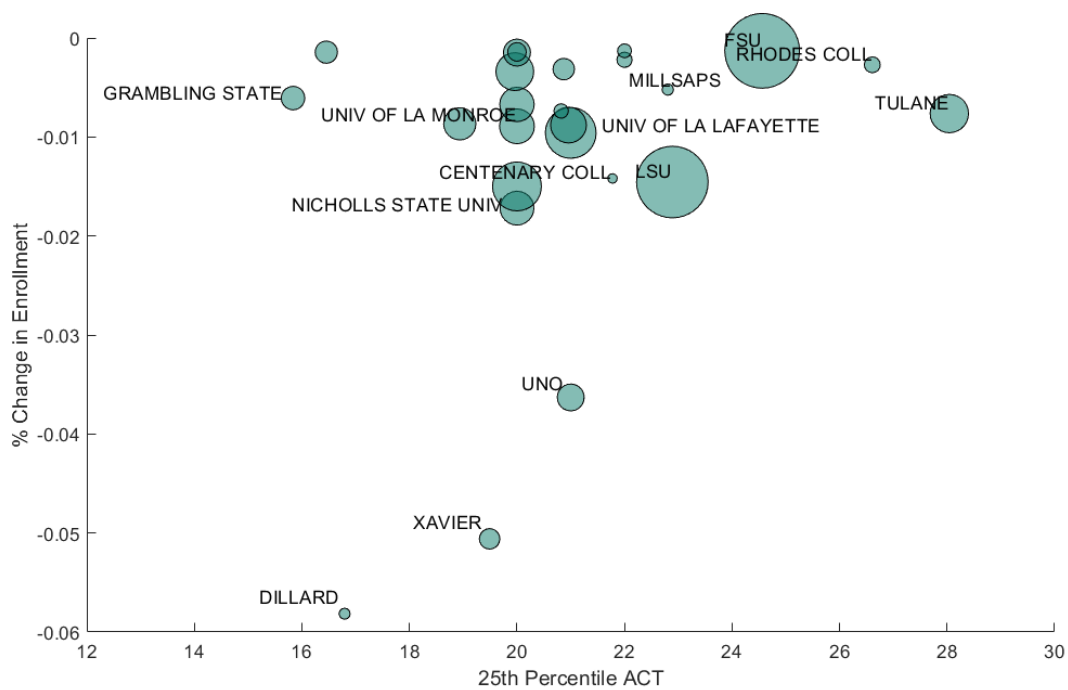
NOTE: The figure shows admission thresholds as estimated by the procedures described in Section 5.2. Example colleges are shown for illustration. Vanderbilt and Tulane are selective private colleges, while Louisiana State University (LSU) is the public flagship of Louisiana. Nicholls State is a regional public college in Louisiana.

Figure 4: Cross-price Elasticities for Example Colleges: Louisiana

(a) LSU



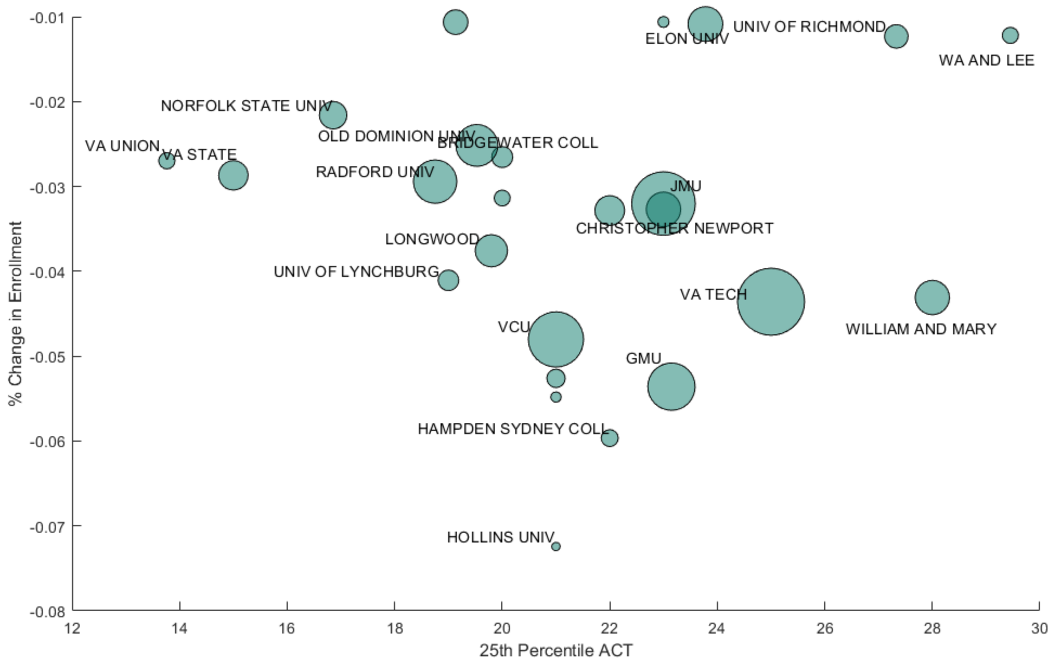
(b) Loyola



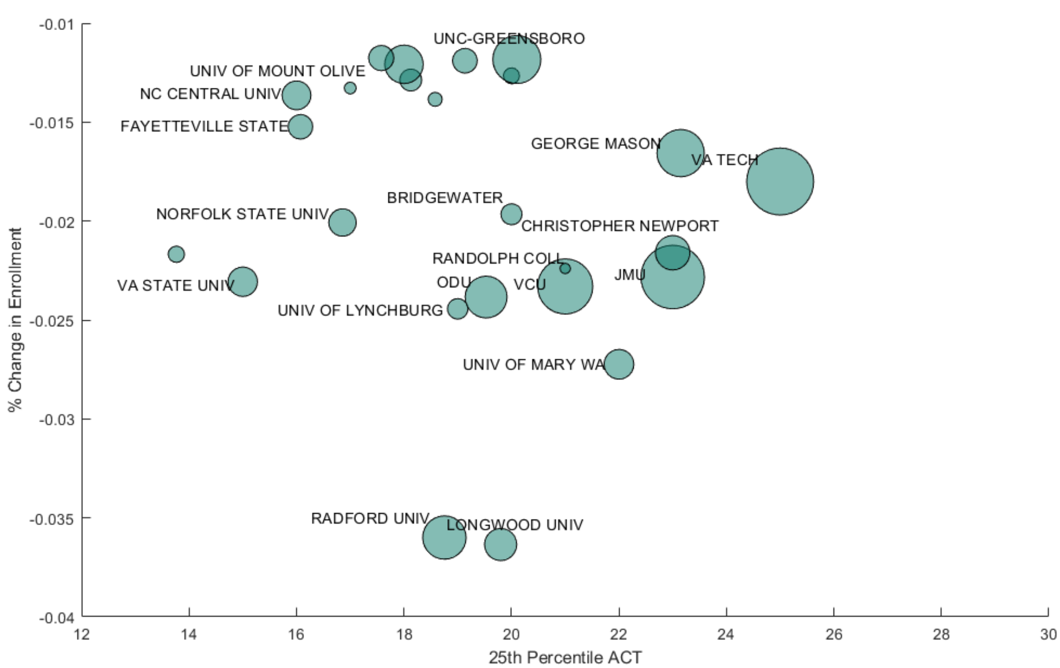
NOTE: Each subfigure shows the percentage change in enrollment at competitor colleges in response to a 1% decrease in tuition for the college in the subfigure title, *ceteris paribus*. Thus, the percentage change in enrollment can be interpreted as a cross-price elasticity. The colleges shown are the top 25 competitors in terms cross-price elasticity with respect to the tuition at LSU (panel a) or Loyola (panel b). College financial aid adjusts proportionally with tuition in these simulations, as in the model.

Figure 5: Cross-price Elasticities for Example Colleges: Virginia

(a) UVA



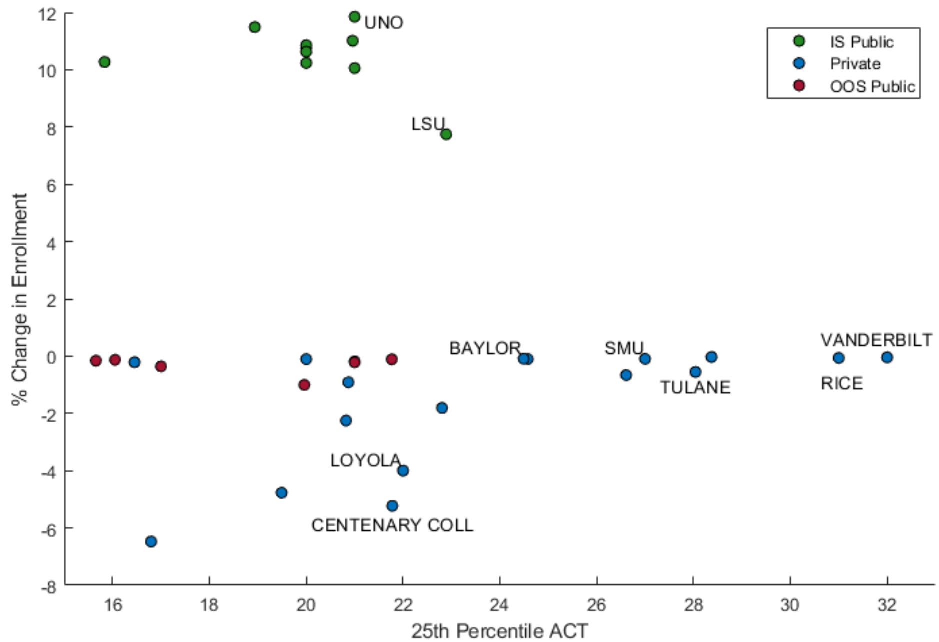
(b) Liberty



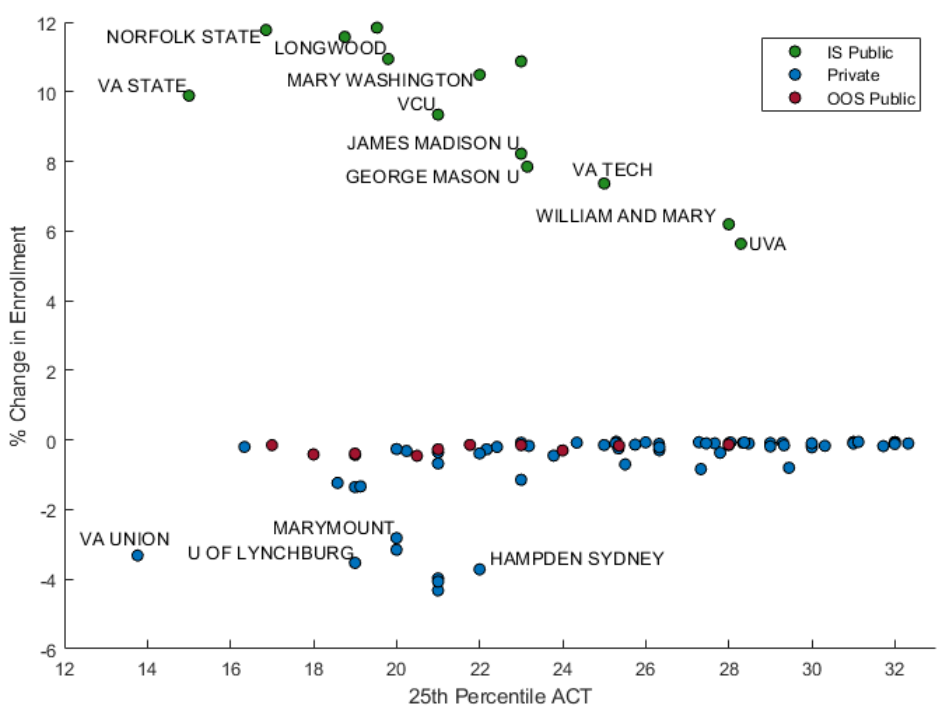
NOTE: Each subfigure shows the percentage change in enrollment at competitor colleges in response to a 1% decrease in tuition for the college in the subfigure title, *ceteris paribus*. Thus, the percentage change in enrollment can be interpreted as a cross-price elasticity. The colleges shown are the top 25 competitors in terms cross-price elasticity with respect to the tuition at UVA (panel a) or Liberty University (panel b). College financial aid adjusts proportionally with tuition in these simulations, as in the model.

Figure 6: Demand-Side Responses to a $1,000 Scholarship for Public Colleges

(a) Louisiana



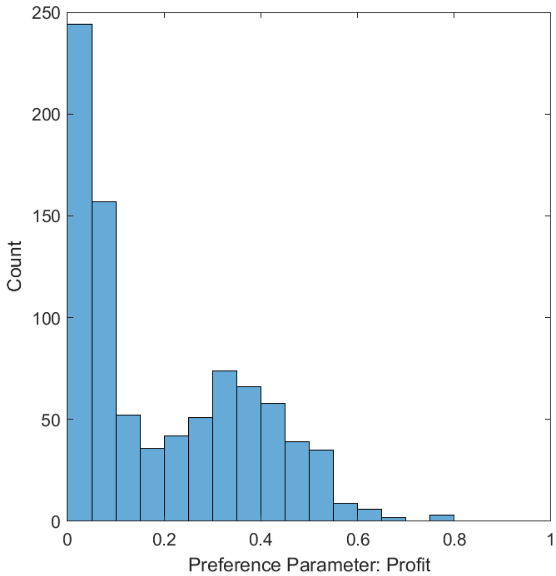
(b) Virginia



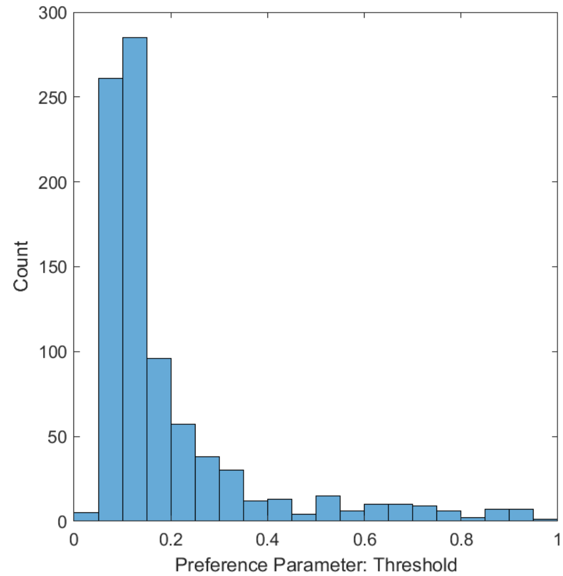
NOTE: Panel A shows the percentage change in enrollment for colleges that compete for Louisiana students when Louisiana introduces a $1,000 scholarship for Louisiana students at in-state public colleges. Panel B shows the same for Virginia. These are demand-side effects only and assume that tuition, aid policies and admissions standards remain fixed.

Figure 7: Supply-Side Results: Distribution of College Preference Parameters

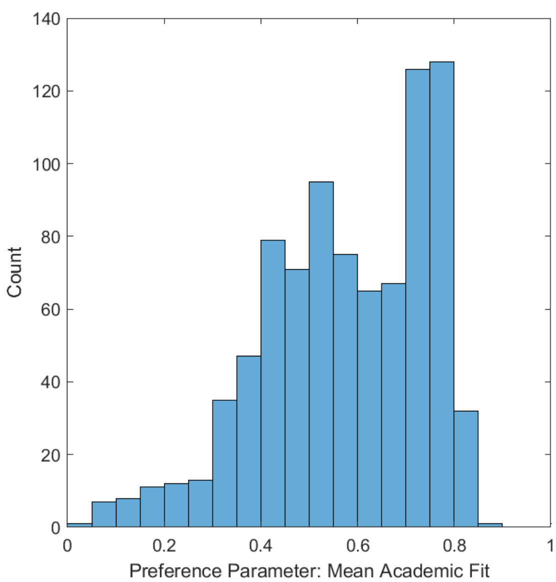
(a) Profit



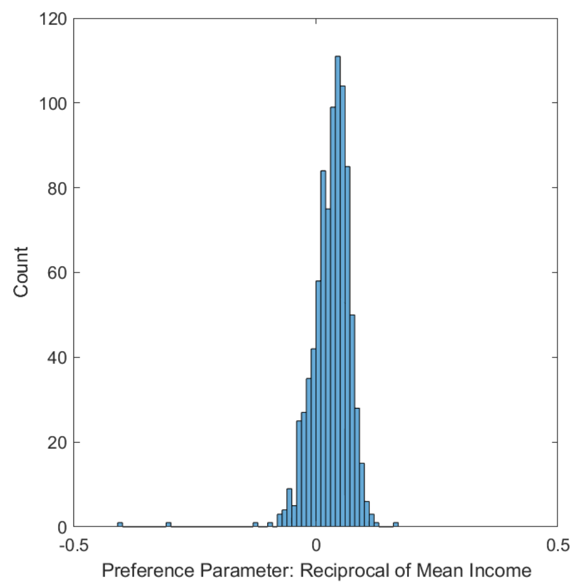
(b) Admission Threshold



(c) Average Academic Fit



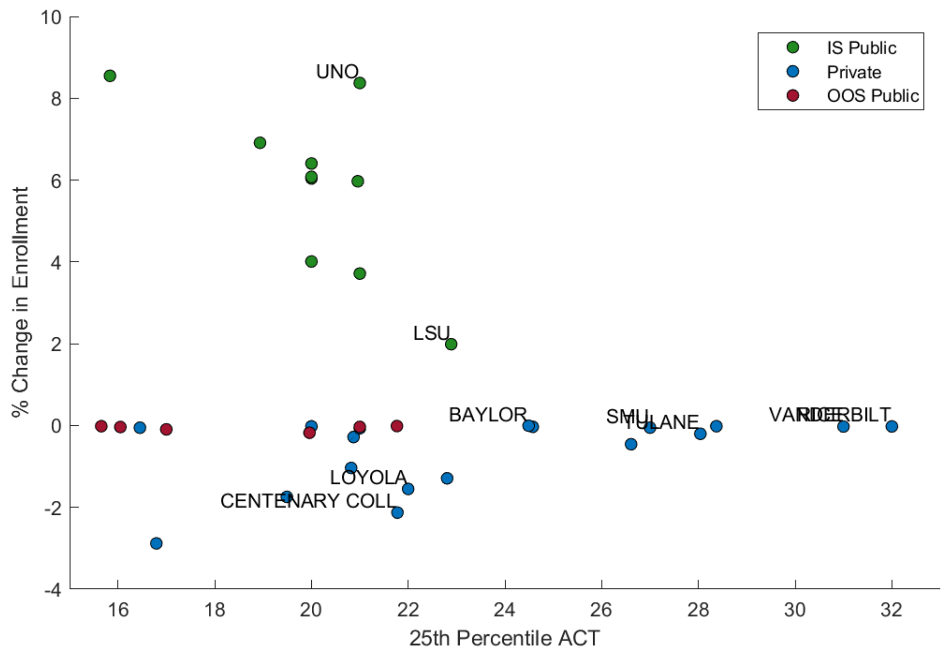
(d) Reciprocal of Mean Income



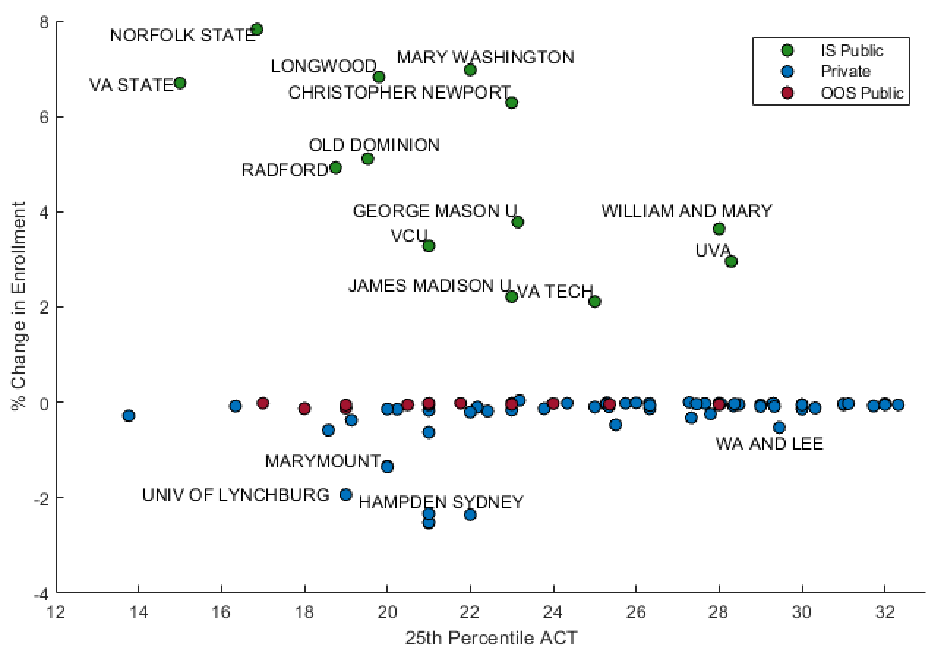
NOTE: The figure shows the distribution of university preference parameters in the supply-side objective function. See the description in the text in Section 5.6.

Figure 8: Equilibrium Responses to a $1,000 Scholarship

(a) Louisiana



(b) Virginia



NOTE: The figure shows the percentage change in enrollment for universities that compete for in-state students when the state offers $1,000 in aid to attend a public in-state institution. Panel A shows the results for Louisiana, and Panel B shows the results for Virginia. These are equilibrium effects.

Appendices

A Constructing a Representative Sample of Students

Estimation of the model requires a sample of students from each state that is representative in terms of income, admission test scores, and location. The Current Population Survey provides family income and location of respondents (up to the Public Use Microdata Area, or PUMA). I use these data to form the sample. Specifically, I use 2012-2014 CPS data for persons aged 17 or 18 who are in twelfth grade at the time of the CPS survey. I sample 1,000 of these individuals, weighting by the CPS person weights in order to ensure the sample is representative. Because PUMAs are sometimes very large, I assign individuals to Census tracts probabilistically, based on the population distribution of Census tracts within each PUMA. I then calculate each student's distance to each college using the centroid of the student's assigned Census tract as the student's location. After this step, I have a representative sample of 1,000 students in each state along with their family income and distance to each college.

It remains to add ACT scores to the simulated sample. CPS does not provide college admission test scores, but I am able to calculate the joint distribution of admission test scores and income in the HSLS:09 data. Specifically, I run an ordered logit regression of ACT scores on a quadratic function of family income, estimating cut points for each ACT score. This regression then provides the probability distribution of ACT scores as a function of family income. I use this probability distribution to simulate ACT scores.

B Tables and Figures

Table B1: Data Example: Score Reports Sent by CT High School Graduates

State of Residence	University	University State	Score Reports Sent	% of HS Grads
CT	UNIV OF CONNECTICUT	CT	12,362	27.86
CT	CENTRAL CONNECTICUT STATE UNIV	CT	5,156	11.62
CT	EASTERN CONNECTICUT STATE UNIV	CT	4,315	9.73
CT	SOUTHERN CONNECTICUT STATE UNIV	CT	4,198	9.46
CT	QUINNIPIAC UNIV	CT	3,631	8.18
CT	UNIV OF RHODE ISLAND	RI	3,081	6.94
CT	NORTHEASTERN UNIV	MA	3,034	6.84
CT	WESTERN CONNECTICUT STATE UNIV	CT	2,889	6.51
CT	UNIV OF MASSACHUSETTS AMHERST	MA	2,542	5.73
CT	BOSTON UNIV	MA	2,455	5.53
CT	UNIV OF HARTFORD	CT	2,348	5.29
CT	UNIV OF VERMONT	VT	2,087	4.70
CT	UNIV OF NEW HAVEN	CT	1,928	4.35
CT	FORDHAM UNIV	NY	1,793	4.04
CT	BOSTON COLL	MA	1,518	3.42

NOTE: This table shows the number of score reports sent and the percent of high school graduates sending reports to the top 15 colleges among high school graduates from Connecticut in 2013.

Table B2: Summary of Application and Enrollment Data

	Application Penetration	Enrollment Share
mean	0.02007	0.00451
min	0.00150	0.00001
p10	0.00214	0.00022
p25	0.00368	0.00040
p50	0.00784	0.00092
p75	0.01842	0.00276
p90	0.04629	0.00963
max	0.38405	0.16671

NOTE: This table shows the mean, min, max and percentiles of the enrollment and application variables used to estimate the college choice model. The application penetration for a college is the proportion of high school graduates from a state market who sent a score report to the college. The enrollment share is the proportion of high school graduates that enrolled at the college.

Table B3: Federal Pell Grant Aid Regression

	coef	se
Income, $100,000s	1.215	2.450
Income Sq.	-12.359	2.960
Constant	4.543	0.464
var(e)	8.33	0.552
Obs.	800	

NOTE: Results from a Tobit regression of the Pell Grant amount on family income among students with family incomes below $75,000. The sample is limited to students with records in the National Student Loan Data System, so it includes all students who apply for federal aid. The Tobit regression has a lower limit of $0 and an upper limit of $5,645, corresponding the the Pell grant cap in the 2013-14 academic year. Income is measured in hundreds of thousands, and the Pell grant amount is measured in thousands.

SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09), Base Year, First Follow-Up, and Student Records.

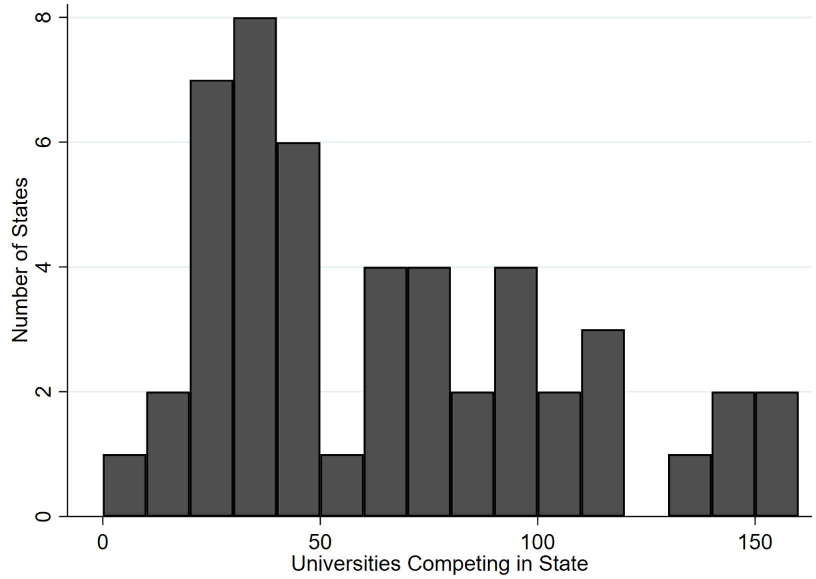
Table B4: State Grant Aid Functions

	(1) States with Large Merit Aid Programs		(2) CA		(3) All Other States	
	coef	se	coef	se	coef	se
Income, $100,000s	-0.788	0.250	-10.698	3.306	-5.241	0.463
Income Sq	0.093	0.032	1.379	0.706	0.481	0.068
ACT	0.649	0.251			0.607	0.384
ACT Sq	-0.010	0.005			-0.014	0.008
State Grant PS	3.900	0.230			7.158	0.950
Public	-1.404	0.346	0.103	4.516	-1.046	0.379
Constant	-8.442	3.013	6.293	4.660	-5.058	4.508
var(e)	4.355	0.380	67.763	19.146	15.678	1.359
Observations	370		80		830	

NOTE: Results from Tobit regressions of the student's state grant aid amount (in thousands) on their family income, ACT score, state grant aid program generosity as measured by the total state grant aid per high school graduate (State Grant PS), and an indicator for whether the student is attending a public or private college. The Tobit regressions each have a lower limit of $0. States are divided by type of grant aid program, as described in the text.

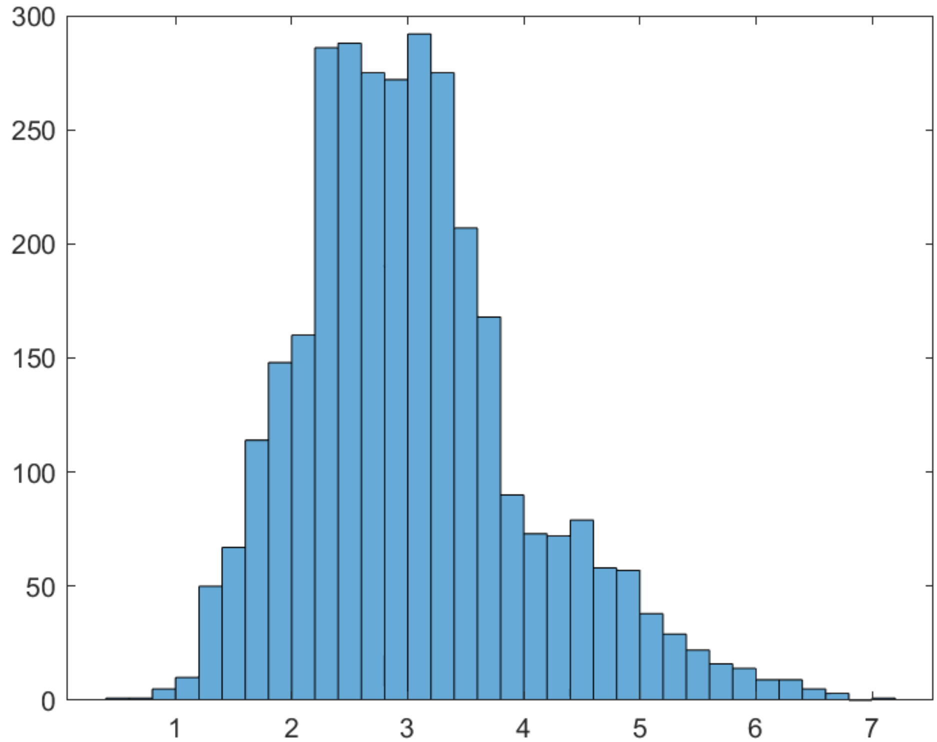
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09), Base Year, First Follow-Up, and Student Records.

Figure B1: Number of Colleges Competing in Each State



NOTE: This graph shows the distribution of the number of colleges competing in each state market.

Figure B2: Distribution of Estimated Application Costs



NOTE: This graph shows the distribution of estimated application costs across state-college pairs.