College Admissions and Inequality*

Ricardo Marto†
University of Pennsylvania

Yaacov Wittman‡
University of Pennsylvania

Job Market Paper

September 1, 2021

Click here for the latest version

Abstract

Empirical studies have found that high achieving, low-income students are less likely to apply to selective colleges despite the generous financial aid typically offered. To reconcile this seeming puzzle, we build a model of the college market featuring tuition discrimination and a decentralized admissions system. Students, who differ in their financial resources and innate ability, apply to a subset of colleges and are uncertain about their prospective admissions and financial aid. Colleges observe only a noisy signal of student ability, and compete by choosing admissions standards and tuition schedules. We find that differences in application rates are due to student expectations over admissions and financial aid, which are consistent with college policies in equilibrium. We address the puzzle by finding that low-income students receive high financial aid at selective colleges because only the highest-ability among them apply, making their signals highly informative. If signals became less informative (e.g. colleges stopped using the SAT), all high-ability students would be worse off and low-ability students would modestly benefit. Finally, we find overall welfare gains from increasing Pell Grants, which would greatly benefit low-income, high-ability students by alleviating credit constraints.

Keywords: College Market, Tuition Discrimination, Admissions, Sorting, Credit Constraints
JEL Classifications: E21, I23, D58, D83

*We thank Damien Capelle, Jesús Fernández-Villaverde, Jeremy Greenwood, Dirk Krueger, and Víctor Ríos-Rull for their insightful comments and suggestions.

†Department of Economics, University of Pennsylvania. Email address: rmarto@sas.upenn.edu.

‡Department of Economics, University of Pennsylvania. Email address: ywittman@sas.upenn.edu.

1 Introduction

Investment in college education is important for human capital accumulation, earnings growth, and can be a powerful tool for social mobility. Despite the benefits of completing college, however, there are well documented gaps in college outcomes across the parental income distribution. Students born to parents from the bottom quartile of family income are much less likely to complete college than students from the top of the distribution (Bailey and Dynarski 2011) and are less likely to be represented at more selective colleges (Chetty et al. 2020). While these gaps can partly be explained by differences in levels of preparedness, gaps remain even for high-achieving students.

One explanation for these income gaps is that they arise from differences in application patterns between low-income and high-income students. Empirical studies such as Hoxby and Avery (2014) and Dillon and Smith (2017) find that there is substantial undermatching at the application stage: low-income students tend to apply to schools that they appear overqualified for relative to their higher-income peers. That is, low-income students appear to be under-represented in the college market not because they are excluded from the colleges directly but because they do not apply in the first place. This finding is troubling because it suggests that student misallocation may exist not only at the extensive margin of college enrollment, but also at the intensive margin of college choice. Moreover, as Hoxby and Avery (2014) note, the gap in application rates is puzzling since many colleges (especially the most selective ones) provide substantial amounts of need-based financial aid for low-income students.

Why then, even after controlling for test scores, do we observe different application patterns for students across the income distribution? The goal of our paper is to answer this question by building and estimating a novel model of the college market featuring an application and admission system similar to the one used by U.S. colleges. Our model recognizes that while colleges would like to enroll all high-ability students, it is expensive for them to offer generous financial aid to their low-income applicants: they face a tradeoff between admitting high-ability, low-income students and admitting lower-ability, high-income students who are willing to pay higher tuition. This trade-off will cause colleges to offer generous financial aid only to the highest-ability students among their low-income applicant pool. Thus, low-income students will find it optimal to apply at lower rates because they have a lower chance of receiving an acceptable level of financial aid.

Our paper proceeds in two parts. First, we empirically study student application and enrollment patterns using the High School Longitudinal Study (HSLS), a recent dataset from the National Center for Education Statistics. Second, we build an equilibrium model of the college market featuring a noisy application and admissions system and estimate it using the HSLS. We show that our model is able to account quantitatively for the application and enrollment patterns observed in the data, with realistic tuition schedules that vary based on students' parental income and test-scores. We use the model to study the role of the admissions system in shaping the allocation of students in the college market, and study the effect of higher education policies in our environment. We find that selective colleges offer high levels of financial aid to low-income students *because* they are less likely to apply: only the highest-ability among them apply, making colleges confident that their low-income applicants are more likely to be higher ability. This justifies the generous financial aid low-income students receive. If the low-income students applied at the same rates as their high-income peers, selective colleges would be less able to distinguish the high-ability students and will therefore lower their financial-aid. Our model suggests that such an increase in tuition for low-income applicants would reduce their enrollment at selective colleges by about a half.

In the empirical part of the paper, we use detailed student-level data from the HSLS to provide descriptive evidence on college application and enrollment patterns, and how they vary by parental income and student test scores. We show that both parental income and test-scores are important predictors in determining students' application behavior, even when controlling for other demographic and economic characteristics. This is true both at the extensive margin, where we see if students applied at all to any 4-year colleges, and at the intensive margin, where we see if students included any highly selective colleges in their application portfolios. In both cases we find that test scores, when measured by SAT/ACT or high school GPA, is the strongest factor in predicting a student's application choice. A one standard deviation increase in the SAT score, for example, is associated with an increase in the average probability of applying at all or applying to a highly selective college by about 10% and 20% respectively (for students with median household income and median SAT scores). While parental income has a lower effect than test scores, it is also an important factor in predicting a student's application choice. Descriptive analysis of our data suggests that for students with median household income and SAT scores in the 75th percentile, a one standard deviation increase in parental income is associated with a 5% increase in the probability of applying at all and a 5% increase in the probability of applying to a selective college.

Motivated by these facts, we then develop an equilibrium model of the college market featuring student heterogeneity, college competition, and a decentralized admissions system. Colleges differ by the value they add to students in the labor market. Students, who differ by their innate ability and financial resources, choose to apply to a subset of colleges or not apply at all. Applications are costly, and if a student decides to apply they then send a noisy signal of their ability to the colleges. Each college has its own threshold for the minimum acceptable signal necessary for admissions, and each college charges a tuition level that varies based on a student's financial resources and signal. Thus admissions *and* financial aid are risky for the student: they make their application decisions based on an expectation over the possible realizations of the signal. Finally, once the signal is realized and students have offers from the colleges they were accepted to, they decide which college to enroll in (if any).

On supply side of the market, there are a discrete number of colleges who differ in their technology, endowment income, and costs. The objective of the colleges is to maximize the value they add to their students, which depends on the level of instructional spending per student and the average ability of the student body. To capture the uncertainty of the admissions process, we assume that colleges are unable to observe the true ability of the students in their pool of applicants but instead observe only the noisy signal mentioned above. Thus, the admissions system is modeled as a signal extraction problem for the colleges. To maximize their objective, colleges choose their minimum acceptable signal (admissions standard) and tuition schedules. When colleges compete, they take as given the admissions standards and tuition schedules of the other colleges. Thus, their pricing decisions are limited by the students' outside options if their applicant pool is likely to contain students who applied to multiple colleges.

Combining the signal extraction problem of the college with their ability to price discriminate is new to our model, and introduces an important mechanism that influences the degree of price discrimination due to the signal. Colleges can fully observe student wealth, and while they do not observe ability directly, we assume they know the joint distribution of wealth and ability among their applicants (they take student application choices as given in equilibrium). Thus, if low-wealth students are more likely to apply only if they are high ability, signals from low-wealth students will be more informative to the college. Since only the highest ability among the low-wealth students apply, the signals the colleges observe from low-wealth applicants will be more likely to have come from high ability students. Colleges will then be happy to offer high tuition discounts to the low-wealth students they enroll. This mechanism helps resolve the puzzle: low-income students are offered substantial financial aid precisely because they are less likely to apply, and only the ones with the highest ability apply find it optimal to apply in equilibrium.

We estimate the model using the HSLS and college level data available from The Integrated Postsecondary Education Data System (IPEDS). We find that the model does a good job accounting for the different application and enrollment rates across the parental income distribution. It is also able to capture the observed variation in tuition both across and within colleges, where higher income students pay more and students with higher test scores pay less. The model is also able to rationalize the result in Dale and Krueger (2002) who find little return to college selectivity when comparing the earnings of students who were accepted to the same colleges but chose to attend different ones. This result at first seems contrary to our model, where the more selective college has higher value-added. However, due to strong selection effects, we find that students who were accepted to both colleges but chose the less selective one are in fact higher ability on average. The reason is that low-income students are more likely to reject offers from the more selective college if they do not receive sufficient financial aid, and instead enroll in the less selective college. But these students had to have been higher ability in order to apply to both colleges in the first place. Thus, this selection effect may counteract the effect of returns to selectivity, which helps explain the small earnings gap observed by Dale and Krueger (2002).

Next, we use the model to study the effect of the application and admissions system on student allocation in the college market. We first examine the effect of low-income student application decisions in order to account for the importance of signal informativeness in our baseline equilibrium. We solve for the equilibrium of a counterfactual economy where low-income students apply as much as their higher income peers. This introduces low-ability students to the low-income applicant pool, which makes the signals of the high-ability, low-income students less informative. Selective colleges will then reduce the financial aid for all low-income students since the applicant pool has worsened. We find that this effective increase in tuition is large, and lowers low-income student enrollment at the selective colleges by about a half. Overall, this finding shows that high-ability, low-income students benefit greatly from more informative signals due the lower application rates of their lower-ability peers. Thus, interventions that encourage low-income students to apply may be harmful for their overall enrollment if the interventions are not targeted by ability.

We also study the effects of application signals becoming *less* informative, motivated by many colleges' decision to pause their use of standardized tests in the admissions process during the Covid-19 pandemic. To do this, we recalculate a counterfactual equilibrium where we increase the noise associated with student applications, making it more difficult for colleges to infer the students' true ability. We find that in the new equilibrium, students at the top of the ability distribution are made worse off from the change since they now find it harder to differentiate themselves from the lower-ability applicants. The only students who have modest gains from the less informative signals are the low-ability students who now find it easier to match with the colleges due to the increase in the signal variance. This is especially beneficial to the high-income, low-ability students: the lowest ability ones can now more easily attend less selective colleges, and the moderately low ability ones can more easily attend the selective colleges. These gains, however, are smaller than the losses of the high ability students who would benefit the most from attending college.

Finally, we study the effects of expanding the Pell Grant program, which would increase the amount of federal grant funding to low income students, and make middle-income students eligible for federal aid. We find that the grants have little effect on enrollment in the selective colleges since they already admit a relatively large fraction of low-income students. There is a larger effect, however, on enrollment of low-income students in the less selective colleges. We find that the policy change is most beneficial to high-ability, low-income students who were before less likely to enroll due to credit constraints. Higher-income students lose from the change due to the higher tax rates and increased competition with lower-income students. The degree to which high-income students lose, however, depends on their ability: those with high ability lose less since it is easier for them to compete with the newly unconstrained low-income students. Overall, we find that the policy has a net-positive effect on welfare. The value of a college education is higher for the high-ability, low-income students who benefit from the policy compared to the low-ability, high-income students who lose from it.

Related literature. This paper builds on three different strands of the literature. The first is the large empirical literature documenting inequality in higher education and the role of applications and the admissions system. The second relates to the structural literature on the college market and the admissions problem. The third relates to the literature which studies the distributional effects of education policies.

Our paper is complementary to the empirical literature on the college market and its outcomes. Recent work by Chetty et al. (2020) document a large degree of income segregation within and across U.S. colleges. Consistent with our empirical facts and model, they find that students from wealthier backgrounds are disproportionately represented at more selective schools. They also find that more selective colleges give higher returns to education, and estimate the causal effect of an individual college in earnings to be around 80%. Differences in application behavior related to income differences or ability have also been documented by Hoxby and Avery (2014), Dillon and Smith (2017), Delaney and Devereux (2020). Dynarski et al. (2018) further study the role of expectations about tuition and admission at the application stage. They conduct an experiment where low-income high-achieving high school students are informed that they will be offered free tuition if admitted at the University of Michigan. They find large increases application and enrollment, consistent with our findings of the importance of tuition in driving these decisions. Our empirical analysis on tuition discrimination also relates to Fillmore (2020), who studies the effect of different FAFSA information disclosure policies on tuition levels.

Our model builds on the work of Epple et al. (2006, 2017), who consider quality-maximizing colleges that price discriminate among students to study the effect financial aid policies. Gordon and Hedlund (2016) build on that framework to study the rise in college tuition, showing that demand forces helps explain most of this increase. Similar to these papers, we assume that colleges compete monopolistically and choose a tuition schedule to maximize their value-added, which depends on the composition of their students. Importantly, our model also draws on Chade et al. (2014), who introduce matching frictions in college admissions problem, and allow students to make multiple college applications. Our model also complements Fu (2014), which jointly models tuition and admissions, but ours adds the important margin of heterogeneity in parental income and credit constraints. Several recent papers have also studied the college market and its interaction with inequality and intergenerational mobility. Cai and Heathcote (2018) study the role of income inequality in explaining the rise in tuition using a novel model that gives rise to an endogenous distribution of colleges. Using a similar framework, Capelle (2019) studies the role of the college market in shaping intergenerational mobility for heterogeneous students. Our paper includes realistic features of the application and admissions problem, and studies their effect on the college market.

Finally, this paper is also related to the literature on the macroeconomic effects of education policies. Several papers have modeled and quantified the effect of policies on school choice, inequality, or labor market returns (Fernandez and Rogerson (1996), Bénabou (2002), Lochner and Monge-Naranjo (2011), Ionescu and Simpson (2016), Krueger and Ludwig (2016), Kotera and Seshadri (2017), Caucutt and Lochner (2017), Abbott et al. (2019)). In particular, our paper complements the analysis of Abbott et al. (2019), who study the effect of financial aid policies and intergenerational transfers on welfare, Ionescu and Simpson (2016) and Lucca et al. (2018), who examine policy changes in student loan limits on college enrollment and tuition, as well as Krueger and Ludwig (2016), who analyze the optimal mix of tax and education subsidies and their impact on human capital accumulation.

Outline. The remainder of the paper is organized as follows. Section 2 presents evidence on application and enrollment patterns among students transitioning from high school to college, Section 3 describes our model of the college market, Section 4 presents our calibration and estimation procedure, Section 5 and 6 discuss our results, Section 7 provides counterfactual policy analysis, and Section 8 concludes.

2 Empirical Evidence

We empirically study the college application and enrollment behavior of U.S. students by using detailed micro-level data on high-school students combined with institutional college level data. We first describe the datasets we use and then document stylized facts on student application and enrollment patterns.

2.1 Data

We use the High School Longitudinal Study of 2009 (HSLS) to study high-school students' outcomes in the college market. Published by the National Center for Education Statistics (NCES), the HSLS consists of a nationally representative sample of more than 23,000 ninth graders from 944 high schools who are followed throughout their secondary and postsecondary education. The students and their parents are first interviewed in 2009, then again in 2013 once the students have graduated and moved on to college or the labor market, and then once more in 2016. The dataset includes rich information about students' test scores, college application and enrollment behavior, as well as demographic and economic characteristics. Our analysis relies on the restricted-use version of the HSLS, which also provides student SAT/ACT scores, lists of colleges applied to and enrolled in, and more detailed information about household economic variables.

Our goal is to study how the students' college application and enrollment varies based on their parental income and college preparedness. In the HSLS parental income is provided by the parents in the survey, where they are asked to provide their households' income from all sources in 2011. For college preparedness, we use the students' high school GPA and their SAT or ACT scores[1]. Each student's GPA is reported directly by their high-school, and is honors-weighted in a procedure by the NCES used to make the GPA comparable across different high schools. The SAT/ACT is reported directly by the students' colleges, and is therefore unavailable for students who did not attend college or did not take the test. For these students, we impute the SAT/ACT score in a similar manner to Chetty et al. (2020). The scores are imputed for these students using the SAT/ACT score of other HSLS students from the same parent income group, level of parental education, race, and GPA decile.

We combine the HSLS data with data from the Institutional Post-Secondary Database (IPEDS) to study the school level characteristics of the colleges that students apply to and choose to attend. We group all colleges according to their selectivity using Barron's selectivity index (Profile of American Colleges, 2015). For simplicity, we classify 4-year non-profit colleges into two broad groups using this index. The first group, which we call "Highly-selective" colleges, corresponds to Barron's Tier 1 and 2 colleges and universities. This group makes up 192 schools. For reference, a list of all colleges and universities in this group is provided in Appendix B.1. All other 4-year non-profit colleges are counted in the second group, which we refer to as "Non-selective".

The HSLS also includes detailed information about each student's application behavior, which we use to categorize the students' application portfolios. In the post-high school follow up survey, students were asked not only which college they were currently attending, but also to list two other colleges they applied to and were seriously considering. The answers to these questions then help us determine the strength of each student's application portfolio. While the data include only the names of the three most relevant schools in the portfolio, most students indicated that they had applied to three or fewer schools, suggesting that the HSLS gives us a good picture of the total application portfolios.

If we do not observe any non-profit 4-year applications among the three schools, we consider the

[1] The NCES converts the ACT score into an SAT equivalent score for students who only took the ACT instead of the SAT. Henceforth, whenever we mention the SAT, we refer to either the SAT or ACT score.

student as not having applied. If we observe at least one non-profit 4-year application but none of the three schools are Highly-selective, we consider the student as only applying to Non-selective schools. If we see at least one Highly-selective and one Non-selective application, we consider the student has having applied to both. Finally, if we see only Highly-selective applications, we count these as having applied to both only if the number of *total* applications is greater than the number of *listed* applications (that is, we assume such students also applied to a Non-selective school as a safety).

2.2 College Attendance in the HSLS

We first use the HSLS to document how different students apply and sort into different types of schools. The left panel Figure 1 shows how enrollment rates vary by parental income and SAT scores, which are broken down by quartile[2]. We see that conditional on test score quartile, student outcomes are correlated with parental income. On the extensive margin of college attendance, we see that even for students in the top quartile, students with parental income that is below the median ($55,000) were more than twice as likely not to attend 4-year non-profit colleges as students in the top parental income. The data also reveal a similar pattern for enrollment at the intensive margin of college selectivity. As can be seen in Figure 1, students from high income families are more likely to attend Highly selective colleges. For students in the top quartile for SAT scores, we again see that students in the highest income group are about twice as likely to attend a Highly-selective school.

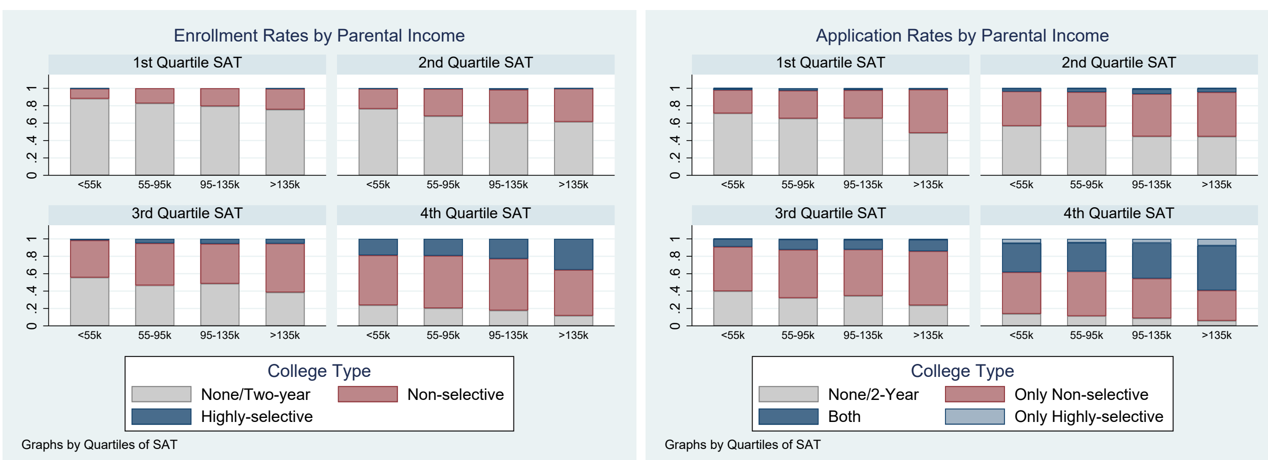


Figure 1: College application and enrollment by parental income and student SAT/ACT score. Left panel: Student enrollment. Right panel: Student applications. Source: HSLS.

Next, we turn to the application behavior of the students in the sample, which is presented in the right panel of Figure 1. This plot similarly shows how application rates vary among students from different parental income and SAT score groups. It shows the fraction of each group of students who do not apply at all, as well as the fractions who apply to the highly selective schools, the non-selective schools, or both. Looking at applications as opposed to enrollment from the left panel of Figure 1

[2] We present the same figures using the students' GPA instead of SAT scores in Figure A.1.

we can that the patterns are very similar, suggesting that the gaps in college enrollment are due to differences in application rates between low and high income students. Thus in order to explain differences in enrollment, it is necessary to understand student application decisions.

2.3 Analysis of College Application Behavior

Motivated by the importance of college application behavior in explaining the differences in college enrollment, this subsection uses a logit model to test whether parental income and test scores have predictive power in determining application behavior while controlling for other factors. We run these statistical models on student-level data from our full HSLS sample to study the factors correlated with application decisions.

In the first logit model we consider, we study the factors correlated with the extensive margin of application choice, where the dependent variable is an indicator for whether or not the student decided to apply to college at all. Table A.1 shows the results of the estimation, where only the coefficients for parental income and test scores are shown. The table shows that regardless of whether we measure student preparedness by GPA or by SAT score (which are standardized in the estimation), we see that both test scores and parental income are strongly correlated with a student's decision to apply to college. This relationship is evident after for controlling for other correlated variables including parental employment and education level. Importantly, we see that there is a stronger relationship between the decision to apply and test scores rather than parental income, and that the interaction between the two is statistically significant whereas parental income alone is not. This suggests that the effect of parental in the application process is strongest for students with high test scores. In Figure 2, we use our estimates to show how application probabilities vary with parental income, even for students at the top of the test-score distribution.

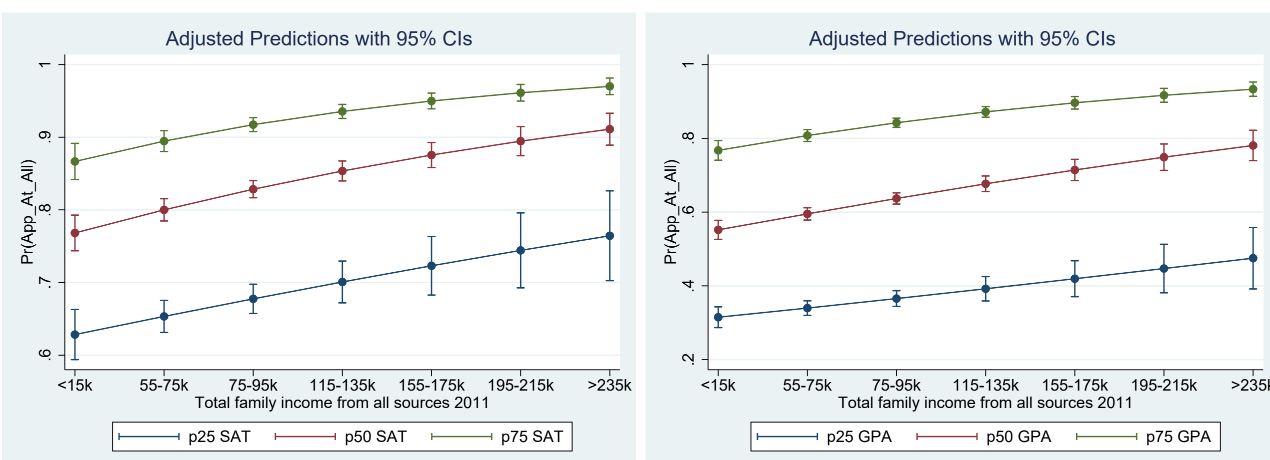
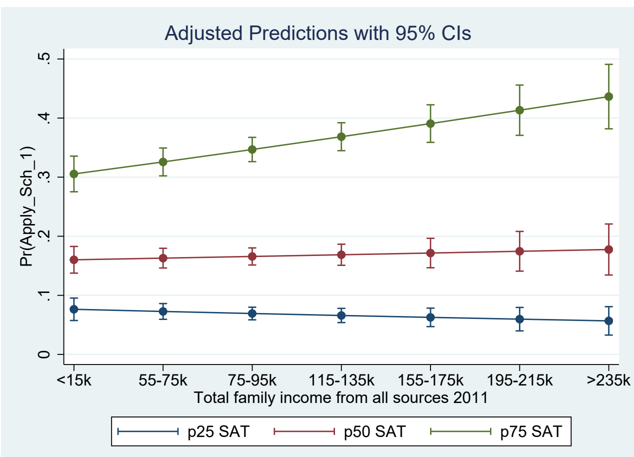


Figure 2: Margin plots for Logit models. Source: HSLS.

We next turn to the intensive margin by studying how parental income and test score are correlated with student application portfolio choice. Restricting our sample only to students who apply at all, we estimate our next set of logit models where the dependent variable is an indicator for whether or not the student included a Highly-selective school in their application portfolio. This includes students who applied to both types of schools and those who applied only to Highly-selective schools, as discussed in the previous section. Table A.2 shows the estimation results from these logit models, which again are estimated separately for GPA and SAT scores.

Table A.2 reveals similar patterns for a student's decision to include Highly-selective colleges in their portfolios as were seen in the previous table. Again, test scores are more strongly correlated with the decision than parental income, and after including the controls, we see that the interaction is significant while parental income on its own becomes slightly negative and insignificant. Again, margin plots are provided in Figure 3 to see how changes in test scores and parental income affect the application probabilities. Overall, these estimates show that parental income is an important factor in determining not only whether or not students will apply, but also which types of college they are targeting.



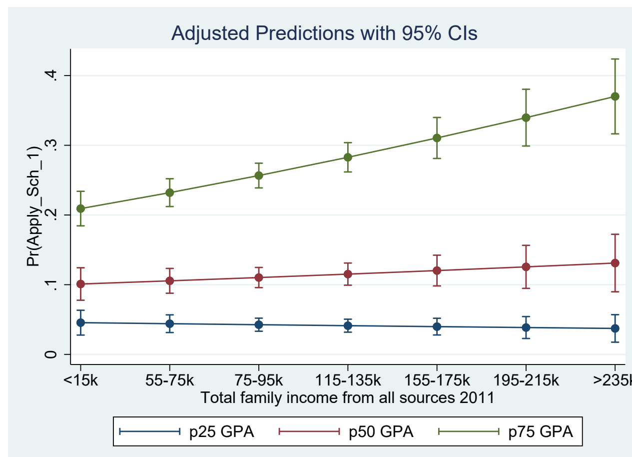


Figure 3: Margin plots for Logit models. Source: HSLS.

2.4 Enrollment Conditional on Admission

Finally, we give suggestive evidence that financial aid is risky for the students who apply. Even if they get in to their top college choice, students may choose to enroll in a different college if they do not receive sufficient financial aid. We use the HSLS to show how students allocate into colleges conditional on being accepted, and the role of costs in guiding their decision. Using the reported application portfolios from the HSLS, Table 1 shows the enrollment patterns for students who were accepted to certain types of colleges. We see that for students accepted to both types of schools, 39% choose to go to a Non-selective college, and for students with the option of attending a Non-selective college, 16% choose not to attend.

The HSLS also helps give us a sense of the extent to which these ex-post decisions are driven by tuition costs, rather than preferences. In the survey, students were asked if they would have attended another school they were accepted to if not for the cost. In the bottom row of Table 1, we see that a significant fraction of students do not attend their preferred college because of costs. Note that this occurs most frequently for students who were accepted to both types of schools. This finding reflects that much of the relatively low rates of attending a Highly-selective school conditional on being accepted to both is due to students not receiving sufficient financial aid from the Highly-selective school.

Accepted to...	Both	Non-selective Only	Highly-selective Only
No College	3.1%	16.2%	3.0%
Non-selective	38.9%	83.8%	–
Highly-selective	58.0%	–	97.0%
% Not attending preferred college because of costs	23.5%	21.6%	3.5%

Table 1: Enrollment patterns among accepted students. Source: HSLS.

Motivated by the empirical findings of this section, we next turn to our model of college application and enrollment. We use the model to rationalize the facts observed in the data, and study the effect of the admissions system in shaping student outcomes in the college market.

3 Model

Overview. The economy is populated by a unit measure of heterogeneous individuals, two colleges of different types, and a government. Individuals live for two periods: young and old. Young students start life with wealth a and ability level ℓ, and decide whether they want to work or invest in their human capital by attending college. If they decide to attend college, they must choose a subset of colleges to apply to. Admissions, however, are risky since colleges can only observe a noisy signal σ of the student's true ability ℓ. We assume that a students' wealth is fully observable to colleges[3]. Students with a high enough realization of σ receive an offer of admission and a college-specific tuition level that depends on a and σ. Once the uncertainty is resolved and students know their admissions and financial aid decisions, they choose which college to enroll in. At any stage, students may choose the outside option of working instead of going to college (for simplicity, we do not consider two-year colleges).

Colleges maximize the value-added they provide to their students on the labor market, denoted Γ_s. Value-added is taken as given by the students but determined endogenously by the colleges. A college's value-added will depend its average instructional spending, and the average ability of its student body. Each college competes by setting different admission standards and tuition levels. Note that there is only one college per type, so competition occurs across types rather than within types. When colleges make an offer of admission, they take into consideration that students may have received other offers

[3]In order to receive federal grants or loans students must complete the Free Application for Federal Student Aid (FAFSA), which states students' parental income and financial assets. The FAFSA is fully observable by the colleges that the student applies to, and most students complete it.

and decide not to enroll. This option value for students makes the choice of tuition depend not only on college-specific characteristics, but also on the pricing and admissions policy of the other colleges. For simplicity, we assume there are two types of colleges $s = \{1, 2\}$ that exogenously differ in their endowment income, efficiency, and costs. We refer to the more competitive elite college as College 1, and the less competitive college as College 2. Finally, the government taxes the working population to subsidize colleges and pay for grant programs.

Model timing. The timing of events in the first period is as follows:

1. Individuals first choose either to apply to college or go straight to the labor market. Those who apply must choose an application portfolio which includes either or both colleges.
2. Colleges receive applications and choose which students to accept by defining their admission standards and tuition schedules.
3. Students make their attendance decision given the admission offers received.
4. Individuals make their consumption and savings decisions.

3.1 Student Application and Enrollment Decisions

We proceed in chronological order by introducing the problem of an applicant, then the decision problem involving the acceptance and rejection of offers, and finally the problem of a matched student. The problem of a worker comes last.

Student's value from applying. A student who decides to apply must choose either to send one application to College $s \in \{1, 2\}$, or apply to both. If a student applies they draw a realization of σ, which is a noisy signal of their ability ℓ and is unknown to the student at the time of their application decision. The signals are drawn from a continuous conditional density $g(\sigma|\ell)$ with cdf $G(\sigma|\ell)$ and support $[0, \infty]$. For simplicity, only one signal is drawn regardless of the application portfolio, so that both colleges observe the same signal[4]. Taking the admissions standard $\underline{\sigma}_s$ as given, a student is admitted to College s if $\sigma > \underline{\sigma}_s$. In the rest of the section, we focus (without loss of generality) on the case where College 1 is at least as selective as College 2 (i.e. $\underline{\sigma}_1 \geq \underline{\sigma}_2$)[5].

Applications are costly, and include a financial cost as well as a 'psychic' cost meant to capture the disutility of completing the applications. We denote these disutilities by $(\phi_1, \phi_2, \phi_{12})$, so they depend on whether the student applies to College 1, 2, or both. For simplicity, we assume that the financial application cost does not depend on the application portfolio, and denote it by ψ. We include

[4]This assumption, while made for tractability, is supported empirically since it is very unlikely that a student admitted to a highly-selective college will not also be admitted to a less-selective college. Most students who were admitted to a highly-selective college (as defined in Section 2.1) were also admitted by a less-selective college (conditional on having applied to both).

[5]While it is possible for College 2 to have a higher admissions standard in equilibrium, we confirm that this is not the case in our baseline estimation and subsequent analysis.

both types of costs since financial application costs alone are not enough to account for the observed application rates[6].

Let $V^{AB}(a,\ell)$ be the expected value from applying to both colleges, and $V^{As}(a,\ell)$ the expected value from applying to College s only, $s \in \{1,2\}$. The expected value of applying to both colleges is given by

$$\begin{aligned} V^{AB}(a,\ell) &= \int_{\underline{\sigma}_1}^{\infty} V^{OB}(a,\ell,T^1(a,\sigma),T^2(a,\sigma))g(\sigma|\ell)d\sigma + \int_{\underline{\sigma}_2}^{\underline{\sigma}_1} V^{O2}(a,\ell,T^2(a,\sigma))g(\sigma|\ell)d\sigma + \\ &\quad G(\underline{\sigma}_2|\ell)V^W(a,\ell,1) - \phi_{12}, \end{aligned} \tag{1}$$

where $V^{OB}(a,\ell,T^1(a,\sigma),T^2(a,\sigma))$ is the value of a student with both offers of admission in hand, each priced at $T^1(a,\sigma)$ and $T^2(a,\sigma)$. $V^{O2}(a,\ell,T^2(a,\sigma))$ is the value of a student with only College 2's offer of admission who was rejected by College 1, and $V^W(a,\ell,1)$ is the value of a worker (the "1" in the value function denotes that the student will still pay the financial application cost).

The expected value from applying only to College s is given by

$$V^{As}(a,\ell) = \int_{\underline{\sigma}_s}^{\infty} V^{Os}(a,\ell,T^s(a,\sigma))g(\sigma|\ell)d\sigma + G(\underline{\sigma}_s|\ell)V^W(a,\ell,1) - \phi_s, \tag{2}$$

where $V^{Os}(a,\ell,T^s(a,\sigma))$ is the value of a student who only applied to College s and was admitted at price $T^s(a,\sigma)$.

Optimal application. The optimal application strategy for a student with characteristics (a,ℓ) solves the following simple discrete choice problem:

$$\max\left\{V^{A1}(a,\ell), V^{A2}(a,\ell), V^{AB}(a,\ell), V^W(a,\ell,0)\right\}. \tag{3}$$

Student's value with offers of admission. After the signal is realized and students know their admissions and financial aid status, they decide which offer, if any, to accept. We assume that admitted students draw idiosyncratic preference shocks ϵ^s over their available alternatives. These shocks are mean zero Type I extreme value shocks with scale parameter $\lambda_c > 0$. Including these shocks simplifies the analytical tractability of the model because they allow for closed form solutions of the students' enrollment functions, which are taken as given by the colleges. The shocks can be interpreted as allowing the students to change their mind about their college preference before enrolling for idiosyncratic reasons (e.g. campus visits, or seeing how their friends enroll).

The value of a student with both offers of admission in hand is the expected value from choosing between accepting College 1's offer, College 2's offer, or working,

$$V^{OB}(a,\ell,T^1(a,\sigma),T^2(a,\sigma)) = \int \max\left\{V^{C1}(a,\ell,T^1(a,\sigma)) + \epsilon^1, V^{C2}(a,\ell,T^2(a,\sigma)) + \epsilon^2,\right. \tag{4}$$

$$\left. V^W(a,\ell,1) + \epsilon^3\right\} dG_\epsilon, \tag{5}$$

[6] In our estimation, we calculate the financial costs directly from the data so we can separately identify the psychic costs.

where $V^{Cs}(a,\ell,T^s(a,\sigma))$ is the value of attending College s paying tuition $T^s(a,\sigma)$. The value of a student with only one offer of admission in hand is the expected value from choosing the maximum value between accepting College s's offer and working, i.e.

$$V^{Os}(a,\ell,T^s(a,\sigma)) = \int \max\left\{V^{Cs}(a,\ell,T^s(a,\sigma)) + \epsilon^s, V^W(a,\ell,1) + \epsilon^3\right\} dG_\epsilon, \tag{6}$$

for $s = \{1,2\}$. Given the extreme value shocks, the value functions simplify to

$$V^{OB}(a,\ell,T^1(a,\sigma),T^2(a,\sigma)) = \frac{1}{\lambda_c}\log\left(e^{\lambda_c V^{C1}(a,\ell,T^1(a,\sigma))} + e^{\lambda_c V^{C2}(a,\ell,T^2(a,\sigma))} + e^{\lambda_c V^W(a,\ell,)}\right) \tag{7}$$

$$V^{Os}(a,\ell,T^s(a,\sigma)) = \frac{1}{\lambda_c}\log\left(e^{\lambda_c V^{Cs}(a,\ell,T^s(a,\sigma))} + e^{\lambda_c V^W(a,\ell,1)}\right) \tag{8}$$

Optimal acceptance. With the Type I extreme value shocks over the students' options conditional on offers of admission, we can solve for the probability that a student with characteristics (a,ℓ), signal σ, and application strategy $i \in \{A1, A2, AB\}$ accepts the offer of College s given tuition levels and admissions thresholds. Given admissions thresholds $\underline{\sigma}_1, \underline{\sigma}_2$, the probabilities are given by

$$q_i^s(a,\ell,\sigma,T^1,T^2) = \begin{cases} \frac{\exp\{\lambda_c V^{Cs}(a,\ell,T^s)\}}{\sum\limits_j \exp\{\lambda_c V^{Cj}(a,\ell,T^j)\} + \exp\{\lambda_c V^W(a,\ell,1)\}} & \text{if } i = AB,\ \sigma \geq \underline{\sigma}_1 \\ \frac{1}{1+\exp\{\lambda_c[V^W(a,\ell,1) - V^{Cs}(a,\ell,T^s)]\}} & \text{if } i = As,\ \ \sigma \geq \underline{\sigma}_s, \\ & \text{or } i = AB,\ \ \underline{\sigma}_1 \geq \sigma \geq \underline{\sigma}_2,\ s = 2 \end{cases} \tag{9}$$

Note that the acceptance probability depends on both tuition levels only if a student chose to apply to both schools and was admitted to both. Students who applied only to one college, or applied to both and were rejected by College 1 will therefore not have an outside option of attending a different college.

Student's value from attending college. Once students have accepted an offer of admission from a particular college, they face a two period consumption-savings problem. The first period corresponds to four years of college and the next period corresponds to the rest of their life. In the college period, the student must finance their consumption, tuition payment, and application costs using initial wealth, grants, and student debt. Students may borrow up to a limit denoted by $\underline{a}_s$[7]. Grants depend on the student's wealth and are separated into grants from outside sources, denoted $\hat{g}(a)$, and grants that are paid for by the government, denoted $\hat{P}(a)$. Outside grants are exogenous in the model and are meant to capture private scholarships or limited public scholarships. Grants from the government are paid for by taxes in the model, and are meant to capture the federal Pell Grant program.

After graduating college and moving to the next period, the student enters the labor market and receives a wage w_o net of taxes τ per unit of human capital acquired in college. The resulting human

[7]This borrowing limit is motivated by the existing limits to federal student loans imposed by the Department of Education. Consistent with the federal limit, the modeled borrowing limit does not depend on the student's earnings potential.

capital is given by $\Gamma_s \ell^{\alpha_s}$, which depends on the college's value added, Γ_s, and parameter α_s that governs the returns to ability. The student also pays her loans back priced at the interest rate R, and discounts the future with discount factor β. Finally, the student has preferences over consumption given by $u(c)$, and we allow for an additional utility benefit (or cost if negative) for attending College s, captured by ν^s.

The value of an individual enrolled in college s, paying tuition T^s, is then given by

$$
\begin{aligned}
V^{Cs}(a, \ell, T^s) &= \max_{c, c', a'} u(c) + \nu^s + \beta u(c') \qquad (10) \\
\text{s.t.} \qquad c + a' &= a - T^s - \psi + \hat{g}(a) + \hat{P}(a) \\
c' &= R\, a' + (1 - \tau) w_o \Gamma_s \ell^{\alpha_s} \\
a' &\geq \underline{a}_s
\end{aligned}
$$

Denote by $a^{s'}(a, \ell, T^s)$ the student's borrowing or savings policy function if studying at College s, after being offered tuition $T^s(a, \sigma)$.Then, the student's optimal consumption-savings decision is governed by the usual Euler equation

$$
u_1(c^s) \;\; \geq \;\; \beta R\, u_1(c^{s'}), \qquad (11)
$$

where $c^s(a, \ell, T^s)$ is the student's consumption. Note that there is a maximum tuition level the student can afford, which is given by

$$
T_{max}(a) = a - \underline{a}_s - \psi + \hat{g}(a) + \hat{P}(a). \qquad (12)
$$

Therefore, any tuition level such that $T^s(a, \sigma) \geq T_{max}(a)$ is automatically rejected by a student of wealth a.

Value from working. Individuals end up as workers either by choosing not to apply, being rejected, or by choosing not to attend college conditional on an offer of admission. The income of an individual working without a college degree is $(1 - \tau) w \ell_w^{\alpha}$, where $\alpha_w < \alpha_s$ reflects that higher ability individuals have a higher return from attending college. $V^W(a, \ell, n)$ is the value of an individual with wealth a and ability ℓ who submitted $n \in \{0, 1\}$ college applications:

$$
\begin{aligned}
V^W(a, \ell, n) &= \max_{c, c', a'} u(c) + \beta u(c') \qquad (13) \\
\text{s.t.} \qquad c + a' &= a + (1 - \tau) w_y\, \ell^{\alpha_w} - \psi \mathbf{1}_{\{n>0\}} \\
c' &= R\, a' + (1 - \tau) w_o\, \ell^{\alpha_w}.
\end{aligned}
$$

3.2 Colleges

There are two colleges who compete for students by setting their tuition schedule and admission standard taking as given the tuition schedule, admission standard, and value-added of the other college. The choices of one college affects the other through the student enrollment function defined in (9). An equilibrium in the college market will then be a fixed point in the set of possible college policies: the optimal choice of each college will be consistent with the optimal choice of the other college.

The objective of each college is to maximize its value-added, which is given by

$$\Gamma_s \equiv \xi_s Q(I_\mu, L_\mu). \tag{14}$$

ξ_s is the efficiency with which quality is transformed into value added and $Q(I_\mu, L_\mu)$ denotes the college's quality, which depends on the average amount of instructional spending per student I_μ and the average ability of the student body L_μ. The dependence of college quality on L_μ accounts for peer effects, where students benefit more from the college when its student body has a higher average ability. We also assume that Q is strictly increasing and differentiable, with $Q_{L_\mu} > 0$ for $L_\mu > 0$, $Q_{I_\mu} > 0$ for $I_\mu > 0$.

In setting their tuition schedules, colleges may offer different prices to students of different types. We assume that a is perfectly observable by the colleges, but that ℓ is unobservable. Instead, colleges can only observe the signal σ of ability described in Section 3.1. Colleges thus face a signal extraction problem: they set their tuition and admissions policy to influence the distribution of their student body, which is determined through Bayesian updating. We assume that colleges know the distribution of initial student characteristics $\mu : \mathbb{R}_+ \times \mathbb{R}_+ \to [0, 1]$, the conditional signal distribution function, and the application decisions for students of each type (which are determined in equilibrium according to 3). Taking these distributions as given, colleges take into account how their choices for tuition and the admissions standard will affect the final distribution of the student body.

To calculate the distribution of characteristics in each college, it is helpful to first define the total probability a student with characteristics (a, ℓ) will enroll in College s. Let $p_i(a, \ell)$ be the probability that such a student chooses application strategy $i \in \{AB, A1, A2\}$. The total probability is then:

$$\begin{aligned}\tilde{q}^s(a, \ell, \sigma, T^s, T^{-s}(a, \sigma)) = {} & q^s_{AB}(a, \ell, \sigma, T^s, T^{-s}(a, \sigma)) p_{AB}(a, \ell) \\ & + q^s_{As}(a, \ell, \sigma, T^s) p_{As}(a, \ell),\end{aligned} \tag{15}$$

where $q^s_i(a, \ell, \sigma, T^s, T^{-s})$ is given in Equation (9). Note that $\tilde{q}^s$ depends on the tuition schedule $T^{-s}(a, \sigma)$ of the other college if the student was accepted to both schools, revealing the nature of competition between both colleges.

Having defined the total probability of enrollment for a student of type (a, ℓ), we can easily calculate the total enrollment, average ability of the student body, and the total tuition revenue for each college. Integrating over all student types and acceptable signals, the total enrollment and average student ability in College s are given by

$$\kappa = \int_{\underline{\sigma}_s}^{\infty} \int \tilde{q}^s(a, \ell, \sigma, T^s(a, \sigma), T^{-s}(a, \sigma))\, g(\sigma|\ell)\, d\mu(a, \ell)\, d\sigma \tag{16}$$

$$L_\mu = \frac{1}{\kappa} \int_{\underline{\sigma}_s}^{\infty} \int \ell\, \tilde{q}^s(a, \ell, \sigma, T^s(a, \sigma), T^{-s}(a, \sigma))\, g(\sigma|\ell)\, d\mu(a, \ell)\, d\sigma \tag{17}$$

Similarly, total tuition revenue is given by

$$\mathcal{T}^s = \int_{\underline{\sigma}_s}^{\infty} \int T^s(a,\sigma)\, \tilde{q}^s(a,\ell,\sigma,T^s(a,\sigma),T^{-s}(a,\sigma))\, g(\sigma|\ell)\, d\mu(a,\ell)\, d\sigma. \tag{18}$$

Colleges balance their budget. Their revenue is derived from their own exogenous endowment income E^s, government subsidies or appropriations Tr^s (which may depend on the fraction of students enrolled, κ), and the total amount of tuition paid by students, $\mathcal{T}^s$. In addition to total instructional spending, $\kappa^s I_\mu$, the college faces operating expenses $C^s(\kappa)$, increasing in the college's enrollment level. For instance, these operating expenses could relate to administrative or maintenance costs that do not increase the value added to students in the labor market, but are covered by the tuition students pay. The budget constraint of College s is thus given by

$$\kappa I_\mu + C^s(\kappa) = E^s(\kappa) + Tr^s(\kappa) + \mathcal{T}^s. \tag{19}$$

College's problem. College s then solves the following problem:

$$\max_{\underline{\sigma}_s, T^s(a,\sigma), \kappa, I_\mu, L_\mu} \quad \xi_s Q(I_\mu, L_\mu) \tag{20}$$

$$\text{s.t.} \qquad T^s(a,\sigma) \leq \bar{T}^s,\ \underline{\sigma}_s \geq 0 \quad \text{and} \quad (16), (17), (18), (19).$$

Note that the problem of College s depends on the tuition and admissions standard chosen by the other college due to the dependence of $\tilde{q}^s$ on the policies of *both* colleges.

Finally, note that we have introduced an upper bound on tuition for each college. We introduce these tuition caps for two reasons. The first is to improve the empirical properties of the model. A well known feature of the higher education market is that colleges post a sticker price tuition level, which a certain fraction of the student body pays[8]. This means that colleges are limited in how much they are able to extract from students with the highest willingness to pay.

The second reason is that this constraint on tuition also induces a non-trivial choice for $\underline{\sigma}_s$. Without limits on tuition, colleges can simply charge the low-signal students enough to compensate for the decrease in average student ability they cause. Thus colleges will not need to use an admissions threshold and will simply set $\underline{\sigma}_s = 0$. When tuition is bounded, however, students with low signals who are willing to pay $\bar{T}^s$ will not be able to compensate the college for lowering its average ability. The college would therefore find it optimal to screen these students out by raising its admissions threshold $\underline{\sigma}_s$.

Optimal tuition. An interior solution for the optimal tuition level applied to students with observable

[8] We rule out the possibility of admissions being influenced by donations for students at the very top of the wealth distribution.

characteristics (a, σ) and positive probability of being accepted $(\sigma \geq \underline{\sigma}_s)$ is given by

$$T^s(a,\sigma) = \underbrace{I_\mu + C^{s'}(\kappa) - E^{s'}(\kappa) - Tr^{s'}(\kappa)}_{\text{Marginal resource cost}} - \underbrace{\frac{\int \tilde{q}^s g(\sigma|\ell)\mu(a,d\ell)}{\int \frac{\partial \tilde{q}^s}{\partial T^s(a,\sigma)} g(\sigma|\ell)\mu(a,d\ell)}}_{\text{(Posterior) markup}}$$
$$- \underbrace{\frac{Q_{L_\mu}}{Q_{I_\mu}} \left(\mathbb{E}[\ell|\sigma,a] - L_\mu\right)}_{\text{(Posterior) ability discount}} \qquad \forall\, a, \sigma \geq \underline{\sigma}_s. \tag{21}$$

The equation is derived in Appendix D.1. It is found by combining the first order conditions for the college-level aggregates κ, I_μ, L_μ with the first order condition for the tuition charged to a student with wealth a and signal $\sigma \geq \underline{\sigma}_s$.

The optimal tuition level can be broken down into three components. First, tuition must cover the marginal resource cost incurred by the college for enrolling an additional student. Note that this cost is common to all students and does not depend on a or σ. Second, since the college has market power, the tuition level also takes into account the student's willingness to enroll in the college conditional on being admitted. This introduces the familiar markup over marginal cost, which is positive since $\partial \tilde{q}^s / \partial T^s(a,\sigma) \leq 0$. The markup will be increasing in a, since wealthier students have a higher willingness to pay that can be captured by the colleges by charging higher tuition. Note that since the enrollment probability depends on the unobservable ℓ, colleges will also need to consider the distribution of possible ℓ values informed by the observable σ and a.

Third, since the college values a higher average ability among its students, there is a discount for students who on average have a higher ability than the current average ability of the student body. Since colleges cannot observe ℓ directly, they use the posterior mean of the distribution of ℓ that is obtained after observing the student's signal σ and wealth a. The posterior distribution of ℓ is given by:

$$f(\ell|a,\sigma) = \frac{\frac{\partial \tilde{q}^s}{\partial T^s(a,\sigma)} g(\sigma|\ell)\mu(a,\ell)}{\int \frac{\partial \tilde{q}^s}{\partial T^s(a,\sigma)} g(\sigma|\ell)\mu(a,d\ell)}. \tag{22}$$

Since higher ability levels yield higher average signals through the signal density g, the (posterior) ability discount will be increasing in σ. Importantly, the distribution also depends on wealth through $\tilde{q}^s$, as defined in (15). This allows the equilibrium application choices of the students to affect the tuition they are offered. For example, if at a given wealth level only high ability students apply, colleges will offer higher discounts to students at that wealth level since they are more likely to be high ability applicants. This mechanism helps explain why low-income students receive high levels of financial aid, since only the highest ability among them apply.

Optimal admission standard. Lastly, the optimal admissions standard (given the policies of the other College) satisfies the following inequality for each College s:

$$\frac{\int T^s(a,\underline{\sigma}_s)\tilde{q}(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)d\mu(a,\ell)}{\int \tilde{q}(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)d\mu(a,\ell)} \geq I_\mu + C'(\kappa) - E'(\kappa) - Tr'(\kappa) - \frac{Q_{L_\mu}}{Q_{I_\mu}}\left(\frac{\int \ell\ \tilde{q}(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)d\mu(a,\ell)}{\int \tilde{q}(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)d\mu(a,\ell)} - L_\mu\right), \tag{23}$$

where $\tilde{q}^s(\underline{\sigma}_s) = \tilde{q}^s(a,\ell,\underline{\sigma}_s,T(a,\underline{\sigma}_s))$. The first order condition in (23) holds with equality when $\underline{\sigma}_s > 0$, and its derivation is provided in Appendix D.2. Equation (23) can be interpreted to mean that the average tuition revenue received from the lowest-signal students must be enough to compensate the College for the resource cost they impose, and the change they bring to the average ability.

Note, however, that the College's tuition policy in Equation (21) guarantees that it will be compensated for admitting the lowest-signal students since tuition depends on both σ *and* a. Thus, without any constraints on tuition, a College will not choose an interior value for $\underline{\sigma}_s$. Proposition 1 below formalizes this by demonstrating that when there are no caps on tuition, (23) will never hold with equality since the College will be able to charge the lowest-signal students enough to compensate for the fact that they lower the average ability of the student body.

Proposition 1. *Suppose $\bar{T}^s$ is large enough so that the constraint $T^s(a,\sigma) \leq \bar{T}^s$ never binds. Then $\underline{\sigma}_s = 0$. That is, if there are no tuition caps, the college does not exclude students using an admissions threshold.*

Proof. See Appendix D.3 □

When there is a binding tuition cap, however, students with very low-signals will not be able to compensate the College for lowering its average ability. The College will then have to raise its admissions standard to exclude these students. This can be seen by examining (23). A low enough tuition cap will lower the left hand side by decreasing tuition revenue. Once the left hand side is sufficiently small, the College will need to increase $\underline{\sigma}_s$ to lower the right hand side and achieve equality. Increasing $\underline{\sigma}_s$ lowers the right hand side because it raises the average ability of the College's lowest-signal students, $\frac{\int \ell\ \tilde{q}(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)d\mu(a,\ell)}{\int \tilde{q}(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)d\mu(a,\ell)}$.

3.3 Government

The government taxes labor income and uses the revenue to finance Pell Grants and college subsidies. The tax base is composed of all workers who did not attend college (including those who applied and those who did not) in both periods, and college educated workers in the second period. The intertemporal government budget constraint must hold according to

$$\sum_s \left[Tr^s(\kappa^s) + \int_{\underline{\sigma}_s}^{\infty} \int \hat{P}(a) \tilde{q}^s g(\sigma|\ell) d\mu(a,\ell) d\sigma \right] =$$

$$\tau\, w \left[(1+R^{-1}) \int \ell^{\alpha_w}\, \tilde{q}^w(a,\ell)\, d\mu(a,\ell) \; + R^{-1} \sum_s \Gamma_s \int_{\underline{\sigma}_s}^{\infty} \int \ell^{\alpha_s} \tilde{q}^s g(\sigma|\ell)\, d\mu(a,\ell) \right], \tag{24}$$

where $\tilde{q}^w$ is the total probability of not attending college:

$$\tilde{q}^w(a,\ell) = 1 - \sum_s \int_{\underline{\sigma}_s}^{\infty} \tilde{q}^s(a,\ell,\sigma,T(a,\sigma)) g(\sigma|\ell) d\sigma. \tag{25}$$

3.4 Equilibrium

An equilibrium in the college market consists of value functions for applicants V^{Aj}, students with offers of admission V^{Oj}, enrolled students V^{Cj}, and workers V^W for $j = \{1,2,B\}$, policy functions $\{d, a^{s'}, c^s, a^{w'}, c^w\}$, probabilities $\tilde{q}^s, \tilde{q}^w$, admissions standards $\underline{\sigma}_s$ and tuition schedules $T^s : R_+ \times R_+ \to (-\infty, \bar{T}^s]$, college choices $\{\kappa^s, I^s_\mu, L^s_\mu, \Gamma_s\}$ for $s = \{1,2\}$, and tax rate τ, such that

1. Given $\{\tau, \Gamma_s\}$, $\{V^{Cs}, a^{s'}, c^s\}$ solves (10) for $s \in \{1,2\}$, and $\{V^W, a^{w'}, c^w\}$ solves (13).
2. Given $\{V^{Cs}, V^W\}$ and T^s for $s \in \{1,2\}$, V^{Oj} satisfy (7)-(8) for $j \in \{1,2,B\}$.
3. Given V^{Oj} and $\{T^s, \underline{\sigma}_s\}$ for $s \in \{1,2\}$, V^{Aj} solve the problem of the applicant in (1)-(2) for $j \in \{1,2,B\}$.
4. $d : R_+ \times R_+ \to \{W, AB, A1, A2\}$ is the application choice which solves (3).
5. $\tilde{q}^s, \tilde{q}^w$ are, respectively, the total probabilities of ending up at College s and ending up as a worker:

$$\begin{aligned} \tilde{q}^s(a,\ell,\sigma,T(a,\sigma)) &= q^s_{AB}(a,\ell,\sigma,T^s(a,\sigma))\mathbf{1}_{\{d(a,\ell)=AB\}} \\ &\quad + q^s_{As}(a,\ell,\sigma,T(a,\sigma))\mathbf{1}_{\{d(a,\ell)=As\}} \\ \tilde{q}^w(a,\ell) &= 1 - \sum_s \int_{\underline{\sigma}_s}^{\infty} \tilde{q}^s(a,\ell,\sigma,T(a,\sigma)) g(\sigma|\ell) d\sigma. \end{aligned}$$

   where $q^s_i, i \in \{AB, As\}$ is defined in (9).
6. Given $\tilde{q}^s$ for $s \in \{1,2\}$, $\{\kappa^s, I^s_\mu, L^s_\mu, \Gamma_s, T^s, \underline{\sigma}_s\}_{s\in\{1,2\}}$ is a solution to the the College game presented in Section 3.2. That is, for each College $s \in \{1,2\}$, given the tuition schedule, admissions threshold and value added of the other college, $\{\kappa^s, I^s_\mu, L^s_\mu, T^s, \underline{\sigma}_s\}$ is a solution to (20), and

$$\Gamma_s = \xi_s Q(I^s_\mu, L^s_\mu).$$

7. The government balances its budget according to equation (24).

3.4.1 Equilibrium Selection

The presence of peer effects introduces the potential for multiple equilibria. We follow Epple et al. (2006) and Epple et al. (2017) in focusing on an equilibrium where the ranking of college quality corresponds to the ranking of the endowment size. In our baseline equilibrium, College 1 will then have the higher value added and attract the higher ability students. This is consistent with its higher endowment which allows for higher amounts of spending per student. We do not explore the possibility that College 2 will have a higher value added despite having lower endowment income.

Appendix G provides details on how the equilibrium of the college market is solved numerically.

4 Estimation

We pick a subset of the model parameters from the data and literature, and estimate the rest using the method of moments. Table 2 summarizes the parameters determined outside the model, while Table 4 summarizes the resulting estimated parameters.

4.1 Students' attributes

We start by describing our choices for the parameters that govern the students' problem. A list of these parameters and their values are given in Table 2.

Data. The HSLS is described in detail in Section 2.1. For the estimation, we use it to calculate key moments regarding student application and enrollment rates. We also use it to estimate the parameters governing the distribution of students characteristics, i.e. the distribution over (a, ℓ).

The BPS is a longitudinal dataset from the National Center for Education Statistics which follows a cohort students who begin college for the first time. The data include details of all financial aid received by a representative cohort of students who began college in the 2011-2012 academic year (close to our HSLS cohort who began college in the fall of 2013). This rich information about enrollment, tuition, and financial aid helps us supplement the HSLS in parameterizing our model. We use the BPS to set the level of grants available to students and determine how grants vary with parental income. We also use it to determine the college tuition caps.

Preferences. Individuals have logarithmic preferences over consumption each period:

$$u(c) = \log(c)$$

In our two period model, we consider the first period to account for 4 years (time in college), and the second period to account for 60 years (time spent working). In order to pick the appropriate values for $\{\beta, R, w\}$, we show in Appendix E how a life-cycle model with $T + 1$ periods maps into our two

Table 2: Exogenous Parameters

Variable	Description	Value	Source
Returns to education			
α_s	Returns to ability (school)	0.78	Abbott et al. (2019)
α_w	Returns to ability (no school)	0.55	Abbott et al. (2019)
Aggregate prices and borrowing constraint			
$\underline{a}_s$	Student borrowing limit	-0.8	USED
R	Gross interest rate	1.0386	USED
w	Wage	2.7	CPS
Budget Parameters			
ψ	Financial application cost	0.00375	IPEDS
g_0	Intercept of grant function	–	BPS
g_1	Slope of grant function	–	BPS
P_{max}	Pell Grant maximum	0.5645	USED
College Parameters			
o_2^s	Cost function	[5.77, 0.43]	IPEDS
E^s	Endowment income	[0.12, 0.06]	IPEDS
Tr_1^s	Gov transfers per student	[1.00, 0.62]	IPEDS
$\bar{T}^s$	Tuition Cap	[2.5, 1.2]	BPS

period model, where we take $T = 15$. In the life-cycle model, we set $\beta = (0.95)^4$. Mapped into our two-period model, this gives us the values $\tilde{\beta} = 4.19$ that correspond to the problem presented in (10).

In order to simplify the computation of the equilibrium, we add a set of Type I extreme value shocks to the discrete choice problem in (3). With the shocks, the problem becomes

$$\max\left\{V^{A1}(a,\ell)+\epsilon_A^1, V^{A2}(a,\ell)+\epsilon_A^2, V^{AB}(a,\ell)+\epsilon_A^3, V^W(a,\ell,0)+\epsilon_A^4\right\}. \tag{26}$$

where we let $\epsilon_A \sim Gumbel(\frac{1}{\lambda_a})$. The shocks simplify the equilibrium computation since they smooth the application probabilities taken as given by the Colleges when setting tuition. The shocks allow these probabilities to respond continuously to changes in the tuition schedule in each iteration, which simplifies the convergence of the solution algorithm described in Appendix G. Since the shocks are used only to simplify the computation, we pick λ_a to be large in order to minimize their variance. We set $\lambda_a = 60$. Note that these shocks are included in addition to the shocks that occur after the application stage, when the students make their enrollment decision. We include the scale parameter for the second-stage shocks in our method of moments estimation, described in more detail below.

Distribution of characteristics. We take the distributions of student characteristics, i.e. (a, ℓ), directly from the HSLS. Since we work directly with the amount of "wealth" the students receive from their parents (a in the model), we abstract from parental transfers so we do not use parental income data directly. Instead we use the student's Expected Family Contribution (EFC), which is determined according to rules set by the Department of Education using the FAFSA filled out by the students and their parents. EFC is a measure of the amount of resources a student reasonably has available to them in order to attend college (before financial aid), and so maps well into our notion of student wealth in the model. Since the students in the HSLS began college in the 2013-2014 academic year, we use 2013 dollars as our numeraire and re-scale by \$40,000. Note that a period in the model corresponds to four years. Thus, for example, $a = 1$ in the model corresponds to an EFC of \$10,000 per year over four years.

In the HSLS, we are able to observe the EFC for all students who completed the FAFSA and attended college, and we calculate it for those who did not using the household income reported by their parents in the survey. Our direct calculation uses the EFC formula established by the Department of Education, and is described in detail in Appendix C. In the HSLS, we find that approximately 30% of the students have an EFC of 0. In fitting our distribution, we therefore assign a mass point of 30% for $a = 0$. For the remaining 70% of the distribution with $a > 0$, we assume that a follows a log-normal distribution, $a \sim LogNormal(\mu_a, \sigma_a^2)$.

We use the students' SAT score in the HSLS as a proxy for ability to determine the distribution of ℓ. Again, we assume that ℓ follows a log-normal distribution, $\ell \sim LogNormal(\mu_\ell, \sigma_\ell^2)$. Since a and ℓ are correlated, we estimate these parameters separately for the case where $a = 0$ and where $a > 0$.

To summarize, the joint distribution of (a, ℓ) is given by:

$$\begin{cases} (a = 0, \ell) \sim LogNormal(\mu_{\ell 0}, \sigma_{\ell 0}^2) & w.p. \;\; 30\% \\ (a > 0, \ell) \sim LogNormal\left(\begin{bmatrix} \mu_a \\ \mu_{\ell 1} \end{bmatrix}, \begin{bmatrix} \sigma_a^2 & \sigma_{a\ell} \\ \sigma_{a\ell} & \sigma_\ell^2 \end{bmatrix} \right) & w.p. \;\; 70\% \end{cases}$$

(The values of these distributional parameters are currently suppressed as they are undergoing disclosure review).

Signal distribution. We assume that signals follow a normal distribution conditional on ability, with mean ℓ, and variance σ_g^2. We also assume that the distribution is truncated below at 0, so that the lower bound of the support of the signals is finite. We include the variance σ_g^2 as part of the joint parameter estimation below.

Aggregate Prices. We choose $w = 2.7$, so that the model produces the average wage calculated from the Current Population Survey (CPS). We set $R = (1.0386)^4$ since annual borrowing rates for undergraduate students was 3.86% in 2013-2014 academic year. This rate was set by the US Department of Education (USED) through the Federal Student Loan Program. Again, we need to adjust these life-cycle level values to fit into our two period model using Appendix E. Doing so gives us $\tilde{w} = 4.03, \tilde{R} = 1.23$.

Financial application costs. The application process involves two types of costs: a psychic cost and a financial cost. We find in the HSLS that students send about 3 applications on average, and from IPEDS that the average application cost is \$50. This total cost corresponds to a value of $\psi = 0.00375$ in the model. Note that for simplicity, the financial cost of applying is the same regardless of whether students send one or two applications. The marginal cost of sending an additional application in the model is instead captured by the psychic costs since $\phi_{12} > \max\{\phi_1, \phi_2\}$. Since the financial cost includes only the direct resource cost of applying to college, all other indirect costs are captured by the psychic costs. These include the disutility of effort that may come from writing essays, or the cost of lost leisure time from applying (we do not include forgone wages as part of the resource cost).

Returns to ability and education. We take the parameters governing the labor market returns to ability α_s, α_w directly from Abbott et al. (2019). They find that for college graduates, the ability gradient is 0.797 for males and 0.766 for females. For high school graduates they find the gradients to be 0.517 and 0.601 for males and females respectively. Hence we simply set $\alpha_s = 0.78$ for graduates of both colleges and set $\alpha_w = 0.55$. The labor market returns for attending the more selective school are then captured only by differences in Γ_s.

Borrowing constraints. According to the Federal Student Loan Program, the aggregate limit for dependent students who are attending an undergraduate degree is \$31,000 in federal loans. Students also have access to private credit markets, and we see in the BPS that many students borrowed approximately \$10,000 in cumulative private loans. For that reason, we set the student borrowing limit to $\underline{a}_s = -1.0$, which corresponds to \$10,000 per year over four years.

Grants and Aid. We consider two types of grants available to students who attend college and allow these to vary by the student's wealth level. The first are grants that are exogenous in the model, denoted $\hat{g}(a)$. These exogenous grants stand in for unmodeled state grants or private grants including scholarships. The second grants stand in for Pell Grants that are paid for by the government and financed through taxes, denoted $\hat{P}(a)$.

To find $\hat{g}$ in the data, we add up all grants and aid that the student received excluding loans, Pell grants, and institutional aid (which are all endogenous in the model or based on government policy). We find that these grants are decreasing in EFC and tend to level off at around \$15,000. We therefore assume that $\hat{g}$ varies with the student's wealth according to:

$$\hat{g}(a) = \max\{g_0 - g_1 a, 0.15\}. \tag{27}$$

We estimate g_0, g_1 using OLS for EFC less than the \$15,000, controlling for SAT score. (The estimated values of g_0, g_1 are currently suppressed as they are undergoing disclosure review).

Pell Grant amounts are set according to the Department of Education. In order to qualify, students must demonstrate sufficient financial need as measured by the difference between their EFC and the Net Cost of Attendance (tuition plus room and board minus institutional financial aid). The Pell Grant makes up this difference, up to a maximum level. In the model, we simply set

$$\hat{P}(a) = \max\{P_{max} - a, 0\}, \tag{28}$$

so that Pell grants effectively give all low wealth students a minimum level of wealth. In 2013, the maximum Pell grant amount was \$5,645 so we set $P_{max} = 0.5645$. We later study the effects of changes to this parameter in the policy section of the paper.

4.2 Colleges' attributes

To pick parameters for the colleges, we rely on data made available through the Integrated Post-secondary Education Data System (IPEDS). This database contains rich information on colleges' enrollment patterns, student compositions, and balance sheets. We provide details about how we construct our sample of colleges from IPEDS in Appendix B.2.

College Types. To separate both types of colleges in the model, we rely on Barron's 2015 rankings of U.S. colleges. This ranking is commonly used in the literature as a measure of school quality. We consider College 1 to be Barron's Tier 1 and 2 schools ("elite" and "highly selective"), and College 2 to be all other four year universities (excluding for-profit schools). Table 3 summarizes key empirical differences across these school types.

College Type	**1**	**2**
Number of Colleges	182	1,621
Undergraduate Enrollment	17%	83%
Fraction Public	53%	75%
Average SAT Score	1297	1056
Rejection Rate	57%	32%
Instructional Expenditures per Student (US$)	22,468	8,846
Average Tuition and Fees (US$)	26,327	12,077
Net Cost for Bottom 20% Income (US$)	8,750	9,646
Median earnings 10 Years After Entry (US$)	58,706	41,906
Endowment Assets per Student (US$)	141,838	12,332

Table 3: Empirical Differences Across College Types. Source: IPEDS, College Scorecard.

Table 3 shows that the differences between each college type are consistent with the model's mechanism. The top colleges enroll a smaller share of the total student population, have higher expenditure per student, higher SAT scores, higher average tuition, while still offering lower tuition to students at the bottom of the income distribution. This suggests that defining school types in the model based on Barron's rankings leads to a natural partition between colleges that can be used for the quantitative analysis. Note that empirically, the top schools as a group still account for students mostly enrolled in public schools. Therefore, unmodeled institutional differences between public and private schools are unlikely to play a big role in explaining differences between the two types of schools since both groups have a similar composition of students within each type.

Technology. Colleges' quality is defined by

$$Q(I_\mu, L_\mu) \quad = \quad I_\mu^{1-\rho_L} L_\mu^{\rho_L},$$

where ρ_L is the share parameter estimated internally.

The ratio of marginal qualities needed to compute net tuition and admission standards are given by

$$\frac{Q_{L_\mu}}{Q_{I_\mu}} = \frac{\rho_L}{1-\rho_L}\left(\frac{I_\mu}{L_\mu}\right).$$

Costs and revenues. We follow Epple et al. (2017) and assume colleges operating costs are a quadratic polynomial in their total enrollment,

$$C^s(\kappa) = o_0^s + o_1^s \kappa + o_2^s \kappa^2. \tag{29}$$

We set the linear cost $o_1^s = 0$, and estimate the fixed cost and quadratic cost terms using our sample of colleges from the IPEDS data. We first compare the share of student enrollment in each type of college in our IPEDS sample to the enrollment shares we observe in the HSLS. We find the following fraction of students who enroll at each type of college in the HSLS:

$$\kappa^1 = 0.069 \qquad \kappa^2 = 0.357.$$

To confirm that these numbers are consistent with our IPEDS data, we first add up all full-time equivalent undergraduate students across all schools in our sample. We then infer the total number of high school graduates using the 42% 4-year enrollment rate reported by NCES[9]. We find that of this population, 7% attend College 1 and 35% College 2, similar to the HSLS enrollment patterns described above.

Confident that our IPEDS sample includes all the schools relevant for our analysis, we use variation in school enrollment size at each college type to estimate the quadratic term in the cost function with the following regression:

$$cost_i^s = O_0^s + \hat{O}_2^s k_i^2 + \varepsilon_i, \tag{30}$$

for college i of college type s, where k is the enrollment level of school i. In order to measure the costs for each college in our sample, we use detailed financial data available from IPEDS. Our cost variable c_i is derived from the budget constraint in Equation (19). We infer these costs by adding up all tuition revenue, net grant revenue, government appropriations, unrestricted revenue from private sources, and subtract off total instructional expenditure.

In order to aggregate our college level estimates of the cost functions and the corresponding variables in the model, we again follow Epple et al. (2006) in their aggregation procedure. If there are n_s colleges in type s, we assume they are identical so that total enrollment is given by $K = n_s \kappa_s$, so we have

$$C(K) = n_s O_0^s + n_s O_2^s k^2$$
$$C(K) = o_0^s + o_2^s K^2$$

where $o_0^s = n_s O_0^s$, and $o_2^s = O_2^s / n_s$. We have 182 colleges in type 1, and 1,483 in type 2, so we find that $(o_2^1, o_2^2) = (5.77, 0.44)$, and $(o_0^1, o_0^2) = (0.15, 0.25)$, based on our estimates of the parameters from (30).

College subsidies and endowments. We assume that government subsidies to colleges simply depend linearly on the size of the school's student body:

$$Tr^s(\kappa) = Tr_1^s \kappa.$$

[9] See here: https://nces.ed.gov/programs/coe/indicator_cpa.asp.

Similar to the cost function estimation above, we estimate Tr_1^s for each College type s using variation in enrollment within each college type from the IPEDS data. We find $Tr_1^1 = 1.0, Tr_1^2 = 0.62$.

We assume that private endowment income at the colleges has a component that depends on enrollment (private donations that are restricted to be used on scholarships per enrolled student), and a fixed component component:

$$E^s(\kappa) = E_0^s + E_1^s \kappa. \tag{31}$$

We make this distinction for endowment income because the enrollment levels for each college are determined largely by the net fixed costs: $o_0^s - E_0^s$[10]. For each college, enrollment will then be determined largely by the fraction of the endowment used to offset fixed costs, and the fraction used to offset per-student costs. This is important for College 1 where the endowment is large, so we estimate E_1^1 to match the observed College 1 enrollment (discussed in the estimation procedure below). For College 2, we find that we are unable to match the observed enrollment given our estimate of o_0^2 even if we assume that $E_0^2 = 0$. We thus leave College 2 enrollment as an un-matched target and simply assume that $E_0^2 = 0$.

Finally, to find the remaining endowment parameters, we use our IPEDS sample to calculate the total level of private endowment income received by each college by adding up all private revenue the college received. This includes unrestricted revenue the school may use from gifts, investment return from their endowments, or contributions from affiliates. We find the *total* endowments to be $E^1 = 0.12, E^2 = 0.06$. Thus for College 1, we set $E_0^1 = 0.12 - E_1^1 \kappa^1$ (using our estimate for E_1^1 and taking κ^1 directly from the data), and for College 2 we have $E_1^2 = 0.06/\kappa^2$.

Constraints. To find the tuition caps $\bar{T}^s$ we first note that in the data, the average level of tuition minus college financial aid is flat in EFC for high-EFC students, reflecting the fact that such students do not receive need-based financial aid. As in the model, where the constraint on tuition binds for students with high wealth, we pick $\bar{T}^1 = 2.5,\ \bar{T}^2 = 1.2$ to be consistent with the data.

Value added. Colleges' efficiency parameter ξ_s is chosen to match the college wage premium in the data. Abbott et al. (2019) estimate an average college wage premium of 0.6. Since 17% of students are enrolled in selective schools, we have that the average wage premium is such that

$$0.17 \log \Gamma_1 + 0.83 \log \Gamma_2 = 0.6.$$

To find the dispersion between Γ_1, Γ_2, Chetty et al. (2020) report that 80% of the difference in median log earnings 10 years after college can be explained by differences in colleges' selectivity. Using the numbers from Table 3, we have

$$\log \Gamma_1 - \log \Gamma_2 = 0.113.$$

This gives us $\Gamma_1 = 2.01, \Gamma_2 = 1.79$. We fix these corresponding values for Γ_s in the baseline estimation below, and find ξ_s so such that the resulting value for I_μ^s, L_μ^s are consistent with the corresponding Γ_s values.

[10] Since colleges only care about *per* student variables, they have an incentive to keep their enrollment levels as small as the net fixed costs permit.

4.3 Method of moments estimation

We estimate the remaining parameters $\Theta = \left\{\phi_{s1}, \phi_{s2}, \phi_2, \lambda_c, \nu^1, \nu^2, E_1^1, \rho_L, \sigma_g^2\right\}$ jointly by minimizing an unweighted quadratic distance criterion function between data moments and calculated model moments. The results of the estimation are presented in Table 4, where we provide description of the data moments used.

Variable	Description	Value	S.E.
ϕ_1	College 1 application cost	0.26	0.26
ϕ_2	College 2 application cost	0.48	0.26
ϕ_{12}	Both application cost	0.70	0.32
ν^1	College 1 preference	1.24	1.15
ν^2	College 2 preference	0.33	0.34
λ_c	EV shock scale parameter	1.24	0.31
σ_g^2	Variance of signal	0.05	0.09
ρ_L	Quality parameter	0.82	0.11
E_1^1	College 1 endowment	1.18	0.57
	Moment	**Model**	**Data**
	% Applying College 1	2.00	2.00
	% Applying College 2	42.0	42.0
	% Applying to both	13.9	14.0
	Coll 1 conditional attend rate	–	–
	Coll 2 conditional attend rate	–	–
	College 2 attendance rate	–	–
	College 1 acceptance rate	0.72	0.73
	Tuition elasticity wrt SAT	-0.08	-0.08
	% Enrolled College 1	7.08	7.00

Table 4: Jointly estimated parameters. A few data moments are currently suppressed as they are undergoing disclosure review.

Psychic application costs. To help identify the psychic application costs, we match the fraction of students applying to each college, and the fraction of those applying to both. In the HSLS, we observe that 2% apply only to College 1, 42% apply only to College 2, and 14% apply to both. Surprisingly, we find that the estimate for the College 1 application is lower than the College 2 application. Since it is risky to apply only to College, a relatively small application cost is required to match the target.

To assess the magnitude of these costs, we report in Table 5 the average psychic application costs in the population as a percentage of average life-cycle consumption. We see that in terms of life-cycle consumption, the application costs to each individual college are similar. Finally, we note that the application costs are considerably smaller than those estimated in the literature. The reason is that our model accounts for the fact that students need to be admitted to college in order to attend, and the probability of not getting in is sufficiently low for many students. Thus our model can help rationalize the relatively high psychic costs of schooling found in the literature: many students do not to enroll not because of preferences, but because they are unlikely to be admitted and therefore do not apply in the first place. A relatively small psychic cost is then sufficient to exclude them from the college market[11].

	Avg psychic cost
College 1 application	2.51%
College 2 application	3.52%
Application to both	4.43%
Abbott et al. (2019)	10.7%

Table 5: Average psychic application costs among the student population expressed as a percentage of average life-cycle consumption (calculated in the model to be about $800,000). The psychic cost of college calculated in Abbott et al. (2019) is provided for reference to the literature.

Extreme value shock. λ_c is the scale parameter for the extreme value shock introduced in (4)-(6) that is realized after a student receives offers from either college. Unlike the shocks that were added to the application choice, these shocks are meaningful because they encourage students to apply more aggressively by increasing the option value of having another college to choose from (compare 7 with 8). Therefore, if λ_c is low then the large fraction of students applying only to College 2 will be unlikely to enroll since their applications were motivated more by the increased option value rather than the value-added from attending College 2. Thus, variation in College 2 attendance conditional on applying only there and being accepted created by variation in λ_c helps us identify the parameter.

College preference parameters. In (10), we introduced preference parameter ν^s to capture additional net benefit (or psychic cost) for attending College s. If positive, these parameters indicate that there is an additional benefit to attending colleg above the labor market return that will help offset the psychic costs of applying. If negative, these parameters will indicate that college is costly to complete. Variation in these parameters leads to changes in the flows of students into colleges conditional on being accepted, which allows help identify them using our data. For ν^1, We find that a large, positive value is necesary to correclty match the fraction of students accepted to both colleges who choose College 1. For ν^2, we match the fraction of students who enroll in College 2 after being rejected by College 1. Perhaps surprisingly, a positive value for ν^2 is required to match this fraction, suggesting that only the labor-market value-added from attending College 2 is not enough to account for its enrollment pattern.

College quality and signal strength. The ability parameter ρ_L in the colleges' quality function and the variance of the signal σ_g^2 are chosen to help the model match the average responsiveness of tuition with respect to ability, and the average acceptance rate to College 1 conditional on applying. As discussed in Section 3, the (posterior) ability discount depends on the strength of the signal, so

[11] Note that Abbott et al. (2019) calculate psychic costs separately for the overall population and among those who attend college. Since psychic costs in their model depend on ability, they are considerably smaller for the college-going population. In our model, we assume that the psychic application costs are common for all students.

that the average responsiveness of tuition to the signal is sensitive to the choice of the variance. In the BPS, we estimate this effect by regressing tuition minus institutional grants on SAT score, controlling for EFC and college fixed effects.

Similarly, matching the average conditional acceptance rate for College 1 helps us identify ρ_L because we find that a higher value for this parameter increases the size of the ability discount, which increases the attractiveness of applying to College 1. All else equal, this will tend to reduce acceptance rate due to the increase in applications with little change to the enrollment.

Endowment and Fixed Costs. Lastly, the fixed cost each college faces is an important variable for determining its enrollment level. As mentioned above, colleges care about *per student* variables, so they will remain as small as their fixed costs allows them to be. Thus, variation in the fixed component of a college's endowment, E_0 in (31), will have an effect on its overall enrollment level. We find that this is particularly important for College 1, who has a high total endowment income relative to its estimated fixed cost. We then estimate E_1^1 to match the College 1 enrollment level observed in the data, and back out the corresponding value for E_0^1 from the observed total endowment income calculated from IPEDS. For College 2, we find that the empirical enrollment level is too large to be explained only by its endowment. We thus leave its enrollment level to be an un-targeted moment.

4.4 Model Fit

College Level Statistics. Table 6 below compares college level statistics produced by the model to ones we calculate in the data. We find that the model somewhat underestimates the average EFC within each college, though it captures the fact that EFC in College 1 is about twice as high as in College 2. In order to compare average student ability in the model and in the data, we report the average SAT and average application signal within each college (since they are both noisy signals of true ability). To put the two measures in similar units, we report the standardized values (where we standardize the signal according to the distribution of applicants). As we see in the second row of the table, the model does a good job of allocating relatively high signal students into College 1, and students with average signals into College 2.

	College 1		**College 2**	
	Model	*Data*	*Model*	*Data*
Avg Student EFC	2.19	3.04	1.08	1.55
Avg Standardized Ability (SAT)	1.53	1.12	0.16	0.05
Instr Spending per Student	2.29	2.25	1.06	0.89
Average Net Tuition	1.79	1.41	0.66	0.50

Table 6: College Level Statistics

Finally, the model is able to capture the large difference in instructional spending per student across each college. The amount spent per student in College 1 is large due to its higher tuition

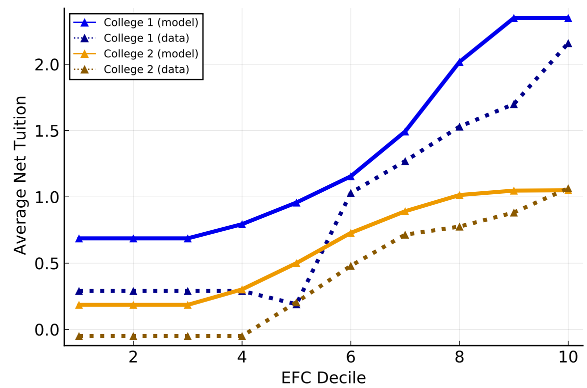


Figure 4: Average net tuition for students by Expected Family Contribution (EFC) decile. Source: BPS.

levels and large endowment. The model is also able to capture the fact that spending per student is significantly higher than average tuition revenue per student, which reflects how government subsidies and grants are able to cover remaining college expenses. Note, however, that average net tuition at each college is slightly exaggerated in the model compared to the data.

Tuition Variation. Figure 4 shows how the average net tuition in the model compares with the data in each college when students are grouped by EFC decile. In the data, net tuition is defined as sticker price tuition minus all grants available to the student. Thus in the model, we set net tuition for a student with wealth a and signal σ to be $T(a, \sigma) - \hat{P}(a) - \hat{g}(a)$. Due to the presence of the tuition caps we have imposed, the model is generally able to capture the tuition for students at the very top of the EFC distribution, especially in College 2. In College 1, we see that the model generally over-predicts the tuition charged to students. This is likely due to the relatively high value for ν^1 in the baseline estimation, which allows College 1 to charge high tuition levels even when competing with College 2. Another reason tuition is high relative to the data is that nearly all students in the model borrow the full amount of student loans available to them, allowing the colleges to charge higher tuition. Student borrow the full amount because the returns to college are sufficiently high and there is no risk in the model, so students borrow heavily to equalize consumption across both periods.

4.5 Relation to Empirical Literature

4.5.1 Returns to Selectivity

Despite the fact that graduates of more selective colleges have higher earnings (see, Table 3), Dale and Krueger (2002) show that these differences disappear when controlling for the set of colleges that each student was accepted to. That is, conditional on being accepted to the same colleges, two students with similar observable characteristics will earn similar wages if they attend different colleges. This result suggests that there are no returns to selectivity, and controlling for the set of colleges students were accepted to helps account for the potential selection bias of higher ability students simply choosing more selective colleges.

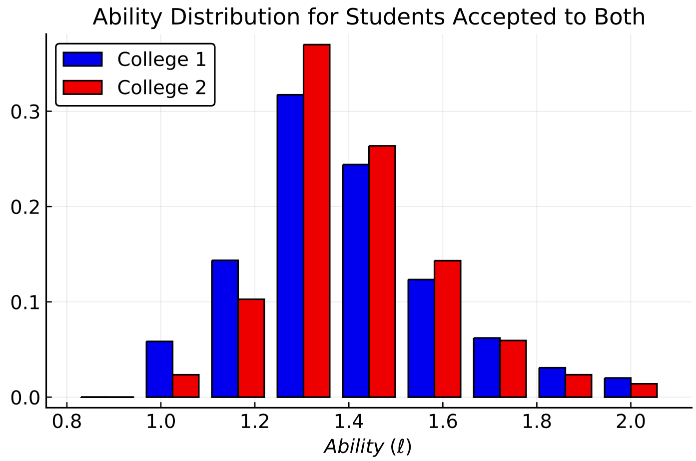


Figure 5: Model implied ability distribution at each college for students who were accepted to both colleges.

Our paper helps shed additional light on this result by considering the admission and enrollment stages sequentially. Since there is still selection at the enrollment stage, it can be potentially misleading to compare students accepted to the same colleges because their decision to attend different colleges is not random. For example, low-income students who are high-ability may choose not to attend the less selective college because they do not receive sufficient financial aid, whereas higher-income, lower-ability students admitted to the same colleges may attend the more selective one. This difference in ability may cancel out potential value-added that the more selective school could have offered the lower-income student.

We use our model to examine the extent to which this selection is present. Figure 5 shows the distribution of ability within each college for students who were accepted to both. It shows that selection effects are strong enough for the students in College 2 to surprisingly be higher ability on average than the students in College 1 (conditional on being accepted to both colleges), despite the fact that College 1 students are on average higher ability overall. This happens because among high-income students, the high ability ones apply *only* to College 1, while the high-ability low-income students apply to both. Thus the pool applicants applying to both colleges consists of relatively low-ability high-income students, and relatively high-ability, low-income students. From this pool, enough low-ability, high-income students are able to enroll at College 1, leading it to have a lower average ability among students who were accepted to both colleges, despite generally being more selective.

This mechanism illustrates the importance of needing to control for parental income when testing for the presence of returns to selectivity. Dale and Krueger (2002) for example rely on student reported parental income data that is missing for many students in their sample, and Dale and Krueger (2014) rely on imprecise proxies for parental income. In more recent work Chetty et al. (2020) estimate that around 80% of the differences in earnings across colleges can be attributed to causal effects, even after controlling for the set of colleges the students applied to. They suggest that their precise measure of parental income may help explain why they find returns to selectivity while the previous papers do not.

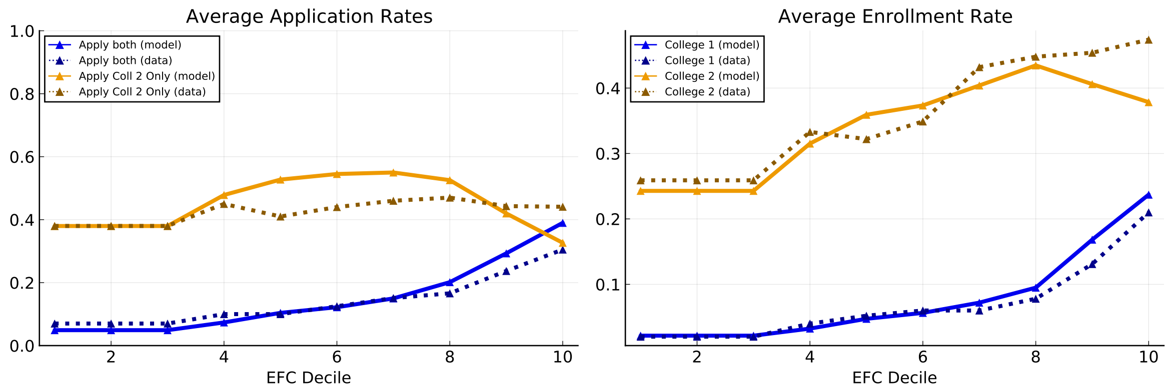


Figure 6: The left panel displays the average application rates for students by Expected Family Contribution (EFC) decile, and the right panel displays the average enrollment rates to each college. In the left panel, we include only applications to College 2 or applications to both since the fraction of students only applying to College 1 is small both in the model and the data. Source: HSLS.

5 Results

5.1 Student Application and Enrollment Patterns

This section shows how the model is able to account for the low application and enrollment rates observed for low-income students despite the presence of substantial financial aid described above. Figure 6 shows the model predicted average application and enrollment rates alongside the ones observed in the HSLS data for students grouped by Expected Family Contribution (EFC). We see that the model is able to account for the fact that applications to the more selective college are generally increasing in EFC (we do not report the fraction only applying to College 1 since it is small). The model somewhat underestimates the fraction of low EFC students who apply, since the ones that do have a relatively high chance of being admitted. The model also overestimates the application rates for students at the very top.

The right panel in Figure 6 shows the model predicted enrollment rates and the ones observed in our data. We find that the model does a good job of capturing the increasing patterns in enrollment for students at both colleges, though the model slightly overestimates enrollment at the top for College 1. Note that in each plot the overall means of enrollment and application are targeted in our baseline estimation, but the distributions are not.

Student Distribution Within Colleges. Next, we show that the model delivers student income distributions within each college that are consistent with the data. Table 7 below shows the joint distribution of student wealth (Expected Family Contribution in the data), and ability (SAT/ACT score in the data) within in each college[12]. We see that in both colleges the EFC distribution is well accounted for, especially in College 1 where the model very closely captures the correct share of low-income students. Normally in this type of college-market model, colleges have a strong incentive to

[12]Note that SAT score is only a noisy measure of true ability, so there is subsantial dispersion in the data compared to the model, where college policies are designed to admit high ability students.

enroll mostly students from the top of the income distribution since they pay higher tuition, and are generally higher ability since income and ability are correlated. Epple et al. (2017), for example, over-predict the share of high-income students, and argue that unmodeled social objectives like affirmative action may help explain the gap in low-income student enrollment.

		Model				Data			
		EFC				EFC			
College 1		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Ability	Q1	0.0	0.0	0.0	0.0	–	–	0.0	0.0
	Q2	0.0	0.0	0.0	0.0	–	0.7	–	0.5
	Q3	0.0	0.0	0.0	0.0	1.8	1.4	5.4	5.2
	Q4	9.3	11.7	25.7	53.3	6.4	9.8	17.3	49.5
	Total	9.3	11.7	25.7	53.3	8.2	11.9	22.7	55.2
		EFC				EFC			
College 2		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Ability	Q1	0.0	0.0	0.0	0.0	4.9	2.0	1.3	0.9
	Q2	0.0	0.0	0.0	0.0	6.6	3.6	5.5	4.1
	Q3	12.4	10.4	13.2	10.9	6.7	5.8	10.1	10.2
	Q4	10.1	11.1	18.2	13.7	5.1	4.8	11.3	17.1
	Total	22.5	21.5	31.4	24.6	23.3	16.2	28.2	32.3

Table 7: Joint Distribution of Student Wealth/Expected Family Contribution and Ability/SAT Score

In our model, the application and admissions system helps us capture the correct share of low- and high- income students at the more selective college without making any additional assumptions about social objectives. This happens for two reasons. The first is that since colleges cannot observe ability perfectly, they will limit the size of the ability-based tuition discounts they would otherwise offer. Thus, in order to match the sensitivity of tuition to SAT observed in the data, we require a relatively large value for ρ_L, the parameter which governs college's willingness to substitute average instructional spending for a higher average ability. This provides the colleges with a smaller motive to raise revenue and instead enroll higher ability students, many of which are lower-income and would otherwise be excluded. The second reason is due to selection effects arising from the application and admissions system. Since low-income students are less likely to apply to the more selective college, the ones who do apply are very high-ability in equilibrium. Thus the college can be confident that enrolling low-income students will help increase its average ability, and it will give large tuition discounts to the low-income students it enrolls. This selection effect due to the application choices of low-income students is explored further in the following section.

5.2 Effect of Applications on Low-Income Student Enrollment

This section illustrates the importance of the application and admissions process in accounting for low-income student enrollment. As discussed before, an important mechanism in the model that accounts for merit-based financial aid at the colleges is the informativeness of the students' signals. If, for example, only the highest ability among low-EFC students apply, then their signals will be highly informative about their ability and they will receive high levels of financial aid since the colleges will be confident that they are high ability.

To illustrate this mechanism, Figure 7 shows what happens to enrollment when all low-wealth students apply in the same way as their wealthier peers. Specifically, we adjust the application probabilities from the baseline equilibrium by requiring that low-EFC students apply to both colleges at the same rate as high-EFC students. We then recompute the equilibrium, holding fixed these new application pools (thus this is a partial-equilibrium analysis since student application behavior is taken as given). Note, that we still allow tuition, admissions policies, and student enrollment behavior to adjust. This exercise helps isolate the effect of student application portfolios on the overall allocation of students.

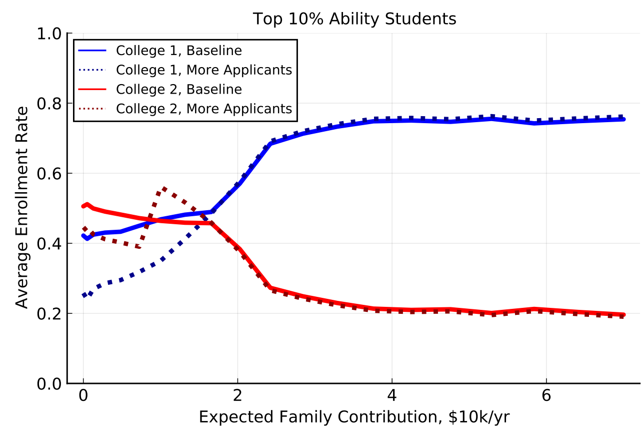


Figure 7: Average enrollment rates for students in the top 10% of the ability distribution by wealth. The plot shows the effect on enrollment of switching to an equilibrium where low-wealth students apply to both Colleges, even if low ability.

As we see in Figure 7, worsening the application pool of low-EFC students reduces the enrollment of high-ability, low-EFC students in College 1 by about a half. We see that with the substantial increase in low-ability, low-EFC applicants, the signals of low-wealth students are now much more likely to have come from low-ability students. This will cause the college to lower the financial aid it offers the low-income students, causing them to enroll at much lower rates. Note that since College 1 is now unable to enroll as many high-ability students, its value-added will also decline due to the lower average ability in its student body. Finally, we see that this effect is present not only for low-income students, but also for middle-income students who enroll in College 2 instead. A summary of the overall changes to the college market in the new equilibrium is provided in Table F.3.

Overall, this section highlights an important insight of the model: the composition of college application pools affects the informativeness of the signals and hence the colleges' tuition levels. This helps resolve the puzzle of low application rates from low-income students despite the presence of high financial aid. According to the model, low-income students receive high financial aid precisely *because* they are less likely to apply. Thus, policies aimed at increasing applications that are not targeted by ability will be harmful to low-income, high-ability students who would otherwise benefit from having their signal be more informative.

6 The Role of Admissions Uncertainty

In this section we study the equilibrium effects of changes to the signal variance. In our first counterfactual exercise we examine the effect of increasing the signal variance, motivated by concerns over the use of the SAT (assuming its use increases application informativeness). In our second exercise, we study the effect of switching to perfectly informative signals by removing admissions uncertainty for students, and allowing colleges to fully observe student ability.

6.1 Effect of Less Informative Signals

What would happen if the application signals were to become less informative? This question is motivated by the fact that many colleges waived SAT or ACT requirements during the Covid-19 pandemic. Moreover, the use of standardized tests for college admissions has recently come under scrutiny as the University of California system and other colleges have begun phasing out their reliance on the SAT/ACT for admissions. If we consider the use of the SAT or ACT as part of the technology that increases the informativeness of the signals, we can model the removal of these tests by making the signals less informative. To study the effects of less informative signals, we increase the variance of the signal by 85% from the baseline estimation and recompute the equilibrium. Note that since the model does not give us a way to quantify the effect of removing the SAT or ACT on the signal variance, this exercise should be interpreted qualitatively rather than quantitatively.

Table F.4 shows the effect of increasing the signal variance on college-level variables. In the new equilibrium the higher variance leads to poorer sorting, which causes the average student ability to drop in both colleges. Less informative signals also reduce the marginal cost of admitting lower signal students, since their signals are now more likely to have come from higher ability applicants. Thus we see increases in the fraction of students enrolled, and decreases in the admissions standards. While there is little change to the EFC distribution within both colleges, the ability distribution has become more diffuse, reflecting the fact that lower ability students can now more easily sort in to the colleges by getting sufficiently high signals by chance. Finally, the colleges endogenously increase their spending per-student, but it is not enough to offset the decrease in average ability, leading to lower value added at both colleges.

To see which students are affected by the decrease in the signal informativeness, Figure 8 plots the percent of consumption in each period different types of students are willing to forgo in order to be born in the new equilibrium. For reference, Figure F.2 plots the changes to student enrollment in the new equilibrium. As expected, the relatively high ability students are the ones who are most hurt from switching to the new equilibrium since the higher variance decreases their likelihood of being matched with the colleges at the expense of lower ability students. Importantly, this effect is strongest for the highest ability, high-EFC students who now compete with lower ability applicants, making the high-EFC signals less informative. This is less of a problem for the high-ability, low-EFC students who generally have a stronger applicant pool as described above. Indeed, Figure F.2 shows that their enrollment patterns are largely unaltered.

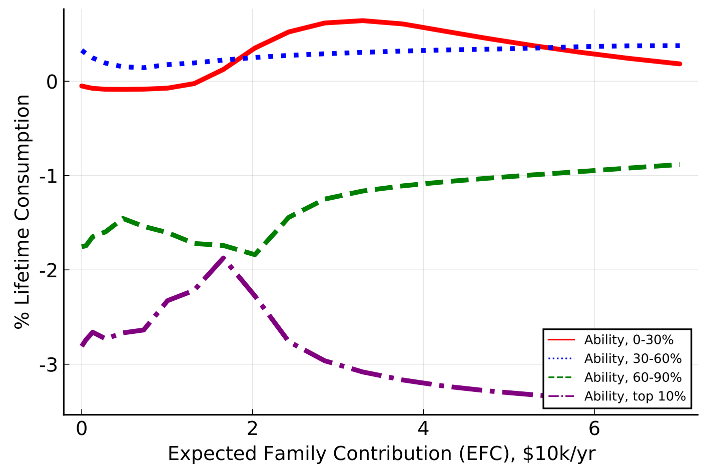


Figure 8: Percent of lifetime consumption students of different EFC and ability are willing to give up to switch to the equilibrium with less informative signals.

The students who mainly benefit from this change are the relatively high-EFC, low-ability students, who can now more easily match with the colleges. As can be seen from the solid line in Figure 8 and from Figure F.2, wealthy students at very bottom of the ability distribution benefit from a less informative signal by being able to match more easily with College 2. Similarly, the wealthiest students in the 30-60% ability range also benefit from being able to more easily sort into College 1, while the lower-EFC students in the 30-60% ability range benefit from more easily sorting into College 2. Overall, the average change in welfare among all students is about -0.51%, since the gains for the lower ability students are not enough to offset the losses of the higher ability students for whom a college education is the most beneficial.

6.2 Perfectly Informative Signals

To further study the effects of the information frictions due to the application and admissions system, we now set the signal to be perfectly informative so that the colleges may observe the student's true ability (i.e. $g(\sigma|\ell) = 1$ for $\sigma = \ell$). In this exercise, colleges set their tuition and admissions standards based on ability instead of the signal, and students konw ex-ante whether or not they will be admitted and exactly how much financial aid they will receive. We then compare the perfect information equlibrium to the baseline equilibrium.

The effect of the new equilibrium on each college is shown in Table F.5. If signals are perfectly informative, the marginal cost of enrolling a high ability student decreases substantially because colleges can perfectly tell them apart. This will lower tuition for high ability students and raise the average ability within the student body, thereby increasing the marginal cost for relatively lower ability students. As the marginal cost of enrolling a wealthy, low ability student rises, colleges will increase their admissions threshold to exclude these students since the tuition cap is not high enough to justify admitting them. The higher average ability makes it more costly to admit lower ability students, leading to a decrease in enrollment. Additionally, the decreases in marginal cost for the new higher ability enrollees causes them to have a lower tuition. This decreases tuition revenue, which lowers the average instructional spending per students. Despite this drop, however, the increase in average student ability at each college causes the peer-effect component of school quality to be high enough to increase the overall value-added at both colleges.

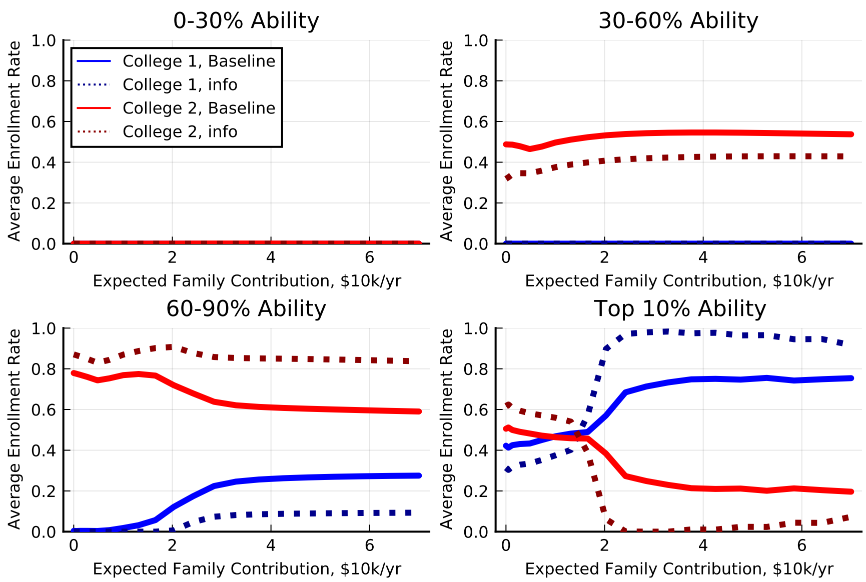


Figure 9: Average enrollment rates for students by EFC in baseline and perfect information equilibria.

The effect of perfectly observable signals on student enrollment at each college is shown in Figure 9 below. We see that sorting based on ability is vastly improved in the new equilibrium, with students at the bottom of the ability distribution substantially reducing their enrollment. There is a large drop in students from the 30-60% ability group enrolling into College 2, replaced by a large increase in students from the 60-90% range and low-EFC students from the top 10% group. College 1 is now able to enroll students almost exclusively from the top 10% of the ability distribution, mostly to the benefit of high-EFC students who no longer have to compete with lower-ability applicants.

Perhaps surprisingly, perfect information actually *reduces* the enrollment of high-ability, low-EFC students into College 1. This seems puzzling at first because tuition actually decreases at College 1 for high ability students since they can now be perfectly sorted. The reason for the decrease in College 1 enrollment is that College 2 now lowers its tuition enough to attract these high ability students so that their College 2 enrollment is even higher than it was in the baseline. College 2 could not offer such high discounts before because the strength of the applicant pool was much weaker for low-EFC students since it included many relatively low-ability students.

To understand the strength of college competition under perfect information, Figure 10 shows College 1 enrollment for top 10% ability students in a partial equilibrium world where uncertainty disappears for students, but tuition and admissions policies remain the same (the dotted line). In this scenario, the high-ability, low-EFC students enroll in College 1 at much higher numbers since they know they will be admitted at a low tuition level. However, when colleges adjust their policies in response to perfect information, the low-EFC students prefer to attend College 2 where tuition is even lower.

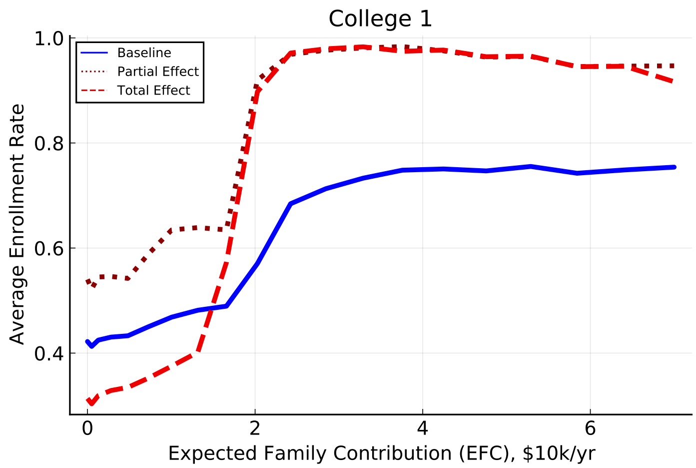


Figure 10: Average College 1 enrollment rate by EFC for students in the top 10% of the ability distribution. The 'Partial Effect' shows how enrollment changes when students have perfect information, but tuition and admissions are the same as in the baseline. The 'Total Effect' shows the perfect-information equilibrium where colleges adjust tuition and admissions.

7 Policy Experiments

7.1 Increasing Pell Grants.

Increasing the maximum level of Pell Grants was part of President Biden's proposal for higher education policy on the campaign trail. It is now being discussed more concretely as part of the recently released "American Families Plan", which calls for the maximum to be raised to $7,895 per year. In this section, we study the effects of more than doubling the maximum from 2013 to $15,000 per year (our model is calibrated to match values measured in 2013 dollars). As with the actual Pell Grant program, students in our model receieve Pell Grant aid if their EFC is lower than the set Pell Grant maximum. The amount they receieve is equal to the difference between the maximum and their EFC (see 28). Note that the increase in Pell Grants is paid for with tax increases, as implied by the government budget in Equation (24).

The effects on colleges of increasing the Pell Grant maximum can be seen in Table F.6, and the effects on student sorting can be seen in Figure 11. In Table F.6 we see that for College 1, there is little change in the distribution of students enrolled. The biggest change is that College 1 is able to raise additional tuition revenue from the low income students who were already enrolled, leading to higher instructional spending and higher value-added. College 2, which now increases enrollment, enrolls a larger fraction of students from the third EFC quartile since they are affected by the substantial increase in the grants. It is slightly able to increase its value-added with modest increases on spending per student and average ability.

Figure 11 illustrates the effect of the higher Pell Grants on student sorting, where enrollment is plotted against EFC for students at different parts of the ability distribution. As expected from Table F.6, we see very little change in enrollment behavior for students at College 1. For College 2, there is an increase in enrollment for students in the bottom three-quarters of the EFC distribution. This is noticeable in the 30-60% ability group, where middle-income students affected by the new Pell Grant maximum largely replace higher-income students. We also see a modest increase in enrollment for low-income students in the top two-thirds of the ability distribution. This is particularly important because these students were previously excluded from the college market, which is evident by noticing that enrollment rates at College 1 are unchanged.

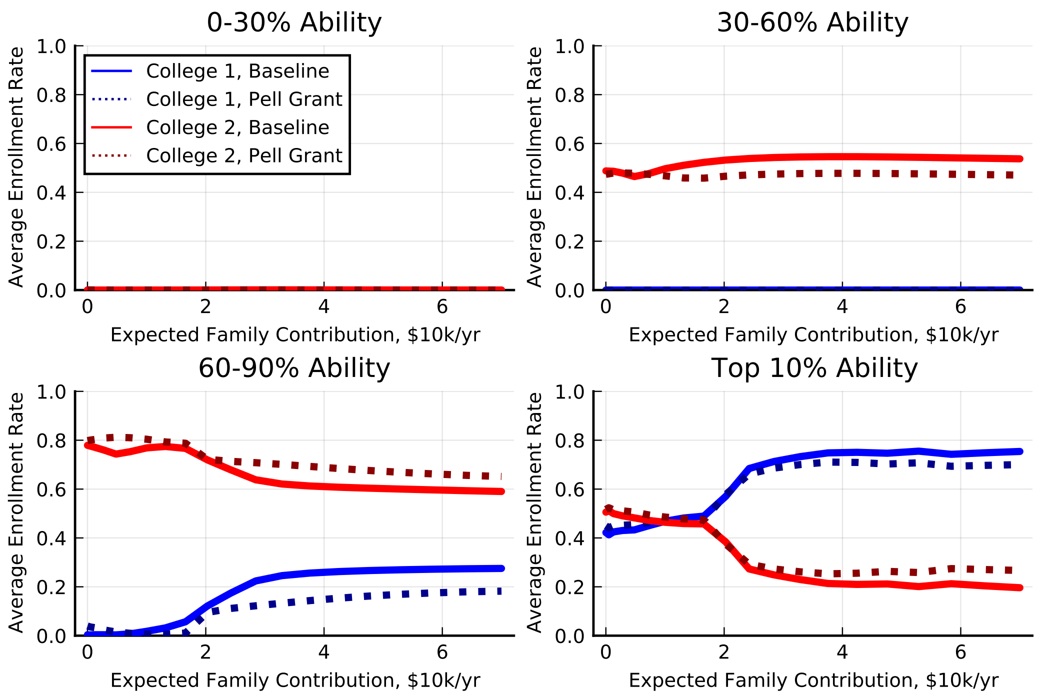


Figure 11: Student sorting by EFC and ability in baseline equilibrium and equilibrium where Pell Grants have doubled.

Finally, it is helpful to study the welfare effects of the policy change. Figure 12 plots the percent change in period consumption students of different EFC and ability group would be willing to give up in order to switch to the equilibrium with higher Pell Grants. As expected, we see that the students who benefit from the policy are the relatively low-income, college-going students. Higher-income students are made worse off from the policy due to the higher taxes, and due to being replaced in the college market by lower-income students who can now attend more easily. Since the lowest and highest ability group are infra-marginal, they will not be made worse off from the policy other than the tax increase, since their college outcome has not changed. Lower ability groups, however, will be worse off due to the tax increase and due to being replace in the college market by the relatively lower-income students.

For students in the 30-60% ability group, we see that the gains for the low-income students who benefit from the Pell Grants is roughly symmetric to the losses for the higher-income students. For the higher-ability low-income students however, we see very large welfare gains due to the fact that college is more valuable to them compared to the lower ability students they replace. We therefore see that the benefits to high-income, high-ability students are enough to compensate the losses which accrue to the high-income students, and the average welfare change of the policy in terms of period consumption is 1.98%.

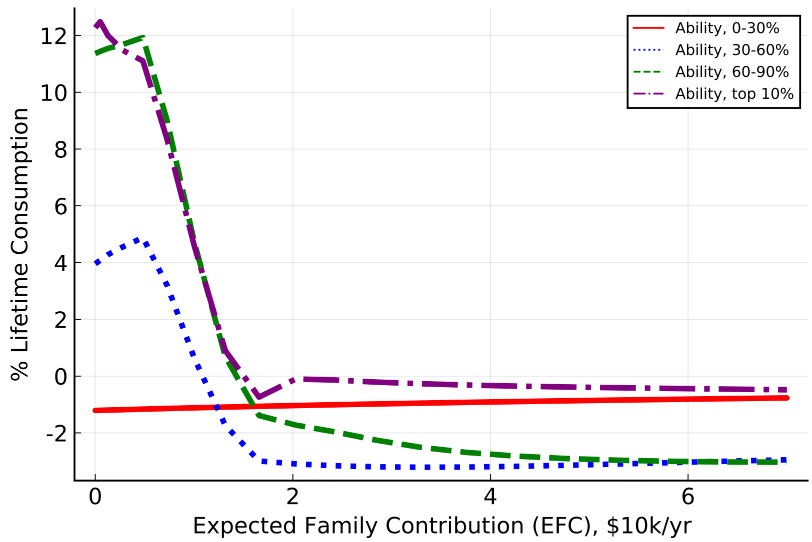


Figure 12: Percent of lifetime consumption students of different EFC and ability are willing to give up to switch to the equilibrium with the higher Pell Grant maximum.

8 Conclusion

This paper studies the role of the admissions system in shaping inequality in the college market. Using micro-level data on high school students transitioning to college, we showed that parental income is correlated with college applications and enrollment. We found that higher-income students are more likely to apply to college not only at the extensive margin, but also at the intensive margin by applying to more selective colleges. This confirms the findings of the empirical literature that has found substantial undermatching of students at the application stage. Why then, despite the presence of financial aid, do low-income students not apply as much their higher-income peers?

To answer this question, we built an equilibrium model of the college market with student heterogeneity and a non-trivial application and admissions system. The model is able to account for income differences in application and enrollment rates on the student side of the market, and high levels of tuition discrimination on the college side of the market. We find that lower-income students apply at lower rates because of expectations that they will not receive sufficient financial aid. Since a college education is complementary with ability, only the highest-ability low-income students apply, making their signals more informative to the colleges. Colleges are then confident that the low-income students are of high ability, justifying the high levels of financial aid they receive.

In focusing on the role of applications, we have abstracted from many important issues in the college market. By assuming there are only two colleges, we were unable to distinguish between public and private colleges. This is an important distinction since the funding for state schools depends on state governments, and these schools are able to offer substantially lower levels of tuition to in-state students. We have also abstracted from the source of student 'wealth', which in practice depends on parental transfers. Many factors determine a parent's willingness to invest in their children's education that we have excluded from our model. Finally, we have assumed that students are guaranteed to graduate and can perfectly forecast their post-college earnings. In reality students may face substantial drop-out and post-college earnings risk, which may be an important factor in determining a student's willingness to pursue a college education. We leave these considerations for future research.

References

B. Abbott, G. Gallipoli, C. Meghir, and G. L. Violante. Education policy and intergenerational transfers in equilibrium. *Journal of Political Economy*, 127(6):2569–2624, 2019.

M. J. Bailey and S. M. Dynarski. *Inequality in Postsecondary Education*, pages 117–132. Russell Sage Foundation, 2011.

R. Bénabou. Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica*, 70(2):481–517, 2002.

Z. Cai and J. Heathcote. College tuition and income inequality. Working paper, July 2018.

D. Capelle. The great gatsby goes to college: Tuition, inequality and intergenerational mobility in the us. *Working Paper*, 2019.

E. Caucutt and L. Lochner. Early and late human capital investments, borrowing constraints, and the family. *Working Paper*, 2017.

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

R. Chetty, J. N. Friedman, E. Saez, N. Turner, and D. Yagan. Income Segregation and Intergenerational Mobility Across Colleges in the United States*. *The Quarterly Journal of Economics*, 02 2020.

S. B. Dale and A. B. Krueger. Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables*. *The Quarterly Journal of Economics*, 117(4):1491–1527, 11 2002.

S. B. Dale and A. B. Krueger. Estimating the effects of college characteristics over the career using administrative earnings data. *Journal of human resources*, 49(2):323–358, 2014.

J. M. Delaney and P. J. Devereux. Choosing differently? college application behavior and the persistence of educational advantage. 2020.

E. Dillon and J. Smith. Determinants of the match between student ability and college quality. *Journal of Labor Economics*, 35:45–66, 2017.

S. Dynarski, C. Libassi, K. Michelmore, and S. Owen. Closing the gap: The effect of a targeted, tuition-free promise on college choices of high-achieving, low-income students. Working Paper 25349, National Bureau of Economic Research, 2018.

D. Epple, R. Romano, and H. Sieg. Admission, tuition, and financialaid policies in the market for higher education. *Econometrica*, 74:885–928, 2006.

D. Epple, R. Romano, S. Sarpca, and H. Sieg. A general equilibrium analysis of state and private colleges and access to higher education in the u.s. *Journal of Public Economics*, 155:164–178, 2017.

R. Fernandez and R. Rogerson. Income distribution, communities and the quality of public education. *The Quaterly Journal of Economics*, 111(1), 1996.

I. Fillmore. Price discrimination and public policy in the u.s. college market. *Working Paper*, 2020.

C. Fu. Equilibrium tuition, applications, admissions and enrollment in the college market. *Journal of Political Economy*, 122(2):225–281, 2014.

G. Gordon and A. Hedlund. Accounting for the rise in college tuition. Working Paper 21967, National Bureau of Economic Research, February 2016.

C. M. Hoxby and C. Avery. The missing "one-offs": The hidden supply of high-achieving, low income students. *Brookings Papers on Economic Activity*, 2014.

F. Ionescu and N. Simpson. Default risk and private student loans: Implications for higher education policies. *Journal of Economic Dynamics and Control*, 64(1):119–147, 2016.

T. Kotera and A. Seshadri. Educational policy and intergenerational policy. *Review of Economic Dynamics*, 25:187–207, 2017.

D. Krueger and A. Ludwig. On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium. *Journal of Monetary Economics*, 77(C):72–98, 2016.

L. J. Lochner and A. Monge-Naranjo. The nature of credit constraints and human capital. *American Economic Review*, 101(6):2487–2529, October 2011.

D. O. Lucca, T. Nadauld, and K. Shen. Credit Supply and the Rise in College Tuition: Evidence from the Expansion in Federal Student Aid Programs. *The Review of Financial Studies*, 32(2):423–466, 2018.

Appendix

A Additional Empirical Results

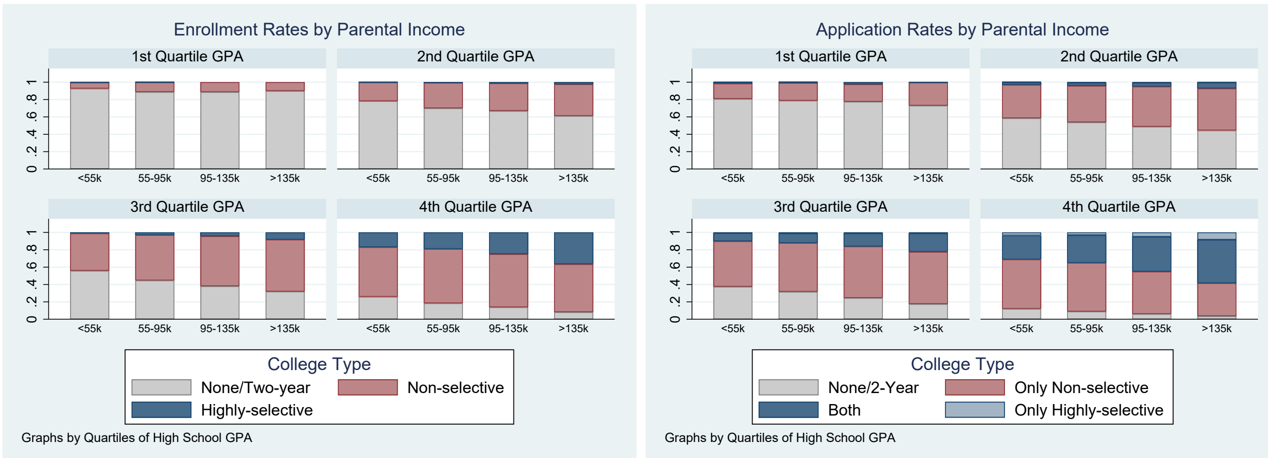


Figure A.1: College application and enrollment by parental income and student GPA. Left panel: Student enrollment. Right panel: Student applications. Source: HSLS.

Dep Var: Apply at all	**(1)**	**(2)**	**(3)**	**(4)**
Parental Income	0.096***	0.094***	0.049	0.070
	(0.012)	(0.014)	(0.039)	(0.050)
GPA (std)	1.182***		1.300***	
	(0.075)		(0.081)	
SAT (std)		0.721***		0.906***
		(0.101)		(0.107)
Parental Inc.*GPA/SAT	0.067***	0.071***	0.053***	0.057***
	(0.016)	(0.019)	(0.017)	(0.020)
Controls	No	No	Yes	Yes
N	15,420	11,820	15,420	11,820
R^2	0.260	0.148	0.283	0.176

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Controls include race, sex, parental education, parental employment status, parental relationship pattern, household size, and high school type.

Table A.1: Logit estimation results for whether or not students applied to any non-profit 4-year colleges. Data: HSLS.

Dep Var: Apply to selective|applied	(1)	(2)	(3)	(4)
Parental Income	0.028	0.001	-0.024	-0.050
	(0.027)	(0.019)	(0.057)	(0.054)
GPA (std)	0.922***		1.070***	
	(0.166)		(0.140)	
SAT (std)		0.955***		1.129***
		(0.123)		(0.121)
Parental Inc.*GPA/SAT	0.081***	0.073***	0.069***	0.053***
	(0.024)	(0.018)	(0.020)	(0.018)
Controls	No	No	Yes	Yes
N	10,000	9,330	10,000	9,330
R^2	0.166	0.206	0.200	0.232

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Controls include race, sex, parental education, parental employment status, parental relationship pattern, household size, and high school type.

Table A.2: Logit estimation results for whether or not students applied to any Highly-selective colleges, conditional on applying to non-profit 4-year colleges. Data: HSLS.

B Data Appendix

B.1 Barron's Selectivity Index

List of Barron's Tier 1 and 2 Colleges and Universities (Alphabetical)

Tier 1: Amherst College, Barnard College, Bates College, Boston College, Bowdoin College, Brown University, Bryn Mawr College, Bucknell University, California Institute of Technology, Carleton College, Carnegie Mellon University, Case Western Reserve University, Claremont McKenna College, Colby College, Colgate University, College of Mount Saint Vincent, College of the Holy Cross, College of William & Mary, Colorado College, Columbia University/City of New York, Connecticut College, Cooper Union for the Advancement of Science and Art, Cornell University, Dartmouth College, Davidson College, Duke University, Emory University, Franklin and Marshall College, George Washington University, Georgetown University, Georgia Institute of Technology, Hamilton College, Hampshire College, Harvard University/Harvard College, Harvey Mudd College, Haverford College, Johns Hopkins University, Kenyon College, Lehigh University, Macalester College, Massachusetts Institute of Technology, Middlebury College, New York University, Northeastern University, Northwestern University, Oberlin College, Ohio State University at Marion, Pitzer College, Pomona College, Princeton University, Reed College, Rensselaer Polytechnic Institute, Rice University, Rose-Hulman Institute of Technology, Santa Clara University, Smith College, Southern Methodist University, Stanford University, Swarthmore College, The Ohio State University, Tufts University, Tulane University, Union College, United States Air Force Academy, United States Military Academy, United States Naval Academy, University of California at Berkeley, University of California at Los Angeles, University of Chicago, University of Miami, University of Missouri/Columbia, University of North Carolina at Chapel Hill University of Notre Dame, University of Pennsylvania, University of Richmond, University of Rochester, University of Southern California, University of Virginia, Vanderbilt University, Vassar College, Villanova University, Wake Forest University, Washington and Lee University, Washington University in St. Louis, Webb Institute, Wellesley College, Wesleyan University, Whitman College, Williams College, Yale University

Tier 2: Allegheny College, American University, Augustana College, Austin College, Babson College, Bard College, Bard College at Simon's Rock, Baylor University, Beloit College, Bennington College, Bentley University, Berea College, Berry College, Binghamton University/The State University of New York, Boston University, Brandeis University, Brigham Young University, California Polytechnic State University, Centre College, Christian Brothers University, Clark University, Clarkson University, Clemson University, College of New Jersey, College of the Atlantic, Colorado School of Mines, Cornell College, CUNY/City College, Denison University, Dickinson College, Drexel University, Elon University, Emerson College, Florida State University, Fordham University, Furman University, Gettysburg College, Gonzaga University, Grinnell College, Grove City College, Gustavus Adolphus College, Hendrix College, Hillsdale College, Illinois Institute of Technology, Indiana University Bloomington, Ithaca College, Kalamazoo College, Kettering University, Lafayette College, Lawrence University, Miami University, Mills College, Mount Holyoke College, Muhlenberg College, New College of Florida, New Mexico Institute of Mining and Technology, North Carolina State University, Pepperdine University, Polytechnic Institute of New York University, Providence College, Purdue University/West Lafayette, Rhodes College, Rollins College, Sarah Lawrence College, Sewanee: The University of the South, Skidmore College, St. John's College, Santa Fe, St. John's College-Annapolis, St. Lawrence University, St. Mary's College of Maryland, St. Olaf College, State University of New York / College of Environmental Science and Forestry, Stevens Institute of Technology, Stony Brook University / State University of New York, SUNY College at Geneseo, Syracuse University, Texas Christian University, Trinity College, Trinity University, Truman State University, United States Coast Guard Academy, United States Merchant Marine Academy, University of California at Davis, University of California at Santa Barbara, University of Connecticut, University of Florida, University of Illinois at Urbana-Champaign, University of Maryland, University of Michigan/Ann Arbor, University of Minnesota/Twin Cities, University of Pittsburgh at Pittsburgh, University of Puget Sound, University of San Diego, University of Texas at Austin, University of Texas at Dallas, University of Tulsa, University of Wisconsin/Madison, Virginia Polytechnic Institute and State University, Westmont College, Wheaton College, Wheaton College, Worcester Polytechnic Institute

B.2 IPEDS

The Integrated Postsecondary Education Data System is made publicly available through the National Center for Education Statistics. We take IPEDS data from 2013-2016 (the relevant 4 year period given that our HSLS cohort begins college in 2013), and make the following restrictions to the universe of colleges in the system:

- U.S. only; Title IV participating; Degree-granting
- Undergraduate enrollment at least 100
- No Theological/faith related institutions
- No 2 year colleges

- No for-profit colleges

C Calibration Appendix

C.1 EFC Calculation

For students who did not fill out the FAFSA, we calculate their EFC directly using the 2013-2014 EFC formula with data from the HSLS survey. To calculate EFC, one must first calculate Adjusted Available Income (AAI), which combines household income net of allowances (which depend on household size) and household assets (excluding the family's home). Since the HSLS does not report assets, we assume that the contribution from assets is 0.

The parents' contribution from AAI is then calculated from a (progressive) non-linear function of AAI, described in the table below:

If parents' AAI is –	**Parents' contribution from AAI is –**
Less than -\$3,409	-\$750
-\$3,409 to \$15,300	22% of AAI
\$15,301 to \$19,200	25% of AAI over \$15,300 + \$3,366
\$19,201 to \$23,100	29% of AAI over \$19,200 + \$4,341
\$23,101 to \$27,000	34% of AAI over \$23,100 + \$5,472
\$27,001 to \$30,900	40% of AAI over \$27,000 + \$6,798
\$30,901 or more	47% of AAI over \$30,900 + \$8,358

EFC is then the parents' contribution divided by the number of children that are enrolled in college. The HSLS specifies only if students have a sibling in college at the same time, so to find EFC we divide the parents' contribution by two if the student does have a sibling in college. We are thus able to construct EFC for the students in our sample even if they did not complete the FAFSA.

C.2 Cost Function Estimation

D Proofs of Propositions

D.1 Derivation of (21)

Let $\lambda_I, \lambda_L, \lambda_\kappa$ be the Lagrange multipliers on the budget constraint (19), average ability identity constraint (17), and enrollment identitiy constraint (16) respectively. First order conditions for $I_\mu, L_\mu, \kappa, T(a, \sigma)$ are then

$$
\begin{aligned}
&[\kappa] && -\lambda_\kappa = \lambda_L L + \lambda_I[I + C'(\kappa) - Tr'(\kappa)] \\
&[I_\mu] && \xi_s Q_I = \lambda_I \kappa \\
&[L_\mu] && \xi_s Q_L = \lambda_L \kappa \\
&[T(a,\sigma)] && T(a,\sigma) = -\frac{\lambda_\kappa}{\lambda_I} - \frac{\lambda_L}{\lambda_I}\frac{\int \ell \frac{\partial \tilde{q}}{\partial T(a,\sigma)} g(\sigma|\ell)\mu(a,d\ell)}{\int \frac{\partial \tilde{q}}{\partial T(a,\sigma)} g(\sigma|\ell)\mu(a,d\ell)} \\
& && \qquad - \frac{\int \tilde{q}\; g(\sigma|\ell)\mu(a,d\ell)}{\int \frac{\partial \tilde{q}}{\partial T(a,\sigma)} g(\sigma|\ell)\mu(a,d\ell)} + \frac{\lambda_T(a,\sigma)}{\lambda_I \int \frac{\partial \tilde{q}}{\partial T(a,\sigma)} g(\sigma|\ell)\mu(a,d\ell)}
\end{aligned} \tag{32}
$$

where $\lambda_T(a,\sigma)$ is the Lagrange multiplier for the constraint on tuition. Combining these equations gives (21).

D.2 Derivation of (23)

Let $\tilde{q}(\underline{\sigma}_s) = \tilde{q}(a,\ell,\underline{\sigma}_s, T^s(a,\underline{\sigma}_s), T^{-s}(a,\underline{\sigma}_s))$. The first order condition the admissions standard is

$$
\frac{\int T(a,\underline{\sigma}_s)\tilde{q}(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)d\mu(a,\ell)}{\int \tilde{q}(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)d\mu(a,\ell)} \geq -\frac{\lambda_\kappa}{\lambda_I} - \frac{\lambda_L}{\lambda_I}\frac{\int \ell\; \tilde{q}(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)d\mu(a,\ell)}{\int \tilde{q}(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)d\mu(a,\ell)}, \tag{33}
$$

which holds with equality when $\underline{\sigma}_s > 0$. Combining with (32), we arrive at (23).

D.3 Proposition 1

Proof. Combining (23) and (21), we find

$$
\begin{aligned}
&\frac{Q_{L_\mu}}{Q_{I_\mu}} \int \left(\ell - \frac{\int \ell \frac{\partial \tilde{q}^s(\underline{\sigma}_s)}{\partial T^s(a,\underline{\sigma}_s)} g(\underline{\sigma}_s|\ell)\mu(a,d\ell)}{\int \frac{\partial \tilde{q}^s(\underline{\sigma}_s)}{\partial T^s(a,\underline{\sigma}_s)} g(\underline{\sigma}_s|\ell)\mu(a,d\ell)} \right) \tilde{q}^s(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)d\mu(a,\ell) \\
&- \int \frac{\left[\int \tilde{q}^s(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)\mu(a,d\ell)\right]^2}{\int \frac{\partial \tilde{q}^s(\underline{\sigma}_s)}{\partial T^s(a,\underline{\sigma}_s)} g(\underline{\sigma}_s|\ell)\mu(a,d\ell)} da + \frac{\int \lambda_T(a,\underline{\sigma}_s)\tilde{q}^s(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)d\mu(a,\ell)}{\lambda_I \int \frac{\partial \tilde{q}^s(\underline{\sigma}_s)}{\partial T^s(a,\underline{\sigma}_s)} g(\underline{\sigma}_s|\ell)\mu(a,d\ell)} \geq 0.
\end{aligned} \tag{34}
$$

Since the constraint on tuition is never binding, we have $\lambda_T(a,\sigma) = 0$. Rewrite the left hand side of (34) to obtain

$$
\begin{aligned}
&\frac{Q_{L_\mu}}{Q_{I_\mu}} \int \left[\int q^s(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)\mu(a,d\ell) \right] \underbrace{\left[\frac{\int \ell q^s(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)\mu(a,d\ell)}{\int q^s(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)\mu(a,d\ell)} - \frac{\int \ell \frac{\partial q^s(\underline{\sigma}_s)}{\partial T^s} g(\underline{\sigma}_s|\ell)\mu(a,d\ell)}{\int \frac{\partial q^s(\underline{\sigma}_s)}{\partial T^s} g(\underline{\sigma}_s|\ell)\mu(a,d\ell)} \right]}_{\geq 0} da \\
&+ \underbrace{\int \frac{\left[\int q^s(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)\mu(a,d\ell;y)\right]^2}{\int -\frac{\partial q^s(\underline{\sigma}_s)}{\partial T^s(a,\underline{\sigma}_s)} g(\underline{\sigma}_s|\ell)\mu(a,d\ell;y)} da}_{>0},
\end{aligned} \tag{35}
$$

where $q^s(\underline{\sigma}_s) = q^s(a, \ell, T(a, \underline{\sigma}_s))$. We show that (35) is strictly positive for all $\underline{\sigma_s} > \underline{\sigma}$, so that the college does not choose an interior admissions standard.

Clearly, the term on the bottom is strictly positive provided there is a positive mass of students for which $q^s(\underline{\sigma}_s) > 0$ and $\frac{\partial q^s(\underline{\sigma}_s)}{\partial T^s} < 0$. To see that the bracketed term on the top is positive (for each a), note that it is the difference in means of ℓ according to two distributions, H_0 and H_1, with associated densities: $h_0(\ell) = \frac{q^s(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)\mu(a,\ell)}{\int q^s(\underline{\sigma}_s)g(\underline{\sigma}_s|\ell)\mu(a,d\ell)}$ and $h_1(\ell) = \frac{\frac{\partial q^s(\underline{\sigma}_s)}{\partial T^s}g(\underline{\sigma}_s|\ell)\mu(a,\ell)}{\int \frac{\partial q^s(\underline{\sigma}_s)}{\partial T^s}g(\underline{\sigma}_s|\ell)\mu(a,d\ell)}$. The term being positive is then implied by the fact that H_0 has first-order stochastic dominance over H_1.

To see that this is the case, define

$$\begin{aligned} F(\hat{\ell}) &= H_0(\hat{\ell}) - H_1(\hat{\ell}) \\ &= \int_0^{\hat{\ell}} h_0(\ell)d\ell - \int_0^{\hat{\ell}} h_1(\ell)d\ell \end{aligned}$$

Note that F is continuously differentiable, and $\lim_{\ell \to 0} F(\ell) = 0$, $\lim_{\ell \to \infty} F(\ell) = 0$. The interior solution to the first order condition $F'(\ell^*) = 0$ uniquely minimizes F:

For College 1,

$$\begin{aligned} F'(\ell) &= h_0(\ell) - h_1(\ell) \\ &= \left\{ q^1(\underline{\sigma}_1) - \left[\frac{\int q^1(\underline{\sigma}_1)g(\underline{\sigma}_1|\ell)\mu(a,d\ell)}{\int \frac{\partial q^1(\underline{\sigma}_1)}{\partial T^1}g(\underline{\sigma}_1|\ell)\mu(a,d\ell)} \right] \frac{\partial q^1(\underline{\sigma}_1)}{\partial T^1} \right\} g(\underline{\sigma}_1|\ell)\mu(a,\ell) \\ &= \Bigg\{ \underbrace{\frac{1}{1-q^1(\underline{\sigma}_1)} - \left[\frac{\int q^1(\underline{\sigma}_1)g(\underline{\sigma}_1|\ell)\mu(a,d\ell)}{\int \frac{\partial q^1(\underline{\sigma}_1)}{\partial T^1}g(\underline{\sigma}_1|\ell)\mu(a,d\ell)} \right] \lambda_C V_T^{C1}(a, \ell^*, T(a, \underline{\sigma}_1))}_{G(\ell)} \Bigg\} q^1(\underline{\sigma}_1)g(\underline{\sigma}_1|\ell)\mu(a,\ell)[1-q^1(\underline{\sigma}_1)] \end{aligned}$$

We have a unique interior minimum since,

1. $\lim_{\ell \to 0} G(\ell) < 0$
2. $\lim_{\ell \to \infty} G(\ell) = \infty$
3. $G'(\ell) > 0$

Thus we have $F(\ell) < 0$ for $\ell \in (0, \infty)$.

Since (35) is strictly positive, the lower bound constraint on $\underline{\sigma}_s$ binds, so the college does not use the admissions threshold to screen students. □

E Lifecycle Model

In this section, we describe how a simple lifecycle model used for the calibration maps easily into the two period model introduced in Section 3. Consider an individual who lives for $T + 1$ periods, where a period is four years and the first years are spent in college:

$$
\begin{aligned}
\max_{\{c_j, a_{j+1}\}_{j=0,\dots,T}} \quad & u(c_0) + \sum_{j=1}^{T} \beta^j u(c_j) && (36)\\
\text{s.t.} \quad & c_0 + a_1 + T = a_0 \\
& c_j + a_{j+1} = a_j R + w(1-\tau)\gamma_s \ell^{\alpha}, \quad j = 1, \dots, T \\
& a_1 \geq \underline{a}_s && (37)
\end{aligned}
$$

CRRA preferences imply that from the Euler equation we can write $u(c_{1+t}) = (\beta R)^{\frac{t(1-\sigma)}{\sigma}} u(c_1)$, which means that

$$
\tilde{\beta} u(c_1) = \sum_{j=1}^{T} \beta^j u(c_j).
$$

We also combine the budget constraints for $j = 1, \dots, T$, and plug in the result from the Euler equation to get

$$
c_1 \sum_{j=1}^{T} \frac{(\beta R)^{\frac{j-1}{\sigma_c}}}{R^{j-1}} = a_1 R + w(1-\tau)\gamma_s \ell^{\alpha} \sum_{j=1}^{T} \frac{1}{R^{j-1}}.
$$

Putting everything together, we have the same two period model from Section 3:

$$
\begin{aligned}
\max_{c_0, c_1} \quad & u(c_0) + \tilde{\beta} u(c_1) && (38)\\
\text{s.t.} \quad & c_0 + a_1 + T = a_0 \\
& c_1 = a_1 \tilde{R} + \tilde{w}(1-\tau)\gamma_s \ell^{\alpha} \\
& a_1 \geq \underline{a}_s && (39)
\end{aligned}
$$

where

$$
\tilde{\beta} = \sum_{j=1}^{T} \beta^j (\beta R)^{\frac{(j-1)(1-\sigma_c)}{\sigma_c}}
$$

$$
\tilde{R} = \frac{R}{\sum_{j=1}^{T} \frac{(\beta R)^{\frac{j-1}{\sigma_c}}}{R^{j-1}}}
$$

$$
\tilde{w} = \frac{w \sum_{j=1}^{T} \frac{1}{R^{j-1}}}{\sum_{j=1}^{T} \frac{(\beta R)^{\frac{j-1}{\sigma_c}}}{R^{j-1}}}
$$

F Additional Model Results

F.1 Effect of Equalizing Application Patterns

		College 1		College 2	
		Baseline	**More Applicants**	**Baseline**	**More Applicants**
L_μ	Avg Student Ability	1.48	1.46	1.01	0.98
I_μ	Instr Spending per Student	2.29	2.27	1.06	1.06
Γ_s	Value-added	2.01	1.98	1.79	1.75
$\underline{\sigma}$	Admissions Threshold	1.25	1.25	0.82	0.75
κ	% Enrolled	7.23	6.81	33.68	33.54
τ	Tax rate (%)	1.83	1.82	1.83	1.82
	% Students Q1 EFC	0.09	0.07	0.22	0.22
	% Students Q2 EFC	0.12	0.09	0.21	0.21
	% Students Q3 EFC	0.26	0.23	0.31	0.31
	% Students Q4 EFC	0.53	0.61	0.25	0.26
	% Students Q1 Ability	0.0	0.0	0.0	0.0
	% Students Q2 Ability	0.0	0.0	0.0	0.0
	% Students Q3 Ability	0.0	0.0	0.47	0.54
	% Students Q4 Ability	1.0	1.0	0.53	0.46

Table F.3: Effect on college market of fixing application choices for all students to be the same as the application choices of high-wealth students in the baseline estimation

F.2 Effect of Less Informative Signals

		College 1		College 2	
		Baseline	**Higher σ Variance**	**Baseline**	**Higher σ Variance**
L_μ	Avg Student Ability	1.48	1.44	1.01	0.99
I_μ	Instr Spending per Student	2.29	2.35	1.06	1.09
Γ_s	Value-added	2.01	1.97	1.79	1.76
$\underline{\sigma}$	Admissions Threshold	1.25	1.24	0.82	0.77
κ	% Enrolled	7.23	7.61	33.68	34.8
τ	Tax rate (%)	1.83	1.92	1.83	1.92
	% Students Q1 EFC	0.09	0.09	0.22	0.23
	% Students Q2 EFC	0.12	0.12	0.21	0.21
	% Students Q3 EFC	0.26	0.26	0.31	0.3
	% Students Q4 EFC	0.53	0.52	0.25	0.25
	% Students Q1 Ability	0.0	0.0	0.0	0.0
	% Students Q2 Ability	0.0	0.0	0.0	0.04
	% Students Q3 Ability	0.0	0.01	0.47	0.47
	% Students Q4 Ability	1.0	0.99	0.53	0.49

Table F.4: Effect on colleges of making signals less informative by increasing the variance of the signal distribution

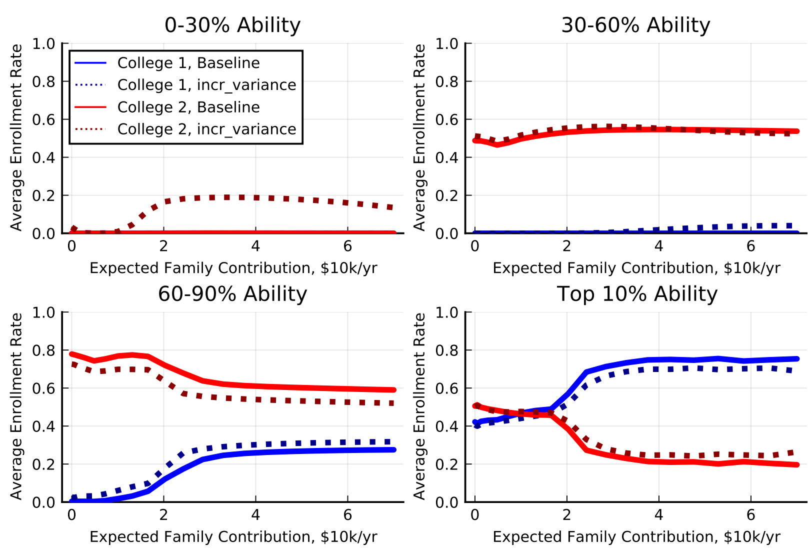


Figure F.2: Student enrollment rates in baseline equilibrium vs. equilibrium with less informative signals.

F.3 Effect of Perfectly Informative Signals

		College 1		College 2	
		Baseline	**Perfect Info**	**Baseline**	**Perfect Info**
L_μ	Avg Student Ability	1.48	1.53	1.01	1.06
I_μ	Instr Spending per Student	2.29	2.22	1.06	1.01
Γ_s	Value-added	2.01	2.05	1.79	1.85
$\underline{\sigma}$	Admissions Threshold	1.25	1.27	0.82	0.9
κ	% Enrolled	7.23	6.66	33.68	31.36
τ	Tax rate (%)	1.83	1.66	1.83	1.66
	% Students Q1 EFC	0.09	0.07	0.22	0.21
	% Students Q2 EFC	0.12	0.09	0.21	0.21
	% Students Q3 EFC	0.26	0.21	0.31	0.33
	% Students Q4 EFC	0.53	0.62	0.25	0.24
	% Students Q1 Ability	0.0	0.0	0.0	0.0
	% Students Q2 Ability	0.0	0.0	0.0	0.0
	% Students Q3 Ability	0.0	0.0	0.47	0.36
	% Students Q4 Ability	1.0	1.0	0.53	0.64

Table F.5: Effect on colleges of making signals perfectly informative

F.4 Policy Experiments: Increasing Pell Grants

		College 1		College 2	
		Baseline	**Pell Grant**	**Baseline**	**Pell Grant**
L_μ	Avg Student Ability	1.48	1.5	1.01	1.02
I_μ	Instr Spending per Student	2.29	2.56	1.06	1.09
Γ_s	Value-added	2.01	2.07	1.79	1.81
$\underline{\sigma}$	Admissions Threshold	1.25	1.31	0.82	0.87
κ	% Enrolled	7.23	7.02	33.68	33.97
τ	Tax rate (%)	1.83	3.02	1.83	3.02
	% Students Q1 EFC	0.09	0.12	0.22	0.22
	% Students Q2 EFC	0.12	0.14	0.21	0.22
	% Students Q3 EFC	0.26	0.27	0.31	0.32
	% Students Q4 EFC	0.53	0.48	0.25	0.24
	% Students Q1 Ability	0.0	0.0	0.0	0.0
	% Students Q2 Ability	0.0	0.0	0.0	0.0
	% Students Q3 Ability	0.0	0.0	0.47	0.44
	% Students Q4 Ability	1.0	1.0	0.53	0.56

Table F.6: Effect on colleges of increasing the Pell Grant maximum.

G Computational Appendix

This section describes the general procedure of how the equilibrium of the model is solved numerically. We start with a guess for College policies $\{\kappa_s, I_s, L_s, \Gamma_s, T_s(a,\sigma), \underline{\sigma_s} = 0\}_{s\in\{1,2\}}$ student application probabilities $\{p_i(a,\ell)\}_{i\in\{AB,A1,A2\}}$, and tax rate τ. We then update all choices so that they are consistent with optimal agent behavior, and iterate until our guess is consistent with itself. Our procedure is detailed in the following steps:

1. Using guesses $\{p_i(a,\ell)\}_{i\in\{AB,A1,A2\}}$, τ, Γ_s, solve for the total enrollment probability functions at each College, $\tilde{q}^s(a,\ell,T^1,T^2)$, $s\in\{1,2\}$. We find these functions using (15) and (9).
2. Given guess $\{\kappa_s, I_s, L_s, T_s(a,\sigma)\}$, $\tilde{q}^s(a,\ell,T^1,T^2)$, for $s\in\{1,2\}$, update the tuition and admissions standards for each college (we iterate until we find a fixed-point in the college policies).
   (a) For each College s, use guess for own aggregates $\{\kappa_s, I_s, L_s\}$, and guess for other Colleges tuition policy, $T_{-s}(a,\sigma)$ to find updated tuition policy $\hat{T}_s(a,\sigma)$ using (21). To handle the tuition cap, set $\hat{T}_s(a,\sigma) = \bar{T}_s$ if $\hat{T}_s(a,\sigma) > \bar{T}_s$.
   (b) For each College s, use updated tuition policy $\hat{T}_s(a,\sigma)$ to solve for the updated admissions standard $\underline{\hat{\sigma}}_s$ in (23)
   (c) Check for convergence
      - If $\sup|T_s(a,\sigma) - \hat{T}_s(a,\sigma)| > 10^{-5}$, set $T_s(a,\sigma) = \hat{T}_s(a,\sigma)$ and repeat (a)

- Otherwise, continue

3. Using updated tuition and admissions, find updated College aggregates $\{\hat{\kappa}_s, \hat{I}_s, \hat{L}_s, \hat{\Gamma}_s\}$ using (16), (17), (19), (14)

4. Using updated guesses, find updated tax rate $\hat{\tau}$ using (24)

5. Using all updated values and functions, find updated application decisions $\{\hat{p}_i(a, \ell)\}_{i \in \{AB, A1, A2\}}$ as in (26)

6. Check for convergence. Let $\hat{X} \equiv \left(\{\hat{\kappa}_s, \hat{I}_s, \hat{L}_s, \hat{\Gamma}_s\}_{s \in \{1,2\}}, \hat{\tau} \right)$, and $X \equiv \left(\{\kappa_s, I_s, L_s, \Gamma_s\}_{s \in \{1,2\}}, \tau \right)$

   - If $\sup |X - \hat{X}| > 10^{-5}$, set $X = \hat{X}$, $\{p_i(a, \ell)\}_{i \in \{AB, A1, A2\}} = \{\hat{p}_i(a, \ell)\}_{i \in \{AB, A1, A2\}}$, and repeat (1)
   - Otherwise, exit