

# RISKY COLLEGE ADMISSIONS AND THE (MIS)ALLOCATION OF TALENT

RICARDO MARTO  
Federal Reserve Bank of St. Louis

YAACOV WITTMAN  
Cornerstone Research

Why are high-achieving, low-income students less likely to apply to selective colleges despite the generous financial aid typically offered? To reconcile this seeming puzzle, an equilibrium model of the U.S. college market featuring tuition discrimination and a noisy application and admissions system is presented and estimated. Students, who differ in their financial resources and innate ability, can apply to multiple colleges and are uncertain about their prospective admissions and financial aid. Colleges in turn only observe a signal of students' ability and compete by choosing admission standards and tuition schedules. Low-income students receive generous 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), high-ability students would be worse off and only high-income, low-ability students would benefit. Finally, there are welfare gains from increasing need-based financial aid (e.g. Pell Grants). Despite its fiscal cost, the policy would greatly benefit low-income, high-ability students.

**KEYWORDS:** college admissions, college enrollment, college market, credit constraints, financial aid, High School Longitudinal Study, information asymmetry, Pell grants, sequential search, sorting, tuition discrimination.

---

Ricardo Marto: [ricardo.marto@stls.frb.org](mailto:ricardo.marto@stls.frb.org)

Yaacov Wittman: [ywittman@sas.upenn.edu](mailto:ywittman@sas.upenn.edu)

We thank Damien Capelle, Jesús Fernández-Villaverde, Jeremy Greenwood, Dirk Krueger, Víctor Ríos-Rull, and Holger Sieg for their insightful comments and suggestions. We are also grateful to seminar participants at the University of Pennsylvania, the Washington University in St. Louis Graduate Student Conference, and the University of Warwick PhD Conference. The views expressed here are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

## 1. INTRODUCTION

Despite the benefits of completing college, 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 far less likely to be represented at more selective colleges (Chetty, Friedman, Saez, Turner, and Yagan (2020)). Although this can partly be explained by differences in levels of preparedness, a prominent reason relates to differences in application rates between low- and high-income students. Low-income students tend to apply less to schools that they appear overqualified for relative to their higher-income peers (Hoxby and Avery (2014) and Dillon and Smith (2017)). That is, low-income students are underrepresented in selective colleges not because they are excluded from schools but because they do not apply in the first place—despite the substantial need-based financial aid offered by selective colleges.

Why then, even after controlling for test scores, do we observe different application patterns for students across the income distribution? This paper addresses this question by examining the hypothesis that college admissions and financial aid policies effectively limit the enrollment of low-income students, making it rational for them to apply at lower rates. How? At the application stage, students do not expect to receive sufficient financial aid for them to attend and therefore refrain from applying altogether. To study this hypothesis, we build and estimate a novel model of the college market featuring an application and admissions system similar to the one used by U.S. colleges. While colleges value high-ability students, it is costly for them to offer generous financial aid to their low-income applicants. Hence, they face a trade-off between admitting high-ability, low-income students and admitting lower-ability, high-income students who are willing to pay full tuition. This trade-off will cause colleges to offer generous aid only to the highest-ability students among their low-income applicant pool. In turn, low-income students will find it optimal to apply at lower rates because they recognize that they have a low chance of receiving sufficient financial aid.

This paper proceeds in two parts. In the first part, we provide evidence on how college application and enrollment patterns vary by parental income and test scores. We rely on detailed student-level data from the High School Longitudinal Study of 2009 (HSLs), a representative survey of students followed throughout their secondary and postsecondary education. This paper shows that low-income students are less likely to apply to highly selective colleges even when they have high test scores. In fact, parental income and test scores are important predictors of application behavior both at the extensive margin (i.e. whether students applied at all to any four-year college) and at the intensive margin (i.e. whether students included any highly selective colleges in their application portfolios). Importantly, we also find that there is substantial risk not only in admissions, but also in financial aid. Almost a quarter of all students admitted to both selective and non-selective colleges do not attend their top choice because of high costs. The risk of not receiving sufficient financial aid after being admitted is much more likely to occur to low-income students, even if their SAT scores are within the top 10% of the distribution.

Motivated by these facts, the second part of the paper presents a novel equilibrium model of the college market featuring student heterogeneity in terms of parental income and innate ability, and a noisy application and admissions system in the tradition of Epple, Romano, and Sieg (2006) and Fu (2014). The model is used to study the role of the admissions system in shaping the allocation of students across selective vs. nonselective colleges, as well as the effect of higher education policies (such as removing SAT scores and increasing federal aid) on application and admission decisions. The estimated model is able to account quantitatively for the application and enrollment patterns observed in the data, featuring realistic tuition schedules

that vary both across and within colleges based on students' parental income and test scores. We find that selective colleges offer high financial aid to low-income students because only the highest ability among them apply, making the colleges confident that their low-income applicants are likely to be high ability. Additionally, we find that making applications less informative (e.g. by removing the SAT) would lower merit-based financial aid and reduce admissions standards, hurting all high-ability students and modestly benefiting low-ability, high-income students.

The college market features a discrete number of colleges who differ in their technology, endowment income, and costs. Their objective is to maximize the value added to their students in the labor market, which depends on the average ability of their student body and the average level of instructional spending per student. Colleges are unable to observe the true ability of the students in their applicant pool and only see a noisy signal of their ability. Parental transfers are public knowledge for colleges, which is in line with the information revealed through the Free Application for Federal Student Aid (FAFSA). To maximize their objective, colleges choose their admissions standards (i.e. the minimum acceptable signal of ability) and tuition schedules that vary based on a student's parental income and signal of ability. Given the sequential nature of the application process, colleges have to factor in the possibility of students applying to other colleges. Combining the signal extraction problem with the college's ability to price discriminate is novel and introduces an important mechanism that influences the sorting of students in equilibrium.

Students choose to apply to a subset of colleges or not apply at all. If a student decides to apply, a noisy signal of her ability together with information on parental income is sent to the colleges of her choice. Once these costly applications are sent, students can receive offers of admission from none, one, or multiple colleges. Students then choose which college to attend among the set they have been accepted to. Given the information asymmetry, both admissions *and* financial aid are risky for the student. In particular, they have to make their application decisions based on their expectations of the possible signal realizations that can be sent to each college. If only the highest ability low-income students choose to apply, signals from low-income students will be more informative to the college because they are more likely to have come from high-ability students. Colleges will then offer high financial aid to the low-income students they enroll because they are confident that such students are likely to be high-ability. This mechanism helps explain why we observe both low application rates and high financial aid among low-income students.

To study the importance of the signal's informativeness, a counterfactual economy in which all low-income students are as likely to apply to selective colleges as their higher income peers is analyzed. By adding low-ability students to the low-income applicant pool, the signals of high-ability low-income students become less informative. In response, selective colleges reduce their financial aid for all low-income students as their applicant pool has worsened. This effective increase in tuition ends up reducing low-income student enrollment at the selective colleges by about a quarter. This finding highlights the benefit high-ability, low-income students derive from the informativeness of their signal when the application rates of low-ability, low-income students is small. Thus, interventions that encourage low-income students to apply may actually *reduce* their overall enrollment in selective colleges if they are not targeted by student ability.

Two policy counterfactuals are then analyzed. Motivated by the decision of many colleges to pause their use of standardized tests during the Covid-19 pandemic, we first study the effects of applicants' signals becoming *less* informative about ability. By increasing the noise associated with applications, it becomes harder for colleges to infer students' true ability. In the new equilibrium, colleges reduce their admissions standards and merit-based financial aid, making

students at the top of the ability distribution worse off as they now have a lower chance of being admitted to selective colleges. High-ability, low-income students are particularly harmed because of the reduction in merit-based financial aid. The only students who gain from the less informative signals are the low-ability, high-income students who now find it easier to be admitted. Low-ability, low-income students have little net-change in welfare because the gains they experience from the increased admissions rates are offset by the losses they experience from lower financial aid. Overall, there are welfare losses as the high-ability students who benefit the most from attending college are made significantly worse off than their peers.

Finally, we study the effects of a large expansion in the federal Pell Grant program, which would increase the amount of grant funding to low-income students and make middle-income students eligible for federal aid. This policy is most beneficial to high-ability low- and middle-income students, who before were less likely to apply and enroll due to credit constraints. This policy considerably reduces the concentration of income in selective colleges, but gaps remain as income and ability are correlated in the student population. In terms of welfare, higher income students are worse off due to the higher tax rate they will have to pay and the increased competition with the newly unconstrained lower income students. Overall, the policy has a net-positive effect on welfare since the value of a college education is higher for the high-ability, low-income students relative to the lower ability, high-income students they replace. Moreover, failing to account for the presence of the admissions system would lead the welfare gains to be overstated by more than a factor of two.

**Related literature.** This paper builds on three different strands of 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 literature using equilibrium models to study the college market and the forces driving the sorting of students. The third relates to the literature that studies the distributional effects of education policies.

This paper is complementary to the empirical literature studying outcomes of college students. Recent work by [Chetty, Friedman, Saez, Turner, and Yagan \(2020\)](#) documents a large degree of income segregation within and across U.S. colleges. Relatedly, work by [Hoxby and Avery \(2014\)](#), [Hoxby and Turner \(2013\)](#), [Dillon and Smith \(2017\)](#), and [Delaney and Devereux \(2020\)](#) document differences in application behavior related to differences in family income and student ability. This paper confirms these results using novel data covering a representative sample of high-school students. The noisy application and admission mechanisms is related to [Dynarski, Libassi, Michelmore, and Owen \(2018\)](#), who study the role of expectations about financial aid at the application stage. In their experiment, low-income, high-achieving high school students were encouraged to apply to the University of Michigan with the *promise* of full tuition scholarships over four years if admitted. They find significantly higher application and enrollment rates among their treated group in contrast to [Bettinger, Long, Oreopoulos, and Sanbonmatsu \(2012\)](#), who find no effect on student applications when information is provided without any commitment. These findings are consistent with our modeling choices about the role of expectations about financial aid in driving student applications.

This paper is also related to the literature on peer effects as the framework developed here inherently features peer effects from attending college. There is ample evidence that students benefit from having better peers. For example, [Sacerdote \(2001\)](#), [Zimmerman \(2003\)](#), and [Carrell, Fullerton, and West \(2009\)](#) find positive effects on grades from randomly assigning students interacting with other high scoring students. The evidence on the effect of better peer groups in terms of labor market outcomes is more nuanced. [Dale and Krueger \(2002, 2014\)](#) and [Mountjoy and Hickman \(2020\)](#) find no returns to college selectivity after controlling for

the set of colleges students applied to and had been accepted to.<sup>1</sup> However, [Hoekstra \(2009\)](#), [Zimmerman \(2014\)](#), [Andrews, Imberman, and Lovenheim \(2020\)](#), and [Bleemer \(2021\)](#) show that academically marginal students have a higher return from attending more selective colleges.<sup>2</sup> In our model, peer effects exist and lead to better wages in the labor market as colleges spend more on education when their pool of students is better.

The framework presented here builds on the work of [Epple, Romano, and Sieg \(2006\)](#) and [Epple, Romano, Sarpca, and Sieg \(2017\)](#), who study equilibrium models of the college market with quality-maximizing colleges that price discriminate among their students. It complements [Fu \(2014\)](#), who jointly models tuition and admissions, by adding the important margin of heterogeneity in parental income and credit constraints. This paper also adds to this framework the sequential noisy application and admissions problem, which draws on [Chade, Lewis, and Smith \(2014\)](#), who introduce matching frictions in the college admissions problem and allow students to make multiple college applications. The signal extraction problem also complements [Fillmore \(2020\)](#), who studies the effect of different FAFSA information disclosure policies on tuition levels. Several recent papers have also studied the college market and its interaction with inequality and intergenerational mobility. For instance, [Cai and Heathcote \(2022\)](#) study the role of income inequality in explaining the recent tuition increases using a model that gives rise to an endogenous distribution of colleges. Similarly, [Gordon and Hedlund \(2016, 2021\)](#) study the rise in college tuition, showing that demand forces help explain much of the increase. [Capelle \(2019\)](#) studies the role of the college market in shaping intergenerational mobility for heterogeneous students.

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 \(1996\)](#), [BÁf nabou \(2002\)](#), [Lochner and Monge-Naranjo \(2011\)](#), [Ionescu \(2009\)](#), [Ionescu and Simpson \(2016\)](#), [Krueger and Ludwig \(2016\)](#), [Kotera and Seshadri \(2017\)](#), [Caucutt and Lochner \(2017\)](#), [Abbott, Gallipoli, Meghir, and Violante \(2019\)](#), [Colas, Findeisen, and Sachs \(2021\)](#)). In particular, our paper complements the analysis of [Abbott, Gallipoli, Meghir, and Violante \(2019\)](#), who study the effect of financial aid policies and intergenerational transfers on welfare, [Ionescu and Simpson \(2016\)](#) and [Lucca, Nadauld, and Shen \(2018\)](#), who examine policy changes in student loan limits on college enrollment and tuition, and [Krueger and Ludwig \(2016\)](#), who analyze the optimal mix of tax and education subsidies and their impact on human capital accumulation. Our paper contributes to this literature by studying the effectiveness of education policies in an environment that takes into account endogenous changes in the colleges market.<sup>3</sup>

**Outline.** The remainder of the paper is organized as follows. Section 2 presents empirical evidence on application and enrollment patterns among students transitioning from high school to college; Section 3 describes the equilibrium model of the college market; Section 4 presents the estimation procedure; Section 5 discusses the model's mechanisms; Section 6 analyzes the effect of removing the SAT in college applications; Section 7 focuses on the effect of increasing need-based financial aid such as the Pell Grants; and Section 8 concludes.

<sup>1</sup>Though [Dale and Krueger \(2014\)](#) do find returns to selectivity among disadvantaged students.

<sup>2</sup>For instance, [Chetty, Friedman, Saez, Turner, and Yagan \(2020\)](#) find that more selective colleges give higher returns to education even after controlling for the set of colleges the students applied to. They estimate the causal effect (due to value-added) of earnings differences across colleges to be around 80%.

<sup>3</sup>The response of colleges to changes in financial aid policy has been shown to be empirically relevant. For example, [Lucca, Nadauld, and Shen \(2018\)](#) and [Turner \(2017\)](#) provide evidence that college tuition increases in response to expansions in federal financial aid.

## 2. EMPIRICAL EVIDENCE

This section studies the college application and enrollment behavior of U.S. high-school students. It shows empirically that income gaps in college enrollment identified in the previous literature persist and that they are mostly explained by differences in application rates. Additionally, this section shows that, conditional on being accepted, low-income students are less likely to attend their preferred college due to costs. This evidence supports the theory that low-income students are less likely to apply because they do not expect to receive sufficient aid even if admitted.

### 2.1. Data

The analysis here presented relies on the High School Longitudinal Study of 2009 (HSLs). Published by the National Center for Education Statistics (NCES) of the U.S. Department of Education, 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 from high-school, 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 other economic characteristics. We use the restricted-use version of the HSLs, which also provides student SAT and ACT scores, lists of colleges applied to and enrolled in, and more detailed information about household-level economic variables.

The focus is on how application and enrollment decisions vary based on parental income and college preparedness. The former is provided directly by the parents in the HSLs survey, where they are asked for their households' income from all sources in 2011.<sup>4</sup> For college preparedness, we use the students' high-school GPAs and SAT or ACT scores.<sup>5</sup> Each student's GPA is reported directly by her high-school and is honors-weighted by the NCES in a procedure used to make the GPA comparable across different high-schools. The SAT is reported directly by each student's college and is therefore unavailable for students who did not attend college or did not take the test. For these students, we simply impute the test scores using the mean SAT score of other HSLs students from the same parental income group, level of parental education, race, and GPA decile.

Colleges are categorized into different selectivity groups using the Barron's selectivity index (Profile of American Colleges, 2015) as done in [Chetty, Friedman, Saez, Turner, and Yagan \(2020\)](#). We restrict our focus to all four-year non-profit colleges and use two selectivity groups. The first group, which we call "Highly-selective", corresponds to Barron's Tier 1 and 2 colleges. This group makes up 192 schools. For reference, a list of all colleges in this group is provided in [Appendix B.1](#). All other four-year non-profit colleges and universities are counted in the second group, which we refer to as "Non-selective." In our HSLs sample, the Highly-selective colleges account for about 16% of all four-year, non-profit enrollment and represent about 7% of students.<sup>6</sup>

<sup>4</sup>Surveyed parents report their income by selecting from a pre-set range (i.e. less than \$15,000, \$15,000 to \$35,000, \$35,000 to \$55,000, etc.). The percentages for each grouping are too coarse to fit neatly into quartiles or quintiles. For reference, the parental income distribution from our HSLs sample is provided in [Figure B.1](#) of the Appendix.

<sup>5</sup>The NCES converts the ACT score into an equivalent SAT 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.

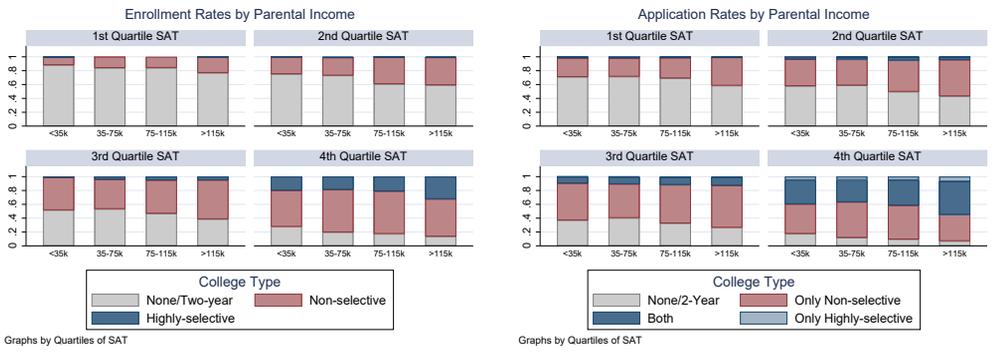
<sup>6</sup>As expected, Highly-selective colleges have higher median SAT scores, higher instructional spending and endowment assets per student, and higher median earnings after graduation. They are generally more expensive on average, but provide generous aid for their low-income students. See [Table B.1](#) in the appendix for a comparison of key statistics between each selectivity group.

The HSLS also includes detailed information about each student’s application portfolio choice. In the follow up survey after completing high-school, students were asked to provide the college they were currently attending and to list two other colleges they had applied to and seriously considered. Additionally, students were asked to provide the *total* number of applications they sent. While the data therefore include only the three most relevant colleges, most students indicated that they had applied to three or fewer schools, suggesting that the HSLS gives a good picture of overall application portfolios.<sup>7</sup>

## 2.2. College applications and enrollment

Figure 2.1 below shows how the observed application and enrollment rates vary by parental income and SAT score quartile. While applications and enrollment are highly correlated with test scores, they also correlate with parental income conditional on SAT scores. Importantly, the left panel shows that there are income gaps in enrollment even for students in the top SAT quartile. For example, students in the bottom third of the parental income distribution were more than twice as likely not to attend college than those in the top sixth, and about half as likely to enroll in a highly-selective college. Turning to the right panel, Figure 2.1 shows that these patterns emerge from student decisions at the *application* stage. For students with top SAT scores, those at the bottom of the income distribution were significantly less likely to apply to any college (and highly-selective colleges in particular) compared to students from the top of the income distribution.

FIGURE 2.1.—College application and enrollment by parental income and student SAT score



(a) Enrollment rates

(b) Application rates

Note: Panel (a) shows college enrollment rates by parental income (x-axis) and SAT score (different subplots) by enrolled college type (none, non-selective, or highly-selective). Panel (b) shows college application rates by parental income (x-axis) and SAT score (different subplots) by applied college types (none, only non-selective, only highly-selective, or both non and highly-selective). Figure A.1 in the Appendix shows similar patterns when GPA is used instead of SAT scores.

Motivated by the importance of application decisions in explaining differences in enrollment, we study whether parental income has predictive power in determining application choices

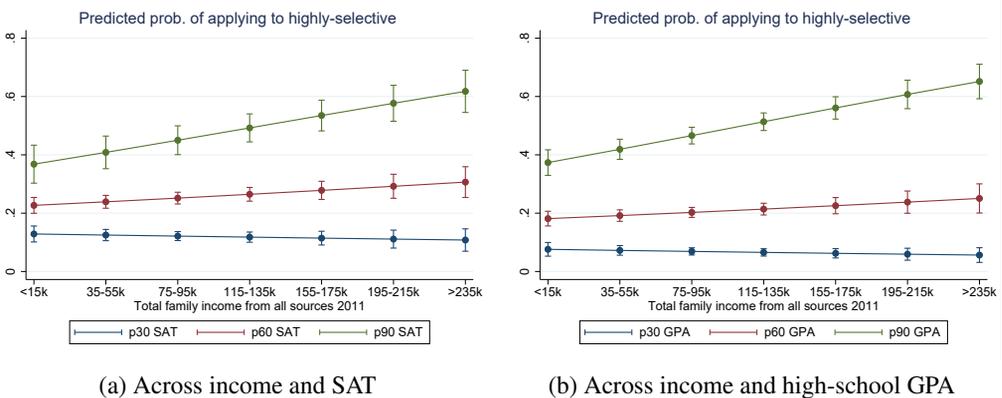
<sup>7</sup>In the preceding analysis, application portfolios are categorized as follows: a student who did not list any non-profit four-year college is considered to not have applied. If at least one non-profit four-year application is listed but none of the three schools are Highly-selective, the student is classified as only having applied to Non-selective colleges. If at least one Highly-selective and one Non-selective application are listed, the student is classified as having applied to both. Finally, if only applications to Highly-selective colleges are listed, the student is counted as having applied to both only if the number of *total* applications is greater than the number of *listed* applications (i.e. students are assumed to have also applied to a Non-selective college as a safety).

while controlling for other factors. In that regard, two sets of logit models are estimated using the HSLs student-level data. The upshot of the analysis is the following: parental income is an important factor in determining not only whether or not a student applies to college, but also in determining which type of colleges a student decides to include in her application portfolio.

In the first logit model, the factors correlated with the extensive margin of the decision to apply are analyzed. The left columns of Table A.1 in the Appendix show the results of the estimation, where the dependent variable is an indicator for whether or not the student decided to apply to any four-year, non-profit college. Both test scores and parental income are strongly correlated with a student's decision to apply. This relationship remains significant after controlling for other potential confounders, including race, sex, parental education level, and high-school level characteristics. To aid in interpreting the results, the predicted application probabilities implied by the logit estimates for different parental income levels and test scores are plotted in Figure A.2 in the Appendix. While the relationship between parental income and applications declines for higher test-scores, there is still a strong effect for students with median SAT scores.

The second logit model analyzes the intensive margin of college applications, i.e. how parental income and test scores are correlated with students' application portfolio choice. The right columns of Table A.1 show the estimation results, where the dependent variable is an indicator for whether or not the student included any highly-selective college in her application portfolio (the sample is restricted to students who apply). Test scores have predictive power and parental income when interacted with test scores is highly significant. The predicted probability of applying to a highly-selective college for different SAT scores and parental income levels is presented in Figure 2.2. Since income by itself is insignificant, there is little difference in application rates to highly-selective colleges for students with relatively low test scores. However, since the interaction between SAT and income is significant, students at the top of the parental income distribution are more likely to apply to a highly-selective college than their lower-income peers who have similar high test scores. For instance, an increase in family income of about \$80,000 is associated with a 10% increase in the probability that a student in the 90<sup>th</sup> percentile of SAT scores applies to a highly-selective college.

FIGURE 2.2.—Predicted probability of applying to highly-selective colleges



Note: Panel (a) shows the predicted probability of applying to highly-selective colleges implied by the estimated logit model (conditional on applying at all) for parental income and SAT scores. Panel (b) shows the predicted probability of applying to highly-selective colleges implied by the estimated logit model (conditional on applying at all) for parental income and high-school GPA. See right columns of Table A.1 for estimation results.

### 2.3. Enrollment patterns conditional on admission

Thus far, the findings of this section are consistent with the previous literature that has documented that college enrollment gaps can be explained in part by differences in application choices between high- and low-income students. This subsection shows that even if low-income students are accepted to their preferred college, they are less likely to enroll because of high tuition costs. This finding suggests that students face the risk of not receiving sufficient financial aid when they apply and provides evidence for the hypothesis that low-income students do not apply in the first place because they do not expect to receive enough aid.<sup>8</sup>

A benefit of the HSLS is that it surveys students about their enrollment decisions conditional on admission. In addition to listing the colleges they applied to, students were also asked whether they were admitted to each school and to identify their top choice. Importantly, students were also asked which college they preferred among those *accepted* to if not for costs, which allows us to determine if students received enough financial aid in order to enroll. Table 2.1 below shows that among students admitted to any four-year non-profit college, most were admitted to their preferred college regardless of their income group. However, lower-income students were less likely to attend their preferred college due to costs. As a result, these students were significantly more likely to be diverted into two-year colleges or chose to go straight into the labor market.

TABLE 2.1  
ADMISSIONS AND ENROLLMENT STATISTICS FOR ADMITTED STUDENTS

	Parental income (US\$)			
	<35k	35-75k	75-115k	>115k
Admitted to preferred college	85.0%	87.0%	86.8%	87.4%
Attending most preferred college admitted to	64.8%	67.7%	72.2%	77.2%
<i>Percent attending four-year non-profit if...</i>				
Enrolled in preferred college	93.5%	95.3%	95.6%	97.0%
Not enrolled in preferred college (due to cost)	55.4%	68.1%	70.0%	73.7%

*Note:* Sample restricted to all students who were admitted to any four-year, non-profit college. The "preferred college" in the bottom two rows refers to the student's top choice among the colleges they were *admitted* to.

Table 2.2 shows that these sorting patterns hold even among high achieving students who are admitted to highly-selective colleges that generally provide significant aid to needy students. While rates of preference for attending a highly-selective college are similar for admitted students with different incomes, low-income students are significantly less likely to attend one because of costs. Conditional on selecting a highly-selective college as their top choice among those admitted to, students in the bottom third of the parental income distribution were twice as likely not to attend one compared to students in the top sixth. Figure 2.3 shows that this negative relationship remains even among admitted students with SAT scores above 1300 (approximately the top 10% of test takers), who are most likely to receive generous financial aid at selective colleges.

<sup>8</sup>Note that in addition to need-based financial aid which can to some extent be predicted by looking at college websites, many selective colleges also use merit-based aid that depends on an evaluation of the student's application and cannot be predicted with certainty. While it is well known that colleges at the very top like Harvard or Yale cover all expenses for low-income students, other selective schools have separate merit-based aid that is granted competitively and thus their low-income students have higher net tuition on average. See the scatterplot in Figure A.3, which shows a positive relationship between admission rates and net tuition for low-income students.

TABLE 2.2

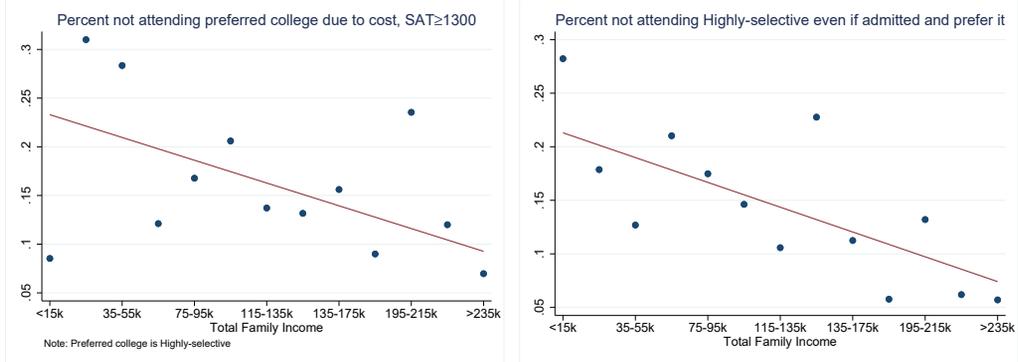
ADMISSIONS AND ENROLLMENT STATISTICS FOR STUDENTS ACCEPTED TO AT LEAST ONE HIGHLY-SELECTIVE COLLEGE

	Parental income (US\$)			
	<35k	35-75k	75-115k	>115k
Preferred college admitted to is HS	65.8%	73.3%	67.8%	71.1%
Not attending HS conditional on preferring it*	21.0%	17.5%	15.9%	11.3%
Attending any HS college	57.6%	65.4%	61.0%	67.2%

Note: Sample restricted to students who identified a Highly-selective (HS) college when asked which college they preferred among those admitted to (if not for costs).

To summarize, we find that gaps in college enrollment are due not just to differences in application, but also to enrollment decisions conditional on being admitted once students know how much financial aid they will receive. Motivated by these facts, the next section builds an equilibrium model of the college market with a realistic application and admissions system to study how admissions and financial aid policies shape outcomes for students across the income distribution. Importantly, students are allowed to make enrollment decisions *after* being admitted. The model is used for two reasons. The first is to create an environment that allows us to test whether our theory that expectations about financial aid is sufficient in explaining the observed college application gaps. The second is to serve as a laboratory for evaluating counterfactual policies, while taking general equilibrium effects into account. In particular, the model is used to study equalizing application rates or removing SAT requirements would affect the allocation of students in general equilibrium.

FIGURE 2.3.—Students not attending a highly-selective college despite preferring one and being admitted to one



(a) Not attending preferred highly-selective college

(b) Not attending any highly-selective college

Note: Panel (a) shows the fraction of students not attending their most preferred highly-selective college despite preferring one (if not for costs) and being admitted to one conditioning on students who scored at least 1300 on the SAT. Panel (a) shows the fraction of students not attending *any* highly-selective college despite preferring one (if not for costs) and being admitted to one.

### 3. MODEL

**Overview.** The economy is populated by a unit measure of heterogeneous individuals, who live for two periods: young and old. Young students start life with parental transfer  $y$  and ability level  $\ell$ , and decide whether to work or invest in their human capital by attending college. As there are multiple colleges, students have to choose which set of colleges they want to apply to.

Admissions, however, are risky as colleges and the student can disagree on her ability. Colleges' perception of a student's ability is  $\sigma$ , which is correlated with the student's ability  $\ell$ .<sup>9</sup> For lack of a better term, we call  $\sigma$  the ability signal. Students with a high enough realization of  $\sigma$  can receive multiple offers of admission and college-specific tuition levels that depend on her income,  $y$ , and ability signal,  $\sigma$ . Once the uncertainty is resolved and students know their admissions and financial aid decisions, they choose which college—among possibly many—to enroll in. At any stage, students may choose the outside option of working instead of going to college.<sup>10</sup> While old, individuals work. Some of them will reap the benefits of attending college.

There are two tiers of colleges: highly selective colleges, indexed by  $H$ , and less selective colleges, indexed by  $L$ . There are  $N_H$  selective colleges and  $N_L$  less selective colleges in the economy. Colleges differ exogenously in their endowment income, fixed costs, and productivity. Colleges maximize the value they add to their students on the labor market. A college's value added is determined *endogenously* in equilibrium and is a function of the average ability of its student body and its average instructional spending per student. The college market is monopolistically competitive. Colleges compete by setting different admission standards and tuition schedules for each student type, but they take into account the admission standards and tuition schedules of the competing colleges. For instance, when competing colleges make an offer of admission, a college takes into consideration that the student may have received other offers and thus 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 policies of all the competing colleges.

Finally, there is a government that taxes the working population to subsidize colleges and pay for grant programs that support low-income students.

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

1. Individuals choose either to apply to college or go straight to the labor market. Those who apply must choose an application portfolio which includes any combination of highly-less selective and less selective colleges.
2. Colleges receive applications and choose which students to accept by setting their admission standards and tuition schedules.
3. Students make their attendance decision given the admission and financial aid offers received.
4. Individuals make their consumption and savings decisions.

### 3.1. Students

We proceed in chronological order by first introducing the decision problem of an applicant, then the decision problem involving the acceptance and rejection of offers of admission, and finally the problem of a student attending college. The problem of a person who did not attend college comes last.

---

<sup>9</sup>Students' parental transfer is assumed to be fully observable to colleges. 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, and is fully observable by the colleges that the student applies to. About 77% of students in the HLSL complete the FAFSA.

<sup>10</sup>For simplicity, we do not separately model two-year colleges.

### 3.1.1. Application stage

At the college application stage, a student can choose among  $2^{(N_H+N_L)}$  possible application portfolios (i.e., the power set), including not applying to college and starting to work.<sup>11</sup> When a student applies to college, colleges draw a realization of the applicant's ability,  $\sigma$ , that is unknown to her. Think of  $\sigma$  as what admission committees think of the applicant's ability based on her essays, SAT scores, high-school grades, and other extra-curricular activities.

The signals are drawn from a continuous conditional density  $g(\sigma|\ell)$  with cumulative density function  $G(\sigma|\ell)$  and support  $(0, \infty)$ . The conditional density function is increasing in  $\ell$  such that high signals are more likely to come from high-ability students. For simplicity, colleges agree on the same signal of ability.<sup>12</sup> A student is admitted to a college  $s \in S = \{H_1, \dots, H_{N_H}, L_1, \dots, L_{N_L}\}$  if her signal is above the college's admission standard, i.e.,  $\sigma \geq \underline{\sigma}_s$ . Without loss of generality, we rank colleges within their tiers in ascending order in terms of admission standards, i.e.,  $\underline{\sigma}_{j_1} \geq \dots \geq \underline{\sigma}_{j_{N_j}}$  for  $j = \{H, L\}$ , and focus on the case in which the worst highly-selective college is at least as selective as the best less selective college, i.e.,  $\underline{\sigma}_{H_{N_H}} \geq \underline{\sigma}_{L_1}$ .<sup>13</sup>

College applications are costly. Applying involves a financial cost as well as a 'psychic' disutility cost meant to capture the effort needed to complete applications.<sup>14</sup> We denote the financial application cost by  $\psi(n)$ , where  $n$  is the number of applications submitted, and the disutility of applying by  $\phi(n, A)$ , which might also depend on the set of colleges the student applied to,  $A$ , as some colleges might require more elaborated applications than others.

**Value of applying to colleges.** There are  $2^{(N_H+N_L)}$  application portfolios that span the power set of colleges available. The expected value of a student with parental income  $y$  and ability  $\ell$  of applying to colleges in the set  $A \subseteq \mathbb{P}(S)$  is given by  $V^A(y, \ell, A)$ , which corresponds to the expected value of being admitted to the  $|A|$  colleges the student applied to.

To give an example, the expected value of applying to all colleges in the economy (i.e.,  $|S| = N_H + N_L$ ) is given by

$$\begin{aligned}
 V^A(y, \ell, S; \mathbf{T}(y, \sigma), \underline{\sigma}) &= \underbrace{\int_{\underline{\sigma}_{H_1}}^{\infty} V^O(y, \ell, N_H + N_L, S; \mathbf{T}(y, \sigma))g(\sigma|\ell)d\sigma}_{\text{Expected value of being admitted to colleges in } S \\ &\quad \text{including an offer from the most selective one}} + \dots + \\
 &\quad \underbrace{\int_{\underline{\sigma}_{L_1}}^{\underline{\sigma}_{H_{N_H}}} V^O(y, \ell, N_H + N_L, S \setminus \{H_1, \dots, H_{N_H}\}; \mathbf{T}(y, \sigma))g(\sigma|\ell)d\sigma}_{\text{Expected value of being admitted to colleges in } S \\ &\quad \text{excluding all highly-selective colleges}} + \dots +
 \end{aligned}$$

<sup>11</sup>For instance, if  $N_H = N_L = 1$ , a student has four options as she can choose to only apply to the highly-selective college, only apply to the less-selective college, apply to both colleges, or not apply at all.

<sup>12</sup>This assumption, while made for tractability, is also supported empirically. We find that in the HSLs, 97% of students who were admitted to a highly-selective college (as defined in Section 2.1) were also admitted by a non-selective college (conditional on having applied to both).

<sup>13</sup>While it is possible for a less selective college to have a higher admissions standard than a highly-selective college, we confirm that this is not the case in our baseline estimation and subsequent analysis.

<sup>14</sup>This cost includes the effort and lost time associated with writing essays and filling application forms, preparing for and taking the SAT (perhaps multiple times), and the time spent researching which colleges are worth applying to.

$$\begin{aligned}
& \underbrace{\int_{\underline{\sigma}_{L_{N_L}}}^{\underline{\sigma}_{L_{N_L}-1}} V^O(y, \ell, N_H + N_L, \{L_{N_H}\}; \mathbf{T}(y, \sigma)) g(\sigma | \ell) d\sigma}_{\text{Expected value of only being admitted to the least selective college}} + \\
& \underbrace{G(\underline{\sigma}_{L_{N_L}} | \ell) V^W(y, \ell, N_H + N_L)}_{\text{Expected value of not being admitted to any college after submitting } (N_H + N_L) \text{ applications}} - \underbrace{\phi(N_H + N_L, S)}_{\text{Psychic application cost}}, \tag{3.1}
\end{aligned}$$

where  $V^O(y, \ell, n, O; \mathbf{T}(y, \sigma))$  is the value of a student with offers of admission in the set  $O \subseteq A$  after submitting  $n$  applications to colleges in the set  $A \subseteq \mathbb{P}(S)$ ,  $\mathbf{T}(y, \sigma)$  is a vector of tuition fees for an admitted student with parental income  $y$  and signal  $\sigma$ , and  $V^W(y, \ell, n)$  is the value of not attending college after applying to  $n$  colleges. To fix ideas, the remaining  $2^{(N_H + N_L)} - 1$  application portfolios look like problem (3.1) for each possible combination of colleges.<sup>15</sup>

**Optimal application.** The optimal application decision for a student with characteristics  $(y, \ell)$  solves the problem of choosing the application set (i.e., which colleges she chooses to apply to) among the subsets of the power set  $\mathbb{P}(S)$ , i.e.,

$$\max \underbrace{\{V^A(y, \ell, \{L_1\}), \dots, V^A(y, \ell, S), V^W(y, \ell, 0)\}}_{\text{Expected values of the } 2^{(N_H + N_L)} \text{ application portfolios}}. \tag{3.2}$$

The solution to this problem yields the student's application set, which can be any subset of the power set  $\mathbb{P}(S)$ —including not applying for college. In that case, the student gets the value from working without submitting any college application  $V^W(y, \ell, 0)$ .

### 3.1.2. Enrollment stage

After the signal is realized and students know their admissions and financial aid status, they decide which offer, if any, to accept. At this stage, students face no other uncertainty regarding their admissions.

**Value with offers of admission.** The value of a student with  $|O|$  offers of admission in hand after submitting  $n$  applications is the expected value from choosing between accepting any offer within the set  $O \subseteq A$  or working after spending financial resources applying for college. Suppose, for instance, that the set of offers the student receives includes the least selective college  $L_1$  through the  $n^{\text{th}}$  highly-selective college, or  $H_n$ . Then, her value after applying to  $(N_L + n)$  colleges is given by

$$\begin{aligned}
V^O(y, \ell, N_L + n, O; \mathbf{T}(y, \sigma)) = \max \underbrace{\{V^{L_1}(y, \ell, N_L + n, T_{L_1}(y, \sigma)), \dots, V^{H_n}(y, \ell, N_L + n, T_{H_n}(y, \sigma))\}}_{\text{Values of attending one of the } |O| \text{ colleges after submitting } N_L + n \text{ applications}}, \\
V^W(y, \ell, N_L + n), \tag{3.3}
\end{aligned}$$

<sup>15</sup>For instance, a student that only applies to the least selective college as the following value

$$\begin{aligned}
V^A(y, \ell, \{L_{N_H}\}) = \underbrace{\int_{\underline{\sigma}_{L_{N_L}}}^{\infty} V^O(y, \ell, 1, \{L_{N_H}\}; \mathbf{T}(y, \sigma)) g(\sigma | \ell) d\sigma}_{\text{Expected value of only being admitted to the least selective college}} + \underbrace{G(\underline{\sigma}_{L_{N_L}} | \ell) V^W(y, \ell, 1)}_{\text{Expected value of not being admitted to college}} - \underbrace{\phi(1, \{L_{N_H}\})}_{\text{Psychic application cost}}.
\end{aligned}$$

where  $V^s(y, \ell, n, T_s(y, \sigma))$  is the value of attending college  $s \in S = \{H_1, \dots, H_{N_H}, L_1, \dots, L_{N_L}\}$  and paying tuition  $T_s(y, \sigma)$  after submitting  $n$  applications. The solution to this problem yields the college the student will attend (or the value of working with a high-school education if she chooses not to attend college).

**Enrollment probability.** We introduce idiosyncratic preference shocks over the students' alternatives when applying for college and when choosing between offers of admission (decision problems (3.2) and (3.3)). Let  $\lambda_a > 0$  and  $\lambda_c > 0$  be the scale parameters of the mean zero Type I extreme value shocks. Including these shocks adds analytical tractability as they allow for closed-form (and differentiable) solutions to the students' enrollment decision, which colleges then take as given.

Let  $q_s(y, \ell, \mathbf{T}(y, \sigma), \underline{\sigma})$  denote the enrollment probability of a student with characteristics  $(y, \ell)$  of the college  $s$  offer given tuition schedules and admission standards. The total probability a student will enroll in college  $s$  is the sum of the product of the probability of applying to a particular set of colleges among  $2^{(N_H + N_L)}$  application portfolios and the probability of accepting the college  $s$  offer given the set of offers  $O \subseteq A$  and the chosen application portfolio  $A \subseteq \mathbb{P}(S)$ , or

$$q_s(y, \ell, \mathbf{T}(y, \sigma), \underline{\sigma}) = \sum_{i \in \mathbb{P}(S)} \Pr(\text{Accepting college } s \text{ offer} | \text{App. portfolio } i \text{ and set of offers } O_i) \times \Pr(\text{Choosing application portfolio } i). \quad (3.4)$$

This enrollment probability is a key input in the colleges' problem as it represents each student's demand for college.

### 3.1.3. The college-educated individual's problem

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. While in college, the student must finance her consumption, tuition fees, and application costs using parental transfers, grants, and student debt. Students may borrow up to a limit, denoted by  $\underline{a}_c$ .<sup>16</sup> Grants depend on the student's income and are separated into grants from outside sources,  $Gr(y)$ , and government-funded grants,  $P(y)$ . Outside grants are meant to capture private scholarships and are taken exogenously. Government-funded grants are meant to capture the federal Pell Grant program and are paid for using labor income taxes on workers.

After graduating from college, the student pays her loans back priced at the interest rate  $R$  and enters the labor market, receiving the wage rate  $w$  net of taxes  $\tau$  per unit of human capital acquired in college. If the student graduates, the resulting human capital is given by  $\delta_s Z_s \ell^{\alpha_c}$ , which depends on the probability of graduating,  $\delta_s$ , the value added by the college she attended,  $Z_s > 1$ , and the exponent  $\alpha_c < 1$  governing the returns to the student's ability. Better colleges offer better prospects by raising the human capital a student has. If the student fails to graduate, her human capital is  $(1 - \delta_s) \ell^{\alpha_w}$ , where  $\alpha_w < \alpha_c < 1$  implies that the return from attending college is larger, the higher is the ability of the individual (for  $\ell > 1$ ).

The student has preferences over consumption and discounts the future at rate  $1/\beta$ . While in college, she faces a utility cost from attending college, denoted by  $\nu_s(\ell) \geq 0$ . This disutility

<sup>16</sup>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.

cost is weakly decreasing in the student's ability as higher-ability students might find attending college easier than lower-ability students. The disutility cost might also differ across colleges as attending a highly-selective college might be more demanding than attending a less selective college. This psychic cost of attendance is motivated by [Cunha, Heckman, and Navarro \(2005\)](#) and [Heckman, Lochner, and Todd \(2006\)](#), who find that such costs are necessary to explain observed college enrollment levels since pecuniary returns alone would predict higher enrollment.

**Value of attending college.** The value of a student enrolled in college  $s$  and paying tuition  $T_s(y, \sigma)$  after submitting  $n$  applications is given by

$$V^s(y, \ell, n, T_s(y, \sigma)) = \max_{c, c', a'} u(c) - \nu_s(\ell) + \beta u(c') \quad (3.5)$$

$$\begin{aligned} \text{s.t.} \quad c + a' &= y + Gr(y) + P(y) - T_s(y, \sigma) - \psi(n) \\ c' &= Ra' + (1 - \tau)w [\delta_s Z_s \ell^{\alpha_c} + (1 - \delta_s) \ell^{\alpha_w}] \\ a' &\geq \underline{a}_c. \end{aligned}$$

### 3.1.4. The high-school-educated individual's problem

Individuals end up as workers either by choosing not to apply to college, being rejected by any college they applied to, or by choosing not to attend college conditional on an offer of admission. A worker does not pay any tuition, but might incur application fees if she submitted college applications.

**Value of working without a college degree.**  $V^W(y, \ell, n)$  is the value of an individual with parental transfers  $y$  and ability  $\ell$  who submitted  $n$  college applications. The decision problem of a worker without a college degree is to choose consumption and savings up to a borrowing constraint  $\underline{a}_w$  according to

$$V^W(y, \ell, n) = \max_{c, c', a'} u(c) + \beta u(c') \quad (3.6)$$

$$\begin{aligned} \text{s.t.} \quad c + a' &= y + (1 - \tau)w \ell^{\alpha_w} - \psi(n) \\ c' &= Ra' + (1 - \tau)w \ell^{\alpha_w} \\ a' &\geq \underline{a}_w. \end{aligned}$$

## 3.2. Colleges

The college market is a monopolistically competitive industry. There are  $(N_H + N_L)$  heterogeneous colleges who compete for students by setting their tuition schedules and admission standards taking as given the tuition schedules, admission standards, and value added of the competing colleges. The choices of one college affect the other colleges through students' attendance decision defined in (3.4), which ultimately affect the final distribution of their student body.

In setting their tuition schedules, colleges offer different prices to students of different types. As students' ability  $\ell$  is unobservable, colleges infer it through Bayesian updating. Colleges

know the distribution of high-school students' characteristics,  $\hat{\mu}(y, \ell)$ , and the conditional signal distribution,  $G(\sigma|\ell)$ . The (posterior) density of students with characteristics  $(y, \ell)$  that colleges perceive as being of ability  $\sigma$  is denoted by  $\mu(y, \ell, \sigma) \equiv g(\sigma|\ell)\hat{\mu}(y, \ell)$ .

**College-specific value added.** The college's value added to students is given by

$$Z = \xi Q(I/\kappa, L/\kappa), \quad (3.7)$$

where  $Q(I/\kappa, L/\kappa)$  denotes the college's quality and  $\xi$  is the efficiency with which quality is transformed into better human capital for its students. A college's quality is increasing in the average amount of instructional spending,  $I/\kappa$ , and the average ability of its students,  $L/\kappa$ . The dependence of college quality on students' average ability allows for peer effects, as students benefit more from college when their peers are of higher ability.

Aggregate instructional spending is directly chosen by the college, while the ability of its students and the number of students enrolled are the outcome of admission standards and tuition schedules of all colleges. Using students' probability of attendance, equation (3.4), we can write the total enrollment in a college as the sum of all students who cleared the bar and accepted the college's offer of admission, i.e.,

$$\kappa = \int_{\underline{\sigma}}^{\infty} q(y, \ell, \mathbf{T}(y, \sigma), \underline{\sigma}) d\mu(y, \ell, \sigma), \quad (3.8)$$

where the integral is taken over all student types  $(y, \ell, \sigma)$  given the college's choice of admission standard  $\underline{\sigma}$ . Note that  $\mathbf{T}(y, \sigma)$  includes the college's tuition schedule as well as the competing colleges' tuition schedules. In turn, the aggregate student ability in the college is the sum of the ability of students who cleared the bar and accepted the college's offer of admission, and it is given by

$$L = \int_{\underline{\sigma}}^{\infty} \ell q(y, \ell, \mathbf{T}(y, \sigma), \underline{\sigma}) d\mu(y, \ell, \sigma). \quad (3.9)$$

**Profits.** Colleges balance their budget. Their revenue is derived from the total tuition paid by students,  $\mathcal{T}$ , their net endowment income  $E$ , and government transfers or appropriations  $Tr$ .<sup>17</sup> The endowment and government transfers might depend on the number of students enrolled  $\kappa$ . In addition to total instructional spending,  $I$ , the college faces other fixed costs that are not directly related to teaching,  $C$  (e.g., administrative or maintenance costs). The budget constraint of a college is thus given by

$$I + C = E(\kappa) + Tr(\kappa) + \mathcal{T}, \quad (3.10)$$

where the total tuition revenue is the sum of the tuition paid by all students who cleared the bar and accepted the college's offer of admission, i.e.,

$$\mathcal{T} = \int_{\underline{\sigma}}^{\infty} T(y, \sigma) q(y, \ell, \mathbf{T}(y, \sigma), \underline{\sigma}) d\mu(y, \ell, \sigma). \quad (3.11)$$

**College's problem.** The objective of a college is to maximize the earnings potential of its students through its value added by choosing its instructional spending, admission standard, and

<sup>17</sup>Here, the net endowment income refers to endowment income net of other operating expenses.

tuition schedule for each student type. The choice of this objective function merits some discussion. Since [Epple, Romano, and Sieg \(2006\)](#), the literature has formalized the college's problem as maximizing its quality as a function of students' ability and/or income and the college's instructional spending subject to a budget constraint.<sup>18</sup> Here, we follow the same strategy. There are two novelties, however. First, the quality of the college a student attends directly influences her earnings potential—as opposed to increasing her utility. Second, ability is unobservable and hence colleges can only indirectly influence the average ability of its student body, which prevents perfect assortative matching across college tiers.

A college solves the following maximization problem

$$\max_{I, \underline{\sigma}, \{T(y, \sigma)\}_{\forall (y, \sigma)}} \xi Q(I/\kappa, L/\kappa) \quad (3.12)$$

$$\text{s.t. } I + C = E(\kappa) + Tr(\kappa) + \mathcal{T}$$

$$T(\bar{y}, \underline{\sigma}) \leq \bar{T} \text{ and } \underline{\sigma} \geq 0.$$

The problem of a college depends on the tuition schedules and admission standards of the competing colleges due to the nature of the students' problem, i.e., students can only choose one offer among possibly many. This translates into having the acceptance probability,  $q$ , being dependent on the policies of all colleges a student applies to. Hence, each college's tuition policy is a best response to the competing colleges' own tuition policy.

We add the tuition cap,  $\bar{T}$ , to avoid extremely wealthy but low-ability students to be able to compensate the college for lowering its average ability. Without limits on tuition, colleges can simply charge the low-signal students enough to compensate for the decrease in average student ability they cause, making the use of admission standards irrelevant.<sup>19</sup>

### 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) while young and old, and college-educated workers while old. The aggregate labor supplied by young and old individuals is given by

$$N = 2 \underbrace{\int \ell^{\alpha_w} q_w(y, \ell) d\hat{\mu}(y, \ell)}_{\text{Labor supplied by young and old high-school workers}} + \underbrace{\sum_{s \in S} \int_{\underline{\sigma}_s}^{\infty} [\delta_s Z_s \ell^{\alpha_c} + (1 - \delta_s) \ell^{\alpha_w}] q_s(y, \ell, \mathbf{T}(y, \sigma), \underline{\sigma}) d\mu(y, \ell, \sigma)}_{\text{Labor supplied by college-educated workers}}, \quad (3.13)$$

<sup>18</sup>[Epple and Romano \(1998\)](#) model colleges as profit maximizers and argue that it leads to predictions similar to the one with quality maximization.

<sup>19</sup>Think of  $\bar{T}$  as the sticker price that some students do pay.

where  $q_w(y, \ell)$  is the total probability of not attending college for an individual with parental income  $y$  and ability  $\ell$ , given by

$$q_w(y, \ell) = 1 - \sum_{s \in S} \int_{\underline{\sigma}_s}^{\infty} q_s(y, \ell, \mathbf{T}(y, \sigma), \underline{\sigma}) g(\sigma | \ell) d\sigma.$$

This probability includes students who never applied to college, those that applied but did not get an offer, and those that got offers but decided not to accept any. The total measure of young and old is normalized to 1.

The government budget constraint must hold according to

$$\tau w N = \underbrace{\sum_{s \in S} \int_{\underline{\sigma}_s}^{\infty} P(y) q_s(y, \ell, \mathbf{T}(y, \sigma), \underline{\sigma}) d\mu(y, \ell, \sigma)}_{\substack{\text{Transfers to students} \\ \text{(Pell grants)}}} + \underbrace{\sum_{s \in S} Tr(\kappa_s)}_{\text{Transfers to colleges}}. \quad (3.14)$$

### 3.4. Equilibrium

An equilibrium in the college market requires that each college's optimal admission standard and tuition schedule are consistent with the optimal choices of the competing colleges. This means that the tuition schedule colleges offer to each student type is a fixed point in the set of all colleges' tuition policies. Note that each student type has  $2^{(N_H + N_L)}$  application portfolios and a set of offers that range from 0 to  $2^{(N_H + N_L)}$ . The definition of the equilibrium follows.

**Definition.** An equilibrium in the college market consists of solutions to: (1) the  $2^{(N_H + N_L)}$  value functions for the application portfolio of each student type,  $\{V^A\}_{\mathbb{P}(S)}$ ; (2) the value functions with offers of admission for each student type given her chosen application portfolio,  $\{V^O\}_{\mathbb{P}(O)}$ ; (3) the enrollment probabilities in college  $s \in S$  of each student type,  $q_s(y, \ell, \mathbf{T}(y, \sigma), \underline{\sigma})$ ; (4) the value functions from attending a college  $s \in S$  for each student type and number of applications submitted,  $V^s$ , and the associated policy functions for savings and consumption,  $a'_s$  and  $c_s$ ; (5) the value functions of working without a college degree for each student type and number of applications submitted,  $V^W$ , and the associated policy functions for savings and consumption,  $a'_w$  and  $c_w$ ; (6) each college  $s \in S$  value added,  $Z_s$ , its associated admission standard  $\underline{\sigma}_s$ , tuition schedule for each student type  $T_s(y, \sigma)$ , instructional spending  $I_s$ , ability level of its students  $L_s$ , and capacity  $\kappa_s$ ; (8) the labor income tax rate  $\tau$ ; such that

1. Given the values of holding offers,  $\{V^O\}_{\mathbb{P}(O)}$ , and each college's admission standard,  $\underline{\sigma}_s$ , the values from applying to different sets of colleges,  $\{V^A\}_{\mathbb{P}(S)}$ , are given by the  $2^{(N_H + N_L)}$  problems that resemble (3.1).
2. Given the values of attending the different colleges  $s \in S$ ,  $V^s$ , the values of holding offers,  $\{V^O\}_{\mathbb{P}(O)}$  are given by decision problems like (3.3).
3. The probability of enrolling in a college a student applied to and got an offer from,  $q_s(y, \ell, \mathbf{T}(y, \sigma), \underline{\sigma})$ , satisfies equation (3.4).
4. Given the tax rate,  $\tau$ , the college value added,  $Z$ , tuition,  $T(y, \sigma)$ , grants,  $Gr(y)$  and  $P(y)$ , and wage rate,  $w$ , the value function from attending a college  $s \in S$ ,  $V_s$ , solves the problem (3.5), and  $a'_s$  and  $c_s$  are the associated policy functions for savings and consumption.
5. Given the tax rate,  $\tau$ , and wage rate,  $w$ , the value from working without a college degree,  $V^W$ , solves the problem (3.6), and  $a'_{w_n}$  and  $c_{w_n}$  are the associated policy functions for savings and consumption.

6. Given students' enrollment probabilities,  $q_s(y, \ell, \mathbf{T}(y, \sigma), \underline{\sigma})$ , admission standards and tuition schedules of all the competing colleges,  $\underline{\sigma}$  and  $\mathbf{T}(y, \sigma)$ , the college value added  $Z_s$  solves the problem (3.12) and  $\sigma_s, T_s(y, \sigma), I_s, L_s$ , and  $\kappa_s$  are the associated policies for the admission standard, tuition schedule for each student type, instructional spending, ability, and capacity of the college.
7. The government balances its budget according to equation (3.14).

### 3.5. Characterizing the equilibrium

To reduce the dimensionality of the problem, we focus on a symmetric equilibrium within each college tier. This assumption simplifies the students' discrete choice problems, which is where the key bottlenecks are when solving for the equilibrium.

**Applicant's problem when submitting college applications.** The dimensionality of the applicant's decision problem in (3.2) simplifies as colleges within each tier are now replicas of each other. As a result, the number of possible application portfolios is now smaller. Each student has now  $(N_H + 1)(N_L + 1)$  possible portfolios, including not applying to college.<sup>20</sup>

We can now formulate the optimal application portfolio (decision problem (3.2)) as

$$\max \underbrace{\{V^A(y, \ell, 1, 0), V^A(y, \ell, 0, 1), V^A(y, \ell, 1, 1), \dots, V^A(y, \ell, N_H, N_L), V^W(y, \ell, 0)\}}_{\text{Expected values of the } (N_H + 1)(N_L + 1) \text{ application portfolios}},$$

where  $V^A(y, \ell, n_H, n_L)$  is the expected value of applying to  $n_H$  highly-selective colleges and  $n_L$  less-selective colleges.

We need to specify three cases for the expected value of applying: when the student applies to both highly-selective and less-selective colleges, when the students only applies to highly-selective colleges, and when the student only applies to less-selective colleges. Let  $\eta_H$  be an indicator function for whether the student got at least one admission offer from a highly-selective college (i.e.  $\eta_H = 1$ ) and  $\eta_L$  for whether the student got at least one admission offer from a less-selective college (i.e.  $\eta_L = 1$ ).<sup>21</sup> The expected value of applying to  $n_H > 0$  highly-selective colleges and  $n_L > 0$  less-selective colleges is then given by

$$\begin{aligned} V^A(y, \ell, n_H, n_L) = & \underbrace{\int_{\underline{\sigma}_H}^{\infty} V^O(y, \ell, n_H + n_L, 1, 1; \mathbf{T}(y, \sigma))g(\sigma|\ell)d\sigma}_{\text{Expected value of being admitted to a highly-selective college}} + \\ & \underbrace{\int_{\underline{\sigma}_L}^{\underline{\sigma}_H} V^O(y, \ell, n_H + n_L, 0, 1; \mathbf{T}(y, \sigma))g(\sigma|\ell)d\sigma}_{\text{Expected value of being admitted to a less-selective college only}} + \\ & \underbrace{G(\underline{\sigma}_L|\ell) V^W(y, \ell, n_H + n_L)}_{\text{Expected value of not being admitted to any college}} - \underbrace{\phi(n_H + n_L, 1, 1)}_{\text{Psychic application cost}}. \end{aligned}$$

<sup>20</sup>Using combinatorics, the number of possible application portfolios when there are  $N_H$  highly-selective colleges and  $N_L$  less-selective colleges is given by  $\left[ \sum_{k=0}^1 C(N_H, k) \right] \left[ \sum_{k=0}^1 C(N_L, k) \right] = (N_H + 1)(N_L + 1)$ .

<sup>21</sup>Recall that as highly-selective colleges are by definition more selective than less-selective colleges, if  $\eta_H = 1$ , then  $\eta_L = 1$ . The reverse is not true as a student can be accepted by a less-selective college but rejected by a highly-selective college.

If the student only applies to one type of college, say to  $n_L > 0$  less-selective colleges, the student's expected value of applying reads

$$V^A(y, \ell, 0, n_L) = \underbrace{\int_{\sigma_L}^{\infty} V^O(y, \ell, n_L, 0, 1; \mathbf{T}(y, \sigma)) g(\sigma|\ell) d\sigma}_{\text{Expected value from being admitted to a less-selective college}} + \underbrace{G(\sigma_L|\ell) V^W(y, \ell, n_L)}_{\text{Expected value from not being admitted to any college}} - \underbrace{\phi(n_L, 0, 1)}_{\text{Psychic application cost}}.$$

The value of applying to only highly-selective colleges can be written in a similar fashion.

**Applicant's problem with offers of admission.** The applicant's problem with offers of admission after submitting  $(n_H + n_L)$  applications (problem (3.3)) can now be written as

$$V^O(y, \ell, n_H + n_L, \eta_H, \eta_L; \mathbf{T}(y, \sigma)) = \max \left\{ \underbrace{\eta_H V^H(y, \ell, n_H + n_L, T_H(y, \sigma))}_{\text{Value from attending a highly-selective college after submitting } n_H + n_L \text{ applications}}, \underbrace{\eta_L V^L(y, \ell, n_H + n_L, T_L(y, \sigma))}_{\text{Value from attending a less-selective college after submitting } n_H + n_L \text{ applications}}, \underbrace{V^W(y, \ell, n_H + n_L)}_{\text{Value from not attending colleges after submitting } n_H + n_L \text{ applications}} \right\}.$$

**Enrollment probability.** The probability of enrolling in a college (equation (3.4)) is the key object that summarizes the two key discrete choices discussed above. It is the product of the probability of a particular application portfolio and the probability of accepting that college's offer given the application portfolio and the offers received.

Start with the probability of submitting  $n_H$  applications to highly-selective colleges and  $n_L$  applications to less-selective colleges. For each student type with income  $y$  and ability  $\ell$ , we have

$$\Pr(\text{Applying to } n_H \text{ and } n_L) = e^{\lambda_a V^A(y, \ell, n_H, n_L)} / \left[ e^{\lambda_a V^A(y, \ell, 1, 0)} + e^{\lambda_a V^A(y, \ell, 0, 1)} + e^{\lambda_a V^A(y, \ell, 1, 1)} + \dots + e^{\lambda_a V^A(y, \ell, N_H, N_L)} + e^{\lambda_a V^W(y, \ell, 0)} \right].$$

Turn to the applicant's problem who submitted  $(n_H + n_L)$  college applications and received offers of admission (i.e.,  $\eta_H$  and  $\eta_L$  are known to the student). For each student type with income  $y$  and ability  $\ell$ , the probability of accepting an offer from a college of type  $s = \{H, L\}$  is given by

$$\Pr(\text{Accept } s \text{ offer} | \text{App. to } n_H + n_L, \eta_H, \eta_L) = e^{\lambda_c \eta_s V^s(y, \ell, n_H + n_L, T_s(y, \sigma))} /$$

$$\left[ e^{\lambda_c \eta_H V^H(y, \ell, n_H + n_L, T_H(y, \sigma))} \right. \\ \left. + e^{\lambda_c \eta_L V^L(y, \ell, n_H + n_L, T_L(y, \sigma))} \right. \\ \left. + e^{\lambda_c V^W(y, \ell, n_H + n_L)} \right].$$

Focus now on the object of interest, the enrollment probability. The total probability of enrolling in a highly-selective college is given by

$$q_H(y, \ell, \mathbf{T}(y, \sigma), \underline{\sigma}) = \underbrace{\frac{1}{N_H} \sum_{n_H=1}^{N_H} \sum_{n_L=1}^{N_L} \Pr(\text{Applying to } n_H \text{ and } n_L)}_{\substack{\text{Probability of applying to a highly-selective college} \\ \text{in addition to less-selective colleges}}} \times \\ \underbrace{\Pr(\text{Accept } H \text{ offer} | \text{App. to } n_H + n_L, \eta_H = \eta_L = 1)}_{\substack{\text{Probability of accepting a highly-selective college offer after applying} \\ \text{to both types of colleges and being accepted}}} \\ + \underbrace{\frac{1}{N_H} \sum_{n_H=1}^{N_H} \Pr(\text{Applying to } n_H \text{ only})}_{\substack{\text{Probability of only applying to a highly-selective college (i.e., } n_L = 0)}} \times \\ \underbrace{\Pr(\text{Accept } H \text{ offer} | \text{App. to } n_H \text{ only, } \eta_H = 1)}_{\substack{\text{Probability of accepting a highly-selective college offer after only applying} \\ \text{to highly-selective colleges and being accepted } (\eta_L = 0)}}.$$

Similarly, we can write the total probability of enrolling in a less-selective college as

$$q_L(y, \ell, \mathbf{T}(y, \sigma), \underline{\sigma}) = \underbrace{\frac{1}{N_L} \sum_{n_H=1}^{N_H} \sum_{n_L=1}^{N_L} \Pr(\text{Applying to } n_H \text{ and } n_L)}_{\substack{\text{Probability of applying to a less-selective college} \\ \text{in addition to highly-selective colleges}}} \times \\ \underbrace{\Pr(\text{Accept } L \text{ offer} | \text{App. to } n_H + n_L, \eta_H = 0, \eta_L = 1)}_{\substack{\text{Probability of accepting a less-selective college offer after applying} \\ \text{to both types of colleges and being accepted only to the less selective}}} \\ + \underbrace{\frac{1}{N_L} \sum_{n_L=1}^{N_L} \Pr(\text{Applying to } n_L \text{ only})}_{\substack{\text{Probability of only applying to a less-selective college (i.e., } n_H = 0)}} \times \\ \underbrace{\Pr(\text{Accept } L \text{ offer} | \text{App. to } n_L \text{ only, } \eta_L = 1)}_{\substack{\text{Probability of accepting a less-selective college offer after only applying} \\ \text{to less-selective colleges and being accepted } (\eta_H = 0)}}.$$

Armed with each student's enrollment probability we can now solve the college problem. We proceed by deriving the optimal tuition schedule and admission standards.

**Optimal tuition.** An interior solution for the optimal tuition level of a college  $s$  for students with observable characteristics  $(y, \sigma)$  and positive probability of being accepted ( $\sigma \geq \underline{\sigma}_s$ ) is equal to a markup over marginal costs according to

$$T_s(y, \sigma) = \underbrace{m_s(y, \sigma)}_{\text{Markup}} \times \underbrace{MC_s(y, \sigma, I, \kappa, L)}_{\text{Marginal cost}}. \quad (3.15)$$

Note the tuition schedule depends on the tuition schedules of the competing colleges as both markups and marginal costs depend on students' enrollment probabilities (the arguments  $T(y, \sigma)$  and  $\underline{\sigma}$  are omitted for clarity). The optimal tuition schedule is derived in Appendix D.1.

In turn, the markup of a college depends on the (posterior) price elasticity of demand of students with parental transfers  $y$  and perceived ability  $\sigma$ ,  $\varepsilon_s(y, \sigma)$ , according to

$$m_s(y, \sigma) = \frac{\varepsilon_s(y, \sigma)}{\varepsilon_s(y, \sigma) - 1},$$

where  $\varepsilon_s(y, \sigma) \equiv -\frac{\partial \int q_s(y, \ell, T(y, \sigma), \underline{\sigma}) \mu(y, d\ell, \sigma)}{\partial T_s(y, \sigma)} \frac{T_s(y, \sigma)}{\int q_s(y, \ell, T(y, \sigma), \underline{\sigma}) \mu(y, d\ell, \sigma)}$ . Since a student's enrollment probability depends on the unobservable ability  $\ell$ , the college needs to infer the possible values of  $\ell$  for each  $\sigma$ . Hence, the college's posterior demand will solely be function of parental income  $y$  and perceived ability  $\sigma$ .

The marginal cost of a college has two components. The first component is independent of a student's characteristics. It is the cost of accepting any additional student as it requires more instructional spending and reduces the average amount of endowment and transfers per student. The second component takes into account how the student affects the average ability within the college. The more a student increases the average ability in the college, the larger the discount she will get. Similarly, there is a penalty for a student that lowers the average ability in the college. Tuition therefore accounts for the externality caused by a peer with an expected ability different from the average. The marginal cost of schooling a student with characteristics  $(y, \sigma)$  is given by

$$MC_s(y, \sigma, I, \kappa, L) = \underbrace{-\frac{Q_{s\kappa}}{Q_{sI}} - (E_{s\kappa} - Tr_{s\kappa})}_{\text{Additional cost per student}} - \underbrace{\frac{Q_{sL}}{Q_{sI}} \widehat{\mathbb{E}}_s[\ell|y, \sigma]}_{\text{Posterior ability discount}},$$

where  $Q_{s\kappa} < 0$ ,  $Q_{sI}, Q_{sL} > 0$ , and  $E_{s\kappa}, Tr_{s\kappa} \geq 0$ . The expected posterior ability of a student is weighted by her demand elasticity according to  $\widehat{\mathbb{E}}_s[\ell|y, \sigma] \equiv \frac{\int \ell [\partial q_s(y, \ell, T(y, \sigma), \underline{\sigma}) / \partial T_s(y, \sigma)] \mu(y, d\ell, \sigma)}{\int [\partial q_s(y, \ell, T(y, \sigma), \underline{\sigma}) / \partial T_s(y, \sigma)] \mu(y, d\ell, \sigma)}$ .

Since higher ability levels yield higher average signals, the (posterior) ability discount will be increasing in  $\sigma$ . Note that the ability discount differs by income level. This has interesting implications. For example, suppose that at a given income level only high-ability students apply. Colleges understand that applicants from such an income level are more likely to be higher ability and will therefore offer them higher levels of financial aid. This novel mechanism helps explain why low-income students receive high levels of financial aid because only the highest ability among them apply.

**Optimal admission standard.** Lastly, the optimal admission standard of a college defines that the tuition revenue received from the lowest-signal students must be enough to compensate the college for the marginal cost they impose and the change they bring to the average ability. The

optimal admissions standard of college  $s$  satisfies the following

$$\underbrace{\frac{\int T_s(y, \underline{\sigma}_s) q_s(y, \ell) \mu(dy, d\ell, \underline{\sigma}_s)}{\int q_s(y, \ell) \mu(dy, d\ell, \underline{\sigma}_s)}}_{\text{Average revenue from lowest-signal students}} = \underbrace{\frac{Q_{s\kappa}}{Q_{sI}} - (E_{s\kappa} - Tr_{s\kappa})}_{\text{Additional cost per student}} - \underbrace{\frac{Q_{sL}}{Q_{sI}} \frac{\int \ell q_s(y, \ell) \mu(dy, d\ell, \underline{\sigma}_s)}{\int q_s(y, \ell) \mu(dy, d\ell, \underline{\sigma}_s)}}_{\text{Posterior average ability discount from lowest-signal students}}, \quad (3.16)$$

where the enrollment probabilities account for the price paid by the lowest-signal student. Its derivation is provided in Appendix D.2

The presence of the tuition cap,  $\bar{T}_s$ , guarantees that the college chooses an interior value for  $\underline{\sigma}_s$ . As a result, students with very low perceived ability will not be able to compensate the college for lowering its average ability. The college will then have to raise its admission standard to exclude these students. This can be seen by examining (3.16). A low tuition cap will lower the left hand side by decreasing the average tuition revenue. As the left hand side is now smaller, the college will need to increase  $\underline{\sigma}_s$  to lower the right hand side for the optimality condition to hold. Increasing  $\underline{\sigma}_s$  lowers the right hand side because it raises the average ability of the college's lowest-signal students.

Proposition 3.1 below formalizes this ideas by demonstrating that when there are no caps on tuition, equation (3.16) will not hold with equality since the college will be able to charge the lowest-signal students enough to compensate for its marginal cost and the change they bring to the average ability of the student body.

**PROPOSITION 3.1:** Suppose  $\bar{T}_s$  is large enough so that the constraint  $T_s(\bar{y}, \underline{\sigma}) \leq \bar{T}_s$  never binds. Then  $\underline{\sigma}_s = 0$ . That is, if there are no tuition caps, the college does not exclude students using an admission threshold.

**PROOF:** See Appendix D.3

*Q.E.D.*

**Equilibrium selection.** The presence of peer effects introduces the potential for multiple equilibria. We follow [Epple, Romano, and Sieg \(2006\)](#) and [Epple, Romano, Sarpca, and Sieg \(2017\)](#) in focusing on an equilibrium where the ranking of college quality corresponds to the ranking of the endowment size. In our baseline equilibrium and subsequent analyses, highly-selective colleges have the higher value added and are therefore able to attract the higher-ability students. This is consistent with its higher endowment, which allows for higher spending per student. Note that our numerical procedure used to compute the equilibrium converges consistently to the same outcome for different initial guesses and is robust to small changes in the parameter space. Appendix G provides details on how the equilibrium of the college market is solved numerically.

#### 4. TAKING THE MODEL TO THE DATA

This section discusses the strategy developed to confront the model with the data presented in Section 2. A subset of parameters are standard in the literature or taken directly from their data counterparts. The remaining parameters are estimated using the simulated method of moments. Table 4.2 summarizes the parameters chosen outside the model, while Table 4.5 summarizes the resulting estimated parameters.

#### 4.1. College tiers

Barron's 2015 rankings of U.S. colleges is used to distinguish the two college tiers in the model. This ranking is commonly used in the literature as a measure of college quality. The highly-selective colleges correspond to Barron's Tier 1 and 2 schools ("elite" and "highly selective"), and the less-selective colleges correspond to all other four-year colleges and universities (excluding for-profit colleges). Table 4.1 summarizes the key empirical differences across these college types using the IPEDS data. The highly-selective colleges enroll a smaller share of the total student population, spend more per student, charge higher tuition, and have higher SAT scores. They, however, offer low tuition to students at the bottom of the income distribution. While highly-selective colleges are more likely to be private, their overall enrollment mostly consists of students in public colleges.

TABLE 4.1  
EMPIRICAL DIFFERENCES ACROSS COLLEGE TIERS

College Tier	H	L
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

*Note:* Statistics are enrollment-weighted averages. *Source:* IPEDS, College Scorecard.

#### 4.2. Functional forms

We describe the functional forms assumed for individuals' preferences, need-based aid, and colleges' technology and budget constraint.

**Preferences.** Individuals have logarithmic preferences over consumption given by

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

College students also face non-pecuniary costs of completing college that are linear in ability and given by

$$\nu_s(\ell) = \nu_{0_s} - \nu_{1_s}\ell,$$

where  $\{\nu_{0_s}, \nu_{1_s}\}$  vary across college tiers  $s = \{H, L\}$ .

**Application costs.** The application process involves two types of costs: a financial cost and a non-pecuniary psychic cost. We consider the financial cost of applying to be the same regardless of the number of applications sent and given by

$$\psi(n) = \psi_0.$$

Two reasons justify this assumption. First, several students get their applications fees waived.<sup>22</sup> Second, it simplifies the problem of students in (3.5) and (3.6) as it removes the need for the number of applications as a state variable.

Similarly, we consider the psychic cost to be independent of the number of applications as essays and other application materials can be reused for other colleges. We do, however, allow the psychic costs to differ by college tier as highly-selective colleges might demand more elaborated essays than less-selective colleges. The psychic cost is given by the following expression

$$\phi(n, H, L) = \begin{cases} \phi_H & \text{if } H = 1 \text{ and } L = 0 \\ \phi_L & \text{if } H = 0 \text{ and } L = 1 \\ \phi_B & \text{if } H = 1 \text{ and } L = 1 \end{cases}$$

**Need-based aid.** The endogenous need-based grants refer to Pell grants. In order to qualify for Pell grants, 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 set by the Department of Education. In the model, Pell grants are given by

$$P(y) = \max\{P_{max} - y, 0\},$$

so that all low-income students are guaranteed a minimum level of transfers. The exogenous private grants and scholarships,  $Gr(y)$ , take the following form

$$Gr(y) = \max\{g_0 - g_1 y, \underline{g}\}.$$

**College technology.** College quality is defined by a Cobb-Douglas function of the average instructional spending per student and the average ability of the student body according to

$$Q(I/\kappa, L/\kappa) = (1/\kappa)I^{1-\rho_L}L^{\rho_L},$$

where  $\rho_L$  is common across both colleges tiers.

**College budget constraint.** The budget constraint differs across college tiers in terms of fixed costs, endowment, and government transfers. College endowment net of operating costs is a quadratic polynomial in enrollment according to

$$E_s(\kappa) = E_{1_s}\kappa + E_{2_s}\kappa^2$$

and the government is assumed to subsidize colleges on a per-student basis according to the tier-specific function

$$Tr_s(\kappa) = Tr_{1_s}\kappa,$$

for  $s = \{H, L\}$ . We drop the constant terms as these would be subsumed in the fixed costs.

**Signal distribution.** Ability signals,  $\sigma$ , follow a normal distribution conditional on ability, with mean  $\ell$  and variance  $\sigma_g^2$ . The distribution is truncated at 0, so that the lower bound of the support of signals is finite.

---

<sup>22</sup>Common App, a platform used to submit college applications, grants fee waivers for students who meet certain criteria (e.g., student participates in the federal free or reduced price lunch program). Currently, almost half of Common App colleges do not charge an application fee.

### 4.3. Data

Three datasets are used to estimate the model: (i) the High School Longitudinal Study of 2009 (HSLs); (ii) the 2012 cohort of the Beginning Postsecondary Students Longitudinal Study (BPS); and (iii) the Integrated Postsecondary Education Data System (IPEDS).

The HSLs consists of a nationally representative sample of high-school students who are tracked over time as they transition from high school to college. A detailed description of the dataset is provided in Section 2.1. The HSLs is used to calculate key moments regarding student application and enrollment rates that are then matched in the estimation. The HSLs is also used to estimate the parameters governing the distribution of students characteristics.

The BPS is a longitudinal survey that follows a representative cohort of students over time starting with their first year of post-secondary studies. The dataset includes details of all financial aid received by the 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 complements the HSLs, which only has self-reported financial-aid data not broken down by source (private institutional aid vs. government grants). The BPS is used to set the level of non-institutional grants available to students, to determine how grants and financial aid vary with SAT scores and parental income, and to compute colleges' tuition caps.

The IPEDS is a dataset that contains rich information on college-level financial variables. The data is used to estimate college-level aggregates as it contains details about college revenue from all sources, including tuition, government appropriation, and endowment income, as well as all costs and other operating expenses. Appendix B.3 provides details on how the sample of colleges is constructed.

### 4.4. Externally estimated

The set of parameters that are either estimated externally or borrowed from the literature are related to the distribution of students characteristics,  $\hat{\mu}(y, \ell)$ , discount factor, returns to ability, and dropout risk,  $\{\beta, \alpha_c, \alpha_w, \delta_H, \delta_L\}$ , borrowing constraints, grants, and financial application costs,  $\{\underline{a}_c, \underline{a}_w, P_{max}, g_0, g_1, g, \psi_0\}$ , and colleges budget constraint  $\{C_H, C_L, E_{1L}, E_{2H}, E_{2L}, Tr_{1H}, Tr_{1L}, \bar{T}_H, \bar{T}_L\}$ . Aggregate prices,  $\{R, w\}$ , are taken exogenously.

**Distribution of characteristics.** The distribution of student characteristics over income and ability,  $\hat{\mu}(y, \ell)$ , is assumed to be lognormal and informed by the HSLs data. Parental income,  $y$ , corresponds to the student's Expected Family Contribution (EFC), which is a measure of the amount of resources a student has in order to attend college (before financial aid) and so maps well into the notion of parental transfers used in the model.<sup>23</sup> In the HSLs, the EFC is available for all students who completed the FAFSA and attended college. For those who did not, the EFC is calculated using the household income reported by parents. The calculation follows the EFC formula established by the Department of Education, described in detail in Appendix C.1. Students ability,  $\ell$ , corresponds to high-school grades. Appendix C.2 provides further details on how the HSLs is used to fit the distribution of student characteristics. Since students in the HSLs began college in the 2013-2014 academic year, 2013 dollars is used as the numeraire. As the first period in the model corresponds to four years, the amount is rescaled by \$40,000.<sup>24</sup>

<sup>23</sup>The EFC is determined according to rules set by the Department of Education using the FAFSA filled out by students and their parents.

<sup>24</sup>For example,  $y = 1$  in the model corresponds to an EFC of \$10,000 per year over four years.

**Returns to ability.** We take the parameters governing the labor market returns to ability,  $\alpha_c$  and  $\alpha_w$ , directly from [Abbott, Gallipoli, Meghir, and Violante \(2019\)](#). 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. We take the midpoint between males and females for each case. We thus set  $\alpha_c = 0.782$  for college graduates (this value is the same across college tiers) and  $\alpha_w = 0.559$  for high-school-educated individuals.

**Dropout risk.** We set the probability of dropping out of college to zero in both college tiers,  $\delta_H = \delta_L = 0$  (work in progress).

**Borrowing constraints.** According to the Federal Student Loan Program, the aggregate loan limit for dependent students who are completing an undergraduate degree is \$31,000 in federal loans. We therefore set the student borrowing limit to  $\underline{a}_c = -0.775$ , which corresponds to \$31,000 over four years. Since there is no risk for workers who did not pursue a college education, the borrowing constraint they face is not binding. Hence, the borrowing limit is set at  $\underline{a}_w = 0$ .

**Need-based grants.** In 2013 the maximum Pell Grant amount was \$5,645 per year, which corresponds to setting  $P_{max} = 0.5645$ . Using the BPS, we set the minimum level of private grants to  $g = 0.15$ , i.e. about \$1,500 per year, and estimate  $g_0 = 0.43$ ,  $g_1 = 0.19$ . for the exogenous grants. The details are provided in [Appendix C.3](#).

**Financial application cost.** The average application in IPEDS is \$50 and students covered in the HSLs send on average about three applications. Since these costs are incurred only once in the four-year period, the financial application cost is set to  $\psi_0 = 0.00375$  in the model.

**Colleges budget constraint.** Colleges fixed costs and endowments are estimated using IPEDS data. We find that the highly-selective college have lower fixed costs than less-selective colleges with  $C_H = 0.15$  and  $C_L = 0.25$ . The linear term in the endowment function of less-selective colleges is estimated to be  $E_{1L} = 0.06$ . The quadratic terms are  $E_{2H} = 5.77$  for highly-selective colleges and  $E_{1L} = 0.44$  for less-selective colleges. The linear term for highly-selective colleges,  $E_{1H}$ , is estimated to match their observed enrollment (discussed in the estimation procedure below).

We estimate government transfers across college tiers using the IPEDS sample by relying on variation in enrollment. The resulting estimates imply that highly-selective colleges receive more government subsidies per enrolled student than less-selective colleges (\$10,000 vs. \$6,200) and therefore set  $Tr_{1H} = 1.0$  and  $Tr_{1L} = 0.62$ . The details of the estimation are provided in [Appendix C.4](#).

BPS data is used to find the tuition caps  $\bar{T}$ . This data has the advantage of including details about all financial aid received by each student. In the data the average tuition minus college-specific financial aid flattens for high-EFC students, as illustrated in [Figure C.1](#). This reflects the fact that high-income students generally do not receive financial aid, as institutional grants are more likely to be given to low-income students. Consistent with the data, tuition caps in the model bind for students with the highest parental transfers. Tuition caps correspond to the average upper bound on tuition paid by high-income students within each college tier (i.e., \$25,000 and \$12,000 per year, respectively), which imply values of  $\bar{T}_H = 2.5$  and  $\bar{T}_L = 1.2$ .

**Aggregate prices.** The two periods in the model are divided as follows: the first period accounts for 4 years (time spent in college) and the second period accounts for 60 years (time spent working). [Appendix E](#) shows how a life-cycle model with  $T + 1$  periods maps into our two-period model and what this means for the aggregate prices,  $\{R, w\}$ . The gross interest rate is set

to  $R = (1.0386)^4$ , which corresponds to the annual borrowing rate for undergraduate students (3.86%) in 2013-2014 academic year. This rate was set by the US Department of Education (USDE) for the Federal Student Loan Program. The discount factor in the life-cycle model is set to  $\beta = (1/R)^4$ , which corresponds to a value of  $\beta = 5.48$ . The wage rate is set to  $w = 2.7$ , so that the model produces the average wage calculated from the Current Population Survey (CPS).

TABLE 4.2  
EXTERNALLY ESTIMATED PARAMETERS

Variable	Description	Value	Source
<i>Preferences</i>			
$\beta$	Discount factor	5.48	USDE
<i>Returns to ability and dropout risk</i>			
$\alpha_c$	Returns to ability (college degree)	0.78	Abbott et al. (2019)
$\alpha_w$	Returns to ability (high-school degree)	0.55	Abbott et al. (2019)
<i>Borrowing constraints, grants and application costs</i>			
$\underline{a}_c$	Student borrowing limit	-0.775	USDE
$\underline{a}_w$	Worker borrowing limit	0	Normalization
$P_{max}$	Pell Grant maximum	0.5645	USDE
$g_0$	Intercept of grant function	0.43	BPS
$g_1$	Slope of grant function	0.19	BPS
$\underline{g}$	Minimum grant level	0.15	BPS
$\psi_0$	Financial application cost	0.00375	IPEDS
<i>College costs and revenues</i>			
$C_H, C_L$	Fixed costs in colleges H, L	0.15, 0.25	IPEDS
$E_{1L}$	Net endowment per student in colleges L	0.06	IPEDS
$E_{2H}, E_{2L}$	Quadratic term in net endowment in colleges H, L	5.77, 0.43	IPEDS
$Tr_{1H}, Tr_{1L}$	Gov transfers per student in colleges H, L	1.00, 0.62	IPEDS
$\bar{T}_H, \bar{T}_L$	Tuition cap in colleges H, L	2.50, 1.20	BPS

#### 4.5. Internally estimated

The remaining parameters relate to the disutility of applying to colleges,  $\{\phi_H, \phi_L, \phi_B\}$ , the disutility of attending college,  $\{\nu_{0H}, \nu_{0L}, \nu_{1H}, \nu_{1L}\}$ , the extreme value scale parameters for applying and enrolling in colleges,  $\{\lambda_a, \lambda_c\}$ , the variance of colleges' perceived student ability, college productivity, quality elasticity, and endowment income,  $\{\sigma_g^2, \xi_H, \xi_L, \rho_L, E_{1H}\}$ . These are jointly estimated by minimizing the unweighted distance between data moments and simulated model moments. The estimated parameters are presented in Table 4.5 and Table 4.6 displays the results of the estimation. The details of the estimation procedure are presented in Section C.6.

**Application disutility costs.** The disutility of applying to colleges,  $\phi_H, \phi_L$ , and  $\phi_B$ , are identified by the fraction of students applying to each college tier only and the fraction of those applying to both college types. In the HSLs, only 2% of high-school graduates apply only to highly-selective colleges, 42% apply only to less-selective colleges, and 14% apply to both highly-selective and less-selective colleges. Note that the estimated cost of applying to highly-selective colleges is lower than the cost of applying to less-selective colleges. The reason is that it is very risky to only apply to highly-selective colleges in the model. Hence, a relatively low application cost is necessary to match the correct fraction of students who would choose only to apply there.

To assess the magnitude of these costs, we calculate the equivalent consumption a student would forgo in order to remove the disutility cost from applying to college. Table 4.3 reports the average consumption equivalent values as a percentage of average life-cycle consumption for all students. The estimated costs are considerably higher than those found in the literature, which is largely due to the high marginal utility of consumption among high-income students and those who choose not to enroll. If we restrict only to lower-income students enrolled in college who have lower consumption due to credit constraints, we see that their application disutility costs are substantially lower.

TABLE 4.3  
AVERAGE APPLICATION DISUTILITY COSTS

	Overall average	Enrollee average (Bottom 50% income)	Fu (2014)
Applying to highly-selective colleges only	0.78%	0.52%	0.30%
Applying to less-selective colleges only	1.19%	0.58%	0.34%
Applying to both highly-selective and less-selective colleges	1.87%	0.89%	0.45%

*Note:* Expressed as a percentage of average life-cycle consumption (calculated in the model to be about \$900,000). The application costs calculated in Fu (2014) are provided for reference to the literature. The costs from Fu (2014) are calculated assuming two applications for highly-selective colleges, three for less-selective colleges, and six for both (which we observed among average applicants in the HSLs).

**Attendance disutility costs.** The psychic disutility cost of attending college varies by college tier and depends linearly on ability. Student enrollment decisions conditional on being accepted helps identify the intercept parameters ( $\nu_{0_H}$  and  $\nu_{0_L}$ ), while variation in student applications across the SAT distribution helps identify the slope parameters ( $\nu_{1_H}$  and  $\nu_{1_L}$ ). The estimated coefficients reveal substantial differences in preferences for each college across the ability distribution. Highly-selective colleges will be less costly to attend than less-selective colleges for high-ability students, while the reverse is true for low-ability students.

To get a sense of the magnitude of these psychic costs of attending college, we again calculate the consumption equivalence for each student and report them as a percentage of average life-cycle consumption in Table 4.4. Psychic costs are smaller than those estimated in the literature (e.g. Abbott, Gallipoli, Meghir, and Violante (2019)), even when the application disutility costs are added. The reason is that the model accounts for the fact that students need to apply and be admitted to college in order to attend. The model can thus help rationalize the relatively high psychic costs of schooling found in the literature: many students choose not to enroll 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 many students from the college market.

**Extreme value shock.** The scale parameter of the extreme value shock when deciding between offers of admission,  $\lambda_c$ , guides the option value of adding other colleges to choose from. If the scale parameter is low, then a large fraction of students applying only to less-selective colleges will be unlikely to enroll. The reason is that their applications were motivated more by the increased option value rather than the value-added from attending less-selective colleges. Thus, variation in attendance in less-selective colleges conditional on only applying to less-selective colleges and being accepted helps identify the parameter value.

The scale parameter of the extreme value shock when applying to college,  $\lambda_a$ , influences how many applications students submit. Variation in the number of college applications submitted identifies the value of the scale parameter.

**Signal strength.** The variance of the ability signal,  $\sigma_g^2$ , is chosen to match the average responsiveness of tuition with respect to ability. As discussed in Section 3, the (posterior) ability

TABLE 4.4  
AVERAGE PSYCHIC COSTS

	Model $\nu_s(\ell)$	Abbott et al. (2019)
Highly-selective colleges	1.23%	10.7%
Less-selective colleges	2.76%	
Among enrollees	1.86%	4.6%

*Note:* Expressed as a percentage of average life-cycle consumption (calculated in the model to be about \$900,000). The costs calculated in [Abbott et al. \(2019\)](#) are provided for reference.

discount depends on the strength of the signal, so that the average responsiveness of tuition to the signal is sensitive to the choice of its variance. This effect is estimated using the BPS data by regressing tuition (minus institutional grants) on SAT score, controlling for EFC and college fixed effects.

**Efficiency.** Colleges' efficiency parameter  $\xi_s$  is chosen to match the college wage premium in the data. [Abbott, Gallipoli, Meghir, and Violante \(2019\)](#) estimate an average college wage premium of 0.6. Since 17% of students are enrolled in selective colleges, we have that the average wage premium is such that

$$0.17 \log Z_H + 0.83 \log Z_L = 0.6.$$

Estimates by [Chetty, Friedman, Saez, Turner, and Yagan \(2020\)](#) report that 80% of the difference in median log earnings 10 years after college can be explained by differences in colleges' selectivity, which helps to pin down the difference between colleges' value added. Using the numbers from [Table 4.1](#), we have

$$\log Z_H - \log Z_L = 0.113.$$

The efficiency parameters,  $\xi_H$  and  $\xi_L$ , are solved to ensure these two equations hold.

**Quality elasticity.** The elasticity of college quality with respect to students' ability,  $\rho_L$ , is chosen to match the average acceptance rate in highly-selective colleges, conditional on having applied. A higher value of  $\rho_L$  increases the size of the ability discount, which increases the attractiveness of applying to highly-selective colleges. All else equal, this will reduce acceptance rates due to the increase in applications with little change on enrollment.

**Endowment.** The linear parameter in the endowment function of highly-selective colleges,  $E_{1H}$ , is estimated to match the enrollment rate in highly-selective colleges observed in the data. For less-selective colleges, the empirical enrollment level is too large to only be explained by endowment and is therefore left untargeted.

TABLE 4.5

INTERNALLY ESTIMATED PARAMETERS			
Variable	Description	Value	Identification
<i>Preferences</i>			
$\phi_H, \phi_L$	Disutility of applying to colleges H, L	0.100, 0.185	Application rates
$\phi_B$	Disutility of applying to both college tiers	0.320	Application rates
$\nu_{0H}, \nu_{0L}$	Disutility of attending colleges H, L (intercept)	2.410, 1.700	Attendance rates
$\nu_{1H}, \nu_{1L}$	Disutility of attending colleges H, L (slope)	-1.467, -0.758	Attendance rates by SAT
$\lambda_a$	Extreme value scale parameter when applying	0.400	Number of applications
$\lambda_c$	Extreme value scale parameter when choosing offers	1.855	Attendance rates
<i>Technology and costs</i>			
$\sigma_g^2$	Variance of ability signal	0.110	Elasticity of tuition wrt SAT
$\xi_H, \xi_L$	Value added efficiency in colleges H, L		College wage premium
$\rho_L$	Quality elasticity of ability	0.854	Acceptance rate
$E_{1H}$	Net endowment per student in colleges H	1.371	Distribution of enrolled students

TABLE 4.6

TARGETED MOMENTS: DATA VS. MODEL

Moment	Model	Data
% Applying to highly-selective colleges	2.03	2.00
% Applying to less-selective colleges	46.1	42.0
% Applying to both highly-selective and less-selective	15.1	14.0
Highly-selective colleges rel app rate (SAT quintile 5/4)	3.43	3.33
Less-selective colleges rel app rate (SAT quintile 4/2)	1.63	1.55
Highly-selective colleges atnd rate (accepted in both tiers)	0.55	0.58
Less-selective colleges atnd rate (app. to L only)	0.79	0.83
Less-selective colleges atnd rate (app. to H and L)	0.94	0.90
$\Delta$ Tuition wrt SAT	-0.08	-0.08
Highly-selective colleges acceptance rate	0.69	0.70
% Enrolled in highly-selective colleges	7.15	7.00

*Note:* All moments were calculated using the HSLS, except the change in tuition with respect to SAT, which is calculated using the BPS. "Highly-selective colleges rel app rate (SAT quintile 5/4)" refers to the relative application rates to highly-selective colleges between students in the 5th vs 4th SAT quintile. For example, students in the 5th quintile are about 3.33 times more likely to apply to highly-selective colleges than students in the 4th SAT quintile. "Less-selective colleges rel app rate (SAT quintile 4/2)" is defined similarly. "Less-selective colleges atnd rate (app. to H and L)" refers to the less-selective colleges attendance rates for students who applied to both college tiers and were only admitted to less-selective colleges. " $\Delta$  Tuition wrt SAT" is determined by regressing net tuition on SAT in our BPS data (controlling for EFC and college-fixed effects).

## 5. RESULTS

### 5.1. Model validation

**College-level statistics.** Table 5.1 below compares college-level statistics produced by the model to the ones calculated in the data. The model captures the fact that EFC in the highly-selective colleges is about twice as high as in the less-selective colleges, though somewhat underestimates the average EFC within each college. The model does a good job of allocating relatively high signal students into highly-selective colleges and students with average signals into less-selective colleges. The model is able to capture the large difference in instructional spending per student across each college. The amount spent per student in highly-selective colleges is large due to its higher tuition 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 endowment income are used to cover remaining college expenses. On the other hand, the average net tuition at each college is slightly higher in the model relative to the data.

TABLE 5.1  
COLLEGE-LEVEL STATISTICS

	Colleges H		Colleges L	
	Model	Data	Model	Data
% Enrolled	7.15	7.00	32.09	36.0
Average Student EFC*	2.22	3.04	1.14	1.55
Average Applicant Signal/SAT	1.45	1.12	0.29	0.05
Instr Spending per Student*	2.10	2.25	0.98	0.89
Average Net Tuition*	1.60	1.41	0.61	0.50
<i>Income Distribution</i>				
Q1 Income	9.8%	9.6%	22.1%	23.4%
Q2 Income	11.6%	12.1%	20.5%	16.4%
Q3 Income	26.4%	23.1%	30.6%	28.1%
Q4 Income	52.2%	55.2%	26.8%	32.2%

*Note:* All statistics are untargeted (except enrollment in highly-selective colleges). Income refers to parental transfers in the model and Expected Family Contribution (EFC) in the data. The average signal (in the model) and SAT score (in the data) among applicants are standardized for comparison. Average Net Tuition refers to the sticker price minus college-specific financial aid and does not include outside grants. \*\$10k/yr.

**Student distribution within colleges.** Table 5.1 shows the distribution of parental transfers (EFC in the data) within in each college. The model is able to deliver student income distributions within each college that are consistent with the data. In particular, the model remarkably captures the correct share of low-income students in highly-selective colleges. In this type of college-market models, colleges have a strong incentive to 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, overpredict 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.

The application and admissions system helps 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, a relatively large value for  $\rho_L$  is needed, which governs a 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 selective college can be confident that enrolling low-income students will help increase its average ability. As a result, the college is willing to 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 next section.

**Tuition by income level.** Figure 5.1 shows the average net tuition in the model and in the data by EFC decile within each college. In the data, net tuition is defined as the sticker-price tuition minus all grants available to the student. In the model, net tuition for a student with parental transfers  $y$  and signal  $\sigma$  corresponds to  $T(y, \sigma) - P(y) - Gr(y)$ . The tuition caps help capture the tuition for students at the very top of the EFC distribution, especially in less-selective colleges. The model is also successful at predicting both the level and change in net tuition across the parental income distribution. One reason net tuition for low-income students is high relative to the data is that nearly all students in the model borrow the full amount of

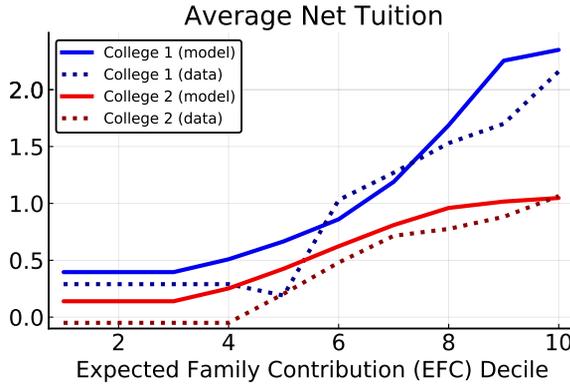


FIGURE 5.1.—Average net-tuition (sticker-price tuition minus all grants) for students by Expected Family Contribution decile. Source: BPS.

student loans available to them, allowing the colleges to charge relatively high tuition. Students borrow the full amount because they have high second period earnings and there is no income risk, so they try to equalize consumption across both periods.

**Student application and enrollment.** The model is able to account for the low application and enrollment rates of low-income students despite the presence of substantial financial aid. Figure 5.2 shows the model predicted average application and enrollment rates alongside the ones observed in the HSLs data by EFC deciles. The model correctly predicts that applications to the more selective college are increasing in EFC (the fraction only applying to less-selective colleges is small and not reported). The model also does a good job of capturing the increasing enrollment by EFC at both colleges. Note that although the overall means of the application rates and enrollment rates in highly-selective colleges are targeted in the baseline estimation, the variation in applications by EFC and the enrollment rates in less-selective colleges are not targeted—which is a notable feature of the model.

### 5.2. Equilibrium effect of applications on low-income student enrollment

As discussed above, the informativeness of the students' signals plays an important role in this model. If, for example, only the highest ability among low-income 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 5.3 shows what happens to high-ability student enrollment when all low-income students apply in the same way as their higher-income peers.<sup>25</sup> Worsening the applicant pool of low-income students reduces the enrollment of high-ability, low-income students in highly-selective colleges by about a quarter. With the substantial increase in low-ability, low-income applicants, the signals of all low-income students are now much more likely to come from low-ability students. This will cause the college to lower the financial aid it offers to low-income admitted students, which in turn significantly reduces their enrollment. Since highly-selective colleges are now unable to enroll as many high-ability students, their value-added will

<sup>25</sup>Specifically, the application rates from the baseline equilibrium are adjusted by requiring that low-income students apply to both colleges at the same rate as the highest-income students. The equilibrium is then recomputed, holding fixed these new application pools (thus this is a partial equilibrium analysis since student application behavior remains fixed). Note, however, 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.

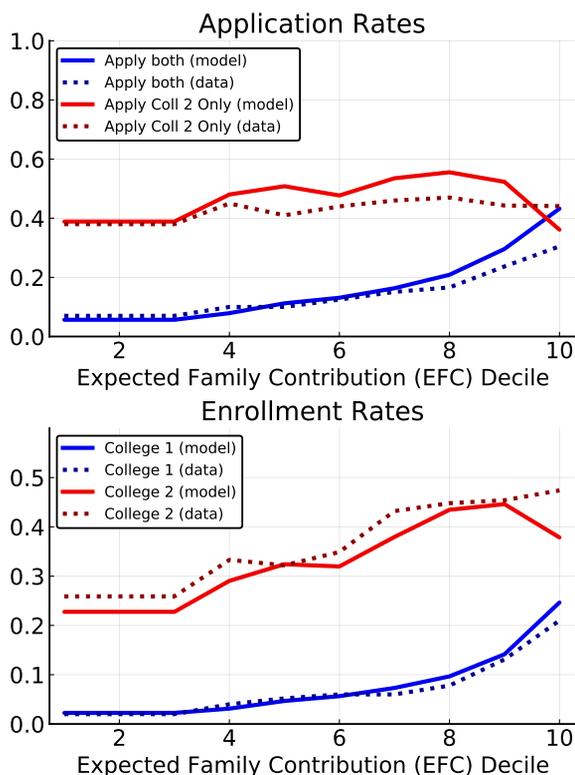


FIGURE 5.2.—The left panel displays the average application rates for students by Expected Family Contribution (EFC) decile, and the right panel displays the average enrollment rates to each college. In the left panel, we include only applications to less-selective colleges or applications to both since the fraction of students only applying to highly-selective colleges is small both in the model and the data. Source: HSLs.

also decline due to the lower average ability of its student body. This effect is not only present for low-income students, but also for middle-income students who enroll in less-selective colleges instead. A summary of the overall changes to the college market in the new equilibrium is provided in Table F.1.

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

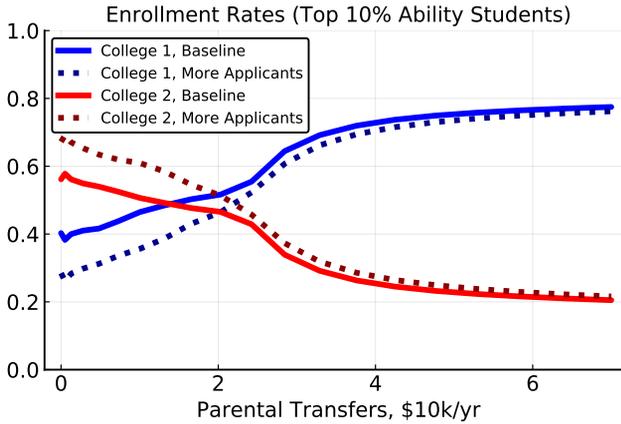


FIGURE 5.3.—Average enrollment rates for students in the top 10% of the ability distribution by parental transfers. The plot shows the effect on enrollment of switching to an equilibrium where low-income students apply to both colleges at the same rates as high-income students conditional on ability.

## 6. THE ROLE OF APPLICATION INFORMATIVENESS

This section studies the general equilibrium effects of changes to the signal’s variance. Motivated by concerns over the distortionary effect of the SAT (assuming its use increases application informativeness), the first counterfactual exercise examines the effect of increasing the signal variance. The second counterfactual exercise studies the effect of switching to perfectly informative signals by removing admissions uncertainty for students and allowing colleges to fully observe their ability.

### 6.1. Ending the SAT

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 has begun phasing out their reliance on the SAT for admissions. Consider the use of the SAT as part of the technology that increases the informativeness of the ability signals of applicants. Then, removing these tests can be interpreted as less informative signals in the model and thus an upper bound for the losses caused from removing the SAT. That is achieved by increasing the variance of the signal by a factor of six from the baseline estimation.<sup>26</sup>

Table F.2 shows the effect of increasing the signal variance on college-level variables. In the new equilibrium the higher variance leads to poorer sorting based on ability, 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. As a result, the fraction of students enrolled increases and the admissions standards decrease. While there is little change to the income distribution within each college, the ability distribution has become more diffuse, reflecting the increased admissions probability for low-ability students and decreased admissions probability for high-ability students.

<sup>26</sup>This value minimizes average welfare in comparison to the baseline equilibrium. Figure F.1 in the appendix plots the change in welfare for a range of possible increases in the signal variance.

To see which students are affected by the decrease in the informativeness of the signals, Figure 6.1 plots the percent of consumption in each period different types of students would be willing to forgo in order to be born in the new equilibrium.<sup>27</sup> As expected, high-ability students are the most hurt from switching to the new equilibrium since the higher signal variance decreases their likelihood of being admitted at the expense of the lower ability students. Importantly, this effect is strongest for the low-income, highest-ability students as they are not wealthy enough to compensate for the risk and highly-selective colleges now offer lower levels of financial aid to the highest signal students.

The students who benefit the most are the high-income, low-ability students, who can now more easily be admitted (solid lines in Figures 6.1 and F.2). Students at the very bottom of the ability distribution with high parental transfers benefit from less informative signals because they can now more easily enroll in less-selective colleges. Similarly, the highest-income students in the 30-60% ability range also benefit from being able to more easily sort into highly-selective colleges, while the lower-income students in the 30-60% ability range benefit from more easily sorting into less-selective colleges. Finally, there is very little change in welfare for low-ability, low-income students as the gains they experience from the increased chance of being admitted are offset by decreases in the financial aid they can expect to receive.

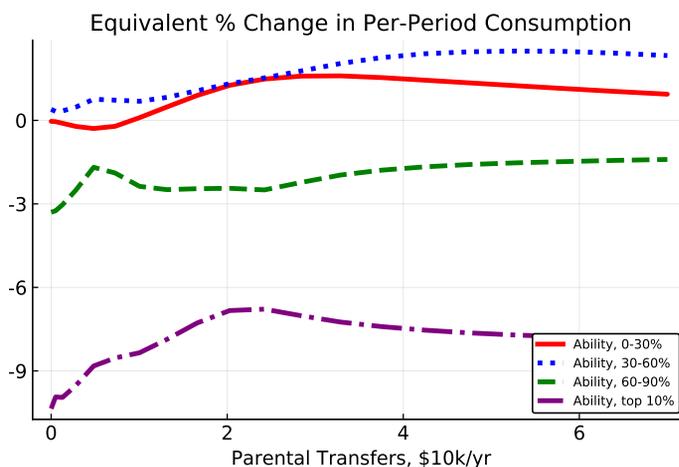


FIGURE 6.1.—Percent of lifetime consumption students of different parental-income and ability groups are willing to give up in order to switch to the equilibrium with less informative signals.

**When signals are correlated with income.** One concern over the use of the SAT is that low-income students tend to score lower than high-income students, suggesting that at the SAT is not as informative about low-income student ability. As a robustness check, we take this consideration into account by allowing the removal of the SAT to decrease the signal informativeness at the top of the parental transfer distribution more than the bottom.<sup>28</sup> The outcome of this experiment is presented in Table F.3 and the welfare effects are presented in Figure F.3.

<sup>27</sup>For reference, Figure F.2 in the Appendix plots the changes to student enrollment resulting from the new equilibrium.

<sup>28</sup>In this experiment, the factor increase in the signal variance is set to grow linearly in the student's parental transfer, with the lowest-income students experiencing their signal variance increase by a factor of 2 and the highest-income students by a factor of 6.

The detrimental effect on the colleges' value-added from making the signals less informative is still present, causing welfare losses for all high-ability students that are significantly larger than the modest gains among the low-ability students. As the signal variance increases with income, the high-income high-ability students are now the ones who are made worse off by the change.

## 6.2. Perfectly informative signals

To further study the effects of the information frictions due to the application and admissions system, the signal is set to be perfectly informative so that the colleges may observe the student's true innate ability (i.e.  $g(\sigma|\ell) = 1$  for  $\sigma = \ell$ ). In this counterfactual, students know ex-ante whether they will be admitted and exactly how much financial aid they will receive. This perfect information equilibrium is then compared to the baseline equilibrium.

Key college-level statistics of the new equilibrium are shown in Table F.4. If signals were perfectly informative, the effective 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 of the student body, thereby increasing the effective marginal cost for relatively low-ability students. As the effective marginal cost of enrolling high-income, low-ability students rises, colleges will increase their admissions threshold to exclude them 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 an overall decrease in enrollment. Additionally, the decline of tuition revenue decreases the average instructional spending per student. However, the increase in average student ability is more than enough to offset the decline in instructional spending, leading to increases in value-added at both colleges.

The effect of perfectly observable signals on student enrollment at each college is shown in Figure F.4 in the Appendix. 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 enrolling in less-selective colleges, from the 30-60% ability group who are replaced by low- and middle-income students from the 60-90% ability group. Finally, highly-selective colleges are now able to enroll students almost exclusively from the top 10% of the ability distribution, mostly to the benefit of high-income students who no longer have to compete with lower-ability applicants.

Perhaps surprisingly, perfect information actually *reduces* the enrollment of high-ability, low-income students in highly-selective colleges. This seems puzzling at first because tuition decreases at highly-selective colleges for high-ability students as they can now be perfectly sorted. The reason for the decrease in enrollment in highly-selective colleges is that less-selective colleges now lower their tuition enough to incentivize many low-income, high-ability students to switch. In the baseline, less-selective colleges could not offer high levels of financial aid because their pool of low-income applicants included many low-ability students, making it hard to sort out the high-ability low-income students. With perfect information, the high-ability low-income students cannot be mistaken for low-ability applicants, allowing less-selective colleges to increase their financial aid and attract them away from highly-selective colleges.

To understand the strength of college competition under perfect information, Figure 6.2 shows enrollment in highly-selective colleges for students in the top 10% of the ability distribution in partial equilibrium, where uncertainty disappears for students, but tuition and admission policies remain the same (the dotted line). In this scenario, the high-ability, low-income students enroll in highly-selective colleges at higher rates since they know they will be admitted at low tuition levels. However, when both colleges adjust their policies in response to signals becoming perfectly informative, the low-income students prefer to attend less-selective colleges where tuition is even lower.

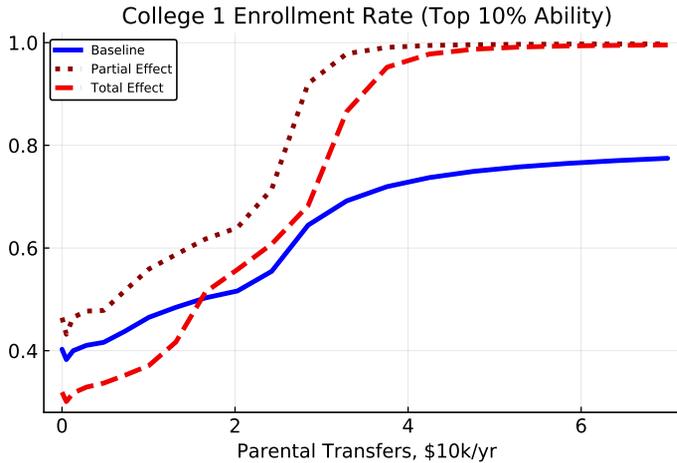


FIGURE 6.2.—Average enrollment rate in highly-selective colleges by parental transfers for students in the top 10% of the ability distribution. The ‘Partial Effect’ shows how enrollment changes when students have perfect information, but tuition and admissions are the same as in the baseline. The ‘Total Effect’ shows the perfect-information equilibrium where colleges adjust tuition and admissions.

## 7. 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 “American Families Plan”, which calls for the maximum to be raised to \$7,895 per year. In this section, we study the effects of increasing the Pell grant maximum to \$25,000 per year, which is equal to the tuition cap at highly-selective colleges. In order to qualify for a Pell grant, a student’s EFC must be lower than their net cost of attendance (tuition plus room and board minus financial aid). The Pell grant then covers this difference up to a maximum level. In the model, the Pell grant is equal to the difference between the set maximum and the student’s parental transfer (see equation (4.1)). In this experiment, the increase in Pell grants is paid for with taxes as implied by the government budget constraint in equation (3.14).

Table F.5 displays the effect of increasing the Pell grant maximum on college statistics. There is a large effect on the income distribution within highly-selective colleges. By increasing the funding available for lower-income students, the grants cause them to enroll in highly-selective colleges at higher rates, making the income distribution less concentrated at the top. Highly-selective colleges will then charge higher tuition to students with increased grant funding. This leads to an overall increase in tuition revenue and instructional spending per student. As high-ability, low-income students enroll at higher rates, the average ability of the student body in highly-selective colleges increases. Figure F.5 illustrates the effect of the higher Pell grants on student sorting, where enrollment is plotted against parental transfers for students at different parts of the ability distribution. For highly-selective colleges, the enrollment profiles in parental income flatten considerably conditional on ability. This is especially important for high-income students in the 60-90% group, who now enroll at highly-selective colleges at much lower rates.

Finally, the welfare effects of the policy change are examined in Figure 7.1. As expected, the students who benefit from the policy are the relatively low-income students. They benefit directly from the increased consumption while in college (which also alleviates the effect of the credit constraint) and from more easily being able to sort into the colleges. Higher-income students are made worse off from the policy due to both the higher taxes when graduating and to higher competition in enrollment as lower-income students can now attend college more

easily. Overall, the average welfare change of the policy in terms of per-period consumption equivalent units is 1.97%.

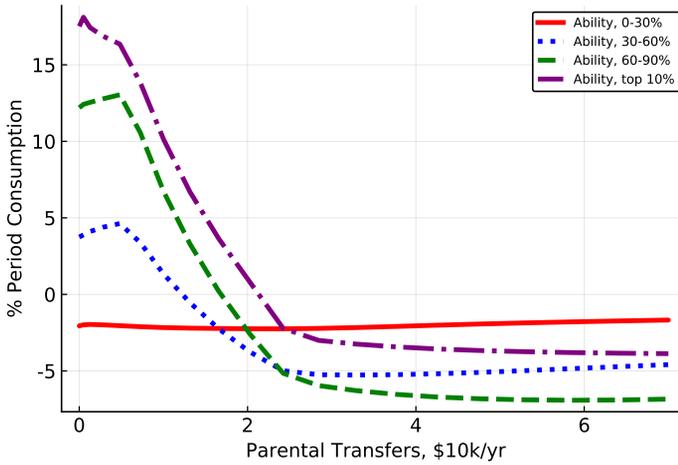


FIGURE 7.1.—Percent of lifetime consumption students of different EFC and ability are willing to give up to switch to the equilibrium with the higher Pell Grant maximum.

### 7.1. Decomposing the effect of the Pell grant increase

To isolate the role of the college market, the first experiment studies the effect of the Pell grant increase in the absence of changes to tuition or admissions. Next, to isolate the effect of the admissions system, the second experiment examines how the change in Pell grants affects students if there were no admissions uncertainty. Finally, the importance of the tuition caps in driving the results is examined by considering how student allocations would change under the Pell grant increase if tuition caps increased as well.

**College market dynamics.** In the absence of adjustments in the college market, the higher Pell grant funding increases total enrollment by 10% in sharp contrast to the negligible effect on enrollment reported in Table F.5. This increase is driven by affected students who enroll at significantly higher rates without any change for students who did not benefit from the Pell grant increase. When tuition and admissions policies are allowed to adjust however, the increase in enrollment for affected students increases by only 1.2%, while unaffected students see a 1.0% decrease in enrollment. Overall, without accounting for changes in the college market, student welfare would increase by 5.2% in consumption equivalent units. This implies that ignoring the effects of changes in the college market would lead to overestimate the welfare gains of the financial aid policy by more than a factor of two.

**No admissions uncertainty.** In this counterfactual, signals are made perfectly informative and the model is then re-estimated to match all targets described in Table 4.5 (except the average admissions rate for selective colleges). Under this new calibration, the welfare increases by 4.9% in consumption equivalent units, which is more than double the increase from the baseline scenario. The noise from the admissions signal thus *dampens* the welfare gains from the grant increase. This happens because low-income students may still draw a low signal and be unable

to benefit from the grant, which leads to a lower ex-ante welfare gain. For high-income students unaffected by the Pell grant change, the welfare losses are larger with signal uncertainty because the grants cause colleges to endogenously increase their standards. This harms even high-ability, high-income students as there is a lower probability of being admitted. If signals were perfectly informative, however, the welfare losses would be concentrated only among the relatively low-ability, high-income students who are replaced.

Another reason the welfare change from the policy is stronger under perfect information is due to higher college value added. In the new calibration the value for  $\rho_L$  is lower, which means that the marginal effect of instructional spending on value added is higher. Overall, this exercise shows that failing to account for the uncertainty associated with the admissions system would lead to exaggerate the positive welfare effects of federal financial aid policy. This is illustrated in Figure F.6, which shows welfare gains from possible increases to the Pell grant maximum. The optimal increase in the Pell grant maximum would be largely overstated in the perfect information calibration as opposed to the baseline.

**Increased tuition caps.** This exercise is motivated by the assumption that tuition caps would remain fixed in response to large increases in federal financial aid. As a robustness check, a 10%, 20%, and 30% increases to the tuition caps are added to account for potential increases in sticker prices. The results are presented in Table F.6. When the tuition caps are higher, colleges respond to the Pell grant increase by raising tuition even more for low-income students but with additional instructional per-student spending. Note that the higher tuition from the increased tuition cap leads to lower enrollment. Next, the adjusted tuition caps lead to more concentration at the top of the income distribution at highly-selective colleges. This suggests that increases in tuition caps that may result from the policy will limit the extent that financial aid will reduce income inequality at selective colleges. Finally, the welfare gains from the policy are not diminished with higher tuition caps as low-income high-ability students would still benefit from the increased financial aid.

## 8. CONCLUSION

This paper studies the role of the admissions system in shaping the allocation of students in the college market. Using micro-level data on high-school students transitioning to college, the analysis shows that parental income is correlated with college applications and enrollment. Higher-income students are more likely to apply to college not only at the extensive margin (i.e. applying or not), but also at the intensive margin (i.e. applying to more selective colleges). Moreover, applicants face risk not only in the college admissions decision, but also in the financial aid decision as many students report being unable to attend their preferred college due to costs.

Motivated by these empirical findings, an equilibrium model of the college market with student heterogeneity and a non-trivial application and admissions system is presented. The model is able to jointly reconcile the income differences in application and enrollment rates together with high levels of financial aid available to low-income students. Lower income students apply to selective colleges at lower rates because of expectations that they will not receive sufficient financial aid. Since higher ability students expect higher average application signals, only the highest ability among the low-income students apply to selective colleges. This makes the selective colleges confident that their low-income applicants are of high ability, justifying the high levels of financial aid we observe in the data.

In focusing on the role of applications and admissions, this paper abstracted from many important other issues in the college market left for future research. An important distinction in the college market is the presence of public and private institutions. For instance, the funding

for state schools depends on state governments, which allows them to offer substantially lower levels of tuition to in-state students. This paper also abstracted from the source of parental transfers. In reality, a parent's willingness to invest in their children's education is conditional on the student's decision to where to enroll and may be an important margin of adjustment in response to policy changes. Finally, students are guaranteed to graduate and can perfectly forecast their post-college earnings. In reality, students face substantial drop-out and post-college earnings risk, which may be important factors in determining a student's willingness to pursue a college education. These considerations are left for future research.

## REFERENCES

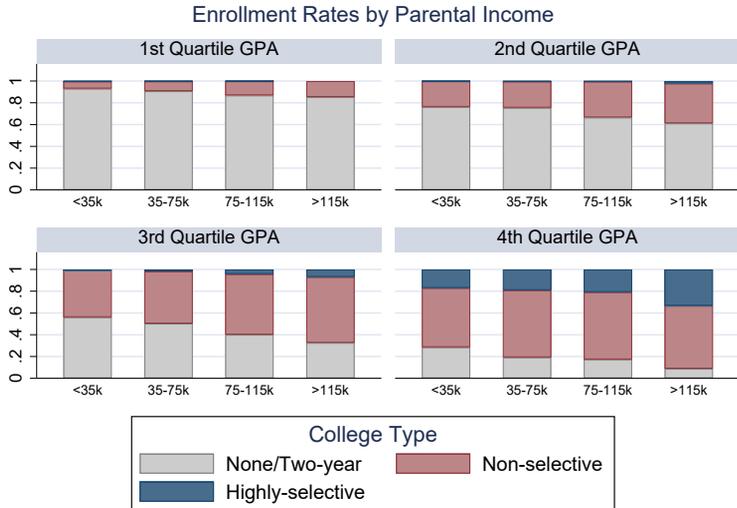
- B. Abbott, G. Gallipoli, C. Meghir, and G. L. Violante. Education policy and intergenerational transfers in equilibrium. *Journal of Political Economy*, 127(6):2569–2624, 2019. [4, 26, 27, 28, 29]
- R. J. Andrews, S. A. Imberman, and M. F. Lovenheim. Recruiting and supporting low-income, high-achieving students at flagship universities. *Economics of Education Review*, 74:101923, 2020. [4]
- M. J. Bailey and S. M. Dynarski. *Inequality in Postsecondary Education*, pages 117–132. Russell Sage Foundation, 2011. [1]
- E. P. Bettinger, B. T. Long, P. Oreopoulos, and L. Sanbonmatsu. The Role of Application Assistance and Information in College Decisions: Results from the H&R Block Fafsa Experiment. *The Quarterly Journal of Economics*, 127(3):1205–1242, 07 2012. [3]
- Z. Bleemer. Top percent policies and the return to postsecondary selectivity. 2021. [4]
- R. BÄġnabou. Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica*, 70(2):481–517, 2002. [4]
- Z. Cai and J. Heathcote. College tuition and income inequality. *American Economic Review*, 112(1):81–121, 2022. [4]
- D. Capelle. The great gatsby goes to college: Tuition, inequality and intergenerational mobility in the us. *Working Paper*, 2019. [4]
- S. E. Carrell, R. L. Fullerton, and J. E. West. Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics*, 27(3):439–464, 2009. [3]
- E. Caucutt and L. Lochner. Early and late human capital investments, borrowing constraints, and the family. *Working Paper*, 2017. [4]
- H. Chade, G. Lewis, and L. Smith. Student portfolios and the college admissions problem. *Review of Economic Studies*, 81:971–1002, 2014. [4]
- 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. [1, 3, 4, 5, 29]
- M. Colas, S. Findeisen, and D. Sachs. Optimal need-based financial aid. *Journal of Political Economy*, 129(2):492–533, 2021. [4]
- F. Cunha, J. Heckman, and S. Navarro. Separating uncertainty from heterogeneity in life cycle earnings. *oxford Economic papers*, 57(2):191–261, 2005. [14]
- 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. [3]
- 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. [3, 4]
- J. M. Delaney and P. J. Devereux. Choosing differently? college application behavior and the persistence of educational advantage. 2020. [3]
- E. Dillon and J. Smith. Determinants of the match between student ability and college quality. *Journal of Labor Economics*, 35:45–66, 2017. [1, 3]
- S. Dynarski, C. Libassi, K. Micheltmore, 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. [3]
- D. Epplé and R. E. Romano. Competition between private and public schools, vouchers, and peer-group effects. *American Economic Review*, pages 33–62, 1998. [16]
- D. Epplé, R. Romano, and H. Sieg. Admission, tuition, and financialaid policies in the market for higher education. *Econometrica*, 74:885–928, 2006. [1, 4, 16, 22, 10]
- D. Epplé, R. Romano, S. Sarpcá, 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. [4, 22, 31]

- R. Fernandez and R. Rogerson. Income distribution, communities and the quality of public education. *The Quarterly Journal of Economics*, 111(1), 1996. [4]
- I. Fillmore. Price discrimination and public policy in the u.s. college market. *Working Paper*, 2020. [4]
- C. Fu. Equilibrium tuition, applications, admissions and enrollment in the college market. *Journal of Political Economy*, 122(2):225–281, 2014. [1, 4, 28]
- G. Gordon and A. Hedlund. Accounting for the rise in college tuition. Working Paper 21967, National Bureau of Economic Research, February 2016. [4]
- G. Gordon and A. Hedlund. Accounting for the rise in college tuition. Working paper, 2021. [4]
- J. J. Heckman, L. J. Lochner, and P. E. Todd. Earnings functions, rates of return and treatment effects: The mincer equation and beyond. *Handbook of the Economics of Education*, 1:307–458, 2006. [14]
- M. Hoekstra. The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach. *The Review of Economics and Statistics*, 91(4):717–724, 11 2009. [4]
- C. Hoxby and S. Turner. Expanding college opportunities for high-achieving, low income students. *Stanford Institute for Economic Policy Research Discussion Paper*, 12:014, 2013. [3]
- 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. [1, 3]
- F. Ionescu. The federal student loan program: Quantitative implications for college enrollment and default rates. *Review of Economic Dynamics*, 12(1):205–231, 2009. [4]
- 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. [4]
- T. Kotera and A. Seshadri. Educational policy and intergenerational policy. *Review of Economic Dynamics*, 25: 187–207, 2017. [4]
- 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. [4]
- L. J. Lochner and A. Monge-Naranjo. The nature of credit constraints and human capital. *American Economic Review*, 101(6):2487–2529, October 2011. [4]
- 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. [4]
- J. Mountjoy and B. Hickman. The returns to college(s): Estimating value-added and match effects in higher education. 2020. [3]
- B. Sacerdote. Peer Effects with Random Assignment: Results for Dartmouth Roommates\*. *The Quarterly Journal of Economics*, 116(2):681–704, 05 2001. [3]
- L. J. Turner. The economic incidence of federal student grant aid. 2017. [4]
- D. J. Zimmerman. Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics*, 85(1):9–23, 2003. [3]
- S. D. Zimmerman. The returns to college admission for academically marginal students. *Journal of Labor Economics*, 32(4):711–754, 2014. [4]

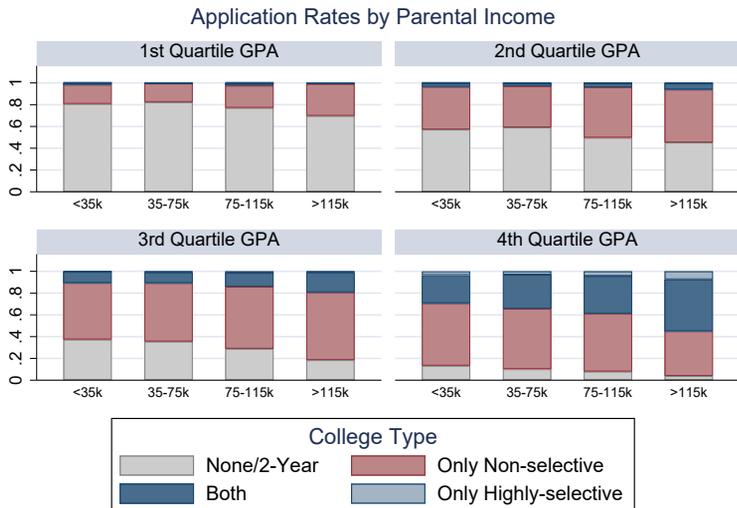
ONLINE APPENDIX

This online appendix contains four sections.

APPENDIX A: ADDITIONAL EMPIRICAL RESULTS



Graphs by Quartiles of High School GPA



Graphs by Quartiles of High School GPA

FIGURE A.1.—College application and enrollment by parental income and student GPA quartile. Note that the “< 35k” group corresponds to the bottom third of the parental income distribution, “35 – 75k” corresponds to the middle third, and both “75 – 115k” and “> 115k” evenly divide the top third. Source: HSLs.

TABLE A.1  
LOGIT ESTIMATION RESULTS FOR STUDENT APPLICATION DECISIONS

Dependent Variable	Apply at all		Apply to Highly-selective	
	Coefficient	SE	Coefficient	SE
Parental Income	1.102***	(0.018)	1.004	(0.019)
SAT (standardized)	1.549***	(0.173)	2.111***	(0.302)
GPA (standardized)	2.925***	(0.195)	2.167***	(0.187)
Parental Income*SAT	1.037*	(0.022)	1.062***	(0.018)
Asian	1.756**	(0.392)	2.358***	(0.369)
Black	3.323***	(0.506)	2.357***	(0.459)
Hispanic	1.768***	(0.215)	2.191***	(0.378)
Other	1.367**	(0.216)	1.259***	(0.190)
Female	1.096	(0.091)	0.802**	(0.073)
Parent has HS diploma or GED	1.241	(0.206)	0.961	(0.300)
Parent has Associate's	1.231	(0.228)	1.023	(0.327)
Parent has Bachelor's	1.337	(0.241)	1.336	(0.418)
Parent has Master's or higher	1.507**	(0.299)	1.533	(0.490)
Number in student's household	0.975*	(0.015)	0.970	(0.029)
High School is Catholic	4.369***	(0.835)	1.175	(0.163)
High School is (other) private	2.283***	(0.443)	1.357**	(0.194)
Percent in HS receiving free lunch	0.989	(0.021)	0.992	(0.024)
Suburban	0.910	(0.110)	0.837	(0.104)
Town	0.559***	(0.078)	0.448***	(0.076)
Rural	0.819*	(0.095)	0.569*	(0.068)
Midwest region	1.016	(0.142)	0.592***	(0.084)
South region	0.826	(0.113)	0.490***	(0.076)
West region	0.684**	(0.108)	0.501***	(0.090)
Constant	1.740*	(0.513)	0.524	(0.241)
N	11,410		8,960	
Pseudo $R^2$	0.247		0.264	

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Note that for the estimation in the right column, the sample is restricted to all students who applied to any four-year, non-profit colleges. The dependant variable is equal to 1 if the student applied to any four-year, non-profit colleges (in the left column) or any Highly-selective colleges (in the right column), and 0 otherwise. Note that all coefficients are expressed as odds ratios. Parental income is coded based on pre-set ranges that parents chose from in the survey: "< \$15k, \$15k to \$35k, \$35k to \$55k, ..., \$215k to \$235k, > \$235k".

FIGURE A.2.—Predicted relationship between SAT scores (left panel) / GPA (right panel), parental income, and application decisions implied by estimated logit model. See left columns of Table A.1 for estimation results.

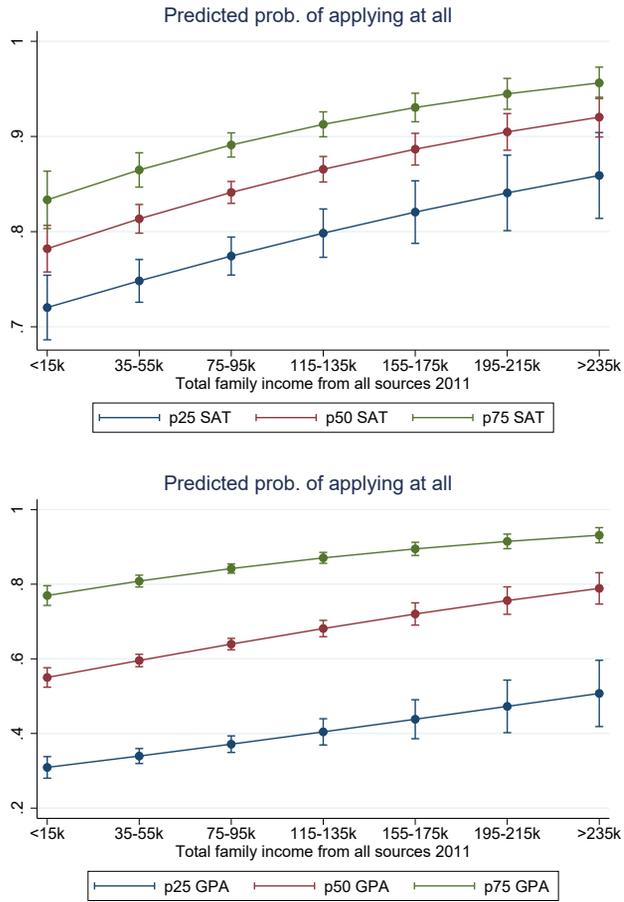
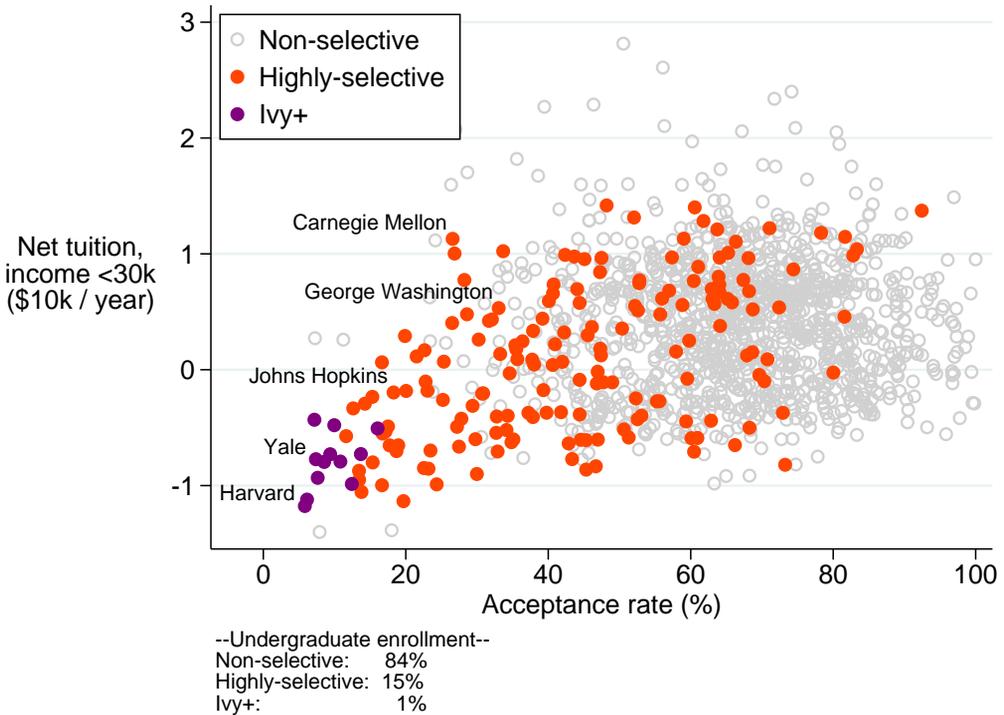


FIGURE A.3.—Scatterplot of net-tuition for low-income students (income <\$30k per year in 2013) and undergraduate acceptance rate



Note: Net tuition data are from the 2013-2014 academic year, and the acceptance rate was calculated by averaging over the 2012 and 2013 admission cycles. This sample consists of all four-year, non-profit Bachelor's degree granting colleges available from IPEDS (excluding theological seminaries and bible colleges). "Highly-selective" colleges correspond to those ranked as "most competitive" and "highly competitive" by *Barron's Profile of American Colleges* (2015), while "Non-selective" corresponds to the remaining colleges in the sample. "Ivy+" corresponds to all Ivy League colleges plus Stanford, MIT, Chicago, and Duke.

## APPENDIX 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

TABLE B.1  
COMPARISON OF COLLEGES IN EACH SELECTIVITY TIER

College characteristics	Highly-selective	Non-selective
Percent public	55%	73%
Acceptance rate	44%	68%
Median SAT	1292	1051
Median Earnings 10 Years After Entry	\$57,803	\$42,286
<b>Attendance Costs (per Year)</b>		
Tuition and fees	\$24,622	\$12,598
Room and Board	\$11,474	\$9,372
<i>Net tuition</i> (\$1,000s)		
Income <30k	-1.85	0.64
Income 30k-48k	-0.02	1.89
Income 48k-75k	4.22	4.85
Income 75k-110k	9.28	7.43
Income >110k	15.92	8.63
<b>Financial variables</b>		
Instructional spending per student	\$20,746	\$8,597
Endowment assets per student	\$131,440	\$13,271
Number of colleges	186	1,577
Fraction of undergraduate enrollment	15.7%	84.3%

Note: All averages are weighted by total undergraduate enrollment at each college. Data are from the 2013-2014 academic year, except the acceptance rate which was calculated by averaging over the 2012 and 2013 admissions cycles. This sample consists of all four-year, non-profit Bachelor's degree granting colleges available from IPEDS (excluding religious colleges). Also note that six Highly-selective colleges from the list in Section B.1 do not appear in the table above because they do not have IPEDS data available. These include: United States Air Force Academy, United States Military Academy, United States Naval Academy, United States Coast Guard Academy, Grove City College, and Hillsdale College.

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. High School Longitudinal Study of 2009 (HSLs:09)

### B.3. IPEDS Sample Restrictions

The Integrated Postsecondary Education Data System is made publicly available through the National Center for Education Statistics. Since we are interested only in four-year nonprofit

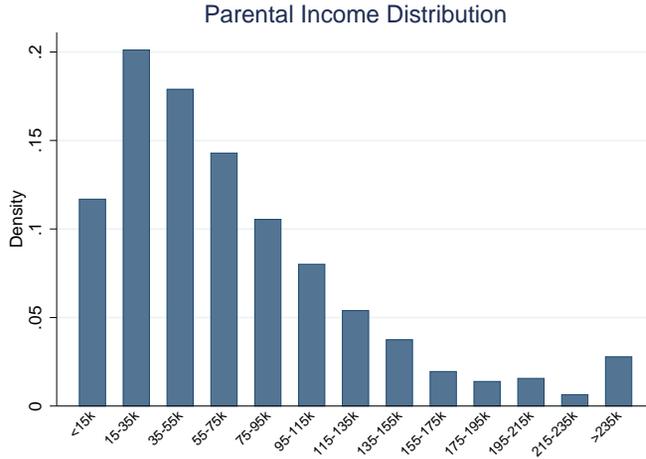


FIGURE B.1.—Distribution of parental income in the HSLs.

U.S. colleges and universities, we 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

We restrict our IPEDS data sample to cover the 2013-2016 time frame, since it is the relevant 4 year period for our HSLs cohort who begin college in 2013. Our final sample includes 1,665 four-year colleges and universities.

## APPENDIX C: ESTIMATION 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:

<b>If parents' AAI is –</b>	<b>Parents' contribution from AAI is –</b>
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. Distribution of Student Characteristics

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  $y = 0$ . For the remaining 70% of the distribution with  $y > 0$ , we assume that  $y$  follows a log-normal distribution,  $y \sim \text{LogN}(\mu_y, \sigma_y^2)$ .

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

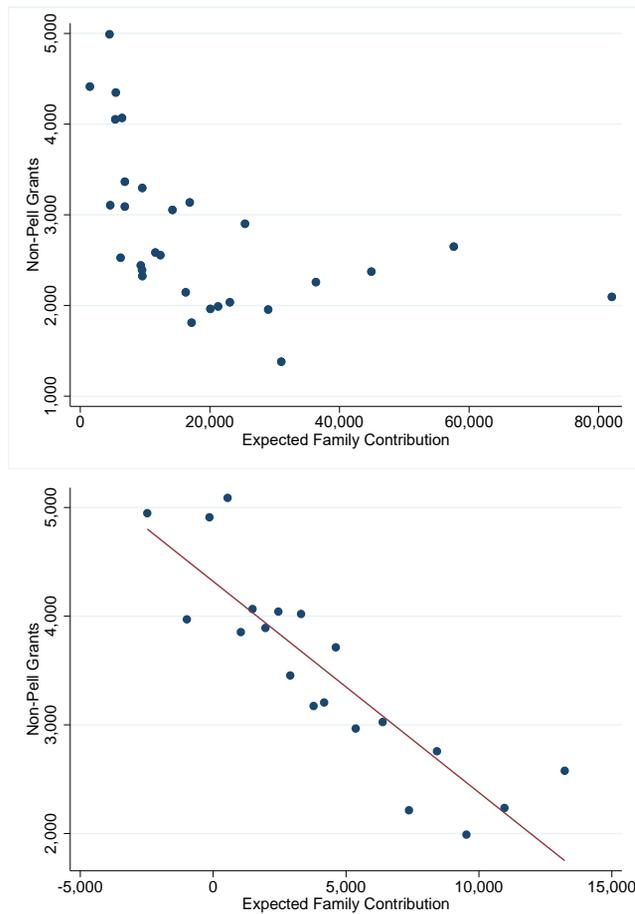
To summarize, the joint distribution of  $(y, \ell)$  is given by:

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

Using the HSLs, we find  $\mu_{\ell 0} = -0.144, \sigma_{\ell 0}^2 = 0.155$  for  $y = 0$ , and  $\mu_{\ell 1} = 0.135, \sigma_{\ell 1}^2 = 0.138$  for  $y > 0$ . We also estimate  $\mu_y = -0.301, \sigma_y^2 = 2.825$ . Finally, to account for this correlation for the case where  $y > 0$ , we estimate the covariance between  $\ell$  and  $a$  to be  $\sigma_{a\ell} = 0.174$ .

### C.3. Grant Function Estimation

Since all college-specific grants (both need-based and merit-based) and federal Pell Grants are accounted for in our model, we add the exogenous grants, denoted by  $Gr(y)$ , to stand in for all other grants that students receive. We identify all such grants in our BPS data by adding up all grants and aid excluding loans, Pell Grants, and institutional aid. As we can see from the binned scatter-plot below, these grants vary substantially with Expected Family Contribution (EFC).



These grants are decreasing in EFC and tend to level off at around \$15,000. For EFC below this level, we estimate a downward sloping linear relationship between EFC and non-Pell Grants using OLS (and controlling for SAT score). We find  $g_0 = 0.43, g_1 = 0.19$  in (4.1), corresponding to right panel of the plot above. We pick  $\underline{g} = 0.15$  to capture that the grants level off at a sufficiently high EFC level.

#### C.4. College Cost Function Estimation

Before estimating the college-specific parameters using our IPEDS sample, we first confirm that the sample includes the relevant colleges relative to the HSLs. We compare the share of total student enrollment accounted for by each type of college<sup>29</sup> in our IPEDS sample with the enrollment shares observed in the HSLs. For students in our HSLs sample, we find the following fractions at each type of college:

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

To confirm that these numbers are consistent with our IPEDS data, we first add up all full-time equivalent undergraduate students across all colleges in our sample. We then infer the total number of high school graduates using the 42% 4-year enrollment rate reported by NCES<sup>30</sup>. We find that of this population, 7% attend College 1 and 35% College 2, similar to the HSLs enrollment patterns.

Confident that our IPEDS sample includes all colleges relevant for our analysis, we use variation in enrollment size for colleges to estimate the parameters of the cost function introduced in (??). We start by estimating the parameters in the following regression separately for each type of college:

$$cost_i^s = \hat{O}_0^s + \hat{O}_2^s(k_i^s)^2 + \varepsilon_i, \quad (\text{C.1})$$

for college  $i$  of college type  $s$ , where  $k$  is the enrollment level of college  $i$ . In order to measure the costs for each college in our sample, we use the detailed financial data available from IPEDS. There are two methods to measure costs: one is by directly adding up all non-instructional expenditure including academic support, student services, and institutional support. The other is by using the budget constraint in Equation (3.10), where we calculate cost by adding up all tuition revenue, net grant revenue, government appropriations, unrestricted revenue from private sources, and subtracting off total instructional expenditure. We rely on both methods by defining costs in the left hand side of equation (C.1) as the maximum over both methods (this helps us deal with cases where the second procedure produces small or negative numbers). The results of the estimation are provided in Table C.1 below.

Next, we follow [Epple et al. \(2006\)](#) in their aggregation procedure to transform our cost function parameter estimates from (C.1) into our corresponding model parameters. If there are  $n_s$  colleges in type  $s$ , we assume they are identical in the model so that total enrollment  $K^s$  simply scales up the individual enrollments:  $K^s = \sum_{i=1}^{n_s} k_i^s = n_s k^s$ . We can then sum up the individual cost functions:

$$C(K^s) = \sum_{i=1}^{n_s} C(k_i^s) = \sum_{i=1}^{n_s} [O_0^s + O_2^s(k_i^s)^2]$$

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

$$C(K^s) = o_0^s + o_2^s (K^s)^2$$

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 (C.1).

<sup>29</sup>Where college types are based on the Barron's ranking as described in Section ??.

<sup>30</sup>See here: [https://nces.ed.gov/programs/coe/indicator\\_cpa.asp](https://nces.ed.gov/programs/coe/indicator_cpa.asp).

	(1) <i>costs</i> <sup>1</sup>	(2) <i>costs</i> <sup>2</sup>
<i>Enrollment squared</i>	1057.6*** (120.4)	624.2*** (14.85)
<i>Private college</i>		-0.0000919*** (0.0000153)
<i>Constant</i>	0.000841*** (0.000111)	0.000222*** (0.0000126)
<i>N</i>	182	1445
<i>adj. R</i> <sup>2</sup>	0.296	0.604

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

TABLE C.1

ESTIMATES FOR THE COST PARAMETERS IN (C.1) FOR EACH COLLEGE TYPE. NOTE THAT ALL VARIABLES HAVE BEEN SCALED DOWN BY THE TOTAL STUDENT POPULATION, AND THE COSTS HAVE BEEN FURTHER SCALED DOWN BY \$40,000 TO FIT THE MODEL UNITS.

**College subsidies.** Similar to the cost function estimation above, we estimate  $Tr_1^s$  for each College type  $s$  using our IPEDS sample. We rely on variation in enrollment within each college type, and measure subsidies to colleges by the level of state government appropriations and federal funding they receive. We simply estimate the slope in the following regression relating government transfers to enrollment levels:

$$transfer_i^s = \hat{Tr}_1^s k_i + \varepsilon_i.$$

The results of the estimation are given in Table C.2 below.

### C.5. Net tuition variation within college types

We use the BPS to show how net tuition, defined as sticker-price tuition minus college-specific grants (need- and merit-based financial-aid) varies with student Expected Family Contribution (EFC). Figure C.1 below provides binned-scatter plots which show how net-tuition levels off for students with high EFC who are less likely to receive financial-aid from the colleges.

	(1)	(2)
	<i>transfer</i> <sup>1</sup>	<i>transfer</i> <sup>2</sup>
<i>Enrollment</i>	1.044*** (0.0516)	0.620*** (0.0112)
<i>N</i>	182	1445
<i>adj. R</i> <sup>2</sup>	0.691	0.679

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

TABLE C.2

ESTIMATES FOR THE RELATIONSHIP BETWEEN GOVERNMENT TRANSFERS TO COLLEGES AND THEIR ENROLLMENT LEVEL.

### C.6. Simulated Method of Moments Procedure

First, let  $M_j^{data}$  denote the  $j$ th empirical moment out of  $J$  total moments. To estimate our model, we first draw  $N = 15,000$  students (consistent with our HSLS sample) across  $S = 100$  simulations using our distribution over  $(y, \ell)$ . Let  $\Theta$  denote the vector of parameters described in Section 4.5, and let  $M_j^s(\Theta)$  denote the  $j$ th model calculated moment implied by  $\Theta$  for simulation  $s = 1, \dots, S$ . We seek  $\hat{\Theta}$  to minimize the following unweighted criterion function:

$$\sum_{j=1}^J \left( \frac{M_j^{data} - \frac{1}{S} \sum_{s=1}^S M_j^s(\Theta)}{M_j^{data}} \right)^2.$$

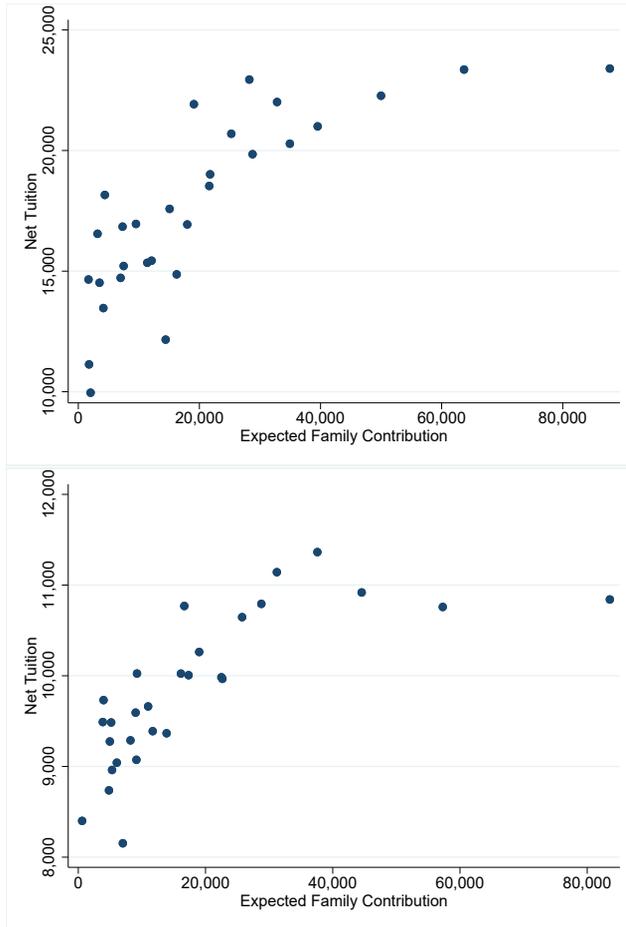


FIGURE C.1.—Binned-scatter plots for net tuition (sticker-price tuition minus college-specific grants) for highly-selective colleges (left panel) and non-selective colleges (right panel). The data come from the 2012 cohort of the Beginning Postsecondary Students Longitudinal Study (BPS).

## APPENDIX D: PROOFS OF PROPOSITIONS

### D.1. Derivation of (3.15)

Let  $\lambda_I, \lambda_L, \lambda_\kappa$  be the Lagrange multipliers on the budget constraint (3.10), average ability identity constraint (3.9), and enrollment identity constraint (3.8) respectively. First order conditions for  $I_\mu, L_\mu, \kappa, T(y, \sigma)$  are then

$$\begin{aligned}
 [\kappa] \quad & -\lambda_\kappa = \lambda_L L + \lambda_I [I + C'(\kappa) - Tr'(\kappa)] \\
 [I_\mu] \quad & \xi_s Q_I = \lambda_I \kappa \\
 [L_\mu] \quad & \xi_s Q_L = \lambda_L \kappa
 \end{aligned}$$

$$\begin{aligned}
[T(y, \sigma)] \quad T(y, \sigma) &= -\frac{\lambda_\kappa}{\lambda_I} - \frac{\lambda_L}{\lambda_I} \frac{\int \ell \frac{\partial \tilde{q}}{\partial T(y, \sigma)} g(\sigma|\ell) \mu(y, d\ell)}{\int \frac{\partial \tilde{q}}{\partial T(y, \sigma)} g(\sigma|\ell) \mu(y, d\ell)} \\
&\quad - \frac{\int \tilde{q} g(\sigma|\ell) \mu(y, d\ell)}{\int \frac{\partial \tilde{q}}{\partial T(y, \sigma)} g(\sigma|\ell) \mu(y, d\ell)} + \frac{\lambda_T(y, \sigma)}{\lambda_I \int \frac{\partial \tilde{q}}{\partial T(y, \sigma)} g(\sigma|\ell) \mu(y, d\ell)}
\end{aligned} \tag{D.1}$$

where  $\lambda_T(y, \sigma)$  is the Lagrange multiplier for the constraint on tuition. Combining these equations gives (3.15).

### D.2. Derivation of (3.16)

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

$$\frac{\int T(y, \underline{\sigma}_s) \tilde{q}(\underline{\sigma}_s) g(\underline{\sigma}_s|\ell) d\mu(y, \ell)}{\int \tilde{q}(\underline{\sigma}_s) g(\underline{\sigma}_s|\ell) d\mu(y, \ell)} \geq -\frac{\lambda_\kappa}{\lambda_I} - \frac{\lambda_L}{\lambda_I} \frac{\int \ell \tilde{q}(\underline{\sigma}_s) g(\underline{\sigma}_s|\ell) d\mu(y, \ell)}{\int \tilde{q}(\underline{\sigma}_s) g(\underline{\sigma}_s|\ell) d\mu(y, \ell)}, \tag{D.2}$$

which holds with equality when  $\underline{\sigma}_s > 0$ . Combining with (D.1), we arrive at (3.16).

### D.3. Proposition 3.1

PROOF: Combining (3.16) and (3.15), we find

$$\begin{aligned}
&\frac{Q_{L\mu}}{Q_{I\mu}} \int \left( \ell - \frac{\int \ell \frac{\partial \tilde{q}^s(\underline{\sigma}_s)}{\partial T^s(y, \underline{\sigma}_s)} g(\underline{\sigma}_s|\ell) \mu(y, d\ell)}{\int \frac{\partial \tilde{q}^s(\underline{\sigma}_s)}{\partial T^s(y, \underline{\sigma}_s)} g(\underline{\sigma}_s|\ell) \mu(y, d\ell)} \right) \tilde{q}^s(\underline{\sigma}_s) g(\underline{\sigma}_s|\ell) d\mu(y, \ell) \\
&- \int \frac{\left[ \int \tilde{q}^s(\underline{\sigma}_s) g(\underline{\sigma}_s|\ell) \mu(y, d\ell) \right]^2}{\int \frac{\partial \tilde{q}^s(\underline{\sigma}_s)}{\partial T^s(y, \underline{\sigma}_s)} g(\underline{\sigma}_s|\ell) \mu(y, d\ell)} dy + \frac{\int \lambda_T(y, \underline{\sigma}_s) \tilde{q}^s(\underline{\sigma}_s) g(\underline{\sigma}_s|\ell) d\mu(y, \ell)}{\lambda_I \int \frac{\partial \tilde{q}^s(\underline{\sigma}_s)}{\partial T^s(y, \underline{\sigma}_s)} g(\underline{\sigma}_s|\ell) \mu(y, d\ell)} \geq 0 \tag{D.3}
\end{aligned}$$

Since the constraint on tuition is never binding, we have  $\lambda_T(y, \sigma) = 0$ . Rewrite the left hand side of (D.3) to obtain

$$\frac{Q_{L\mu}}{Q_{I\mu}} \int \left[ \int q^s(\underline{\sigma}_s) g(\underline{\sigma}_s|\ell) \mu(y, d\ell) \right] \underbrace{\left[ \frac{\int \ell q^s(\underline{\sigma}_s) g(\underline{\sigma}_s|\ell) \mu(y, d\ell)}{\int q^s(\underline{\sigma}_s) g(\underline{\sigma}_s|\ell) \mu(y, d\ell)} - \frac{\int \ell \frac{\partial q^s(\underline{\sigma}_s)}{\partial T^s} g(\underline{\sigma}_s|\ell) \mu(y, d\ell)}{\int \frac{\partial q^s(\underline{\sigma}_s)}{\partial T^s} g(\underline{\sigma}_s|\ell) \mu(y, d\ell)} \right]}_{\geq 0} dy$$

$$+ \underbrace{\int \frac{\left[ \int q^s(\underline{\sigma}_s) g(\underline{\sigma}_s | \ell) \mu(y, d\ell; y) \right]^2}{\int -\frac{\partial q^s(\underline{\sigma}_s)}{\partial T^s(y, \underline{\sigma}_s)} g(\underline{\sigma}_s | \ell) \mu(y, d\ell; y)} dy}_{>0} \quad (\text{D.4})$$

where  $q^s(\underline{\sigma}_s) = \tilde{q}^s(y, \ell, T(y, \underline{\sigma}_s))$ . We show that (D.4) 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  $\ell$  according to two distributions,  $H_0$  and  $H_1$ , with associated densities:  $h_0(\ell) = \frac{q^s(\underline{\sigma}_s) g(\underline{\sigma}_s | \ell) \mu(y, \ell)}{\int q^s(\underline{\sigma}_s) g(\underline{\sigma}_s | \ell) \mu(y, d\ell)}$  and  $h_1(\ell) = \frac{\frac{\partial q^s(\underline{\sigma}_s)}{\partial T^s} g(\underline{\sigma}_s | \ell) \mu(y, \ell)}{\int \frac{\partial q^s(\underline{\sigma}_s)}{\partial T^s} g(\underline{\sigma}_s | \ell) \mu(y, d\ell)}$ . 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  $\lim_{\ell \rightarrow 0} F(\ell) = 0$ ,  $\lim_{\ell \rightarrow \infty} F(\ell) = 0$ . The interior solution to the first order condition  $F'(\ell^*) = 0$  uniquely minimizes  $F$ :

For College 1,

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

We have a unique interior minimum since,

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

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

Since (D.4) is strictly positive, the lower bound constraint on  $\underline{\sigma}_s$  binds, so the college does not use the admissions threshold to screen students. Q.E.D.

## APPENDIX 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:

$$\max_{\{c_j, a_{j+1}\}_{j=0, \dots, T}} u(c_0) + \sum_{j=1}^T \beta^j u(c_j) \quad (\text{E.1})$$

$$\text{s.t. } c_0 + a_1 + T = a_0$$

$$c_j + a_{j+1} = a_j R + w(1 - \tau)\gamma_s \ell^\alpha, \quad j = 1, \dots, T$$

$$a_1 \geq \underline{a}_s \quad (\text{E.2})$$

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:

$$\max_{c_0, c_1} u(c_0) + \tilde{\beta} u(c_1) \quad (\text{E.3})$$

$$\text{s.t. } c_0 + a_1 + T = a_0$$

$$c_1 = a_1 \tilde{R} + \tilde{w}(1 - \tau)\gamma_s \ell^\alpha$$

$$a_1 \geq \underline{a}_s \quad (\text{E.4})$$

where

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

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

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

## APPENDIX F: ADDITIONAL MODEL RESULTS

F.1. *Effect of Equalizing Application Patterns*

		<b>College 1</b>		<b>College 2</b>	
		<i>Baseline</i>	<i>More Applicants</i>	<i>Baseline</i>	<i>More Applicants</i>
$L_\mu$	Avg Student Ability	1.86	1.82	1.28	1.3
$I_\mu$	Instr Spending per Student	2.1	2.04	0.98	0.97
$\Gamma_s$	Value-added	2.01	1.96	1.79	1.81
$\kappa$	% Enrolled	7.15	6.8	32.09	31.09
<i>Income Distribution</i>					
	Q1 Income	0.1	0.09	0.22	0.22
	Q2 Income	0.12	0.11	0.2	0.21
	Q3 Income	0.26	0.25	0.31	0.31
	Q4 Income	0.52	0.55	0.27	0.26
<i>Ability Distribution</i>					
	0-30% Ability	0.0	0.0	0.01	0.02
	30-60% Ability	0.0	0.01	0.41	0.39
	60-90% Ability	0.15	0.25	0.4	0.38
	Top 10% Ability	0.84	0.74	0.17	0.21

TABLE F.1

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*

		<b>College 1</b>		<b>College 2</b>	
		<i>Baseline</i>	$6 * \sigma_g^2$	<i>Baseline</i>	$6 * \sigma_g^2$
$L_\mu$	Avg Student Ability	1.86	1.67	1.28	1.19
$I_\mu$	Instr Spending per Student	2.1	2.4	0.98	1.14
$\Gamma_s$	Value-added	2.01	1.87	1.79	1.71
$\kappa$	% Enrolled	7.15	8.93	32.09	39.72
<i>Income Distribution</i>					
	Q1 Income	0.10	0.10	0.22	0.25
	Q2 Income	0.12	0.12	0.20	0.20
	Q3 Income	0.26	0.27	0.31	0.29
	Q4 Income	0.52	0.51	0.27	0.26
<i>Ability Distribution</i>					
	0-30% Ability	0.0	0.0	0.01	0.14
	30-60% Ability	0.0	0.09	0.41	0.43
	60-90% Ability	0.15	0.30	0.40	0.29
	Top 10% Ability	0.84	0.61	0.17	0.14

TABLE F.2

EFFECT ON COLLEGES OF MAKING SIGNALS LESS INFORMATIVE BY INCREASING THE VARIANCE OF THE SIGNAL DISTRIBUTION

		<b>College 1</b>		<b>College 2</b>	
		<i>Baseline</i>	<i>Incr. signal variance</i>	<i>Baseline</i>	<i>Incr. signal variance</i>
$L_\mu$	Avg Student Ability	1.86	1.77	1.28	1.24
$I_\mu$	Instr Spending per Student	2.1	2.24	0.98	1.03
$\Gamma_s$	Value-added	2.01	1.94	1.79	1.75
$\kappa$	% Enrolled	7.15	7.93	32.09	34.48
<i>Income Distribution</i>					
	Q1 Income	0.1	0.1	0.22	0.24
	Q2 Income	0.12	0.12	0.2	0.2
	Q3 Income	0.26	0.27	0.31	0.3
	Q4 Income	0.52	0.51	0.27	0.26
<i>Ability Distribution</i>					
	0-30% Ability	0.0	0.0	0.01	0.06
	30-60% Ability	0.0	0.04	0.41	0.43
	60-90% Ability	0.15	0.23	0.4	0.34
	Top 10% Ability	0.84	0.73	0.17	0.16

TABLE F.3

EFFECT ON COLLEGES OF MAKING SIGNALS LESS INFORMATIVE, BUT MAKING THE LOSS OF INFORMATION LARGER FOR HIGH-INCOME STUDENTS.

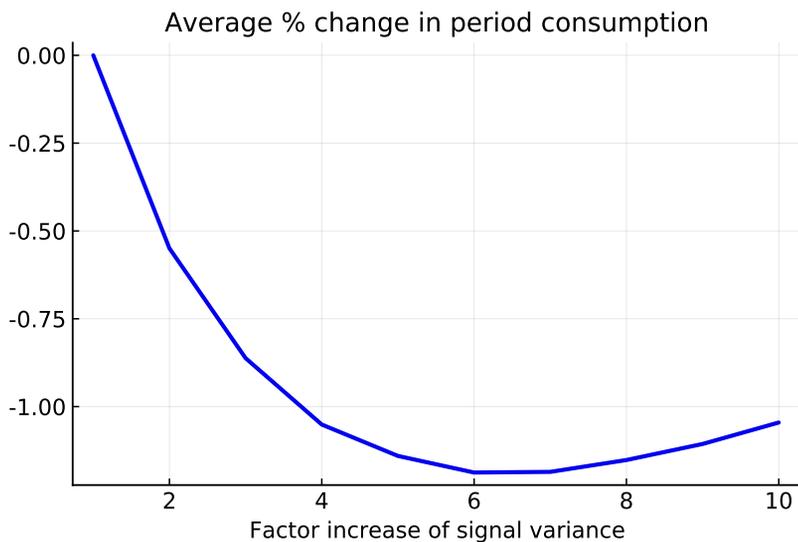


FIGURE F.1.—Welfare changes as a function of increases to the signal variance.

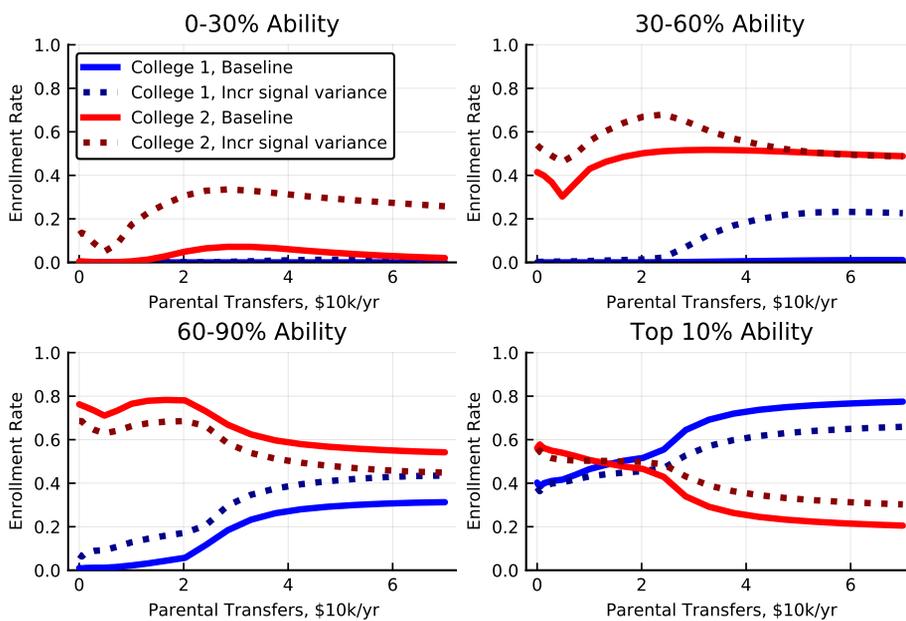


FIGURE F.2.—Student enrollment rates in baseline equilibrium vs. equilibrium with less informative signals.

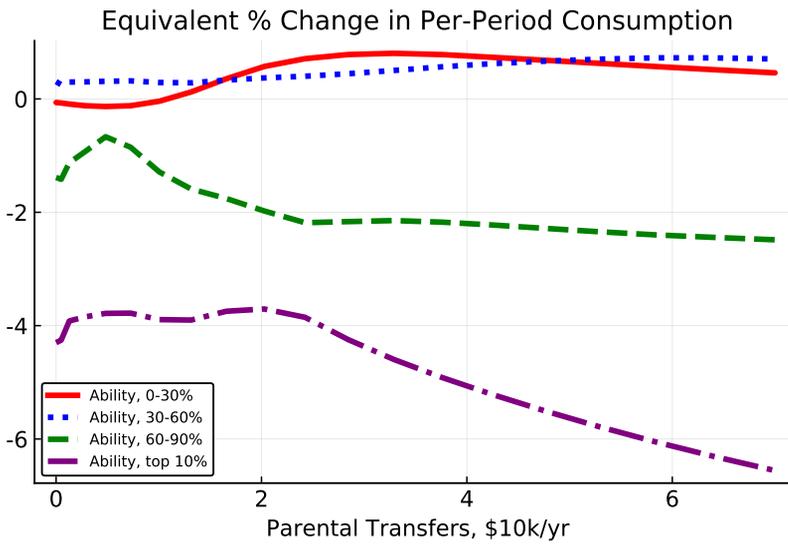


FIGURE F.3.—Percent of lifetime consumption students of different parental-income and ability groups are willing to give up in order to switch to the equilibrium with less informative signals. In this experiment, the loss of information is larger for high-income students.

F.3. *Effect of Perfectly Informative Signals*

		<b>College 1</b>		<b>College 2</b>	
		<i>Baseline</i>	<i>Perfect Info</i>	<i>Baseline</i>	<i>Perfect Info</i>
$L_\mu$	Avg Student Ability	1.86	1.91	1.28	1.36
$I_\mu$	Instr Spending per Student	2.1	1.87	0.98	0.93
$\Gamma_s$	Value-added	2.01	2.02	1.79	1.86
$\kappa$	% Enrolled	7.15	6.5	32.09	29.8
<i>Income Distribution</i>					
	Q1 Income	0.1	0.08	0.22	0.2
	Q2 Income	0.12	0.1	0.2	0.21
	Q3 Income	0.26	0.22	0.31	0.33
	Q4 Income	0.52	0.6	0.27	0.26
<i>Ability Distribution</i>					
	0-30% Ability	0.0	0.0	0.01	0.0
	30-60% Ability	0.0	0.0	0.41	0.27
	60-90% Ability	0.15	0.08	0.4	0.53
	Top 10% Ability	0.84	0.92	0.17	0.2

TABLE F.4

EFFECT ON COLLEGES OF MAKING SIGNALS PERFECTLY INFORMATIVE

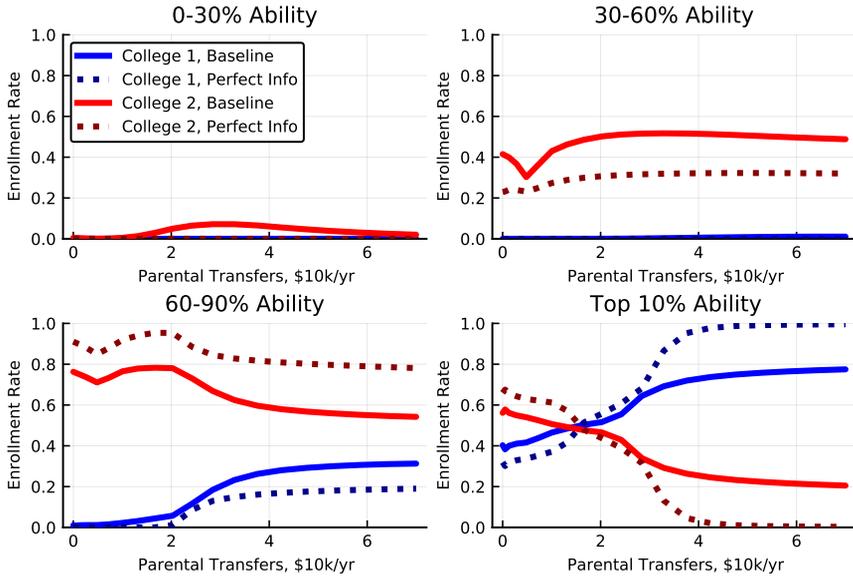


FIGURE F.4.—Average enrollment rates for students by parental transfers in the baseline equilibrium and perfect information equilibrium.

#### F.4. Policy Experiments: Increasing Pell Grants

		College 1		College 2	
		Baseline	Pell Grant	Baseline	Pell Grant
$L_\mu$	Avg Student Ability	1.86	1.87	1.28	1.26
$I_\mu$	Instr Spending per Student	2.1	2.67	0.98	1.04
$\Gamma_s$	Value-added	2.01	2.09	1.79	1.77
$\kappa$	% Enrolled	7.15	7.25	32.09	32.14
$\tau$	Tax rate (%)	1.51	3.93	1.51	3.93
<i>Income Distribution</i>					
	Q1 Income	0.1	0.14	0.22	0.23
	Q2 Income	0.12	0.16	0.2	0.22
	Q3 Income	0.26	0.31	0.31	0.31
	Q4 Income	0.52	0.38	0.27	0.24
<i>Ability Distribution</i>					
	0-30% Ability	0.0	0.0	0.01	0.06
	30-60% Ability	0.0	0.0	0.41	0.38
	60-90% Ability	0.15	0.14	0.4	0.39
	Top 10% Ability	0.84	0.86	0.17	0.17

TABLE F.5

EFFECT ON COLLEGES OF INCREASING THE PELL GRANT MAXIMUM.

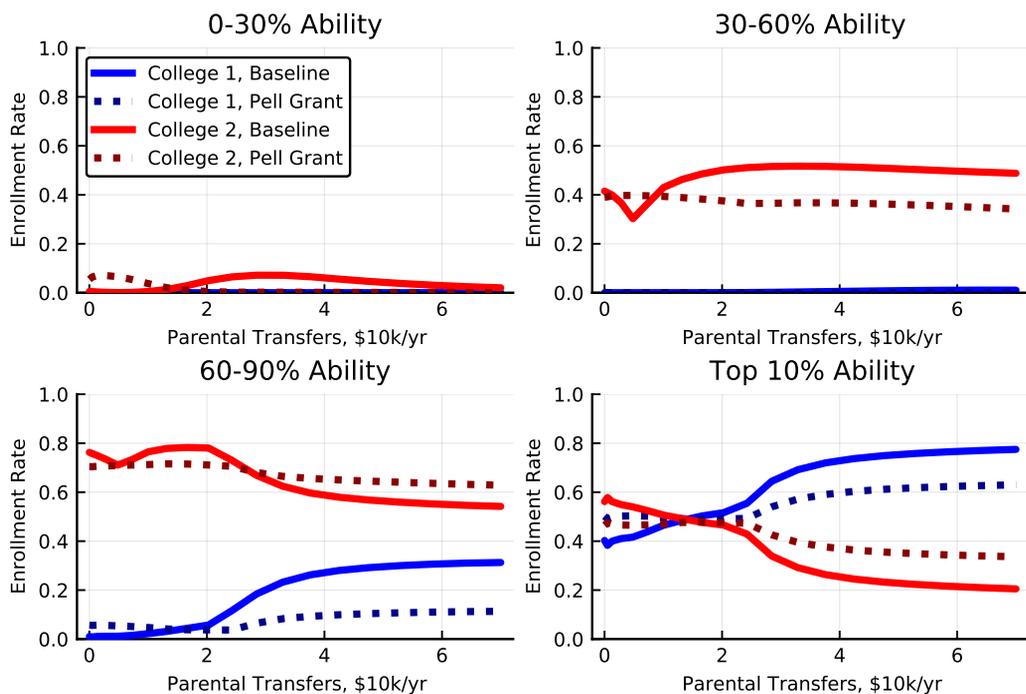


FIGURE F.5.—Student sorting by parental transfers and ability group both in the baseline equilibrium, and in an equilibrium where the Pell Grants maximum has increased from 5.65k per year to 25k per year.

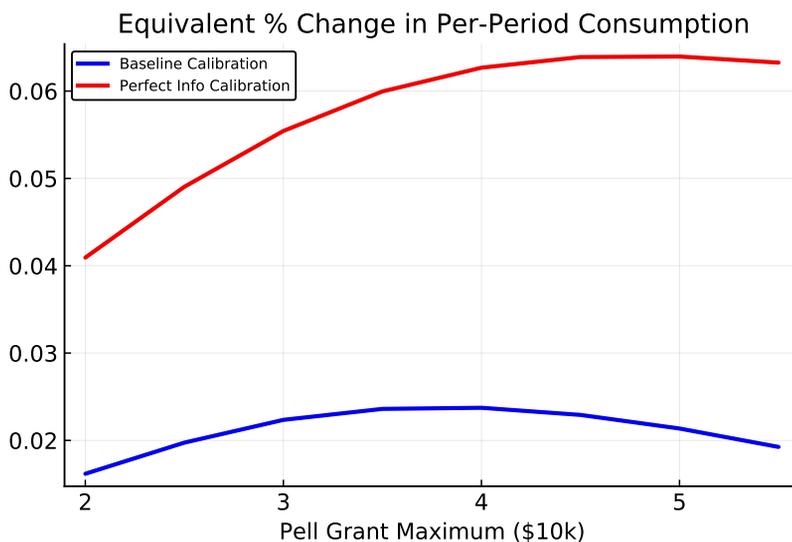


FIGURE F.6.—Average increases in welfare due to increases in the Pell Grant maximum both under the baseline parameter estimates, and under a perfect-information calibration. Failing to account for admissions uncertainty would cause the model to overstate the welfare gains from the policy.

TABLE F.6

EFFECT OF CHANGES TO PELL GRANT MAXIMUM WHEN TUITION CAPS ADJUST. WE COMPARE THE BASELINE SCENARIO TO SCENARIOS WHERE THE PELL GRANT MAXIMUM INCREASES TO 2.5, AND THE TUITION CAPS INCREASE BY 10%, 20%, AND 30%.

		<b>College 1</b>				
		<i>Baseline</i>	<i>Pell Grant</i>	<i>10% Adjust</i>	<i>20% Adjust</i>	<i>30% Adjust</i>
$L_\mu$	Avg Student Ability	1.86	1.87	1.88	1.89	1.91
$I_\mu$	Instr Spending per Student	2.10	2.67	2.81	2.86	2.84
$\Gamma_s$	Value-added	2.01	2.09	2.11	2.13	2.14
$\underline{\sigma}$	Admissions Threshold	1.52	1.76	1.71	1.65	1.58
$\kappa$	% Enrolled	7.15	7.25	6.84	6.38	6.02
<i>Income Distribution</i>						
	Q1 Income	0.10	0.14	0.14	0.13	0.11
	Q2 Income	0.12	0.16	0.16	0.15	0.14
	Q3 Income	0.26	0.31	0.30	0.29	0.29
	Q4 Income	0.52	0.38	0.40	0.44	0.46
<i>Ability Distribution</i>						
	0-30% Ability	0.0	0.0	0.0	0.0	0.0
	30-60% Ability	0.0	0.0	0.0	0.0	0.0
	60-90% Ability	0.15	0.14	0.13	0.10	0.09
	Top 10% Ability	0.84	0.86	0.87	0.90	0.91
		<b>College 2</b>				
		<i>Baseline</i>	<i>Pell Grant</i>	<i>10% Adjust</i>	<i>20% Adjust</i>	<i>30% Adjust</i>
$L_\mu$	Avg Student Ability	1.28	1.26	1.27	1.29	1.31
$I_\mu$	Instr Spending per Student	0.98	1.04	1.13	1.21	1.30
$\Gamma_s$	Value-added	1.79	1.77	1.82	1.82	1.90
$\underline{\sigma}$	Admissions Threshold	1.02	1.17	1.20	1.23	1.25
$\kappa$	% Enrolled	32.09	32.14	30.9	29.66	28.53
<i>Income Distribution</i>						
	Q1 Income	0.22	0.23	0.23	0.22	0.22
	Q2 Income	0.20	0.22	0.21	0.21	0.21
	Q3 Income	0.31	0.31	0.31	0.32	0.32
	Q4 Income	0.27	0.24	0.25	0.25	0.25
<i>Ability Distribution</i>						
	0-30% Ability	0.01	0.06	0.05	0.04	0.03
	30-60% Ability	0.41	0.38	0.37	0.36	0.35
	60-90% Ability	0.40	0.39	0.40	0.41	0.41
	Top 10% Ability	0.17	0.17	0.18	0.19	0.21
<b>Overall welfare change</b>		-	1.97%	2.03%	2.09%	2.17%

## APPENDIX G: COMPUTATIONAL APPENDIX

This section briefly describes the general procedure for 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  $\tau$ . 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  $\{p_i(a, \ell)\}_{i \in \{AB, A1, A2\}}$ ,  $\tau$ ,  $\Gamma_s$ , 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 (3.4) and (??).
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 (3.15). 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  $\hat{\sigma}_s$  in (3.16)
  - (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 (3.8), (3.9), (3.10), (3.7)
4. Using updated guesses, find updated tax rate  $\hat{\tau}$  using (3.14)
5. Using all updated values and functions, find updated application decisions  $\{\hat{p}_i(a, \ell)\}_{i \in \{AB, A1, A2\}}$  as in (??)
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