RISKY COLLEGE ADMISSIONS AND THE (MIS)ALLOCATION OF TALENT

RICARDO MARTO
Federal Reserve Bank of St. Louis

YAACOV WITTMAN
Cornerstone Research

PRELIMINARY

Why are high-achieving, low-income students not attending selective colleges despite the generous financial aid typically offered? Using restricted-use data for the United States, we show that (1) low attendance rates are mostly explained by differences in application rates; (2) admission rates in selective colleges are U-shaped in parental income for high-achieving students; (3) even if admitted to selective colleges, low-income students are less likely to attend one. We build and estimate an equilibrium model of the college market to rationalize this evidence. Colleges compete by choosing admission standards and tuition schedules. Students can in turn apply to multiple colleges and are uncertain about their prospective admissions and financial aid. This uncertainty is the result of asymmetric information about a student's ability. We find that low-income students receive generous financial aid at selective colleges because only the highest-ability among them apply. Those who are not at the very top of the ability distribution apply less to selective colleges as they expect to either be rejected or receive little financial aid. If signals of ability became less informative (e.g. colleges stopped using the SAT), tuition would increase and high-ability students, irrespective of income, would be worse off—only high-income, low-ability students benefit.

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

JEL CLASSIFCATIONS: D43, D82, E21, E62, H52, I24.

Ricardo Marto: ricardo.marto@stls.frb.org.
Yaacov Wittman: ywittman@sas.upenn.edu.
We thank Lisa Barrow, Zachary Bleemer, Damien Capelle, Ciprian Domnisoru, Jesús Fernández-Villaverde, Grey Gordon, Jeremy Greenwood, Dirk Krueger, Oksana Leukhina, Alex Monge-Naranjo, Víctor Ríos-Rull, and Holger Sieg for their insightful comments and suggestions. We are also grateful to seminar participants at the FRB Chicago Applied Microeconomics conference, FRB Minneapolis OIG Institute research conference, FRB NY Equitable Growth conference, FRB St. Louis, Goethe University Frankfurt, Midwest Macro conference in Richmond, SED conference in Barcelona, Toulouse School of Economics Economics of Education Workshop, University of Pennsylvania, University of Warwick PhD conference, Washington University in St. Louis Graduate Student 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.[1] 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.[2] 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), Bénabou (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.[3]

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 8 focuses on the effect of increasing need-based financial aid such as the Pell Grants; and Section 9 concludes.

[1]Though Dale and Krueger (2014) do find returns to selectivity among disadvantaged students.

[2]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%.

[3]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 decisions of high-school students in the United States. We show that (1) low attendance rates of high-ability low-income students in selective colleges are mostly explained by differences in application rates; (2) admission rates in selective colleges are U-shaped in parental income for high-achieving students, i.e., higher for both lower and higher income students, and lower for students who are middle class; (3) even if admitted to selective colleges, low-income students are less likely to attend one.

2.1. *Data*

The analysis here presented relies on the High School Longitudinal Study (HSLS) of 2009. 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, including both public and private schools, who are followed throughout their secondary and postsecondary education. The students and their parents are first interviewed in 2009, then again in 2012 and 2013 once students are applying to college and graduating from high-school, and then once more in 2016. Math and science teachers, the school administrator, and the lead school counselor also completed surveys. This data 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 financial 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 (see Figure A.1 in the Appendix for the parental income distribution). For college preparedness, we use the students' high-school GPAs and SAT or ACT scores.[4] Each student's GPA is reported directly by the high school attended 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 the student's college and is therefore unavailable for students who did not attend college or did not take the test (see Figure A.2 for the grade distributions and Figure A.3 for their correlation with parental income, both in the Appendix).

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 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. Details about the HSLS sample are discussed in Appendix A.1.

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 refer to as "highly selective", corresponds to Barron's Tier 1 and 2 colleges. All other four-year non-profit colleges are counted in the second group, which we refer to as "non selective."

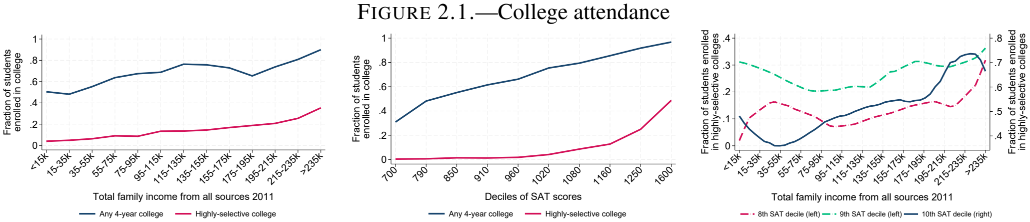
[4]The NCES converts the ACT score into an equivalent SAT score for students who only took the ACT instead of the SAT. Henceforth, whenever we mention the SAT, we refer to either the SAT or ACT score.

Table A.2 in the Appendix summarizes the key differences across these college types using the Integrated Postsecondary Education Data System (IPEDS) data. The highly selective colleges account for 16% of all four-year, non-profit enrollment and comprise 186 colleges. There are 1,577 less selective colleges that make up for the remainder of enrollment. 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. Additional details about each college type are presented in Appendix A.2.[5]

2.2. *College attendance*

Figure 2.1 shows how college attendance varies by parental income and SAT score. Panel (a) shows that wealthier students are more likely to attend any four-year non-profit college in general and a highly-selective college in particular than their less well-off peers. Similarly, students who score higher in standardized tests are more likely to attend a highly-selective college (see panel (b)). Panel (c) shows that top students (here, those who scored in the top 3 deciles of the SAT distribution) who are wealthier are more represented in highly-selective colleges than their peers whose parents earn less.

FIGURE 2.1.—College attendance



(a) By parental income (b) By SAT score (c) By parental income and SAT

Note: Panel (a) shows the fraction of students attending any four-year non-profit college (blue) and highly-selective colleges (pink) across students' parental income. Panel (b) shows the fraction of students attending any four-year non-profit college (blue) and highly-selective colleges (pink) across SAT score deciles. Panel (c) shows the fraction of students attending a highly-selective college across students' parental income and for different deciles of the SAT score (students in the 8th decile in pink, 9th decile in green, and 10th decile in blue). Figure A.5 in the Appendix replicates panel (b) and (c) using high-school GPA.

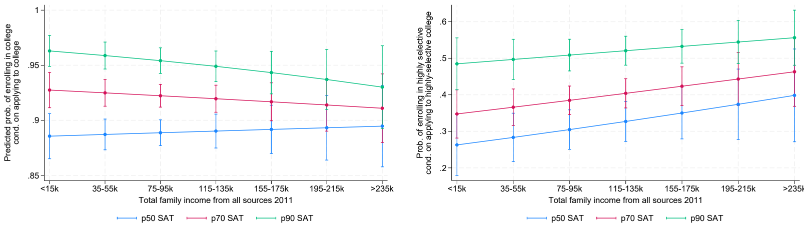
Figure 2.2 plots the expected probability of attending any four-year college (panel (a)) and a highly-selective college (panel (b)) across the parental income distribution and for different percentiles of the SAT distribution. After controlling for differences in students' demographic characteristics and conditional on applying to college, low-income high-ability students are less likely to attend a highly-selective college than their higher income peers—despite the similar probability of attending *any* college.[6] For instance, a student in the percentile 90 of the SAT score distribution whose parents make more than $200,000 is 7 percentage points more likely to attend a highly-selective college than a peer with the same SAT score but with parents who earn

[5]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.

[6]Table A.3 in the Appendix shows the results of the estimation of the logit model where the dependent variable is an indicator for whether or not the student is attending any four-year, non-profit college, conditional on applying to college (left column) and where the dependent variable is an indicator for whether or not the student is attending a highly-selective college, conditional on applying to a highly-selective college (right column).

less than $35,000 (conditional on both applying to a highly-selective college). That difference is more than 10 percentage points larger if we condition on students who applied to any college. Next, we explore the potential reasons for why low-income high-ability students are less likely to attend a highly-selective college.

FIGURE 2.2.—Probability of attending college



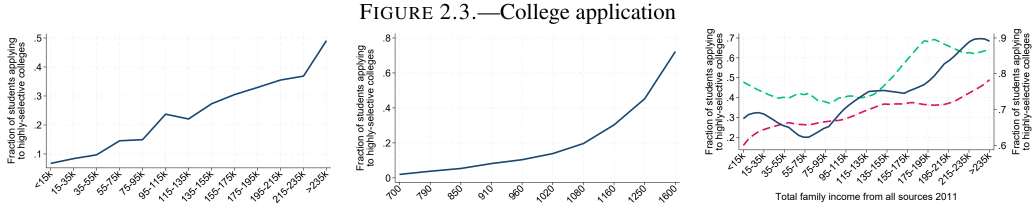
(a) Any four-year college (b) Highly-selective college

Note: Panel (a) shows the predicted probability of attending a four-year non-profit college conditional on applying to one across the parental income distribution for students in the percentile 50 (blue), 70 (pink), and 90 (green) of the SAT distribution. Similarly, panel (b) shows the predicted probability of attending a highly-selective college conditional on applying to a highly-selective college. Controls: high-school GPA, gender, race, parents' education, living situation, and employment, number of siblings, high-school type, urban, region, share of free-lunch. Sample is restricted to students who apply to a four-year college. Table A.3 in the Appendix presents the regression estimates.

2.3. *College applications*

The first step in understanding why low-income high-ability students are attending selective colleges at lower rates than their wealthier peers is to look at how they apply. Figure 2.3 (panel (a)) shows that about 40% of students whose parents earn more than $200,000 have applied to a highly-selective college, while less than 10% of students whose parents earn less than $55,000 did. Panel (b) shows that most students who score highly in the SAT apply to a least one highly-selective college. While applications are highly correlated with test scores, they also correlate with parental income conditional on SAT scores. As panel (c) shows, more than 80% of students in the top decile of the SAT distribution whose parents earn more than $175,000 applied to a highly-selective college, while less 67% of lower-income students did. These differences in shares across parental income are more noticeable for students in the 8th and 9th decile of the SAT score distribution.

FIGURE 2.3.—College application

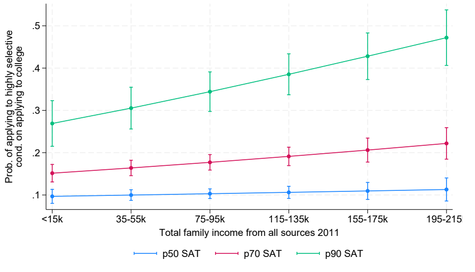


(a) By parental income (b) By SAT score (c) By parental income and SAT

Note: Panel (a) shows the fraction of students who applied to a highly-selective college across students' parental income. Panel (b) shows the fraction of students who applied to a highly-selective college across SAT score deciles. Panel (c) shows the fraction of students who applied to a highly-selective college across students' parental income and for different deciles of the SAT score (students in the 8th decile in pink, 9th decile in green, and 10th decile in blue). Figure A.6 in the Appendix replicates panel (b) and (c) using high-school GPA.

The probability of applying to a highly-selective college—conditional on applying to college—for a student in the 90th percentile whose parents earn more than $200,000 is 20 percentage points higher than a student with the same SAT score whose parents make less than \$35,000 as Figure 2.4 displays. This difference in application rates cannot be explained by demographic characteristics, such as high-school GPA, gender, or race.[7]

FIGURE 2.4.—Probability of applying to a highly-selective college

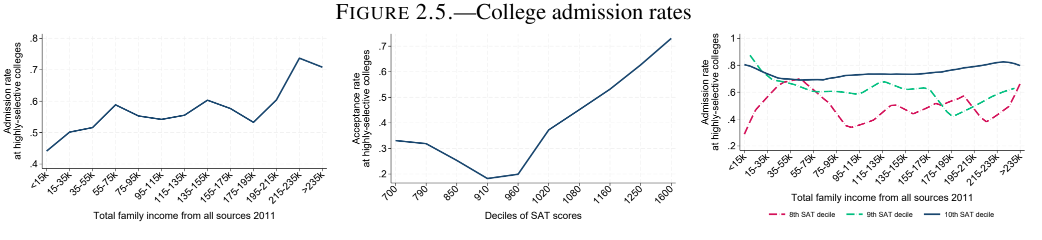


Note: The figure shows the predicted probability of applying to a highly-selective college conditional on applying to any college across the parental income distribution for students in the percentile 50 (blue), 70 (pink), and 90 (green) of the SAT distribution. Controls: high-school GPA, gender, race, parents' education, living situation, and employment, number of siblings, high-school type, urban, region, share of free-lunch. Sample is restricted to students who apply to a four-year college. Table A.4 in the Appendix presents the regression estimates.

2.4. *College admissions*

Admissions rates at highly-selective colleges are higher for wealthier students and for students with better SAT scores (viz. panels (a) and (b) of Figure 2.5). The admission rates are somewhat U-shaped in parental income for top SAT scorers, i.e., they are higher for both lower and higher income students, and lower for middle-class students. Controlling for demographic characteristics shows that there are no statistical differences in the probability of being admitted for top students across the parental income distribution. The upshot of the findings thus far is that differences in attendance in highly-selective colleges is driven by differences in application choices between high- and low-income students rather than differences in colleges' admission rates.

FIGURE 2.5.—College admission rates



(a) By parental income (b) By SAT score (c) By parental income and SAT

Note: Panel (a) shows the admission rates at highly-selective colleges for students who applied to a highly-selective college across students' parental income. Panel (b) shows the admission rates at highly-selective colleges for students who applied to a highly-selective college across SAT score deciles. Panel (c) shows the admission rates at highly-selective colleges for students who applied to a highly-selective college across students' parental income and for different deciles of the SAT score (students in the 8th decile in pink, 9th decile in green, and 10th decile in blue). Figure A.7 in the Appendix replicates panel (b) and (c) using high-school GPA.

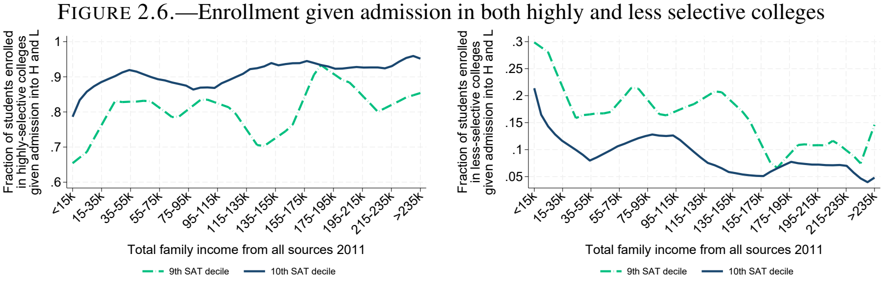
[7]Table A.4 in the Appendix shows the results of the estimation of the logit model where the dependent variable is an indicator for whether or not the student decided to apply to any four-year, non-profit college (left column) and where the dependent variable is an indicator for whether or not the student included any highly-selective college in her application portfolio conditional on students applying to college (right column).

2.5. *Enrollment conditional on being admitted*

We now show that even if low-income students are accepted to a highly-selective college, they are less likely to enroll in one. 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. The cost of attendance is claimed as one of the main reasons for not enrolling in a highly-selective college, which suggests that students were unaware or misinformed of the cost of attendance when they first applied or that they received an attractive offer from a less-selective college.[8]

Figure 2.6 presents the fraction of students enrolled in a highly-selective college but focuses on top students who were admitted at both highly and less selective colleges. Among students who were admitted to both a highly and a less-selective college, lower-income students are less likely to attend the highly-selective college and instead end up choosing to attend less-selective colleges.

FIGURE 2.6.—Enrollment given admission in both highly and less selective colleges



(a) Enroll in highly-selective college (b) Enroll in less-selective college

Note: Panel (a) shows the fraction of students attending a highly-selective college conditional on being admitted into both a highly and less selective college across students' parental income and for different deciles of the SAT score (students in the 9th decile in green and 10th decile in blue). Panel (b) shows the fraction of students attending a less-selective college given that they were admitted into both a highly and less selective college across students' parental income and for different deciles of the SAT score (students in the 9th decile in green and 10th decile in blue). Figure A.8 in the Appendix replicates panels (a) and (b) using high-school GPA.

3. MODEL

Motivated by the facts presented above, this section presents an equilibrium model of the college market with a realistic application and admissions system to study whether expectations about financial aid and admissions can explain the observed gaps in attendance in highly-selective colleges across the income distribution.

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

[8]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.4, which shows a positive relationship between admission rates and net tuition for low-income students.

level ℓ, 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 perceive ability differently. Colleges' perception of a student's ability is σ.[9] For lack of a better term, we call σ the ability signal. Students with a high enough realization of σ can receive multiple offers of admission and college-specific tuition levels that depend on her income, y, and ability signal, σ. Once the uncertainty is resolved and students know their admission 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.[10] While old, individuals work. Some of them will reap the benefits of attending college.

There are two tiers of colleges: highly selective, indexed by H, and less selective, indexed by L. There are N_H highly-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.[11]

The college market is monopolistically competitive. Colleges compete by setting different admission standards and tuition *schedules* for each student type (y, σ) they observe. Given the sequential nature of the college application process, colleges take into account the admission standards and tuition schedules of the other 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 if given an offer. 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-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.

[9]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 HSLS complete the FAFSA.

[10]For simplicity and in line with our empirical motivation, we do not separately model two-year colleges.

[11]The choice of a college's objective function is discussed below.

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.[12] When a student applies to college, colleges draw a realization of the applicant's ability, σ, that is unknown to her. Think of σ 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 the conditional density $g(\sigma|\ell)$ with cumulative distribution function $G(\sigma|\ell)$. The conditional density function is increasing in ℓ such that high signals are more likely to come from high-ability students.[13] A student is admitted to a college $s \in S = \{H_1, ..., H_{N_H}, L_1, ..., L_{N_L}\}$ if her signal is above the college's admission standard, i.e., $\sigma \geq \underline{\sigma}_s$.

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.[14] 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 ℓ 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 (i.e., $|S| = N_H + N_L$) is given by[15]

$$V^A(y,\ell,S;\boldsymbol{T}(y,\sigma),\underline{\boldsymbol{\sigma}}) = \underbrace{\int_{\underline{\sigma}_{H_1}}^{\infty} V^O(y,\ell,N_H+N_L,S;\boldsymbol{T}(y,\sigma))g(\sigma|\ell)d\sigma}_{\substack{\text{Expected value of being admitted to colleges in } S \\ \text{including an offer from the most selective one}}} \quad + ... +$$

$$\underbrace{\int_{\underline{\sigma}_{L_1}}^{\underline{\sigma}_{H_{N_H}}} V^O(y,\ell,N_H+N_L,S\setminus\{H_1,...,H_{N_H}\};\boldsymbol{T}(y,\sigma))g(\sigma|\ell)d\sigma}_{\substack{\text{Expected value of being admitted to colleges in } S \\ \text{excluding all highly-selective colleges}}} \quad + ... +$$

[12]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.

[13]For simplicity, colleges agree on the same signal of ability. 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).

[14]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.

[15]Without loss of generality, we rank colleges within their tiers in ascending order in terms of admission standards, i.e., $\underline{\sigma}_{j_1} \geq ... \geq \underline{\sigma}_{j_{N_j}}$ for $j = \{H, L\}$ and consider the case in which the worst highly-selective college is at least as selective as the best less selective college, i.e., $\underline{\sigma}_{H_{N_H}} \geq \underline{\sigma}_{L_1}$. 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.

$$\underbrace{\int_{\underline{\sigma}_{L_{N_L}}}^{\underline{\sigma}_{L_{N_L-1}}} V^O(y,\ell,N_H+N_L,\{L_{N_H}\};\boldsymbol{T}(y,\sigma))g(\sigma|\ell)d\sigma}_{\substack{\text{Expected value of only being admitted}\\ \text{to the least selective college}}} \quad +$$

$$\underbrace{G\left(\underline{\sigma}_{L_{N_L}}|\ell\right)V^W(y,\ell,N_H+N_L)}_{\substack{\text{Expected value of not being admitted to any}\\ \text{college after submitting }(N_H+N_L)\text{ applications}}} \quad -\underbrace{\phi(N_H+N_L,S)}_{\text{Psychic application cost}}, \tag{3.1}$$

where $V^O(y,\ell,n,O;\boldsymbol{T}(y,\sigma))$ 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 $A \subseteq \mathbb{P}(S)$, $\boldsymbol{T}(y,\sigma)$ is a vector of tuition fees for an admitted student with parental income y and signal σ, 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.[16]

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

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

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 admission and financial aid offer, all uncertainty is resolved. At that stage, students decide which offer, if any, to accept.

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 despite having paid the application fees. Suppose, for instance, that the set of offers the student receives includes the least selective college L_1 through the n^{th} highly-selective college, or H_n. Then, her value after applying to (N_L+n) colleges is given by

$$V^O(y,\ell,N_L+n,O;\boldsymbol{T}(y,\sigma)) = \max\Big\{\underbrace{V^{L_1}(y,\ell,N_L+n,T_{L_1}(y,\sigma)),...,V^{H_n}(y,\ell,N_L+n,T_{H_n}(y,\sigma)),}_{\substack{\text{Values of attending one of the }|O|\text{ colleges}\\ \text{after submitting }N_L+n\text{ applications}}}$$

$$V^W(y,\ell,N_L+n)\Big\}, \tag{3.3}$$

---

[16]For instance, a student that only applies to the least selective college as the following value

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

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

Demand for college. 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)), where $\lambda_a > 0$ and $\lambda_c > 0$ are the scale parameters of the mean zero Type I extreme value shocks. Including these shocks adds analytical tractability as they allow demand for college to be continuously differentiable. That is particularly useful as colleges take demand as given when setting their tuition schedules.

Demand for college s of a student with characteristics (y, ℓ) is given by the enrollment probability $q_s(y, \ell, \boldsymbol{T}(y, \sigma), \underline{\boldsymbol{\sigma}})$. That probability depends on in turn on the probability of applying to a particular set of colleges i 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

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

Section 3.5 will revisit this demand function in greater detail.

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 working life.

A college-educated individual has preferences over consumption and discounts the future at rate $1/\beta$. There is a utility cost from attending college, denoted by $\nu_s(\ell) \geq 0$, which is weakly decreasing in the student's ability as higher-ability students might find attending college easier than lower-ability students. The utility 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.

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 $\underline{a}_s$.[17] Grants depend on the student's income and can be funded by the government, $P(y)$, or outside sources, $Gr(y)$. Government-funded grants are meant to capture the federal Pell Grant program and are paid for using labor income taxes on workers, while outside grants represent private scholarships taken exogenously.

After graduating from college, the student enters the labor market and pays back her student loan priced at the interest rate R. Labor earnings correspond to the product of the wage rate w net of taxes τ and the individual's human capital. If the student graduates, which happens with probability δ_s, her human capital is $Z_s \ell^{\alpha_s}$, where $Z_s > 1$ denotes the value added of

[17]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.

the college attended and the exponent $\alpha_s < 1$ governs the returns to ability ℓ. If the student fails to graduate, which happens with probability $(1-\delta_s)$, her human capital is ℓ^{α_w}, where $\alpha_w < \alpha_s < 1$ implies that the return from attending college is larger, the higher is the ability of the individual (for $\ell > 1$).

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

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

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

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 applied to college.

Value of working without a college degree. $V^W(y,\ell,n)$ is the value of an individual with parental transfers y and ability ℓ 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

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

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

3.2. *Colleges*

The college market is a monopolistically competitive industry. Colleges 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' demand for college (equation (3.4)).

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

$$Z = \xi\, Q(I/\kappa, L/\kappa), \tag{3.7}$$

where $Q(I/\kappa, L/\kappa)$ denotes the college's quality and ξ 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/κ, and the average ability of its students, L/κ. 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, I, is directly chosen by the college, while the ability of its students, L, and the number of students enrolled, κ, are the outcome of admission standards and tuition schedules of all colleges, $\boldsymbol{T}(y,\sigma)$ and $\underline{\boldsymbol{\sigma}}$. Using students' demand for college (equation (3.4)), we can write the number of students enrolled in a college as the sum of all students who cleared the bar and accepted the college's offer of admission, i.e.,

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

where the integral is taken over all student types (y,ℓ,σ) given the college's choice of admission standard $\underline{\sigma}$. 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

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

Budget constraint. 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. 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). Endowment income, government transfers, and operating expenses increase with the number of students enrolled κ. The budget constraint of a college is thus given by

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

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.,

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

College 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 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.[18] 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. 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

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

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

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

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

The tuition cap, $\overline{T}$, is the maximum price students have to pay (the sticker price). Its introduction avoids 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.

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

$$N = \underbrace{2\int \ell^{\alpha_w}\, q_w(y,\ell)\, d\hat{\mu}(y,\ell)}_{\substack{\text{Labor supplied by young and old}\\ \text{high-school workers}}} +$$

$$\underbrace{\sum_{s\in S}\int_{\underline{\sigma}_s}^{\infty} \left[\delta_s Z_s \ell^{\alpha_c} + (1-\delta_s)\ell^{\alpha_w}\right] q_s(y,\ell,\boldsymbol{T}(y,\sigma),\underline{\boldsymbol{\sigma}})\, d\mu(y,\ell,\sigma)}_{\text{Labor supplied by college-educated workers}}, \tag{3.13}$$

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

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

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

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

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. 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, $q_s(y, \ell, \boldsymbol{T}(y, \sigma), \underline{\boldsymbol{\sigma}})$; (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 κ_s; (8) the labor income tax rate τ; 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, $q_s(y, \ell, \boldsymbol{T}(y, \sigma), \underline{\boldsymbol{\sigma}})$, satisfies equation (3.4).
4. Given the tax rate, τ, the college value added, Z_s, tuition, $T_s(y, \sigma)$, 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, τ, 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, $q_s(y, \ell, \boldsymbol{T}(y, \sigma), \underline{\boldsymbol{\sigma}})$, admission standards and tuition schedules of all the competing colleges, $\underline{\boldsymbol{\sigma}}$ and $\boldsymbol{T}(y, \sigma)$, the college value added Z_s solves the problem (3.12) and $\underline{\sigma}_s, T_s(y, \sigma), I_s, L_s$, and κ_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 tuition schedules (recall each tuition level is a best response to the tuition levels offered by the competing colleges).

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.[19]

[19]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 $\left[\sum_{k=0}^{1} C(N_H, k)\right]\left[\sum_{k=0}^{1} C(N_L, k)\right] = (N_H + 1)(N_L + 1)$.

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

$$\max \underbrace{\left\{V^A(y,\ell,1,0), V^A(y,\ell,0,1), V^A(y,\ell,1,1), ..., V^A(y,\ell,N_H,N_L), V^W(y,\ell,0)\right\}}_{\text{Expected values of the } (N_H+1)(N_L+1) \text{ application portfolios}}, \quad (3.15)$$

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 η_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 η_L for whether the student got at least one admission offer from a less-selective college (i.e. $\eta_L = 1$).[20] The expected value of applying to $n_H > 0$ highly-selective colleges and $n_L > 0$ less-selective colleges is then given by

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

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

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

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

$$V^O(y,\ell,n_H+n_L,\eta_H,\eta_L;\boldsymbol{T}(y,\sigma)) = \max \left\{ \underbrace{\eta_H V^H(y,\ell,n_H+n_L,T_H(y,\sigma)),}_{\substack{\text{Value from attending a highly-selective college} \\ \text{after submitting } n_H+n_L \text{ applications}}} \right.$$

[20] 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.

$$\underbrace{\eta_L V^L(y,\ell,n_H+n_L,T_L(y,\sigma)),}_{\substack{\text{Value from attending a less-selective college}\\ \text{after submitting } n_H+n_L \text{ applications}}}$$

$$\left.\underbrace{V^W(y,\ell,n_H+n_L)}_{\substack{\text{Value from not attending colleges}\\ \text{after submitting } n_H+n_L \text{ applications}}}\right\}. \tag{3.16}$$

Demand for college. 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 ℓ, we have

$$\Pr(\text{Applying to } n_H \text{ and } n_L) = e^{\lambda_a V^A(y,\ell,n_H,n_L)} / \Big[e^{\lambda_a V^A(y,\ell,1,0)} + e^{\lambda_a V^A(y,\ell,0,1)} + e^{\lambda_a V^A(y,\ell,1,1)}$$
$$+ \ldots + e^{\lambda_a V_L^A(y,\ell,N_H,N_L)} + e^{\lambda_a V^W(y,\ell,0)}\Big]. \tag{3.17}$$

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

$$\Pr(\text{Accept } s \text{ offer}|\text{App. to } n_H+n_L,\eta_H,\eta_L) = e^{\lambda_c \eta_s V^s(y,\ell,n_H+n_L,T_s(y,\sigma))} /$$
$$\Big[e^{\lambda_c \eta_H V^H(y,\ell,n_H+n_L,T_H(y,\sigma))}$$
$$+e^{\lambda_c \eta_L V^L(y,\ell,n_H+n_L,T_L(y,\sigma))}$$
$$+e^{\lambda_c V^W(y,\ell,n_H+n_L)}\Big]. \tag{3.18}$$

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

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

$$\underbrace{\Pr(\text{Accept } H \text{ offer}|\text{App. to } n_H \text{ only}, \eta_H = 1)}_{\substack{\text{Probability of accepting a highly-selective college offer after only applying} \\ \text{to highly-selective colleges and being accepted } (\eta_L = 0)}}. \tag{3.19}$$

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

$$q_L(y, \ell, \boldsymbol{T}(y,\sigma), \underline{\boldsymbol{\sigma}}) = \underbrace{\frac{1}{N_L} \sum_{n_H=1}^{N_H} \sum_{n_L=1}^{N_L} \Pr(\text{Applying to } n_H \text{ and } n_L) \times}_{\substack{\text{Probability of applying to a less-selective college} \\ \text{in addition to highly-selective colleges}}}$$

$$\underbrace{\Pr(\text{Accept } L \text{ offer}|\text{App. to } n_H + n_L, \eta_H = 0, \eta_L = 1)}_{\substack{\text{Probability of accepting a less-selective college offer after applying} \\ \text{to both types of colleges and being accepted only to the less selective}}}$$

$$+ \quad \underbrace{\frac{1}{N_L} \sum_{n_L=1}^{N_L} \Pr(\text{Applying to } n_L \text{ only})}_{\text{Probability of only applying to a less-selective college (i.e., } n_H = 0)} \quad \times$$

$$\underbrace{\Pr(\text{Accept } L \text{ offer}|\text{App. to } n_L \text{ only}, \eta_L = 1)}_{\substack{\text{Probability of accepting a less-selective college offer after only applying} \\ \text{to less-selective colleges and being accepted } (\eta_H = 0)}}. \tag{3.20}$$

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

$$T_s(y, \sigma) = \underbrace{m_s(y, \sigma)}_{\text{Markup}} \times \underbrace{MC_s(y, \sigma)}_{\text{Marginal cost}}. \tag{3.21}$$

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

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

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

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, σ) is given by

$$MC_s(y,\sigma) = \underbrace{C_{s_\kappa} - \frac{Q_\kappa}{Q_I} - E_{s_\kappa} - Tr_{s_\kappa}}_{\substack{\text{Costs net of transfers} \\ \text{per student}}} - \underbrace{\frac{Q_L}{Q_I}\widehat{\mathbb{E}}_s[\ell|y,\sigma]}_{\text{Posterior ability discount}},$$

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

Since higher ability levels yield higher average signals, the (posterior) ability discount will be increasing in σ. 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 admissions standard. Lastly, the optimal admission standard of a college defines that the average tuition revenue received from the lowest-signal students must be enough to compensate the college for the marginal cost of schooling them. The optimal admissions standard of college s satisfies the following condition

$$\underbrace{\frac{\displaystyle\int T_s(y,\underline{\sigma}_s)\, q_s(y,\ell,\boldsymbol{T}(y,\underline{\sigma}_s),\underline{\boldsymbol{\sigma}}_s)\mu(dy,d\ell,\underline{\sigma}_s)}{\displaystyle\int q_s(y,\ell,\boldsymbol{T}(y,\underline{\sigma}_s),\underline{\boldsymbol{\sigma}}_s)\mu(dy,d\ell,\underline{\sigma}_s)}}_{\text{Average revenue from lowest-signal students}} = \underbrace{C_{s_\kappa} - \frac{Q_\kappa}{Q_I} - E_{s_\kappa} - Tr_{s_\kappa}}_{\substack{\text{Costs net of transfers} \\ \text{per student}}} - \underbrace{\frac{Q_L}{Q_I}\mathbb{E}_s[\ell|\underline{\sigma}_s]}_{\text{Posterior average ability}}, \tag{3.22}$$

where $\mathbb{E}_s[\ell|\underline{\sigma}_s] \equiv \frac{\int \ell q_s(y,\ell,\boldsymbol{T}(y,\underline{\sigma}_s),\underline{\boldsymbol{\sigma}}_s)\,\mu(dy,d\ell,\underline{\sigma}_s)}{\int q_s(y,\ell,\boldsymbol{T}(y,\underline{\sigma}_s),\underline{\boldsymbol{\sigma}}_s)\,\mu(dy,d\ell,\underline{\sigma}_s)}$ is the average innate ability of the lowest-signal students. Its derivation is provided in Appendix B.2.

Solution algorithm 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 B.3 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 taken directly from their data counterparts, while the remaining parameters are estimated using the simulated method of moments. Table 4.1 summarizes the parameters chosen outside the model and Table 4.4 summarizes the resulting estimated parameters.

In terms of timing, we let the first period account for 4 years (time spent in college) and the second period for 60 years (time spent working). Appendix C.1 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 and w.[21]

4.1. *Functional forms*

We describe the functional forms assumed for individuals' preferences, application costs, 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.[22] 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. This in turn allows to make the value of an applicant with offers of admissions (equation (3.16)) and the probability of accepting an offer (equation (3.18)) independent of the number of applications submitted.

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

$$\phi(n, H, L) = \begin{cases} \phi_H & \text{if only applied to highly-selective colleges} \\ \phi_L & \text{if only applied to less-selective colleges} \\ \phi_B & \text{if applied to both highly and less-selective colleges.} \end{cases}$$

[21] As the first period in the model corresponds to four years, the amounts computed are rescaled by \$40,000. For example, $y = 1$ in the model corresponds to an Expected Family Contribution of \$10,000 per year over four years.

[22] 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.

This assumption makes the value of applying to college independent of the number of applications submitted, which collapses the optimal application portfolio (equation (3.15)) into four discrete alternatives: only applying to highly-selective colleges, only applying to less-selective colleges, applying to both college types, and not applying at all. As a result, the number of applications and the number of colleges within each type is irrelevant when solving for the enrollment probabilities (equations (3.19) and (3.20)). However, it is still necessary to keep track of which college tier students apply to and receive offers from.

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 ρ_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. Endowment income 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, which in turn are given by

$$C_s(\kappa) = C_{0_s}.$$

Need-based aid. In order to qualify for Pell grants, students must demonstrate sufficient financial need as measured by the difference between their expected family contribution (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. Pell grants are modeled as follows

$$P(y) = \max\{P_0 - y, 0\}.$$

The exogenous private grants and scholarships, $Gr(y)$, take the following form

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

4.2. *Data*

Three datasets are used to estimate the model: (i) the Integrated Postsecondary Education Data System (IPEDS); (ii) the High School Longitudinal Study of 2009 (HSLS); (iii) the 2012 cohort of the Beginning Postsecondary Students Longitudinal Study (BPS). The first two datasets were discussed in Section 2.1.

The IPEDS dataset is used to estimate college-level aggregates. 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. The HSLS dataset is used to infer the distribution of students' characteristics and to calculate key moments regarding application and enrollment into highly and less-selective colleges.

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 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.

4.3. *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, $\{\beta, \alpha_s, \alpha_w, \delta_H, \delta_L\}$, borrowing constraints, grants, and financial application costs, $\left\{\underline{a}_s, \underline{a}_w, P_0, g_0, g_1, \underline{g}, \psi_0\right\}$, and colleges budget constraint $\left\{C_{0_H}, C_{0_L}, E_{1_H}, E_{1_L}, E_{2_H}, E_{2_L}, Tr_{1_H}, Tr_{1_L}, \overline{T}_H, \overline{T}_L\right\}$. Aggregate prices, $\{R, w\}$, are taken exogenously.

Distribution of characteristics. The distribution of student characteristics over income and ability, $\hat{\mu}(y,\ell)$, is informed by the HSLS data. Parental income, y, corresponds to the student's Expected Family Contribution (EFC). The EFC 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.[23] 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.2. Students' innate ability, ℓ, corresponds to the residuals of a regression of SAT scores on demographic variables. Appendix C.3 provides additional details on the joint distribution of parental income and ability.

Returns to ability. We take the parameters governing the labor market returns to ability, α_s and α_w, directly from Abbott, Gallipoli, Meghir, and Violante (2019). They find that the ability gradient for wages is higher for college graduates than for high school graduates (0.797 vs. 0.517 for men and 0.766 vs. 0.601 for women). We take the midpoint between males and females for each case. We set $\alpha_s = 0.782$ for college graduates and let this value be the same across college tiers, and set $\alpha_w = 0.559$ for students who did not attend college.

Dropout risk. We set the probability of dropping out of college to zero in both college tiers, so that $\delta_H = \delta_L = 1$ (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 $\underline{a}_s = -0.775$, which corresponds to \$31,000 over four years. Since there is no risk for workers who did not pursue a college

[23]The EFC is determined according to rules set by the Department of Education using the FAFSA filled out by students and their parents.

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 $P_0 = 0.5645$. Using the BPS, we set the minimum level of private grants to $\underline{g} = 0.15$, 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.4.

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 highly-selective colleges have lower fixed costs than less-selective colleges with $C_{0_H} = 0.13$ and $C_{0_L} = 0.25$. The linear terms in the tier-specific endowment function are $E_{1_H} = 1.37$ for highly-selective colleges and $E_{1_L} = 0.16$ for less-selective colleges, which implies that highly-selective colleges earn higher endowment income per enrolled student than less-selective colleges. The quadratic terms in the endowment function are $E_{2_H} = 5.77$ for highly-selective colleges and $E_{2_L} = 0.44$ for less-selective colleges.

We estimate the linear relationship between government transfers and student enrollment across college tiers using the IPEDS sample data. The resulting estimates imply that highly-selective colleges receive more government subsidies per enrolled student than less-selective colleges (about \$10,440 vs. \$6,200) and therefore set $Tr_{1_H} = 1.04$ and $Tr_{1_L} = 0.62$.

The BPS data is used to pin down the tuition caps across college types $\overline{T}_s$. We find the average tuition net of college-specific financial aid flattens for higher-income students at about \$25,000 per year at highly-selective colleges and at \$12,000 per year at less-selective colleges. Tuition caps correspond to that upper bound and thus set $\overline{T}_H = 2.5$ and $\overline{T}_L = 1.2$.

Appendix C.5 provides additional details about the costs, endowments, government transfers, and tuition caps.

Aggregate prices. The gross interest rate is set to $R = (1.0386)^4$, which corresponds to the annual borrowing rate for undergraduate students in the 2013-2014 academic year (3.86%). This rate was set by the U.S. Department of Education 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.

4.4. *Internally estimated*

The remaining parameters relate to the disutility of applying to colleges, $\{\phi_H, \phi_L, \phi_B\}$, the disutility of attending college, $\{\nu_{0_H}, \nu_{0_L}, \nu_{1_H}, \nu_{1_L}\}$, the extreme value scale parameters for applying and enrolling in colleges, $\{\lambda_a, \lambda_c\}$, the distribution of colleges' perceived student ability, $g(\sigma|\lambda)$, college productivity and quality elasticity, $\{\xi_H, \xi_L, \rho_L\}$. These parameters, let Θ denote them, are jointly estimated by minimizing the unweighted distance between data moments and simulated model moments according to

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

TABLE 4.1

EXTERNALLY ESTIMATED PARAMETERS

Variable	Description	Value	Source
Preferences			
β	Discount factor	5.48	USDE
Returns to ability and dropout risk			
α_s	Returns to ability (college degree)	0.78	Abbott et al. (2019)
α_w	Returns to ability (high-school degree)	0.55	Abbott et al. (2019)
Borrowing constraints, grants and application costs			
$\underline{a}_s$	Student borrowing limit	-0.775	USDE
$\underline{a}_w$	Worker borrowing limit	0	Normalization
P_0	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
ψ_0	Financial application cost	0.00375	IPEDS
College costs and revenues			
$C0_H, C0_L$	Fixed costs in colleges H, L	0.13, 0.25	IPEDS
$E1_H, E1_L$	Net endowment per student in colleges H, L	1.37, 0.16	IPEDS
$E2_H, E2_L$	Quadratic term in net endowment in colleges H, L	5.77, 0.43	IPEDS
$Tr1_H, Tr1_L$	Gov transfers per student in colleges H, L	1.04, 0.62	IPEDS
$\overline{T}_H, \overline{T}_L$	Tuition cap in colleges H, L	2.50, 1.20	BPS

where M_j^{data} denotes the jth empirical moment out of J moments and $M_j^s(\Theta)$ denotes the jth model-calculated moment implied by parameters Θ for simulation $s = \{1, \ldots, S\}$, with $S = 100$. The model is estimated using $N = 15,000$ students drawn from the distribution of student characteristics. Table 4.4 presents the estimated parameters and Table 4.5 contrasts the moments implied by the estimated parameters with their data counterparts.

Application disutility costs. The disutility of applying to colleges, ϕ_H, ϕ_L, and ϕ_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 4.2 reports the average consumption equivalent values as a percentage of average life-cycle consumption for all students. The estimated costs are 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.

Attendance disutility costs. The 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 (ν_{0_H} and ν_{0_L}), while variation in student applications across the SAT distribution helps identify the slope parameters (ν_{1_H} and ν_{1_L}). The estimated coefficients reveal substantial differences in preferences for each college across the ability distri-

TABLE 4.2

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).

bution. 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 calculate the consumption equivalence for each student and report them as a percentage of average life-cycle consumption in Table 4.3. Attendance 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, λ_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 collleges. 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, λ_a, influences how many applications students submit. Variation in the number of college applications submitted identifies the value of the scale parameter.

TABLE 4.3

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.

College perceived ability. Ability signals colleges receive follow a normal distribution conditional on ability, centered around it and with variance σ_g^2, i.e. $\sigma|\ell \sim N(\ell, \sigma_g^2)$. The distribution is truncated at 0, so that the lower bound of the support of signals is finite. The variance of the ability signal, σ_g^2, is chosen to match the average responsiveness of tuition with respect to ability. This effect is estimated using the BPS data by regressing tuition (minus institutional grants) on SAT score, controlling for EFC and college fixed effects.

College efficiency. Colleges' efficiency parameter, ξ_H and ξ_L, are chosen to match a college wage premium of 0.6, in line with Abbott, Gallipoli, Meghir, and Violante (2019), and the estimates of Chetty, Friedman, Saez, Turner, and Yagan (2020) showing that 80% of the difference in median log earnings 10 years after college can be explained by differences in colleges' selectivity.

College quality elasticity. The elasticity of college quality with respect to students' ability, ρ_L, is chosen to match the average enrollment rate in highly-selective colleges, conditional on having applied. A higher value of ρ_L increases the size of the ability discount, which increases the attractiveness of applying to highly-selective colleges.

TABLE 4.4

INTERNALLY ESTIMATED PARAMETERS

Variable	Description	Value	Identification
Preferences			
ϕ_H, ϕ_L	Disutility of applying to colleges H, L	0.100, 0.185	Application rates
ϕ_B	Disutility of applying to both college tiers	0.320	Application rates
$\nu 0_H, \nu 0_L$	Disutility of attending colleges H, L (intercept)	2.410, 1.700	Attendance rates
$\nu 1_H, \nu 1_L$	Disutility of attending colleges H, L (slope)	-1.467, -0.758	Attendance rates by SAT
λ_a	Extreme value scale parameter when applying	0.400	Number of applications
λ_c	Extreme value scale parameter when choosing offers	1.855	Attendance rates
Technology and costs			
σ_g^2	Variance of ability signal	0.110	Elasticity of tuiton wrt SAT
ξ_H, ξ_L	Value added efficiency in colleges H, L	1.060, 1.453	College wage premium
ρ_L	Quality elasticity of ability	0.854	Acceptance rate

TABLE 4.5

TARGETED MOMENTS: DATA VS. MODEL

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

Note: "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 attnd 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. "Δ 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

		Highly-selective colleges		**Less-selective colleges**	
		Model	*Data*	*Model*	*Data*
κ_s	% Enrolled	7.2	7.0	32.1	36.0
$\mathcal{T}/\kappa_s$	Average net tuition ($1k/year)	16.0	14.1	6.1	5.0
I/κ_s	Instr. spending per student ($1k/year)	21.0	22.5	9.8	8.9
Y/κ_s	Average student EFC ($1k/year)	22.2	30.4	11.4	15.5
Σ/κ_s	Average applicant SAT (std)	1.45	1.12	0.29	0.05
y	*Income Distribution (%)*				
	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

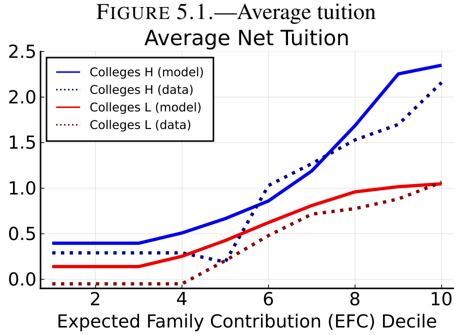
Note: All statistics are untargeted (except enrollment in highly-selective colleges). Income refers to parental transfers in the model and Expected Family Contribution (EFC) in the data. The average signal (in the model) and SAT score (in the data) among applicants are standardized for comparison. Average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

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 ρ_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 σ 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 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.

FIGURE 5.1.—Average tuition



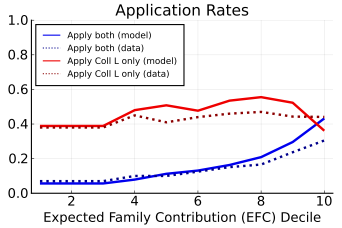
Note: The figure shows the average net tuition (sticker-price tuition minus all grants) for students by Expected Family Contribution decile. Source: BPS.

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.

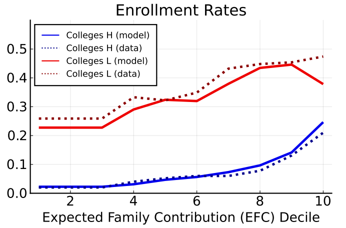
5.2. *Equilibrium effect of applications on low-income student enrollment*

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

FIGURE 5.2.—College application, admission, and enrollment rates



(a) Application rates

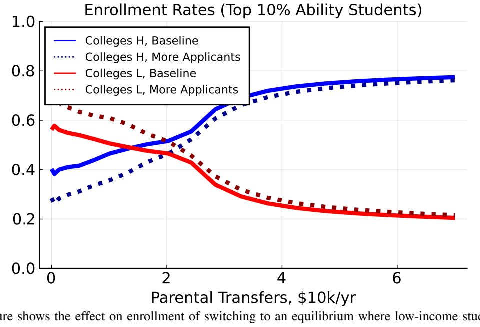


(b) Enrollment rates

Note: Panel (a) displays the average application rates for students by Expected Family Contribution (EFC) decile. It only includes 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. Panel (c) displays the average enrollment rates to each college. Source: HSLS.

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.[24] 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 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 D.1.

FIGURE 5.3.—Average enrollment rates for students in the top 10% of the ability distribution



Note: The figure 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.

[24]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.

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.

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.[25]

Table D.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.

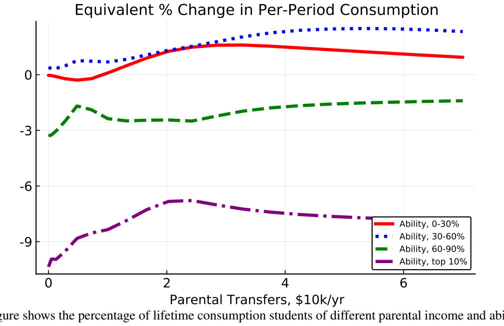
To see which students are affected by the decrease in the informativeness of the signals, Figure 6.1 plots the percent of consumption in each period different types of students would be willing to forgo in order to be born in the new equilibrium.[26] As expected, high-ability students are the most hurt 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.

[25]This value minimizes average welfare in comparison to the baseline equilibrium. Figure D.1 in the appendix plots the change in welfare for a range of possible increases in the signal variance.

[26]For reference, Figure D.2 in the Appendix plots the changes to student enrollment resulting from the new equilibrium.

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 D.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.—Equivalent variation when the variance of the ability signal increases



Note: The figure shows the percentage 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.[27] The outcome of this experiment is presented in Table D.3 and the welfare effects are presented in Figure D.3. 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 D.4. If signals were perfectly informative, the effective marginal cost of enrolling a high-ability student decreases

[27]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.

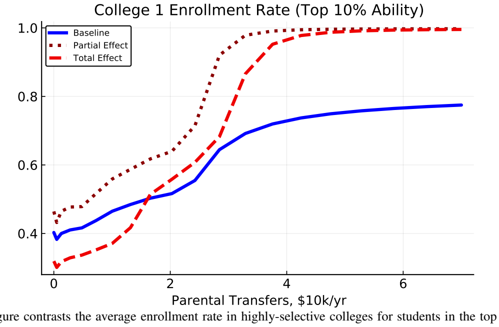
substantially because colleges can perfectly tell them apart. This will lower tuition for high-ability students and raise the average ability of the student body, thereby increasing the effective marginal cost for relatively low-ability students. As the effective marginal cost of enrolling high-income, 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 D.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 for students in the top 10% of the ability distribution



Note: The figure contrasts the average enrollment rate in highly-selective colleges for students in the top 10% of the ability distribution with perfect vs. imperfect information. 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. RELAXING BORROWING CONSTRAINTS

The role of the borrowing constraint on students' ability to finance college is now analyzed. The Federal Student Loan Program is capped at $31,000, on average $7,750 per year, for undergraduate dependent students. That constraint binds for several students. In this section, we study the effect of relaxing the borrowing constraint on the college market equilibrium by setting the borrowing constraint at $20,000 per year (i.e., $\underline{a}_s = -2.0$, while keeping all other parameters at their baseline).

Relaxing the borrowing constraint, increases the value of attending college for students that were previously financially constrained. This makes attending college better than working right after high school, i.e. V^H and V^L become larger relative to V^W. This means that the probability of accepting a college offer increases. As attending college becomes better, the value of an offer in hand is also more valuable, which in turn increases the value of applying to college, i.e. the difference between V^A and V^W is now larger. In turn, this increases the probability of applying to college for students that were not applying for college because of the borrowing constraint. The increase in the probability of applying and of accepting an offer if given one lead to an increase in the demand for college, i.e., the enrollment probability q_H and q_L go up.

Table 7.1 shows the quantitative effect of relaxing the borrowing constraint on college-level variables. By allowing students to borrow more, demand for college increases. As a result, lower income students are now more likely to enroll in both highly and less-selective colleges. Students below the median parental income make up 22.2% and 44.8% of students in highly and less-selective colleges (as opposed to 21.4% and 42.6%). Despite this compositional change, the fraction of students who enroll in college stays about the same. As a result of the increased ability to borrow of lower-income students, the ability distribution in highly-selective colleges improves. This leads to an increase of the average ability of the student body of highly-selective colleges. The complementarity between ability and instructional spending leads to an increase in the average amount spent educating students in highly-selective college. This in turn is compensated by in an increase in tuition.

TABLE 7.1

EFFECT OF RELAXING STUDENTS' BORROWING CONSTRAINT

		Highly-selective colleges		Less-selective colleges	
		Baseline	*↓ Borrowing Const.*	*Baseline*	*↓ Borrowing Const.*
κ_s	% Enrolled	7.2	7.1	32.1	32.2
$\mathcal{T}/\kappa_s$	Average net tuition ($1k/year)	16.0	20.9	6.1	6.5
I/κ_s	Instr spending per student ($1k/year)	21.0	25.8	9.8	10.4
Z_s	Value added	2.01	2.08	1.79	1.80
y	*Income distribution (%)*				
	Q1 Income	9.8	10.3	22.1	23.0
	Q2 Income	11.6	11.9	20.5	21.8
	Q3 Income	26.4	28.7	30.6	31.5
	Q4 Income	52.2	49.1	26.8	23.7
ℓ	*Ability distribution (%)*				
	0-30% Ability	0.0	0.0	1.1	2.7
	30-60% Ability	0.4	0.1	41.2	40.0
	60-90% Ability	15.1	13.3	40.3	40.1
	Top 10% Ability	84.5	86.6	17.4	17.2

8. 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 8.1 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 D.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.

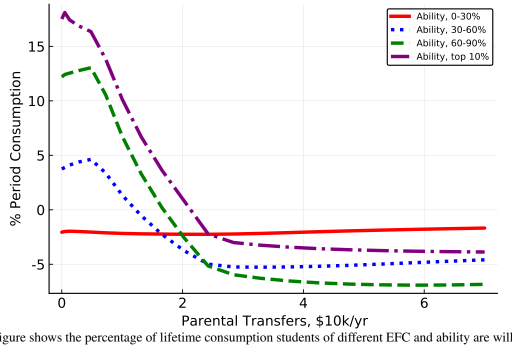
Finally, the welfare effects of the policy change are examined in Figure 8.1. As expected, the students who benefit from the policy are the relatively low-income students. They benefit directly from the increased consumption while in college (which also alleviates the effect of the credit constraint) 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%.

TABLE 8.1

EFFECT OF INCREASING PELL GRANTS

		Highly-selective colleges			Less-selective colleges		
		Baseline	$\downarrow \underline{a}_s$	$\uparrow$ *Pell grants*	*Baseline*	$\downarrow \underline{a}_s$	$\uparrow$ *Pell grants*
κ_s	% Enrolled	7.2	7.1	7.2	32.1	32.2	32.1
I/κ_s	Instr spending per student ($1k/year)	21.0	25.8	26.7	9.8	10.4	10.4
Z_s	Value added	2.01	2.08	2.09	1.79	1.80	1.77
τ	Tax rate (%)	1.51	1.51	3.93	1.51	1.51	3.93
y	*Income distribution (%)*						
	Q1 Income	9.8	10.3	14.1	22.1	23.0	23.0
	Q2 Income	11.6	11.9	16. 2	20.5	21.8	21.4
	Q3 Income	26.4	28.7	31. 3	30.6	31.5	31.1
	Q4 Income	52.2	49.1	38.3	26.8	23.7	24.4
ℓ	*Ability distribution (%)*						
	0-30% Ability	0.0	0.0	0.0	1.1	2.7	6.1
	30-60% Ability	0.4	0.1	0.1	41.2	40.0	38.4
	60-90% Ability	15.1	13.3	14. 2	40.3	40.1	38.8
	Top 10% Ability	84.5	86.6	85.8	17.4	17.2	16.7

FIGURE 8.1.—Equivalent variation when Pell grants increase



Note: The figure shows the percentage of lifetime consumption students of different EFC and ability are willing to give up to switch to the equilibrium with the higher Pell Grant maximum.

8.1. *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 8.1. 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 to match all targets described in Table 4.4 (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 ρ_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 D.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 D.5. 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.

9. 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, 24, 26, 27, 28]

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. Bénabou. 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, 28]

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. [13]

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. Michelmore, and S. Owen. Closing the gap: The effect of a targeted, tuition-free promise on college choices of high-achieving, low-income students. Working Paper 25349, National Bureau of Economic Research, 2018. [3]

D. Epple and R. E. Romano. Competition between private and public schools, vouchers, and peer-group effects. *American Economic Review*, pages 33–62, 1998. [15]

D. Epple, R. Romano, and H. Sieg. Admission, tuition, and financialaid policies in the market for higher education. *Econometrica*, 74:885–928, 2006. [1, 4, 15, 21, 12]

D. Epple, R. Romano, S. Sarpca, and H. Sieg. A general equilibrium analysis of state and private colleges and access to higher education in the u.s. *Journal of Public Economics*, 155:164–178, 2017. [4, 21, 29]

R. Fernandez and R. Rogerson. Income distribution, communities and the quality of public education. *The Quaterly 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, 27]

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. [13]

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]

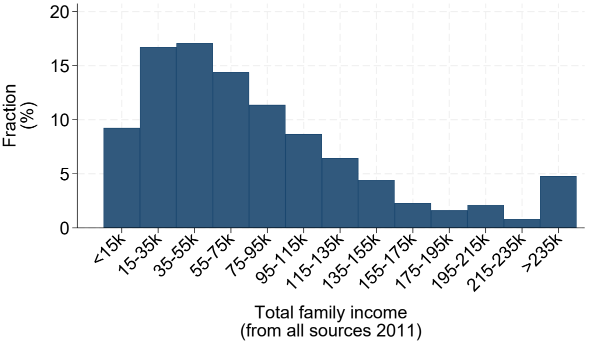
ONLINE APPENDIX

This online appendix contains four sections.

APPENDIX A: EMPIRICAL ANALYSIS

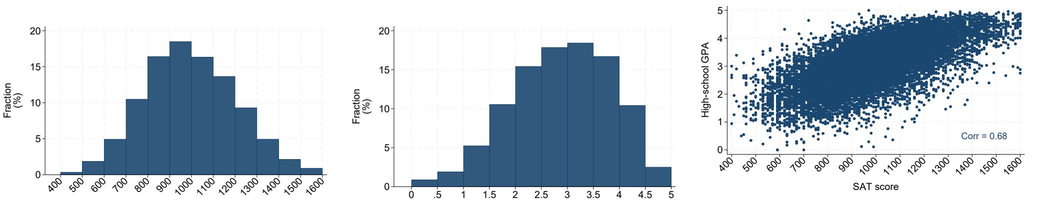
A.1. *HSLS summary statistics*

FIGURE A.1.—Distribution of parental income



Note: The figure shows the distribution of parental income in the HSLS.

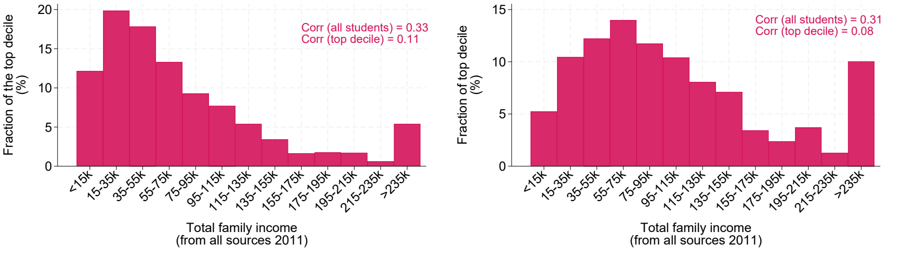
FIGURE A.2.—Distribution of SAT scores and high-school GPA



(a) SAT score (b) High-school GPA (c) SAT score and high-school GPA

Note: Panel (a) shows the distribution of SAT scores in the HSLS. Panel (b) shows the distribution of high-school GPA in the HSLS. Panel (c) shows the joint distribution of SAT scores and high-school GPA in the HSLS.

FIGURE A.3.—Distribution of parental income for students in the top decile of SAT scores/high-school GPA



(a) Top decile of SAT scores

(b) Top high-school GPA

Note: Panel (a) shows the distribution of parental income for students in the top decile of the SAT distribution. Panel (b) shows the distribution of parental income for students in the top decile of the high-school GPA distribution.

TABLE A.1

HSLS SUMMARY STATISTICS

Demographic characteristics (%)

White	52.5
Asian	3.5
Black/African-American	13.4
Hispanic	21.9
Other race	8.7
Male	49.7
Female	50.3
Parent has less than high school	9.9
Parent has HS diploma or GED	44.5
Parent has Associate's	15.9
Parent has Bachelor's	20.0
Parent has Master's or higher	9.6
Lives with 2 parents	68.4
Lives with 1 parent or none	31.6
At least 1 parent is working	89.1
Average number of household members	10.9
High school is public	93.1
High school is private	6.9
High school is in a city	30.6
High school is in a suburb	28.5
High school is in a town	12.1
High school is in a rural area	28.7
High school is in the Northeast	17.8
High school is in the Midwest	22.2
High school is in the South	37.6
High school is in the West	22.4
Observations	18,913

A.2. *Highly-selective vs. non-selective colleges*

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

Tier 1: Amherst College, Barnard College, Bates College, Boston College, Bowdoin College, Brown University, Bryn Mawr College, Bucknell University, California Institute of Technology, Carleton College, Carnegie Mellon University, Case Western Reserve University, Claremont McKenna College, Colby College, Colgate University, College of Mount Saint Vincent, College of the Holy Cross, College of William & Mary, Colorado College, Columbia University/City of New York, Connecticut College, Cooper Union for the Advancement of Science and

Art, Cornell University, Dartmouth College, Davidson College, Duke University, Emory University, Franklin and Marshall College, George Washington University, Georgetown University, Georgia Institute of Technology, Hamilton College, Hampshire College, Harvard University/Harvard College, Harvey Mudd College, Haverford College, Johns Hopkins University, Kenyon College, Lehigh University, Macalester College, Massachusetts Institute of Technology, Middlebury College, New York University, Northeastern University, Northwestern University, Oberlin College, Ohio State University at Marion, Pitzer College, Pomona College, Princeton University, Reed College, Rensselaer Polytechnic Institute, Rice University, Rose-Hulman Institute of Technology, Santa Clara University, Smith College, Southern Methodist University, Stanford University, Swarthmore College, The Ohio State University, Tufts University, Tulane University, Union College, United States Air Force Academy, United States Military Academy, United States Naval Academy, University of California at Berkeley, University of California at Los Angeles, University of Chicago, University of Miami, University of Missouri/Columbia, University of North Carolina at Chapel Hill, University of Notre Dame, University of Pennsylvania, University of Richmond, University of Rochester, University of Southern California, University of Virginia, Vanderbilt University, Vassar College, Villanova University, Wake Forest University, Washington and Lee University, Washington University in St. Louis, Webb Institute, Wellesley College, Wesleyan University, Whitman College, Williams College, Yale University.

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

IPEDS sample. The Integrated Postsecondary Education Data System (IPEDS) is made publicly available through the National Center for Education Statistics. We restrict the IPEDS sample to cover the 2013-2016 time frame, which is the relevant 4 year period for the HSLS cohort who begin college in 2013. The analysis is restricted to four-year nonprofit U.S. colleges and universities that satisfy the following conditions

- 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.

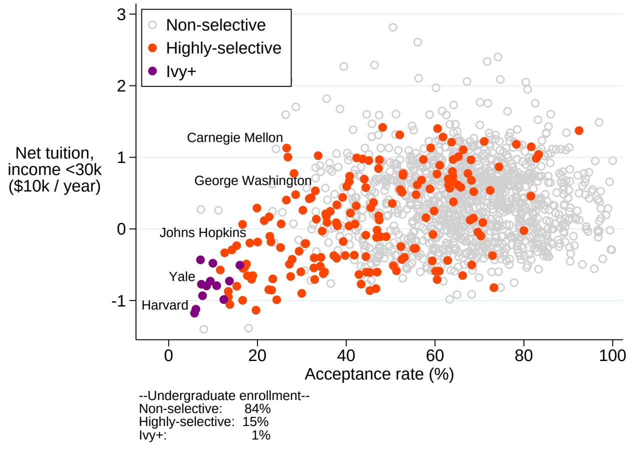
TABLE A.2

COMPARISON OF COLLEGES IN EACH SELECTIVITY TIER

College characteristics	Highly selective	Less selective
Number of colleges	186	1,577
Fraction of undergraduate enrollment	16%	84%
Fraction public	55%	73%
Average acceptance rate	44%	68%
Average rejection rate	*56%*	*32%*
Average SAT score	1,292	1,058
Median earnings 10 years after entry	$57,803	$42,228
Application fees	$59	$40
Tuition and fees	$24,576	$12,662
Room and board	$11,474	$9,372
Net tuition		
Income <$30k	-$1,848	$645
Income $30k-$48k	-$19	$1,888
Income $48k-$75k	$4,218	$4,847
Income $75k-$110k	$9,280	$7,429
Income >$110k	$15,924	$8,626
Instructional expenditures per student	$20,746	$8,597
Endowment assets per student	$131,440	$13,271

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 Barron's list 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.
Source: IPEDS, College Scorecard.

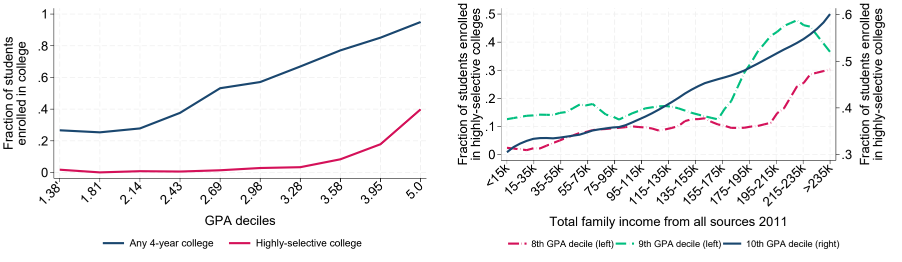
FIGURE A.4.—Net tuition for low-income students and undergraduate acceptance rate



Note: The figure shows the relationship between net tuition paid and acceptance rate for low-income students (those whose income <$30k per year in 2013). 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.

A.3. *Additional figures*

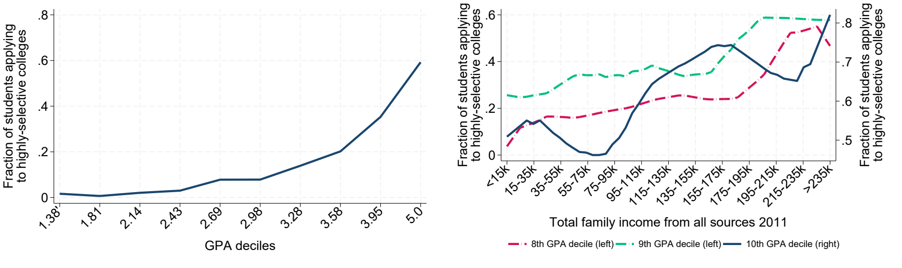
FIGURE A.5.—College attendance by high-school GPA



(a) By GPA (b) By parental income GPA

Note: Panel (a) shows the fraction of students attending any four-year non-profit college (blue) and highly-selective colleges (pink) across high-school GPA deciles. Panel (b) shows the fraction of students attending a highly-selective college across students' parental income and for different deciles of high-school GPA (students in the 8th decile in pink, 9th decile in green, and 10th decile in blue).

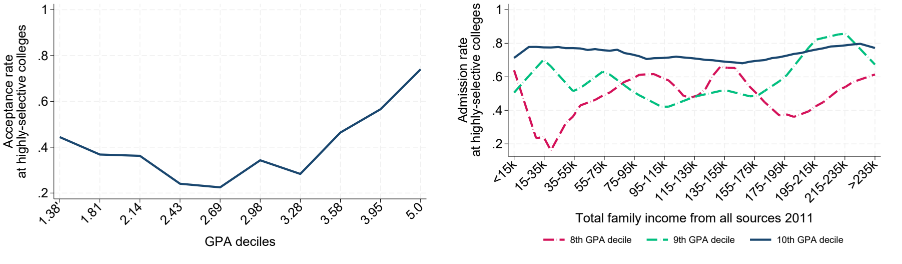
FIGURE A.6.—College application by high-school GPA



(a) By GPA (b) By parental income GPA

Note: Panel (a) shows the fraction of students attending any four-year non-profit college (blue) and highly-selective colleges (pink) across high-school GPA deciles. Panel (b) shows the fraction of students attending a highly-selective college across students' parental income and for different deciles of high-school GPA (students in the 8th decile in pink, 9th decile in green, and 10th decile in blue).

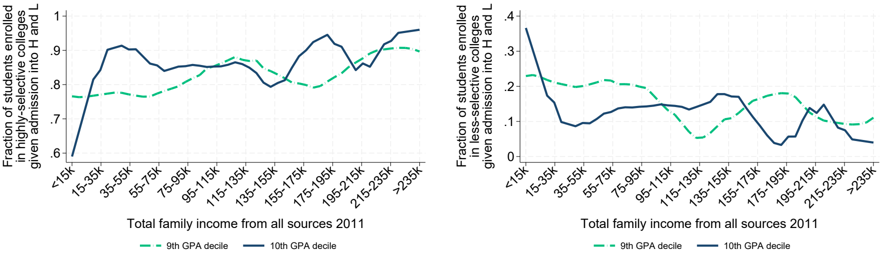
FIGURE A.7.—College admissions by high-school GPA



(a) By GPA (b) By parental income GPA

Note: Panel (a) shows the admission rates at highly-selective colleges for students who applied to a highly-selective college across high-school GPA deciles. Panel (b) shows the admission rates at highly-selective colleges for students who applied to a highly-selective college across students' parental income and for different high-school GPA deciles (students in the 8th decile in pink, 9th decile in green, and 10th decile in blue).

FIGURE A.8.—Enrollment given admission in both highly and less selective colleges



(a) Enroll in highly-selective college

(b) Enroll in less-selective college

Note: Panel (a) shows the fraction of students attending a highly-selective college conditional on being admitted into both a highly and less selective college across students' parental income and for different high-school GPA deciles (students in the 9th decile in green and 10th decile in blue). Panel (b) shows the fraction of students attending a less-selective college given that they were admitted into both a highly and less selective college across students' parental income and for different high-school GPA deciles (students in the 9th decile in green and 10th decile in blue).

A.4. *Additional tables*

TABLE A.3

LOGIT ESTIMATION RESULTS FOR WHETHER STUDENTS ARE ATTENDING COLLEGE (CONDITIONAL ON APPLYING)

	Enrolled in any college		**Enrolled in highly-selective college**	
	Coefficient	SE	Coefficient	SE
Parental income	0.005	(0.023)	0.050***	(0.034)
SAT (standardized)	0.915***	(0.150)	0.715***	(0.198)
Parental income*SAT	-0.046*	(0.023)	-0.020	(0.024)
GPA (standardized)	0.624***	(0.095)	0.862***	(0.124)
Asian	0.396	(0.298)	0.462**	(0.197)
Black/African-American	0.338*	(0.199)	-0.152	(0.303)
Hispanic	-0.054	(0.202)	0.105	(0.245)
Other race	0.417**	(0.204)	-0.371	(0.246)
Female	-0.173	(0.117)	0.017	(0.132)
Parent has HS diploma or GED	0.106	(0.272)	-1.350**	(0.524)
Parent has Associate's	0.487*	(0.289)	-1.091**	(0.539)
Parent has Bachelor's	0.618**	(0.287)	-1.281**	(0.519)
Parent has Master's or higher	0.550*	(0.309)	-0.843	(0.527)
Lives with 2 parents	0.183	(0.128)	-0.387**	(0.160)
At least 1 parent is working	-0.386*	(0.232)	-0.221	(0.317)
Number of household members	-0.080**	(0.039)	-0.061	(0.050)
High school is private	0.364*	(0.211)	0.651***	(0.178)
Fraction in HS receiving free lunch	0.019	(0.030)	-0.012	(0.038)
High school is in a suburb	0.511***	(0.169)	0.284*	(0.164)
High school is in a town	-0.297	(0.212)	-0.151	(0.298)
High school is in a rural area	0.011	(0.157)	0.067	(0.185)
High school is in the Midwest	0.263	(0.191)	0.566***	(0.179)
High school is in the South	0.183	(0.189)	0.167	(0.178)
High school is in the West	-0.587***	(0.209)	0.386*	(0.232)
Constant	1.782***	(0.469)	-1.081	(0.715)
N	8,414		2,712	
Psuedo R^2	0.159		0.161	

Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. The dependent variable is equal to 1 if the student enrolled in any four-year, non-profit colleges (in the left column) or any highly-selective colleges (in the right column), and 0 otherwise. The sample is restricted to all students who applied to any four-year, non-profit college (left column) and to students who applied to a highly-selective college (right column).

TABLE A.4

LOGIT ESTIMATION RESULTS FOR WHETHER STUDENTS APPLIED TO COLLEGE

	Apply to any college		**Apply to highly-selective college**	
	Coefficient	SE	Coefficient	SE
Parental income	0.096***	(0.017)	0.020	(0.019)
SAT (standardized)	0.443***	(0.112)	0.834***	(0.135)
Parental income*SAT	0.035*	(0.021)	0.051***	(0.019)
GPA (standardized)	1.091***	(0.068)	0.956***	(0.081)
Asian	0.558**	(0.220)	0.866***	(0.153)
Black/African-American	1.211***	(0.153)	0.991***	(0.187)
Hispanic	0.586***	(0.123)	0.856***	(0.170)
Other race	0.322**	(0.157)	0.221	(0.146)
Female	0.090	(0.083)	-0.200**	(0.089)
Parent has HS diploma or GED	0.192	(0.168)	-0.061	(0.308)
Parent has Associate's	0.181	(0.187)	-0.005	(0.313)
Parent has Bachelor's	0.263	(0.182)	0.264	(0.309)
Parent has Master's or higher	0.370*	(0.200)	0.378	(0.314)
Lives with 2 parents	-0.124	(0.091)	-0.121	(0.106)
At least 1 parent is working	0.131	(0.140)	-0.302	(0.243)
Number household members	-0.067**	(0.028)	-0.081***	(0.031)
High school is private	1.163***	(0.152)	0.297**	(0.121)
Fraction in HS receiving free lunch	-0.010	(0.021)	-0.015	(0.024)
High school is in a suburb	-0.100	(0.121)	-0.173	(0.121)
High school is in a town	-0.588***	(0.140)	-0.865***	(0.164)
High school is in a rural area	-0.208*	(0.117)	-0.586***	(0.116)
High school is in the Midwest	0.013	(0.139)	-0.523***	(0.134)
High school is in the South	-0.205	(0.136)	-0.694***	(0.142)
High school is in the West	-0.381**	(0.157)	-0.698***	(0.174)
Constant	0.587*	(0.325)	-1.080**	(0.460)
N	11,399		11,399	
Psuedo R^2	0.247		0.322	

Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. The dependent 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. The sample is restricted to all students who applied to any four-year, non-profit college in the right column (right column).

APPENDIX B: MODEL APPENDIX

B.1. *Tuition schedule (equation (3.21))*

Consider the college problem (3.12). Let λ_κ be the Lagrange multiplier on the number of students identity constraint (3.8), λ_L the Lagrange multiplier on the student's body ability identity constraint (3.9), λ_I the Lagrange multiplier on the budget constraint (3.10), and λ_T the Lagrange multiplier on the tuition cap constraint. The first-order conditions $\kappa_s, I_s, L_s, T_s(y,\sigma)$ are given by

$$
\begin{aligned}
&[\kappa_s] && \xi_s Q_{s_\kappa} = \lambda_\kappa + \lambda_I[C_s'(\kappa_s) - E_s'(\kappa_s) - Tr_s'(\kappa_s)] \\
&[I_s] && \xi_s Q_{s_I} = \lambda_I \\
&[L_s] && \xi_s Q_{s_L} = \lambda_L
\end{aligned}
$$

and

$$
\begin{aligned}
[T_s(y,\sigma)] \qquad & \lambda_I \left[\int q_s(y,\ell,\boldsymbol{T}(y,\sigma),\underline{\boldsymbol{\sigma}})\,\mu(y,d\ell,\sigma) + T_s(y,\sigma)\int \frac{\partial q_s(y,\ell,\boldsymbol{T}(y,\sigma),\underline{\boldsymbol{\sigma}})}{T_s(y,\sigma)}\,\mu(y,d\ell,\sigma)\right] \\
& + \lambda_\kappa \int \frac{\partial q_s(y,\ell,\boldsymbol{T}(y,\sigma),\underline{\boldsymbol{\sigma}})}{T_s(y,\sigma)}\,\mu(y,d\ell,\sigma) + \lambda_L \int \ell \frac{\partial q_s(y,\ell,\boldsymbol{T}(y,\sigma),\underline{\boldsymbol{\sigma}})}{T_s(y,\sigma)}\,\mu(y,d\ell,\sigma) - \lambda_T = 0.
\end{aligned}
$$

Rearrange these equations to obtain (3.21) for the case where the tuition cap constraint does not bind. For the case in which the tuition cap constraint binds, we have

$$\overline{T}_s < m_s(y,\sigma) MC_s(y,\sigma).$$

B.2. *Admissions standard (equation (3.22))*

Let λ_σ be the Lagrange multiplier on the admissions standard constraint. The first-order condition with respect to the admissions standard is given by

$$
\begin{aligned}
& \lambda_I \int T_s(y,\underline{\sigma}_s) q_s(y,\ell,\boldsymbol{T}(y,\underline{\sigma}_s),\underline{\boldsymbol{\sigma}}_s)\mu(dy,d\ell,\underline{\sigma}_s) + \lambda_\kappa \int q_s(y,\ell,\boldsymbol{T}(y,\underline{\sigma}_s),\underline{\sigma}_s)\mu(dy,d\ell,\underline{\sigma}_s) \\
& \qquad \lambda_L \int \ell q_s(y,\ell,\boldsymbol{T}(y,\underline{\sigma}_s),\underline{\sigma}_s)\mu(dy,d\ell,\underline{\sigma}_s) - \lambda_\sigma = 0.
\end{aligned}
$$

Replace the Lagrange multipliers defined above to arrive at the optimal admissions standard (equation (3.22)) when the admissions standard is positive. When $\underline{\sigma}_s = 0$, then the optimal condition is

$$
\underbrace{\frac{\int T_s(y,\underline{\sigma}_s)\, q_s(y,\ell,\boldsymbol{T}(y,\underline{\sigma}_s),\underline{\boldsymbol{\sigma}}_s)\mu(dy,d\ell,\underline{\sigma}_s)}{\int q_s(y,\ell,\boldsymbol{T}(y,\underline{\sigma}_s),\underline{\boldsymbol{\sigma}}_s)\mu(dy,d\ell,\underline{\sigma}_s)}}_{\text{Average revenue from lowest-signal students}} > \underbrace{C_{s_\kappa} - \frac{Q_{s_\kappa}}{Q_{s_I}} - E_{s_\kappa} - Tr_{s_\kappa}}_{\substack{\text{Costs net of transfers}\\ \text{per student}}} - \underbrace{\frac{Q_{s_L}}{Q_{s_I}}\mathbb{E}_s[\ell|\underline{\sigma}_s]}_{\text{Posterior average ability}}.
$$

B.3. *Equilibrium algorithm*

This section outlines the procedure to solve for the symmetric equilibrium numerically using a nested fixed point algorithm.

1. Start with a guess of college-level aggregates, $\{\kappa_s, I_s, L_s, T_s(y,\sigma), \underline{\sigma_s}\}$ for $s \in \{H, L\}$, the government tax rate, τ, and the probabilities of applying to highly-selective colleges, less-selective colleges, or both for each student type, $\{p_i(y,\ell)\}$ for $i \in \{\text{apply to H only, apply to L only, apply to both H and L}\}$.

2. Given these guesses, compute demand for college as the enrollment probabilities defined in equations (3.19) and (3.20), $q_H(y,\ell,\boldsymbol{T}(y,\sigma),\underline{\boldsymbol{\sigma}})$ and $q_L(y,\ell,\boldsymbol{T}(y,\sigma),\underline{\boldsymbol{\sigma}})$ for each student type (y,ℓ,σ).

3. Given demand for college and the aggregates, we solve for the tuition schedule, $T_H(y,\sigma)$ and $T_L(y,\sigma)$, across all student types and admission standards, $\underline{\sigma}_H$ and $\underline{\sigma}_L$. This portion of the algorithm is a fixed point that proceeds as follows
   (a) For each college type s, update the tuition policy $\hat{T}_s(y,\sigma)$ using (3.21) by markups and marginal costs.

   (b) For each college type s, use the updated tuition policy, $\hat{T}_s(y,\sigma)$, to solve for the updated admissions standard $\hat{\underline{\sigma}}_s$ in (3.22).

   (c) Check for convergence
      - If $\sup|T_s(y,\sigma)-\hat{T}_s(y,\sigma)|>10^{-5}$, set $T_s(y,\sigma)=\hat{T}_s(y,\sigma)$ and repeat step (a).
      - Otherwise, continue.

4. Given enrollment probabilities, tuition policies, and admission standards, update the college aggregates $\{\hat{\kappa}_s,\hat{I}_s,\hat{L}_s\}$ for each college type s using (3.8), (3.9), and (3.10). Update the government tax rate τ using (3.14) as well as the application probabilities $\{p_i(y,\ell)\}$ according to (3.17).

5. Check for convergence
   - Let $\hat{X}\equiv\left\{\{\hat{\kappa}_s,\hat{I}_s,\hat{L}_s\}_{s\in\{\text{H,L}\}},\hat{\tau},\{\hat{p}_i(y,\ell)\}_{i\in\{\text{H only, L only, Both H and L}\}}\right\}$, and $X\equiv\{\{\kappa_s,I_s,L_s\}_{s\in\{\text{H,L}\}},\tau,\{p_i(y,\ell)\}_{i\in\{\text{H only, L only, Both H and L}\}}\}$. If $\sup|X-\hat{X}|>10^{-5}$, set $X=\hat{X}$, and repeat step (1).
   - Otherwise, exit.

APPENDIX C: ESTIMATION APPENDIX

C.1. *Lifecycle Model*

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

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

Assume that at the terminal period individuals do not save or leave bequests. This allows us to rewrite her budget constraint $j=1,\dots,T$ as

$$c_1+\sum_{j=1}^{T}\frac{1}{R^j}c_{1+j}=a_1R+w(1-\tau)Z\ell^{\alpha}\sum_{j=1}^{T}\frac{1}{R^{j-1}}.$$

Now, the Euler equation for $j = 1, \ldots, T$ is such that $u'(c_1) = (\beta R)^t u'(c_{1+t})$. Let the utility function be represented by CRRA preferences with σ_c as the intertemporal elasticity of substitution. Then, the Euler equation can be rewritten as

$$c_{1+j} = (\beta R)^{j/\sigma_c} c_1.$$

Replace this last equation into the budget constraint above yields the following

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

Putting everything together, we can write the two-period model from Section 3 as

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

where we replaced $\tilde{\beta} u(c_1) = \sum_{j=1}^{T} \beta^j u(c_j)$ and defined

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

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

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

C.2. *Expected Family Contribution (EFC)*

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. We first compute the Adjusted Available Income (AAI), which corresponds to 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 they are 0. The family contribution out of the AAI is calculated from a (progressive) non-linear function of AAI according to Table C.1 below.

A student's EFC is then the family contribution divided by the number of children that are enrolled in college. The HSLS asks if students have a sibling in college at the same time. If the answer is yes, we considered the number of children enrolled in college to be two. We are thus able to construct EFC for the students in our sample even if they did not complete the FAFSA.

TABLE C.1

FAMILY CONTRIBUTION OUT OF THEIR ADJUSTED AVAILABLE INCOME

Family's AAI	Family contribution
< -\$3,409	-\$750
-\$3,409 to \$15,300	22% of AAI
\$15,301 to \$19,200	\$3,366 + 25% of AAI over \$15,300
\$19,201 to \$23,100	\$4,341 + 29% of AAI over \$19,200
\$23,101 to \$27,000	\$5,472 + 34% of AAI over \$23,100
\$27,001 to \$30,900	\$6,798 + 40% of AAI over \$27,000
\$30,901 or more	\$8,358 + 47% of AAI over \$30,900

Source: Department of Education.

C.3. *Distribution of student characteristics*

In the HSLS, approximately 30% of the students have an EFC of 0. In fitting the 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. We also assume that ℓ follows a log-normal distribution. Since y and ℓ are correlated, we estimate the parameters of the distribution for the case in which $y = 0$ and $y > 0$ separately. To summarize, the joint distribution of income and ability is given by

$$\hat{\mu}(y,\ell) = \begin{cases} LogN(\mu_{\ell_0}, \sigma^2_{\ell_0}) & \text{if } y = 0 \\ LogN\left(\begin{bmatrix} \mu_y \\ \mu_{\ell_1} \end{bmatrix}, \begin{bmatrix} \sigma^2_y & \sigma_{y\ell} \\ \sigma_{y\ell} & \sigma^2_{\ell_1} \end{bmatrix} \right) & \text{if } y > 0. \end{cases}$$

Using the HSLS, we find $\mu_{\ell_0} = -0.144$ and $\sigma^2_{\ell_0} = 0.155$ for $y = 0$. For $y > 0$, we find $\mu_y = -0.301$, $\mu_{\ell_1} = 0.135$, $\sigma^2_y = 2.825$, and $\sigma^2_{\ell_1} = 0.138$. The covariance between income and ability is estimated to $\sigma_{y\ell} = 0.174$.

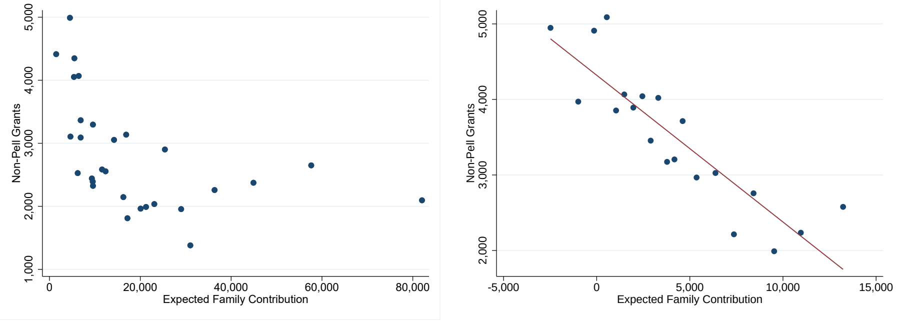
C.4. *External grants*

We use the BPS dataset to identify the external grants received by students, $Gr(y)$, that are not college-specific (need-based or merit-based) grants and federal Pell Grants. We do so by adding up all grants and aid excluding loans, Pell Grants, and institutional aid. The binned scatterplots below, Figure C.1, show how these grants vary with parental income (here, Expected Family Contribution). These external grants are decreasing in EFC and tend to level off at around \$15,000 (panel a). For an EFC below \$15,000, we find a downward sloping relationship between EFC and non-Pell Grants using OLS and controlling for SAT scores (panel b).

C.5. *Colleges' budget constraint*

Costs and endowment. We use IPEDS data to estimate the relationship between non-instructional costs and enrollment for each college selectivity tier. 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 in IPEDS. 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 posit the following relationship between

FIGURE C.1.—External grants received by students



(a) All students

(b) Low-income students

Note: Panel (a) shows the relationship between external grants received by students of different income levels (measured as expected family contribution). Similarly, panel (b) shows this relationship for low-income students.

non-instructional costs and enrollment for each college tier $s = \{H, L\}$

$$OpCosts_{i_s} = O_{0_s} + O_{2_s} \times Enrollment^2_{i_s} + \varepsilon_{i_s},$$

where the costs in the left hand side are the maximum over both methods (this helps us deal with cases where the second procedure produces small or negative numbers). Table C.2 presents the results of the estimation.

We follow Epple, Romano, and Sieg (2006) to transform these estimates into our model parameters. Let $K_s = N_s\kappa_s$ denote total enrollment in colleges of type $s = \{H, L\}$, where it is assumed that colleges within each type are the same. Then summing up individual cost functions yields

$$\begin{aligned} OpCosts(K_s) &= \sum_{i=1}^{N_s}\left[O_{0_s} + O_{2_s} \times \frac{K_s^2}{N_s^2}\right] \\ &= N_s O_{0_s} + \frac{O_{2_s}}{N_s} K_s^2. \end{aligned}$$

IPEDS data is also used to calculate the total level of private endowment income received by each college by adding up all private revenue received over the 2013-2016 sample period, which corresponds to unrestricted revenue from gifts, investment return from their endowment, and contributions from affiliates. The total endowment income of highly-selective colleges is twice as large as that of less-selective colleges. In model units, this corresponds to $E_H(\kappa) = 0.12$ and $E_L(\kappa) = 0.06$.

To get the fixed costs in the model, C_{0_H} and C_{0_L}, we take the estimated constant from the regression, O_{0_H} and O_{0_L}, and multiply it by the number of colleges in our IPEDS sample that are present in the HSLS sample, i.e. there are $N_H = 182$ highly-selective colleges and $N_L = 1,483$ less-selective colleges, net of the constant term in endowment income, i.e.,

$$C_{0_s} = N_s O_{0_s} - E_{0_s},$$

where $E_{0_H} = 0.12 - E_{1_H}\kappa_H$ and we assume $E_{0_L} = 0$.

To proceed we compute the enrollment shares in each college selectivity tier observed in the HSLS data. These are are $\kappa_H = 0.069$ and $\kappa_L = 0.357$. These values are consistent with what we find in IPEDS data. We then infer the total number of high school graduates using the 42% 4-year enrollment rate reported by NCES.[28] and then add up all full-time equivalent undergraduate students across all colleges. We find that 7% of high school graduates attend highly-selective colleges and 35% attend less-selective colleges—similar to the enrollment shares computed with the HSLS data.

This procedure translates into $C_{0_H} = 0.13$ and $C_{O_L} = 0.25$. The linear term in the endowment function for less-selective colleges is given by $E_{1_L} = 0.06/\kappa_L = 0.16$ and that of highly-selective colleges is chosen to match the enrollment share in highly-selective colleges so that $E_{1_H} = 1.37$. Without loss of generality, we let the quadratic term related to enrollment be in the endowment function in the model, E_{2_H} and E_{2_L}. In this case, we take the estimated coefficient on the squared enrollment and divide it by the number of colleges. This yields $E_{2_H} = 5.77$ and $E_{2_L} = 0.44$. The linear term in the enrollment function

TABLE C.2

OLS ESTIMATION RESULTS FOR COSTS

	Highly-selective colleges	**Less-selective colleges**
Enrollment^2	1057.6***	624.2***
	(120.4)	(14.85)
Private college dummy		-0.0000919***
		(0.0000153)
Constant	0.000841***	0.000222***
	(0.000111)	(0.0000126)
N	182	1445
Adj. R^2	0.296	0.604

Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors in parentheses. Variables have been scaled down by the total student population. Costs have also been scaled down by $40,000 to fit the model units.

Government transfers. We estimate the government transfer functions for highly and less-selective colleges using the IPEDS data. Government transfers correspond to state government appropriations and federal funds colleges received. Table C.3 presents the results of the estimation of the linear relationship between government transfers and student enrollment, which help identify Tr_{1_H} and Tr_{2_L}.

TABLE C.3

OLS ESTIMATION RESULTS FOR GOVERNMENT TRANSFERS

	Highly-selective colleges	**Less-selective colleges**
Enrollment	1.044***	0.620***
	(0.0516)	(0.0112)
N	182	1445
Adj. R^2	0.691	0.679

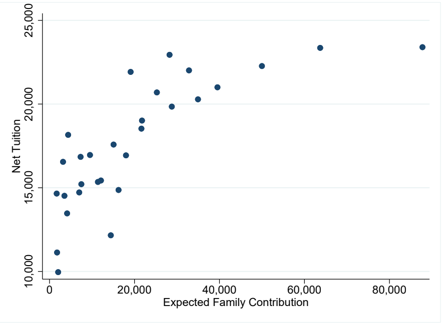
Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors in parentheses.

Sticker prices. We use the BPS data to assess how net tuition (defined as sticker-price tuition minus college-specific need and merit-based grants) varies with student's parental income (here measured as the EFC). Figure C.2 provides binned-scatter plots of that relationship for highly-selective and less-selective colleges. We see a leveling off of net tuition for higher-income

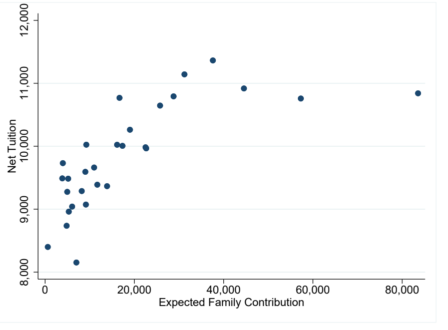
[28] See here: https://nces.ed.gov/programs/coe/indicator_cpa.asp.

students. This helps us set the tuition cap across the two college types, $\overline{T}_H$ and $\overline{T}_L$, which we set at \$25,000 for highly-selective colleges and at \$12,000 for less-selective colleges.

FIGURE C.2.—Net tuition paid by students



(a) Highly-selective colleges



(b) Less-selective colleges

Note: Panel (a) shows the relationship between net tuition (sticker-price tuition minus college-specific grants) paid by students of different income levels (measured as expected family contribution) at highly-selective colleges. Similarly, panel (b) shows this relationship for students at less-selective colleges.

APPENDIX D: ADDITIONAL QUANTITATIVE ANALYSIS

D.1. *Effect of Equalizing Application Patterns*

		College 1		**College 2**	
		Baseline	*More Applicants*	*Baseline*	*More Applicants*
L_μ	Avg Student Ability	1.86	1.82	1.28	1.3
I_μ	Instr Spending per Student	2.1	2.04	0.98	0.97
Γ_s	Value-added	2.01	1.96	1.79	1.81
κ	% Enrolled	7.15	6.8	32.09	31.09
	Income Distribution				
	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
	Ability Distribution				
	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 D.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

D.2. *Effect of Less Informative Signals*

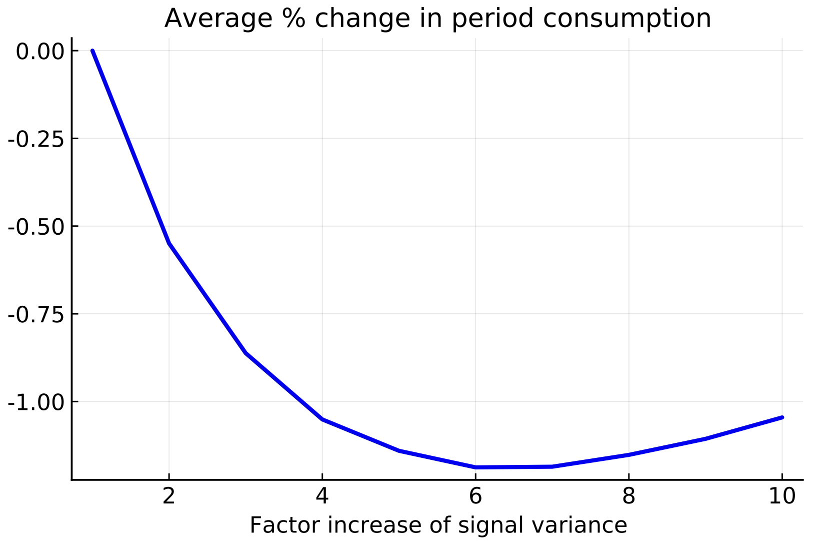


FIGURE D.1.—Welfare changes as a function of increases to the signal variance.

		College 1		College 2	
		Baseline	$6 * \sigma_g^2$	*Baseline*	$6 * \sigma_g^2$
L_μ	Avg Student Ability	1.86	1.67	1.28	1.19
I_μ	Instr Spending per Student	2.1	2.4	0.98	1.14
Γ_s	Value-added	2.01	1.87	1.79	1.71
κ	% Enrolled	7.15	8.93	32.09	39.72
	Income Distribution				
	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
	Ability Distribution				
	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 D.2

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

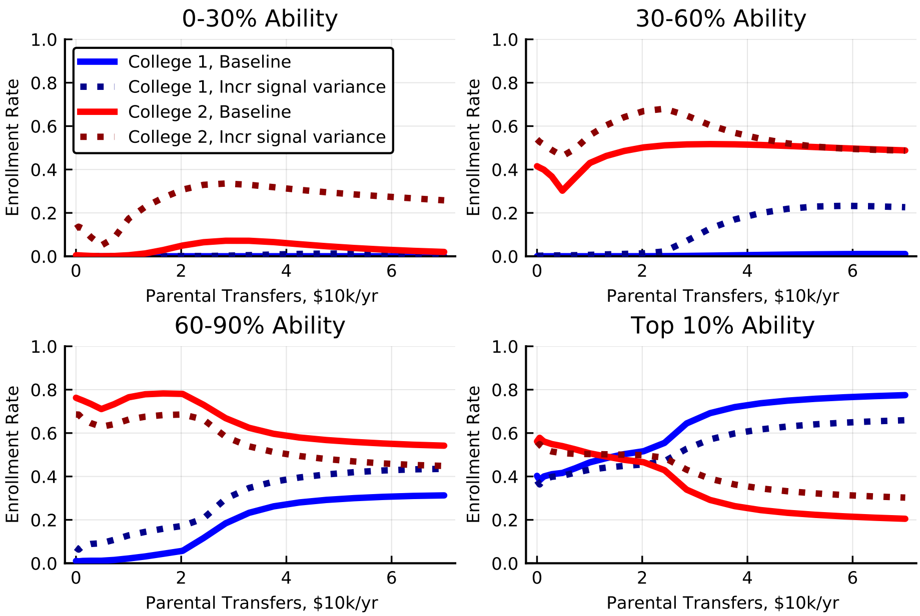


FIGURE D.2.—Student enrollment rates in baseline equilibrium vs. equilibrium with less informative signals.

		College 1		College 2	
		Baseline	*Incr. signal variance*	*Baseline*	*Incr. signal varianc*
L_μ	Avg Student Ability	1.86	1.77	1.28	1.24
I_μ	Instr Spending per Student	2.1	2.24	0.98	1.03
Γ_s	Value-added	2.01	1.94	1.79	1.75
κ	% Enrolled	7.15	7.93	32.09	34.48
	Income Distribution				
	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
	Ability Distribution				
	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 D.3

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

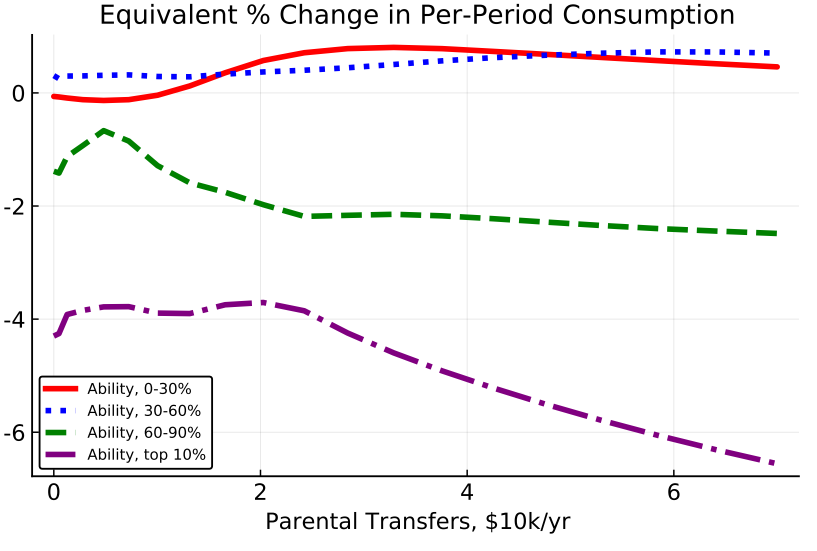


FIGURE D.3.—Percent of lifetime consumption students of different parental-income and ability groups are willing to give up in order to switch to the equilibrium with less informative signals. In this experiment, the loss of information is larger for high-income students.

D.3. *Effect of Perfectly Informative Signals*

		College 1		College 2	
		Baseline	*Perfect Info*	*Baseline*	*Perfect Info*
L_μ	Avg Student Ability	1.86	1.91	1.28	1.36
I_μ	Instr Spending per Student	2.1	1.87	0.98	0.93
Γ_s	Value-added	2.01	2.02	1.79	1.86
κ	% Enrolled	7.15	6.5	32.09	29.8
	Income Distribution				
	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
	Ability Distribution				
	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 D.4

EFFECT ON COLLEGES OF MAKING SIGNALS PERFECTLY INFORMATIVE

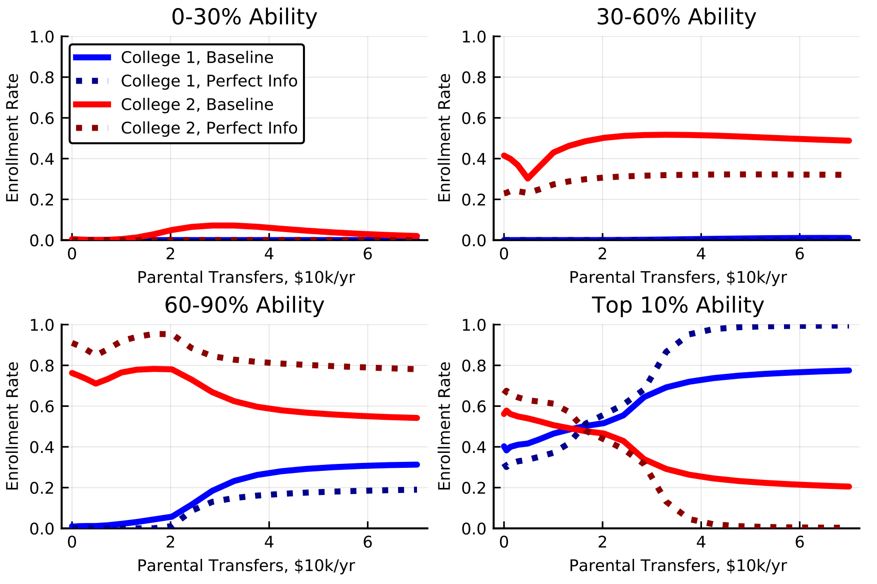


FIGURE D.4.—Average enrollment rates for students by parental transfers in the baseline equilibrium and perfect information equilibrium.

D.4. *Policy Experiments: Increasing Pell Grants*

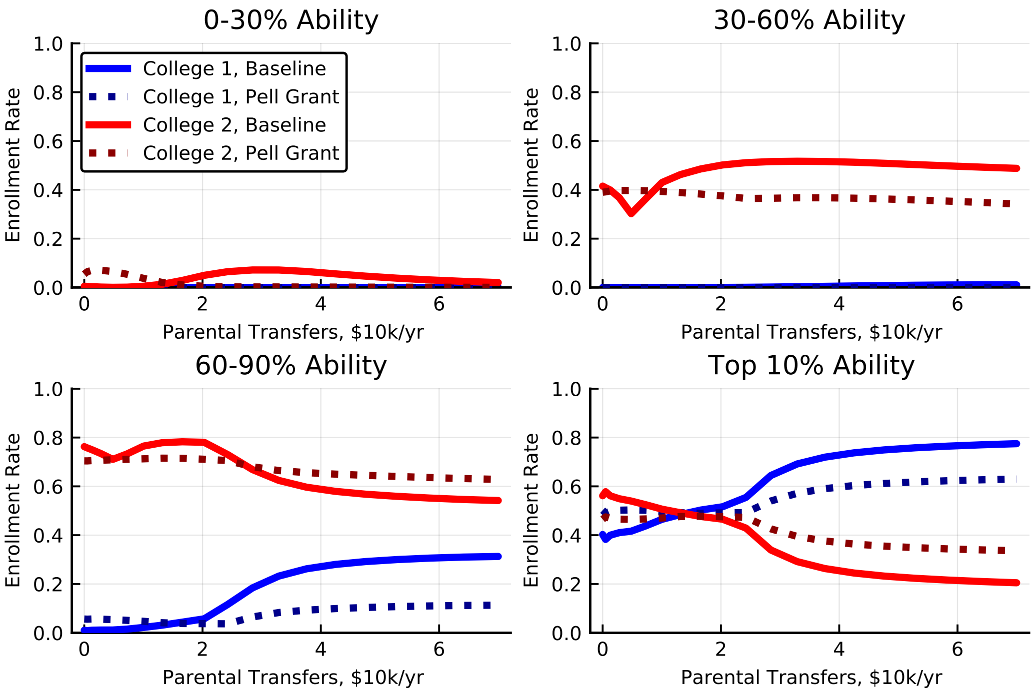


FIGURE D.5.—Student sorting by parental transfers and ability group both in the baseline equilibrium, and in an equilibrium where the Pell Grants maximum has increased from 5.65k per year to 25k per year.

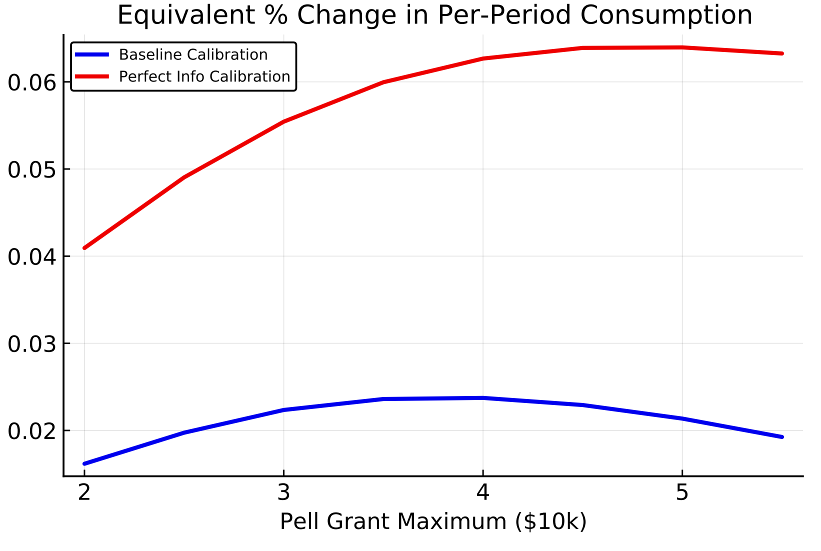


FIGURE D.6.—Average increases in welfare due to increases in the Pell Grant maximum both under the baseline parameter estimates, and under a perfect-information calibration. Failing to account for admissions uncertainty would cause the model to overstate the welfare gains from the policy.

TABLE D.5

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%.

		College 1				
		Baseline	*Pell Grant*	*10% Adjust*	*20% Adjust*	*30% Adjust*
L_μ	Avg Student Ability	1.86	1.87	1.88	1.89	1.91
I_μ	Instr Spending per Student	2.10	2.67	2.81	2.86	2.84
Γ_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
κ	% Enrolled	7.15	7.25	6.84	6.38	6.02
	Income Distribution					
	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
	Ability Distribution					
	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

		College 2				
		Baseline	*Pell Grant*	*10% Adjust*	*20% Adjust*	*30% Adjust*
L_μ	Avg Student Ability	1.28	1.26	1.27	1.29	1.31
I_μ	Instr Spending per Student	0.98	1.04	1.13	1.21	1.30
Γ_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
κ	% Enrolled	32.09	32.14	30.9	29.66	28.53
	Income Distribution					
	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
	Ability Distribution					
	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
	Overall welfare change	–	1.97%	2.03%	2.09%	2.17%