NBER WORKING PAPER SERIES

# TRANSPARENCY AND PERCENT PLANS

Adam Kapor

Working Paper 32372
http://www.nber.org/papers/w32372

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2024, revised June 2024

First draft October 2015. This paper was previously circulated under the title "Distributional Effects of Race-Blind Affirmative Action". I thank Phil Haile, Joseph Altonji, Costas Meghir, and Steve Berry for their guidance and support. I thank Sofia Moroni, Chris Neilson, Marta Tienda, Paulo Somaini, and Seth Zimmerman for helpful discussions and comments. I thank seminar participants at Boston College, Columbia, the Federal Reserve Bank of New York, Michigan, Notre Dame, NYU Stern, Penn, Princeton, Toulouse, Stanford GSB, UCL, UIUC, Wisconsin, WUSTL, and Yale. I gratefully acknowledge financial support from the Cowles Foundation, the Jack Kent Cooke Foundation, and a Yale University Dissertation Fellowship. These data require the following acknowledgment: This research uses data from THEOP, a project directed by Marta Tienda, at Princeton University, in collaboration with Teresa A. Sullivan, formerly at the University of Texas at Austin. This research [THEOP] was supported by grants from the Ford, Mellon and Hewlett Foundations and NSF (Grant SES-0350990). THEOP gratefully acknowledges institutional support from Princeton University's Office of Population Research (NICHD Grant R24 H0047879). Special Acknowledgement is due to Dawn Koffman, for preparing and documenting data for public use, and to Sunny X. Niu, who has been a stalwart collaborator and co-investigator. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Transparency and Percent Plans
Adam Kapor
NBER Working Paper No. 32372
April 2024, revised June 2024
JEL No. I21,I23,I24,I26,I28


**ABSTRACT**

Transparency vs. opacity is an important dimension of college admission policy. Colleges may gain useful information from a holistic review of applicants' materials, but in doing so may contribute to uncertainty that discourages potential applicants with poor information. This paper investigates the impacts of admissions transparency in the context of Texas' Top Ten Percent Plan, using survey and administrative data from Texas and a model of college applications, admissions, enrollment, grades, and persistence. I estimate that two thirds of the plan's 9.1 point impact on top-decile students' probability of attending a flagship university was due to information rather than mechanical effects. Students induced to enroll are more likely to come from low- income high schools, and academically outperform the students that they displace. These effects would be larger if complemented by financial-aid information, and are driven by transparency, not misalignment between the rules used for automatic and discretionary admissions.



Adam Kapor
Department of Economics
Princeton University
280 Julis Romo Rabinowitz Building
Princeton, NJ 08544
and NBER
akapor@princeton.edu


A data appendix is available at http://www.nber.org/data-appendix/w32372

# 1 Introduction

Transparency vs. opacity is an important dimension of college admission policy. Admissions offices may extract useful information from subjective or holistic reviews of students' essays, letters, and other materials. Requiring these materials, however, imposes costs on applicants, and their use may contribute to uncertainty about admissions chances. This uncertainty may discourage applications among some highly qualified potential applicants who in fact would likely be admitted. The effect may be more severe for those with less information about the process. When the broader process is opaque, policies that commit to admit some students transparently—as a known function of known inputs—may encourage strong students with poor information to submit applications.

The goal of this paper is to quantify the impacts of transparency in college admissions when college seats are scarce, applications are costly, and potential applicants may have noisy and heterogeneous information about admissions and financial aid. I do so in the context of Texas' "top ten percent" plan. This policy guaranteed admission to all Texas public colleges and universities to all in-state students in the top decile of their high school class. Importantly, the relevant class-rank measure was made known to applicants before applications were due, providing a guarantee of admission to students who might otherwise have faced subjective uncertainty about their chances.

I focus on four-year and flagship university enrollment in Texas, estimating a model that considers the demographics, class rank, and exam scores of colleges' entering cohorts, and these students' post-enrollment outcomes: persistence, GPA, and major choice. To isolate the impact of information, I first decompose the impacts of the percent plan into mechanical (via increases in the objective admissions chances of top-decile applicants), informative (via application decisions), and equilibrium channels. Next, I use the model to consider a hypothetical financial-aid awareness intervention [Dynarski et al., 2018] and quantify complementarities between admissions transparency and financial-aid information. Finally, class rank alone is just one way to provide transparency, and under the percent plan the universities considered a broader set of characteristics for discretionary admissions. To the extent that the percent plan "pulled in" able students [Black et al., 2023], did it do so because class rank is a better predictor of academic success than the indices used for discretionary admissions, or was it the transparency that mattered? I isolate the role of transparency by considering alternative percent plans in which the weights on grades, test scores, and other observables used for discretionary and guaranteed admissions are aligned.

My main finding is that the informative effect of the percent plan on the diversity and academic outcomes of flagships' entering classes is large, comprises the majority of the total effect, and is a robust feature of the informational environment. According to estimates, students from poorer high schools had little information about how the subjective components of their applications would be evaluated, and faced greater uncertainty than did students from affluent high schools, leading to large effects of the percent plan: a 9.1 point increase in top-decile students' probability of attending a flagship university—two thirds of which is due to information—or 9.6 points if complemented by financial-aid information. Moreover, students induced to apply by transparent admission commitments earn higher grades, and are more likely to persist and to major in STEM, than the students they displace. The presence and relative importance of informative effects, and the impacts on enrolled students' academic outcomes, are not driven by the use of class rank alone nor by misalignment between the rules used for automatic and non-automatic admissions. A "percent plan" that uses the same weights on observables as in discretionary admission decisions would also improve academic outcomes at flagships, but would do so without lowering colleges' payoffs.

Addressing the research questions requires isolating the informative effects of the Top Ten Percent plan and then extrapolating to related policies. To do so, I construct and estimate a model of supply and demand for seats in four-year colleges with heterogeneous imperfect information. The model incorporates colleges' admissions rules, students' choice of application portfolios, financial-aid applications, and matriculation decisions, and students' academic performance in college. On the supply side, admissions offices value students' class rank, exam scores, and high-school characteristics, as well as a term not observed by the econometrician, which I term "caliber". Admissions standards adjust to satisfy capacity constraints when policies change, leading to spillovers to non-top-decile students. On the demand side, three frictions may lead to unequal access to colleges: application costs may discourage some students from applying; individuals may be unaware of financial aid opportunities; and individuals may be imperfectly informed about their admissions chances. Low-income and minority applicants may face higher (or lower) effective application costs, and may have worse (or better) information about aid availability and their admissions "caliber".

I estimate the model using a survey of high-school seniors in Texas public high schools from 2002—when the percent plan was in place—together with administrative data from Texas colleges and universities. I then perform counterfactual experiments, solving for the new admissions rules that colleges would choose, the applications that students would submit, the choices that they would make, and the distributions of post-enrollment outcomes that would realize.

The research design uses variation induced by the top 10% threshold, together with data on application portfolios, admissions outcomes, and matriculation decisions, to learn about key objects of interest. The Texas Top Ten policy introduces a discontinuity in admissions chances—and in subjective perceptions of these chances—as students just above the 10% threshold face a different distribution over sets of offers than those just below. The locally-exogenous variation in admissions chances around the top-decile cutoff shifts the distribution of choice sets available to otherwise similar students. I use it to identify applicants' preferences.

Identifying the role of information makes use of applicants' application portfolios, admissions outcomes, and admissions-relevant observables. Students' exam scores and high-school characteristics shift the extent to which top-decile threshold-crossing effects are due to information. For instance, the average student with top SAT scores has an admissions chance close to 100% even if his GPA is just outside the top decile, while a student with lower exam scores has a larger gain in admissions chances. For the former student, for whom the impact on objective admission chances is small, threshold-crossing effects reveal the extent of imperfect information about admissions chances.

Data on application portfolios reveals heterogeneity in taste for characteristics of colleges, as well as the extent of application costs. Within-portfolio correlation in colleges' characteristics pins down the variance of random tastes for characteristics. The statistical relationship between the "aggressiveness" of students' application portfolios and their admissions outcomes reveals the extent of private information about admissions. For example, students applying to UT-Austin and highly selective institutions reveal either a strong preference for those institutions or a belief that they are likely to be admitted. To the extent that these students are in fact more likely to be admitted to UT-Austin than those applying to portfolios consisting of UT-Austin and less selective programs, one may infer that applicants hold private information about admissions chances.

I model awareness of financial aid in the spirit of "consideration sets" [Goeree, 2008]. Identification exploits variation in application decisions, together with the fact that these consideration sets are partially observed, separately from application decisions. In particular, I see financial-aid application portfolios, including decisions not to apply for aid, at schools to which students have applied for admission. For students likely to receive some aid, the latter reveal imperfect awareness of aid.

I do not assume that admissions offices maximize grades, nor that they ought to do so. To understand the determinants of students' grades, persistence, and major, and how these factors relate to what admissions offices value, the paper again makes use of the 10% threshold. Students just below this threshold have flagship admissions probabilities that are increasing in "caliber", in contrast to those just above who are always admitted. Therefore, conditional on enrolling in a flagship university, caliber is higher on average for students below the cutoff. Additionally, I use the availability of a scholarship at Texas A&M as an excluded shifter of student unobservables conditional on enrolling at UT Austin, and vice versa. This variation reveals the extent to which caliber and other admissions-relevant variables also contribute to grades, persistence, and major choice.

Crucially, the model allows colleges to adjust their admissions policies in response to policy changes. In contrast to most prior race-based affirmative action programs, Texas Top Ten had large effects on admissions chances for all students, not just the "treated" top-decile students. In 2009, 81% of students attending the University of Texas at Austin were admitted automatically under Texas Top Ten, while students outside the top decile competed for fewer seats. Indeed, the main criticism made by opponents of the plan is that the rule displaced many non-automatically-admitted students from the University of Texas.[1] In the model, colleges make admissions decisions subject to bounds on the size of their entering classes and the constraints of the Top Ten law.

This paper builds on prior work on Texas Top Ten and contributes to literatures on "percent plans", affirmative action, information problems in higher education, and market equilibrium in higher education markets. RDD and difference-in-differences designs indicate that, consistent with a reduction in uncertainty, Texas Top Ten made top-decile students less likely to submit large application portfolios and more likely to apply to UT Austin [Andrews et al., 2010, Daugherty et al., 2014], leading to discontinuities in the characteristics of applicants by class rank [Fletcher and Mayer, 2013]. These effects led to positive enrollment impacts for Hispanic students [Niu and Tienda, 2010], and enrollment and persistence gains for students from an urban school district [Daugherty et al., 2014]. I show that the model's predictions match the key estimates of overall impacts.

In contemporaneous work, Black et al. [2023] provide evidence that Texas Top Ten increased enrollment, graduation rates, and earnings of top students at poor schools, but did not decrease college enrollment, graduation, or earnings of students from "feeder" schools

[1] See e.g. "New Law in Texas Preserves Racial Mix in State's Colleges", New York Times, November 24 1999.

who were displaced to out-of-state and non-flagship programs. While I lack earnings data, I use data on application portfolios and admissions outcomes to decompose these effects.[2] In subsequent work, Bleemer [2024] pursues related questions in California.

This paper builds on prior studies of education markets in equilibrium [Arcidiacono, 2005, Howell, 2010, Epple et al., 2006, 2008, 2014]. I extend Howell [2010] and Arcidiacono [2005] by endogenizing colleges' admissions thresholds. Like Fu [2014], I consider selection on applicants' private information, but I allow the informativeness of the signals to vary across students. Methodologically, my demand system is closest to Walters [2018] on charter schools. A key difference is that charter school admissions are determined by independent random lotteries, about which private information is impossible. I extend that paper's approach to allow for private information about admissions chances.

My model of applicants' beliefs and awareness draws on a literature documenting applicants' imperfect information about admissions and aid and responsiveness to interventions [Avery and Hoxby, 2003, Card and Krueger, 2005, Pallais, 2009, Bowen et al., 2009, Thibaud, 2019]. Many high-achieving low-income high school students do not apply to selective colleges [Hoxby and Avery, 2012], but information and application-fee waivers [Hoxby and Turner, 2014] and financial-aid assistance [Dynarski and Scott-Clayton, 2006, Bettinger et al., 2012] can lead them to do so. Dynarski et al. [2018] find large enrollment impacts in a large-scale RCT of an intervention that encouraged students to apply to the University of Michigan largely by informing students about the availability of aid. Embedding these frictions in a model allows investigating their interactions and joint effects.

Finally, this paper relates to a literature on affirmative action [Antonovics and Backes, 2013a, Backes, 2012, Antonovics and Backes, 2013b, Arcidiacono et al., 2014, Arcidiacono and Lovenheim, 2016] and the use of imperfect proxies for race [Fryer Jr et al., 2008, Ellison and Pathak, 2016]. Consistent with descriptive and difference-in-differences evidence [Cortes, 2010, Flores and Horn, 2016, Niu et al., 2006], my estimates show that Texas Top Ten did not provide meaningful affirmative action for URM applicants.

## 2 Background

While elite universities in many countries rank all or most applicants according to linear functions of a small set of scores and grades known to the applicants, selective U.S. univer-

[2] For estimates of labor-market returns to Texas higher-education institutions, see Andrews et al. [2012] and Mountjoy and Hickman [2020].

sities evaluate applicants' materials according to criteria that are not explicitly disclosed. Moreover, policies intended to promote equitable access to these institutions, such as holistic admission and affirmative action policies, typically involve opacity as well—as required by law if they considered applicants' race.[3] The percent plan is a notable exception.

**College Admissions Policy in Texas:** The Top Ten Percent policy was introduced in response to a court decision banning the consideration of race in college admissions. In 1996, the Fifth Circuit Court of Appeals ruled in *Hopwood v. Texas* that race could not be used as a factor in admissions decisions at the University of Texas School of Law nor, by extension, at any Texas public universities. The "Texas Top Ten" law, formally Texas House Bill 588, was passed in 1997 and first took effect for the cohort entering college in Fall 1998. Importantly, the relevant class-rank measure for Texas Top Ten is that of the final grading period before applications are due. In addition, the law required schools to construct class-rank measures and to inform the top 25% of their rank. As a result, Top Ten status was known at the time of applications.

Flagship universities—The University of Texas at Austin and Texas A&M University—ranked non-top-decile applicants on the basis of both personal and academic characteristics. Here I provide details about UT Austin, the most selective institution. In 1996 and previous years, the university admitted students on the basis of predicted first-year grades, together with a race-based affirmative action policy. Since the initial prohibition of affirmative action in 1997, this institution has computed two scores for each applicant, an "academic index" and a "personal achievement index". The academic index is a linear function of grades and class rank, with parameters that depend on the applicant's intended major. The "personal achievement index" reflects personal characteristics including the extent to which the applicant took advantage of opportunities, and includes the student's essays as well as quantitative factors including the ratio of the applicant's SAT scores to his high school's mean score.[4] After admitting top-decile students, the university assigns the remaining positions using (up to discretization into "cells") a cutoff in a weighted sum of the two indices.[5]

[3] *Gratz v. Bollinger, 539 U.S. 244*. In the years 2003-2023 it was illegal for institutions that considered applicants' race to do so via a publicly known formula or points system. On the use of private criteria at elite universities, see also Karabel [2005].

[4] This personal score was introduced in 1997, one year before Texas Top Ten took effect. In 2001 the Texas legislature amended the law to include coursework requirements, but these rules first took effect with the 2004-2005 ninth-grade class. The data in this paper predate this change.

[5] See 570 U.S. ___ (2013) page 4, for a brief discussion. The court's decision is available at http://www.supremecourt.gov/opinions/12pdf/11-345_l5gm.pdf.

Importantly, non-top-decile admissions involved uncertainty. In particular, UT Austin listed the components of the personal index, but did not publicly specify the formula or weights used to construct a score from these inputs in any year.

Texas Top Ten had been the subject of political controversy because of its scope. In the years after 1997, the University of Texas at Austin expanded in order to meet rising demand caused in part by the program. However, this expansion stopped by 2002 while demand continued to rise. For this reason I take capacities in 2002 as given. In 2009, a legislative compromise capped automatic admissions UT Austin at 75% of its entering class beginning in 2010 (it had been 81% in 2009).[6] The data in this paper predate this change.

When the 2002 cohort that is the focus of this paper entered college, the flagship institutions did not consider race. In 2003, a pair of Supreme Court decisions, *Grutter v. Bollinger* and *Gratz v. Bollinger* respectively, allowed the use of race in admissions, provided that race was not part of an explicit formula. I include URM status in my model of admissions, although universities could not explicitly condition on it, as they may have attempted to proxy for it.

**Prices, Scholarships, Financial Aid:** While not my focus, I account for the presence of scholarship and mentoring programs that existed at the time. The Longhorn and Century Scholars programs were introduced in 1999 and 2000, respectively, and operated at partially overlapping sets of predominantly low-income urban high schools, 110 in total. Both programs provided grants of $4000-$5000 to students from designated high schools who enrolled in the relevant flagship university. These grants typically substituted one-for-one with loans. In addition to financial aid, the programs involved mentorship and academic support on campus, as well as information sessions in high schools and visits to high schools by current college students.[7]

In each year, 90% of matriculating students at UT Austin are categorized as "in-state" and pay in-state tuition. Texas' flagship universities' admissions offices did not choose among in-state students on the basis of ability to pay. Moreover, in the year that the surveyed students entered college, public institutions' tuition levels were set by the state.[8] My

[6] See the university's announcement, http://www.utexas.edu/news/2009/09/16/top8_percent/ posted on Sept. 16, 2009.

[7] The Longhorn Scholarship increased enrollment, graduation rates, and earnings of scholarship recipients at UT Austin [Andrews et al., 2020]. Estimated impacts of the Century Scholars program were not significant. As with similar programs outside of Texas [Angrist et al., 2009, Clotfelter et al., 2018, Page et al., 2019], neither program affected admissions chances or otherwise interacted with the admissions process.

[8] In 2003, Texas House Bill 3015 deregulated tuition in Texas universities beginning in the 2003-04 academic year. See http://www.utexas.edu/tuition/history.html.

model is designed to match the public institutions' policies in 2002. For this reason, unlike Fu [2014] and Fillmore [2016], I take list prices as given and do not consider "optimal" pricing.

This paper focuses on high school seniors applying as first-time freshman applicants. I abstract from transfers [Andrews et al., 2011] and from pre-college investments and choices [Hickman and Bodoh-Creed, 2018]. The percent plan provided incentives for students to choose high schools strategically [Cullen et al., 2013], but the share of students who moved in response to incentives was small relative to the direct and informative effects.[9] I abstract from this channel, taking students' high schools as exogenous.

The data in this paper contain outcomes from the first few years of college: GPA, persistence, and major choice. In 2002, certain majors at flagships, such as engineering and business at the University of Texas at Austin, required their own applications. The percent plan did not provide automatic admission to these majors. I revisit this issue when discussing impacts on major choice.

# 3 Data

This paper uses administrative records of Texas colleges and universities, and a set of surveys of high school students conducted by the Texas Higher Education Opportunity Project. My primary dataset is a survey of seniors (final-year students) in Texas public high schools. I use a followup survey conducted in 2003 as well. I supplement this data with administrative data from the state's flagship universities on admissions, matriculation, and enrolled students' grades, persistence, and major choice.

**THEOP survey:** THEOP selected 105 Texas public high schools at random in the spring of 2002. Of these, 86 high schools gave permission for in-class surveys; all seniors and sophomores who were present in school on "survey" day filled out a paper-and-pencil survey. These "survey days" took place between March 4, 2002 and May 27, 2002. At 12 additional schools, students completed surveys by mail in May 2002.

The initial wave surveyed 13,803 seniors: for cost reasons, the survey designers followed up with a randomly selected subsample, interviewing 5,836 of the original seniors

[9] Cullen et al. [2013] focus on the small fraction of students who have multiple available public schools, and would stand to increase their chances of guaranteed admission the most by changing schools without changing their residential location. A statistically significantly higher fraction of such students indeed change schools post-reform. Effects are likely smaller for students for whom changing schools would require moving residences.

the following year in the survey's second wave.[10]

The data include the year and term an applicant desired to enroll, demographics including gender, ethnicity, citizenship and Texas residency, and academic characteristics including high school class rank (by decile), SAT/ACT score, high-school average SAT scores, and whether a student has ever qualified for free or reduced-price lunch ("ever FRPL").

As outcomes, the student-level data include up to five college applications, together with indicators, for each application, for financial aid applications and for admissions. I observe whether students receive financial-aid offers, but do not observe individual financial-aid amounts. I do not see the student's matriculation choice in wave one, but if a student appears in wave two I see which institution, if any, she is currently attending, as well as any college or university the student attended within the past year.

The data do not provide household income, but I observe parents' education and occupational category. In estimation I treat income and expected family contribution (EFC), which is used for determining family aid, as unobserved random variables. I sample household income from the March CPS distribution conditional on living in Texas and the household head's education and occupation. I then construct each household's EFC using the official formula. I provide details in the data appendix.

I restrict the dataset to students who have taken the SAT or ACT. The final dataset consists of $4,143$ high school seniors. Of these, $1,975$ were tracked in wave two of the survey and hence have observable college choices.

Table 1 provides summary statistics at the student level.[11] On average, each student submits 1.4 applications, and 1.0 applications for aid. Roughly 85% of students submit at least one application. Because I condition on students who have taken a college entrance exam, average class rank (with lower numbers better) is 27%. I rescale SAT scores (raw scores go from 200 to 1600) so that the maximum possible score is 1.0. Roughly 18% of students come from schools with greater than 60% of peers ever qualifying for free or reduced-price lunch. About 9% come from schools that participate in the Longhorn Opportunity Scholars (LOS) program, which provides funding at UT Austin. About 1/4 of students come from affluent schools with low poverty rates. Of students whose matriculation decision is observed because they are surveyed in wave two, about 70% enroll in a

[10]The survey designers included in wave two all Black and Asian students in the original sample, as well as random samples of Hispanics and non-Hispanic whites. The 5,836 students who completed the wave two survey represent a 70 percent response rate. See the "Senior Wave 2 Survey Methodology Report" at http://theop.princeton.edu/surveys/senior_w2/senior_w2_methods_pu.pdf for details.

[11]This table reports unweighted means. The survey intentionally oversampled high-poverty schools. In table S1 in the online appendix, I report a version which uses the survey's population weights.

Table 1: Sample Means: Students (Unweighted)

| Variable | Mean | Std. Dev. | Min. | Max. |
|---|---|---|---|---|
| Participates in Century Scholars Program | 0.057 | 0.232 | 0 | 1 |
| Participates in Longhorn Opp. Scholars Program | 0.087 | 0.282 | 0 | 1 |
| White | 0.572 | 0.495 | 0 | 1 |
| Black | 0.095 | 0.293 | 0 | 1 |
| Latino | 0.201 | 0.401 | 0 | 1 |
| Asian | 0.063 | 0.244 | 0 | 1 |
| Female | 0.562 | 0.496 | 0 | 1 |
| Num. Guardians | 1.711 | 0.560 | 0 | 2 |
| Class Rank | 0.268 | 0.181 | 0.1 | 1 |
| SAT* | 0.666 | 0.119 | 0.006 | 1 |
| HS mean SAT | 0.676 | 0.051 | 0.431 | 0.806 |
| SAT / HS mean SAT | 0.985 | 0.167 | 0.008 | 1.553 |
| HS Poverty (share ever FRPL) | 0.345 | 0.253 | 0.023 | 1 |
| High Poverty HS | 0.181 | 0.385 | 0 | 1 |
| Affluent HS | 0.263 | 0.44 | 0 | 1 |
| Applied Anywhere | 0.854 | 0.353 | 0 | 1 |
| Applications | 1.437 | 1.138 | 0 | 5 |
| Admissions Offers | 1.06 | 0.993 | 0 | 5 |
| Aid Applications | 1.034 | 1.053 | 0 | 5 |
| Enrolled | 0.334 | 0.472 | 0 | 1 |
| Did Not Enroll in 4-year Institution | 0.143 | 0.35 | 0 | 1 |
| Not in Wave 2 | 0.523 | 0.5 | 0 | 1 |
| N | | 4143 | | |

four-year college.

On average, students live approximately 190 miles from colleges to which they could potentially apply. However, actual applications are to closer universities, approximately 105 miles from home. Matriculations in four-year institutions are closer still, with students traveling approximately 90 miles from home on average.[12]

I examine differences in demographics and behavior by top-decile status, participation in wave 2, and participation in the Longhorn Opportunity Scholars program. Results are given in Appendix tables S3 through S4, respectively. Wave-2 students are disproportionately Black and Asian. These differences reflect an intentional decision by survey designers to oversample Black and Asian students. They are similar to wave-1 students in SAT scores and high school poverty rates. Top-decile applicants are about 6 percentage points more

[12] A full set of results are available in appendix table S2, which shows sample means at the level of potential applications, observed applications, admissions offers, and matriculations.

female than non-top-decile students. They are more likely to be Asian, and less likely to be Black, than non-top-decile students, but are equally likely to be White and/or Hispanic. They are much more likely to apply to UT-Austin, to apply for aid there, and to be admitted there. LOS participants are more likely to be Black, equally likely to be Hispanic, and less likely to be White or Asian, than the general population. They come from poorer schools than average and have lower SAT scores.

**Comparison to administrative data:** In addition to the survey, THEOP provides admissions, enrollment, and demographic data for the universe of first-time freshman applicants to nine colleges and universities in Texas, including the two flagship institutions from the early 1990s (before Texas Top Ten) through the year of the survey. Table A1 in the Appendix provides the list of colleges and years covered. I restrict attention to the flagship institutions, at which I observe more detailed data, including exact class rank. There I observe the year and term an applicant desired to enroll. I also see demographics: gender, ethnicity, citizenship, TX residency; and academic characteristics: high school class rank, SAT/ACT score, AP classes taken, as well as characteristics of the student's school including high school mean SAT scores and fraction eligible for free/reduced lunch. By semester, for enrolled students, I observe credit hours earned, semester GPA, department and field of study.

# 4 Model

This section describes the model of applications, admissions, enrollment and outcomes. Timing is as follows. First, high-school seniors who have taken a college entrance exam simultaneously choose which colleges to apply to and whether to apply for financial aid at those colleges. Second, colleges observe applications and choose whom to accept. Third, students observe their admissions outcomes, as well as matriculation-time preference shocks, and choose where to matriculate among the colleges that admit them. Finally, persistence/dropout, major choice, and college grades realize.

I begin by presenting the matriculation stage, then work backwards, describing admissions and applications. Finally, I discuss post-enrollment outcomes. Let $i = 1\ldots, I$ denote the set of students, and $j = 1, \ldots, J$ the set of four-year colleges and universities. In estimation I consider $J = 7$ inside options, consisting of the two flagship universities and five aggregate options spanning the set of other four-year institutions. Section 5.2 gives details.

## 4.1 Students' preferences and matriculation decisions

At the time of matriculation decisions, student $i$ picks the college $j \in B_i \subseteq \{0, 1, \ldots, J\}$ offering the highest value $U_{ij}$, where

$$U_{ij} \equiv \delta_j + w_j \beta_i^w + x_{ij} \beta^x + z_i^{\text{admit}} \beta^z + p_{ij} \beta_i^p (y_i) + \nu_{ij} + \varepsilon_{ij}^{\text{enroll}}. \tag{1}$$

The choice set $B_i$ consists of all four-year programs to which $i$ has applied and gained admission, as well as the outside option of not attending any four-year college immediately, which gives utility $U_{i0} = \varepsilon_{i0}^{\text{enroll}}$.

For "inside" options $j \in \{1, \ldots, J\}$, the mean utility term $\delta_j$ captures college-level demand shifters. The variables $w_j$, $z_i^{\text{admit}}$, and $x_{ij}$ are observables, discussed below. Student-level characteristics $z_i^{\text{admit}}$ enter students' preferences for colleges. They will also enter colleges' preferences over students as described in the next section. In the empirical specification, we have

$$z_i^{\text{admit}} = (SAT_i, classrank_i, SAT_i / \overline{SAT}_{h(i)}, \text{poverty}_{h(i)}, \text{urm}_i), \tag{2}$$

where $h(i)$ denotes $i$'s high school, $SAT_i$ is student $i$'s standardized college entrance exam score, $SAT_i / \overline{SAT}_{h(i)}$ the ratio of $i$'s exam score to the mean score at $h(i)$, $\text{urm}_i$ is an indicator equal to 1 if the student is Black or Hispanic, and $\text{poverty}_{h(i)}$ is the share of $i$'s high school peers who have ever qualified for free or reduced-price lunch, a common proxy for poverty in U.S. educational settings.

Preference shifters $x_{ij}$, which do not affect admissions decisions, consist of the presence of targeted scholarships $LOS_i \times 1(j = \text{UT Austin})$ and $Century_i \times 1(j = TAMU)$, distance to college, and an indicator for nearby (distance $< 25$ miles) colleges. In addition, to allow flexible substitution patterns under changes in the percent plan, $x_{ij}$ includes interactions of indicators for each flagship with high school poverty, with minority status, and with SAT scores, as well as an interaction between the student's exam score and an indicator for private institutions.

Terms $w_j$ with random coefficients are student/faculty ratio ($\text{S/F ratio}_j$), an indicator for all inside options, and a term equal to 1 for UT Austin and -1 for Texas A&M, allowing negative correlation.[13] Without loss $E(\beta_i^w) = 0$, as mean effects of college-level observables $w_j$ are subsumed by college effects $\delta_j$. I assume that the random coefficients $\beta_i^w$ are

[13] I have also estimated specifications with an indicator for public institutions, and have found very small variance of the associated random coefficient.

independently normally distributed with variance $\sigma_{rc}^2$.

The term $p_{ij}$ captures the net price that $i$ expects to pay if enrolled in $j$. Person $i$'s distaste for price, $\beta_i^p(y_i)$, varies with their family income, $y_i$, as $\beta_i^p(y) = -\log(1+\exp(\overline{\beta}_0^p + \overline{\beta}_y^p y))$. This functional form guarantees that price is a "bad".

There are two unobserved preference shocks, $\nu_{ij}$ and $\varepsilon_{ij}^{\text{enroll}}$. The key distinction is that applicants are endowed with knowledge of $\nu_{ij}$ before they choose applications, but observe the vector $\varepsilon_i^{\text{enroll}} \equiv (\varepsilon_{i0}^{\text{enroll}}, \varepsilon_{i1}^{\text{enroll}}, \ldots, \varepsilon_{iJ}^{\text{enroll}})$ only once they have received admission offers. Two shocks are needed to fit data that include application and enrollment decisions; sufficiently high draws of $\varepsilon_{i0}^{\text{enroll}}$ or low draws of $\varepsilon_{ij}^{\text{enroll}}$ for $j > 0$ rationalize the event that a person applies to and receives offers from colleges but then does not enroll.

The application-time shocks $\nu_{ij}$ are normally distributed, independently across people and colleges, with mean zero and heteroskedastic variance $\sigma_j^2$. I assume that they are independent of $(z_i, x_{ij})$ and of potential financial-aid awards, discussed below. This restriction is innocuous for my counterfactuals, which hold $z_i$, $x_{ij}$, $w_j$, and list prices fixed, but would be restrictive if I were to consider changes in colleges' prices or characteristics, or endogenous choices of high school or study effort by students that might affect elements of $z_i$.

I model $\varepsilon_i^{\text{enroll}}$ as a nested logit, with all inside goods in a common nest with parameter $\lambda$, and the outside option as its own nest. This specification allows for common shocks to all inside goods relative to the outside good, such as a shock to the value of entering the labor force immediately or staying home to care for a family member.

Let $C_i \in \{0, \ldots, J\}$ denote the element of $B_i$ that maximizes $i$'s utility $U_{ij}$. If offer sets were randomly assigned, one could estimate demand using the choice of $C_i$ only. A challenge is that the choice set $B_i$ may be selected on applicants' beliefs about net prices and admissions chances, and on applicants' preferences, via application decisions. To address this challenge, I model these sources of selection, as described next.

## 4.2 Prices and financial aid

Colleges post list prices but offer discounts—financial aid—to low-income applicants as follows. In order to complete Free Application for Federal Student Aid (FAFSA), a form that is required for applications for need-based financial aid, an applicant must calculate their *expected family contribution* $EFC_i \geq 0$. The quantity $EFC_i$ is a nonlinear function of the income and assets of $i$'s parents, $i$'s own income if any, and characteristics of $i$'s household. Federal student aid eligibility in the U.S. is based on the difference between the cost of attendance and $EFC_i$. In practice, many universities, including the University of Texas at Austin, determine institutional financial aid awards as a function of students' federal $EFC$. In doing so, however, universities may fail to provide sufficient aid to match the student's need as specified by the federal formula.

The total aid that $i$ may receive in the model is given by

$$Aid_{ij} = \max\{\alpha_j^{\text{aid}}(p_j^{max} - EFC_i), p_j^{max}, 0\},$$

where $p_j^{max}$ is the list price. The term $\alpha_j^{\text{aid}}$ captures the amount of "need" that the college meets. I assume that colleges do not price-discriminate on the basis of students' preferences.

When $i$ is choosing where to enroll, $i$ observes the price $p_{ij}$, net of financial-aid, for all $j \in B_i$. In the event that $i$ has not applied for financial aid, or is not eligible for aid, $i$ will pay the full list price if enrolled in $j$. Thus we have:

$$p_{ij} = \begin{cases} p_j^{max} - Aid_{ij} & \text{if } i \text{ receives aid} \\ p_j^{max} & \text{otherwise.} \end{cases}$$

Because Longhorn and Century Scholarship programs involved mentoring and on-campus programs as well as the possibility of grants that typically substituted one for one with aid, I model these programs as directly entering payoffs. In particular, I include an indicator for the "own" program (e.g. the Century Scholarship, for students enrolling in Texas A&M) as an element of $x_{ij}$. As discussed below, these programs may also directly affect students' grades, persistence, and major choice. Later, in estimation, I will use the availability of the "opposite" program (e.g. a scholarship at UT Austin, for students who enroll at Texas A&M) as a shifter of enrollment that is excluded from potential outcomes at the chosen program.

## 4.3 Colleges' preferences

Each college $j$ maximizes the expected quality of its entering class subject to the constraint that its expected enrollment is less than its capacity $k_j$. From the perspective of college $j$, the quality of student $i$ is given by

$$\pi_{ij} = z_i^{\text{admit}} \gamma^{\text{admit}} + q_i + \mu_{ij}^{admit}, \tag{3}$$

where $z_i^{\text{admit}}$ are the admissions-relevant observables defined in equation (2), and $q_i$ is a student-level "caliber" term, commonly observed by colleges and capturing the general quality of $i$'s background, essays, transcript, and other application materials. It is not essential that the admissions shifters $z_i^{\text{admit}}$ do not vary at the college-by-individual $(ij)$ level. The current specification is chosen to match UT Austin's practices, but one could include match-level observables. The term $\mu_{ij}^{\text{admit}} \sim N(0,1)$ is a match-level shock, independent across $(i, j)$, and privately observed by college $j$ in the event that $i$ applies to $j$.

I assume that each college admits those applicants whose quality $\pi_{ij}$ is greater than a cutoff $\underline{\pi}_j$. The Top Ten Percent policy is a constraint on admissions at in-state public institutions: if student $i$ applies to $j$, $i$ is in the top decile of his class, and $j$ is a Texas public institution, then $i$ must receive an offer from $j$. Otherwise, if $\pi_{ij} < \underline{\pi}_j$, the applicant is rejected.

A college is said to be "selective" if its cutoff is determined by a capacity constraint. For each selective college $j$, the cutoff $\underline{\pi}_j$ must be such that, given rivals' cutoffs $\underline{\pi}_{-j}$, the expected number of students matriculating at $j$ is equal to its capacity $k_j$. I take $k_j$ as exogenous, but colleges' cutoffs $\underline{\pi}$ are equilibrium objects. Under a shift in demand or a change in the rules, these cutoffs will adjust so as to just satisfy capacity constraints. In Appendix B.1, under the assumption that colleges' preferences are additively separable across students, I prove that cutoff strategies are optimal among a broad class of strategies.

I maintain the following assumptions. First, as in Avery and Levin [2010] I assume that colleges cannot commit to ex-post suboptimal admissions rules. A college might want to publicly commit to favoring certain applicants in order to induce them to apply, or to not apply to competitors. In this case, the weights $\gamma^{\text{admit}}$ need not reflect its preferences. Other than via a law such as Texas Top Ten, it's not clear that U.S. colleges are able to credibly do so. In particular, since the "soft" information that I am modeling with caliber $q$ and shocks $\mu$ is not verifiable, it seems implausible to commit to using it in a specific ex-post-suboptimal way. Second, I abstract from "yield management", a preference for admitting students who are likely to accept an offer. A college that wants to do so might deviate from a cutoff rule by rejecting applicants it believes are unlikely to matriculate. The flagship institutions at the time did not engage in this practice.

Third, for the main results, I treat parameters $\gamma^{\text{admit}}$ as structural. A college might want to fill some share of seats with students with high personal scores, however, but hold other seats for high academic achievers. The percent plan forces colleges to admit students with high class rank, an academic characteristic. Universities with such preferences might down-weight academic characteristics for the remaining applicants. If there were no percent plan, they might place greater weight on academic characteristics and less weight on personal characteristics.[14] In Appendix E.6 I simulate counterfactuals in which colleges do so.

## 4.4 Students' information and application decisions

**Information about Admissions Chances:** Students have limited information about their admissions chances. They observe their own objective characteristics $z_i^{\text{admit}}$, but do not observe match-level shocks $\mu_{ij}^{\text{admit}}$, and observe only a noisy, potentially uninformative signal of $q_i$, denoted $s_i$.

The signal $s_i \in \mathbb{R}$ is distributed jointly normally with $q_i$ for each applicant and is independent across applicants, according to:

$$\begin{pmatrix} q_i \\ s_i \end{pmatrix} \sim N\left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_q^2(z_i^{\text{info}}) & \sigma_s^2(z_i^{\text{info}}) \\ \sigma_s^2(z_i^{\text{info}}) & \sigma_s^2(z_i^{\text{info}}) \end{pmatrix} \right),$$

where $z_i^{\text{info}}$ is a vector of observables affecting the quality of students' information. Here $\sigma_s^2(z_i^{\text{info}})$ is the variance of the component of $i$'s admissions caliber that is known to $i$, and $\sigma_q^2(z_i^{\text{info}})$ is the full variance of $q_i$. Because the informational content of $s$ is invariant to any monotone transformation, it is without loss to fix the covariance equal to the variance of $s$.[15] This normalization admits a convenient interpretation, in which we partition $q_i$ into the component that is observed by $i$ and the residual, which is not. We may write the marginal distribution of $s$ and the conditional distribution of $q$ as follows:

$$s_i \sim N(0, \sigma_s^2(z_i^{\text{info}}))$$
$$q_i|s_i \sim N(s_i, \sigma_{q|s}^2(z_i^{\text{info}})),$$

where

$$\sigma_{q|s}^2(z_i^{\text{info}}) = \sigma_q^2(z_i^{\text{info}}) - \sigma_s^2(z_i^{\text{info}}).$$

$\sigma_s^2$ can be interpreted as the strength of $i$'s private information about his admissions chances. When $\sigma_s^2(z_i^{\text{info}})$ is large, there is a large component of $i'$s admissions chances (relative to the impacts of observables such as SAT scores and to idiosyncratic shocks $\mu_{ij}^{\text{admit}}$) which is known to $i$ but not to the econometrician. When $\sigma_{q|s}^2(z_i^{\text{info}})$ is large, in contrast, $i$ faces a

[14] Antonovics and Backes [2013a] argue that the University of California changed the relative importance of academic characteristics after the end of affirmative action.

[15] One may work instead with $\tilde{s} \equiv s/\sigma_s$, which has variance 1 and covariance with $q$ given by $\sigma_s$.

large amount of uncertainty about admissions chances after conditioning on his signal. An informative signal structure is one in which $\sigma_s^2$ is large and $\sigma_{q|s}^2$ is small.

Importantly, the quality of students' information, as well as the importance of unobserved quality relative to the information provided by SAT scores and grades, may vary across students. In the data, we have[16]

$$z_i^{\text{info}} \equiv (1, \text{poverty}_{h(i)}, \text{urm}_i).$$

I specify the following functional forms, which guarantee that variances remain positive but are approximately linear in $z_i^{\text{info}}$ for large values of the index:

$$\sigma_s(z_i^{\text{info}}) = \log(1 + \exp(\gamma^s z_i^{\text{info}}))$$
$$\sigma_{q|s}(z_i^{\text{info}}) = \log(1 + \exp(\gamma^{q|s} z_i^{\text{info}})).$$

This information structure allows for correlation in admissions outcomes conditional on observed characteristics and, through the signal, it allows for selection on an econometrician-unobserved admissions-relevant characteristic in the application decision.

I assume $(q_i, s_i)$ are orthogonal to the idiosyncratic admissions preferences $\mu_{ij}^{\text{admit}}$. The restriction $q_i \perp \mu_i^{\text{admit}}$ imposes a factor structure on admissions chances with one latent common factor. This restriction to a single factor is not essential, but it is important that students' information $s_i$ is about common terms rather than idiosyncratic match values. The model rules out private information about admission to a specific university that is unrelated to admissions chances elsewhere, for example via a legacy preference.

I make the additional restriction that $(\mu_i^{\text{admit}}, q_i, s_i)$ are mean-independent of $(\beta_i^w, \nu_i, \varepsilon_i^{\text{enroll}})$. That is, conditional on observables, applicants who prefer college $j$ are not more likely to be admitted. This restriction rules out legacy status at elite private institutions, for example, which may enter students' and colleges' preferences. Fewer than 2% of students in the sample, weighted by population weights, matriculate at elite private institutions.

**Financial Aid Awareness:** Many students fail to complete financial aid applications, even where it appears that those applications would be likely to succeed and to have large effects on net price. Accordingly, the model allows imperfect awareness of financial aid.

$Aware_{ij}$ is a latent variable which determines whether $i$ is able to apply for financial

[16] In principle one may include all elements of $z_i^{\text{admit}}$. Finite data motivate the inclusion of fewer variables in indices for variances, such as $z_i^{\text{info}}$.

aid at college $j$. If $Aware_{ij} = 1$ then $i$ is said to be *aware* of financial aid at college $j$. Conditional on applying to $j$ a student completes an application for financial aid at $j$ if and only if $Aware_{ij} = 1$.

$Aware_{ij} = 0$ need not be taken literally as unawareness, but may reflect difficulty in completing the FAFSA, unawareness of deadlines, difficulty obtaining prior-year tax returns, or other frictions. Students are aware of financial aid at college $j$ with probability

$$Pr(Aware_{ij} = 1) = \Phi(\alpha_i^{\text{aware}} + \alpha_j^{\text{aware}} + (x_{ij}, z_i^{\text{admit}})'\alpha_{x,z}^{\text{aware}} + y\alpha_y^{\text{aware}}), \quad (4)$$

where $\alpha_i^{\text{aware}} \sim N(0, \sigma^2_{\text{aware}})$ is an individual random effect. Awareness at $j$ may vary with all variables that enter preferences for $j$.

A student's awareness of financial aid at $j$ means that she understands that the college's list price is not necessarily the price she will pay if she matriculates. Students who are not aware of financial aid at college $j$, in contrast, assume that they will pay the full cost of attendance when calculating the expected value of application portfolios. Students who would qualify for non-zero financial aid will therefore be less likely to apply to colleges where they are unaware of aid. An implication is that matriculation choices $C_i$ and choice sets $B_i$, described next, may be selected on awareness of aid.

**Application Costs and Optimal Application Portfolios:** At the time of applications, students know all terms that will determine their utilities from attending each college except for their matriculation-time shocks $\varepsilon_i^{\text{enroll}}$. Let $u_{ij} \equiv U_{ij} - \varepsilon_{ij}^{\text{enroll}}$ denote the portion that is known. The value of a choice set $B$ before these shocks are known[17] is given by

$$U_B \equiv \log\left(1 + \left(\sum_{j\in B} \exp(u_{ij}/\lambda)\right)^{\lambda}\right).$$

Students trade off the gain in expected utility from college applications against the applications' cost. The cost of applications depends on the number of colleges to which a student applies. Let

$$C(A, z_i^{\text{info}}, \varepsilon_i^{\text{app}}) = c^{fixed}_{z_i^{\text{info}}} + c^{var}_{z_i^{\text{info}}}|A| - \varepsilon_{iA}^{\text{app}}, \quad (5)$$

[17] The expression for $U_B$ follows from the nested logit error specification. See Train, "Discrete Choice Methods with Simulation" (2009), chapter 4. Without loss I subtract Euler's constant, $k$, from the value of each portfolio including the empty portfolio.

be the cost of applying to portfolio $A \subseteq \{1,\ldots,J\}$ for a student with observables $z_i^{\text{info}}$ and application-cost shocks $\varepsilon_i^{\text{app}} \sim Gumbel(0,\lambda^{\text{app}})$. In estimation, I assume

$$c^{fixed}_{z_i^{\text{info}}} = \gamma_z^{fixed} z_i^{\text{info}}$$
$$c^{var}_{z_i\text{info}} = \gamma_z^{var} z_i^{\text{info}}$$
$$\lambda^{\text{app}} = \log(1+\exp(\gamma_z^{\text{shock}} z_i^{\text{info}}))$$

That is, there is a fixed cost of submitting any applications, plus a marginal cost per application that does not vary with the size of the portfolio, and a portfolio-level shock. To capture heterogeneity in access to college counselors or other resources, I allow the application costs and variance of shocks to vary with the same parameters $z_i^{\text{info}}$ that shift the information structure.

In the main specification, for each applicant, these shocks are independent and identically distributed across portfolios. Specifications in which $\lambda$ is fixed at a small value give similar results. From the student's perspective, the chance of admission to a set $B \subseteq A \subseteq \{1,\ldots,J\}$, having applied to $A$, is:

$$P_i(B;A) = \int_{-\infty}^{\infty} \prod_{j\in B} \left(\Phi(z_{ij}^{\text{admit}}\gamma_j^{\text{admit}} + q_i - \underline{\pi}_j)\right) \prod_{j\in A\setminus B} \left(1-\Phi(z_{ij}^{\text{admit}}\gamma_j^{\text{admit}} + q_i - \underline{\pi}_j)\right) \phi(q_i; s_i, \sigma^2_{q|s}) dq_i,$$

where $\Phi(\cdot)$ is the standard normal CDF, and $\phi(\cdot;\mu,\sigma^2)$ the density of a normal random variable with mean $\mu$ and variance $\sigma^2$. Students do not observe their residual quality $q_i$ or match-specific shocks $\mu_{ij}^{\text{admit}}$, but integrate over the conditional distribution of $q_i$ given their signals. The expected value of an application portfolio $A \subseteq \{1,\ldots,J\}$ is given by

$$V_i(A) \equiv \sum_{B\subseteq A} P_i(B;A) \log\left(1+\left(\sum_{j\in B}\exp(u_{ij}/\lambda)\right)^{\lambda}\right).$$

Students choose a portfolio to maximize $V_i(A) - C(A, z_i^{\text{info}}, \varepsilon_i^{\text{app}})$.

Conceptually, application cost terms $(c^{\text{fixed}}, c^{\text{var}}, \varepsilon_i^{\text{app}})$ differ from preferences in that the costs are sunk at the time of matriculation decisions. The shocks $\varepsilon_i^{\text{app}}$ may be interpreted as capturing optimization noise, behavioral heuristics, and limited strategic sophistication, as well as any preferences over applications per se that lead individuals to depart from the application that maximizes the expected value of the offer set $B_i$, provided that these are invariant under counterfactuals. To the extent that disadvantaged students may be more (or less) subject to frictions and heuristics, we may expect the scale parameter $\lambda$ to vary.

## 4.5 Equilibrium

An equilibrium can be written as a tuple, $\left\{\{A_i\}_{i\in I},\{\underline{\pi}_j\}_{j\in J}\right\}$, where $A_i \in \mathscr{A}$ is the application portfolio that solves $\max_{A\in\mathscr{A}} V_i(A)$, given cutoffs $\underline{\pi}$, for each student $i \in I$, and the cutoffs $\{\underline{\pi}_j\} \in \mathbb{R}^J$ are those that solve each college's problem given the matriculation probabilities induced by $\{A_i\}_{i\in I}$, $\underline{\pi}_{-j}$, and students' preferences.

In appendix D, I provide a formal definition of a rational-expectations equilibrium of the college market which does not impose cutoff strategies, prove that an equilibrium exists, and show that all equilibria can be written as tuples of this form.

Equilibrium need not be unique. The estimation strategy described in the next section conditions on the equilibrium that is being played in the data. Thus non-uniqueness is not a problem for estimation, but may present challenges in counterfactuals. In the Appendix I prove that, conditional on applications, there is a unique vector of cutoffs $\left\{\underline{\pi}_j\right\}_{j\in J}$ that solves each college's problem. This "limited multiplicity" result shows that strategic interaction among colleges, conditional on applications, is not itself a source of multiplicity.

## 4.6 Post-Enrollment Outcomes

If student $i$ enrolls in college $j$, he obtains outcomes of the form

$$\text{outcome}_{ij} = z_i^{\text{outcome}}\gamma_{z,j}^{\text{outcome}} + q_i\gamma_{q,j}^{\text{outcome}} + \mu_{ij}^{\text{outcome}}. \tag{6}$$

The outcomes of interest consist of cumulative GPA, persistence after two years, and an indicator for choosing a STEM major.

Observables $z_i^{\text{outcome}}$ consist of all elements of $z_i^{\text{admit}}$, as well as an indicator for the availability of a LOS/Century scholarship (which is also an element of $x_{ij}$): we have $\text{scholarship}_{ij} = 1$ if $i$'s high school has a targeted scholarship at college $j$. The term $\mu_{ij}^{\text{outcome}}$ is an idiosyncratic shock which has mean zero conditional on covariates.

Coefficients $\gamma_{z,j}^{\text{outcome}}$ and $\gamma_{q,j}^{\text{outcome}}$ differ by outcome, and by college $j$ among the two flagship institutions, allowing for match effects on observables. The model allows for selection on ability levels. For instance, outcome indices may be identical to the weights on $z$ and $q$ in admissions, in which case students with higher potential outcomes will be sorted to more-selective institutions. The model is also consistent with assortativity—the impacts of $q$ on grades or persistence being larger at the most selective institutions—or the reverse, or no such pattern.

An important restriction, however, is that potential outcomes at flagship universities are correlated with application and admissions outcomes via a single index, $q_i$. I rule out selection on idiosyncratic factors privately observed by admissions offices, such as a perception that a student is uniquely suited to Texas A&M. In addition, I rule out selection on students' preferences, as might arise from students preferring institutions where they know that they will perform especially well, or via a causal effect of (un)happiness.

In Appendix E.7 I relax this second restriction, estimating an extension in which outcomes may be selected on students' preference shocks as well as their caliber. If students value grades, for instance, know that they will earn higher grades at some institutions than others, and choose accordingly, then we will recover a positive relationship between preference shocks and grades conditional on caliber. I find that predictions under counterfactuals in this extension are essentially unchanged.

Other empirical studies of Texas colleges (Mountjoy and Hickman (2020)) and of K-12 schools (Abdulkadiroglu, Pathak, Schellenberg, Walters) find that application portfolios and colleges' admissions decisions contain information about future outcomes beyond that provided by other observables, but consistently with my assumptions and findings they do not find evidence of significant unobserved "match effects" on which applicants are able to select.

Finally, Equation (6) is consistent with selection on students' gains from flagship enrollment that is due to heterogeneity in the institutions that students would attend otherwise. For example, high-SES students, who may be more likely to attend flagships if there is no percent plan, may also have fallback options with relatively high returns in the event they do not attend a flagship university. I discuss this issue further in Appendix E.3.

# 5 Estimation

I estimate cutoffs $\left\{\underline{\pi}_j\right\}_{j=1,\ldots,J}$ under the percent plan, and all parameters except the outcome production function parameters, jointly via the generalized method of moments. In a second stage, I estimate outcome equations.

I use two sets of moments in the first stage of the procedure. The first set of moments is the score of the likelihood of all observables in the THEOP survey. For students in the first wave of the survey only, we see application sets $A_i$, admission sets $B_i$, and financial aid applications. For students who are in both waves of the survey, we observe matriculation decisions as well. The second set of moments matches average aid awards in IPEDS to values predicted by the model at each college.[18]

In the second stage, I require that the observed distributions of outcomes in the flagship administrative data match model-predicted outcomes of surveyed students conditional on the event that those students enroll in flagships. I do so with indirect inference strategy.

## 5.1 Identification in practice

This paper faces three key empirical challenges: (1) disentangling students' payoffs over colleges, colleges' preferences over students, and the quality of students' and colleges' information when these objects jointly determine applications, admissions, and matriculation decisions, (2) identifying aid (un)awareness and its impact on students' payoffs, and (3) estimating outcome production functions in the presence of selection into college enrollment.

**Students' preferences and costs:** To identify demand parameters, I use students' application and matriculation decisions. The initial choice of application portfolios depends on students' utilities and beliefs about admissions chances, as each application portfolio induces a lottery over choice sets. To pin down preferences it is helpful to have a source of exogenous variation in choice sets. The Texas Top Ten law serves this purpose, as it exposes similar students on either side of the class rank cutoff to different lotteries over admissions outcomes.

Importantly, the size of the discontinuity in "true" admissions chances at the 10% threshold depends on students' characteristics. At many schools the median second-decile student would be very likely to gain admission to UT Austin, even in the presence of the Top Ten Percent plan. Differences in application rates across the threshold reveal imperfect information about admissions chances.

In addition, in order to identify the parameters that affect substitution patterns, it is especially useful to observe application portfolios. If a particular form of preference heterogeneity (such as heterogeneity in taste for low student/faculty ratio) is important, there will be correlation in the relevant characteristic within application portfolios.

Texas-top-ten-induced variation in admissions chances, and the existence of multiple measurements of preferences—I observe initial application and final enrollment decisions—

[18] I construct average awards for each college using the IPEDS 2003 financial aid supplement. Aid totals include loans as well as federal, state, and institutional grant aid.

also helps distinguish application costs from utilities. Intuitively, costs are incurred by all applicants, but utilities can be realized only if applicants are admitted. Enrollment decisions do not depend on costs, or on information about admissions chances, except via selection.

**Admissions parameters:** To identify colleges' preferences, I make use of admissions outcomes. A challenge is that application decisions depend in part on students' private information about admissions chances, and I observe admissions outcomes only at those colleges to which students apply. Distance to college shifts application decisions but (following the literature, and based on my knowledge of flagship universities' admissions rules) is excluded from admissions offices' preferences. Moreover, because colleges are substitutes, distance to other colleges is also an excluded shifter of applications to college $j$.

The presence of a targeted scholarship, which may affect grades directly, is excluded from admissions offices' decisions, which are made independently from financial-aid decisions. This restriction is consistent with Texas flagships' admissions policies.

**Information structure:** Data on applications portfolios and associated admissions outcomes make it possible to allow for private information about admissions chances. Marginals of admissions probabilities at particular colleges, taken as a function of $z$, provide information about admissions parameters $\gamma$. The importance of unobserved student-level attributes (summarized by $q$) is recovered from the correlation of admissions outcomes within an individual's application portfolio that remains after conditioning on observables. If outcomes remain highly correlated, then caliber is important.

The joint distribution of applications and admissions is used to recover the quality of students' signals $s$. If students have private information about their admissions chances, students with aggressive application portfolios will have relatively high probabilities of admission conditional on observables. To illustrate, consider a portfolio consisting of a flagship and a highly selective private university, and an alternative consisting of a flagship and a regional public campus. The strength of the signal $s$ is pinned down by the extent to which observably identical students who apply to the riskier portfolio are more likely to gain admission to schools that are common to both portfolios.

Moreover, the relation between application decisions and matriculation choices provides information about beliefs. For intuition, consider a student considering applying to a college at which the econometrician predicts that he has a very low chance of admission. If the econometrician is correct, the student must have had a very strong taste for this college in order for the application to have been worthwhile, and if admitted, should therefore be very likely to attend. In contrast, if the variance of $s$ is high, some students with weak preferences may apply to apparent "reach" schools because these schools are not in fact unlikely for them given their information.

**Awareness of Aid:** As mentioned in the introduction, I exploit the fact that awareness of aid is partially observed. In the event that a school belongs to a student's application set, I observe the decision to apply, or not apply, for aid. Identification exploits this data together with variation in application decisions that is excluded from awareness.

To illustrate, consider the case of independence. If application sets $A_i$ were independent of latent aid awareness $Aware_i$ conditional on observables, then one could simply observe the share of students with a given value of observables who are aware of aid at college $j$. The key identification challenge is that students who choose a given application set $A_i$ may select into that set on the basis of the latent financial-aid awareness.[19]

Identification of the marginal distribution of awareness of aid at college $j$ exploits variation in admissions applications $A_{ij}$ that is excluded from $Aware_{ij}$, analogous to identification in the Roy Model via a choice shifter that is excluded from outcomes (in this case, aid applications). This variation comes from policy variation and from the nature of the portfolio problem. For instance, local to the top 10% cutoff, crossing the threshold raises admissions chances at flagship institutions, making it more attractive to apply to them—and via the nature of the portfolio problem, less attractive to apply to private institutions—but does not directly affect aid awareness. More generally, as institutions tend to be substitutes in the portfolio problem, preference shifters for school $k$ that are excluded from preferences and awareness at $j \neq k$, such as distance to $k$, are also implicitly used in the estimation procedure.

Correlation in awareness across institutions then pins down the scale of random effects in aid awareness. Because I may observe a student applying to an institution without applying for aid, the setting differs from the classic "consideration set" setup (e.g. Goeree [2008]) in which belonging to a consideration set is a necessary condition for a product to be chosen. Identification in such settings typically relies on a shifter of "consideration" that is assumed to be excluded from payoffs, such as advertising [Goeree, 2008]. I do not assume such an exclusion restriction.

[19] One may think of awareness of aid at college $j$, denoted $Aware_{ij}$, as a potential outcome that is observed in the event $A_{ij} > 0$. The terms $A_{ij}$ and $Aware_{ij}$ may be correlated via mutual dependence on unobserved income $y_i$, as a lower price is desirable.

**Post-Enrollment Outcomes:** To recover outcome parameters, I rely on variation induced by Texas Top Ten, together with individual outcomes. I assume that grading policies, persistence rates, and major choice probabilities are continuous in high school class rank, conditional on enrolling in college $j$ and on all admissions-relevant variables including caliber $q$. Because caliber is relevant for discretionary admissions only, top-decile students have systematically lower caliber than those just below the top decile, conditional on having enrolled in a public institution.As a result, differences between the marginal automatically-admitted student and the marginal non-automatically-enrolled student provide information about the impact of admissions caliber on outcomes.

In addition, to estimate an outcome equation at program $j$ in the presence of selection into attending $j$, I exploit the presence of characteristics that enter students' preferences for other programs but are excluded from outcomes at program $j$. In particular, the Longhorn Scholarship raises a student's payoff from attending UT Austin, but is excluded from preferences for, and outcomes at Texas A&M. Students who enroll in Texas A&M despite attending a Longhorn-Scholarship high school may be selected on caliber and tastes. The case of the Century Scholarship and UT Austin is analogous.

**Other restrictions:** Some other restrictions are not essential. One could extend the empirical model to allow richer heterogeneity in colleges' and students' preferences. It is in principle possible to let $\gamma$ coefficients differ by college, and/or to allow a full set of interactions of $z$ with dummies for each college in preferences. Limited data motivate the current specification. Intuitively, $\gamma$ is pinned down by the marginal distribution of admissions offers at each college, while information about the variance of $q$ and the quality of students' signals $s$ come from the covariance of admissions offers within a portfolio and from covariances of admissions offers and students' behavior, respectively.

I have chosen to include interaction terms in students' payoffs which help with predicting demand and substitution patterns at flagships. Importantly, all terms which enter information, cost, admissions or outcome indices also enter preferences, and all terms that enter preferences also enter aid awareness indices.

## 5.2 Choice set

I restrict portfolios to contain a maximum of five colleges. In the final survey dataset, most individuals apply to no more than five colleges; among students in our final dataset, 97.9% submit five or fewer applications.

I aggregate colleges which receive relatively few applications. I consider the two flagship universities separately, as well as five aggregate institutions: non-flagship Texas public institutions, secular private institutions, religious institutions, out-of-state public colleges and universities, and highly selective institutions. I provide details and lists of institutions in data appendix B. Students may apply to multiple copies of an aggregate institution. For instance a student may submit five applications to in-state non-flagship colleges and universities.

To further speed computation, I restrict the possible set of portfolios: while I allow all portfolios of up to three colleges, I restrict the set of possible large portfolios. Among portfolios containing five applications, I allow only those that some applicant in the data was observed to have chosen. I allow portfolios of four colleges that are subsets of the allowed 5-college portfolios as well as all portfolios of four colleges which appeared in the data. This restriction reduces the choice set from $|A| = 560$ to $|A| = 281$ possible portfolios. It rules out certain unlikely combinations of colleges; for instance a student cannot apply to three highly selective private institutions and two of the smaller regional non-flagship public colleges.

## 5.3 First-stage GMM estimator

The first stage of estimation use two sets of moments. The first set of moment conditions in the first stage is given by the score of the likelihood of all data in the survey:

$$g^{surv}(\theta) = \frac{1}{N_{survey}} \nabla_\theta \sum_i \log \ell_i(\theta),$$

where $N_{survey}$ denotes the number of surveyed students, and

$$\ell_i(\theta) = \int_s \int_{\omega_i} \ell_i^A(\theta, \omega_i, s) \ell_i^{Aware|A}(\theta, \omega_i) \ell_i^{B|A}(\theta, s) \ell_i^{C|B}(\theta, \omega_i) dF_i(s; \theta) dG_i(\omega_i; \theta)$$

is the likelihood of individual $i$'s application set $A_i \subset \{1, \ldots, J\}$, aid applications $\text{Aware}_i \subset A_i$, offer set $B_i \subset A_i$, and choice $C_i \in B_i \cup \{0\}$ given parameters $\theta$, integrating out the signals $s$ and other terms—income, EFC, random coefficients, and random aid-awareness shocks—denoted $\omega$> I provide details in Appendix C.1.

The second set of moment conditions matches moments of the aid distribution in IPEDs to model-generated values conditional on enrolling in each college. In particular, I match average aid awards at each college $j = 1, \ldots, J$. I provide details in Appendix D.

Evaluating the likelihood and moments requires computing the value of each application portfolio, which in turn requires computing admissions chances $\ell_i^{B|A}(\theta, s)$ for portfolios $A$ and subsets $B$, for each $s$ and $\theta$. Doing so may be computationally expensive. In Appendix D, I provide a method for computing admissions chances for all $A$ and all $B \subseteq A$, based on the inclusion-exclusion principle, which requires integrating over shocks only for the "diagonal" terms $\ell_i^{A|A}(\theta, s)$.

## 5.4 Outcome moments

In the second stage, I estimate the parameters that affect post-enrollment outcomes via indirect inference. I first define the auxiliary regression specification:

$$\text{outcome}_{ij} = w_{ij}^{\text{outcome}} \tilde{\gamma}_{wj}^{\text{outcome}} + \tilde{\mu}_{ij}^{\text{outcome}}, \tag{7}$$

where $w_i$ is a vector of variables including all the variables except $q$ which enter the outcome equation (6), as well as indicators for top-decile status and for a scholarship at the other flagship (i.e., if $j$ is UT Austin, then the school's participation in the Century Scholars program which funds students to attend Texas A&M), and quadratic and cubic terms in class rank. I estimate the auxiliary models separately by college (UT Austin and Texas A&M) and outcome (GPA, persistence, and STEM major) that I consider, comprising six specifications in total.

Let $\widehat{\text{outcome}}_{ij}(\gamma_j^{\text{outcome}}, \theta)$ denote the expected value of $outcome_{ij}$ for person $i$, conditional on outcome parameters $\gamma_j^{\text{outcome}}$ first-stage estimates $\theta$, and the event $i$ enrolls in $j$.[20] I compute the value of $\gamma_j^{\text{outcome}}$ that minimizes the GMM criterion,

$$G^{outcome}(\gamma_j^{\text{outcome}}) = \sum_r \left[ \sum_i weight_{ij}(\theta) \left( \widehat{\text{outcome}}_{ij}(\gamma_j^{\text{outcome}}, \theta) - w_{ij}^{\text{outcome}} \tilde{\gamma}_{wj}^{\text{outcome}} \right) w_{ijr} \right]^2,$$

where $w_{ijr} = 1, \ldots, R$ denotes elements of $w_{ij}^{\text{outcome}}$, and weights satisfy

$$\text{weight}_{ij}(\theta) = (\text{pop. weight})_i Pr(enroll_{ij}; \theta).$$

[20] My approach is related to control function approaches to RDDs. For instance if $outcome_{ij}$ is a linear function of $q$, then computing $\widehat{\text{outcome}}_{ij}(\gamma_j^{\text{outcome}}, \theta)$ consists of plugging in the value of $E(q|enroll_{ij})$ implied by the model at parameters $\theta$. However, for my purposes it is important to fit well globally. Accordingly, I include a polynomial in the running variable and an indicator for threshold crossing, but the sample contains all students not in the bottom decile. I do not restrict to a narrow bandwidth around the discontinuity. See also Bertanha and Imbens [2020].

The coefficients $\gamma_j^{\text{outcome}} = [\gamma_{z,j}^{\text{outcome}}, \gamma_{q,j}^{\text{outcome}}]$ may also differ arbitrarily between the two flagships and across the three outcomes. To understand this criterion function, observe that, for each outcome, flagship university $j$ and element $\tilde{\gamma}_{wjr}^{\text{outcome}}$, the OLS estimates $\hat{\tilde{\gamma}}_{wjr}^{\text{outcome}}$ satisfy

$$\sum_{\{i:C_i=j\}} \left(\text{outcome}_{ij} - w_{ij}^{\text{outcome}} \hat{\tilde{\gamma}}_{w,j}\right) w_{ijr} = 0,$$

for each element $w_{ijr}$ of the vector $w_{ij}^{\text{outcome}}$, where the summation is over students matriculating in $j$. The moments impose that these OLS orthogonality conditions hold in the simulated data as well. Intuitively, the weights $weight_{ij}(\theta)$ make the sample of surveyed potential applicants comparable to the administrative dataset of enrollees.

# 6 Results

This section proceeds in three steps. I begin with descriptive analyses of Texas Top Ten and causal impacts of "threshold crossing". I revisit results from the literature to describe the total impacts of the policy, and to show that the model matches key patterns in the data. Next I present key parameters relevant for decomposing the percent plan, isolating informative effects, and extrapolating to other rules. Finally, I turn to the main results.

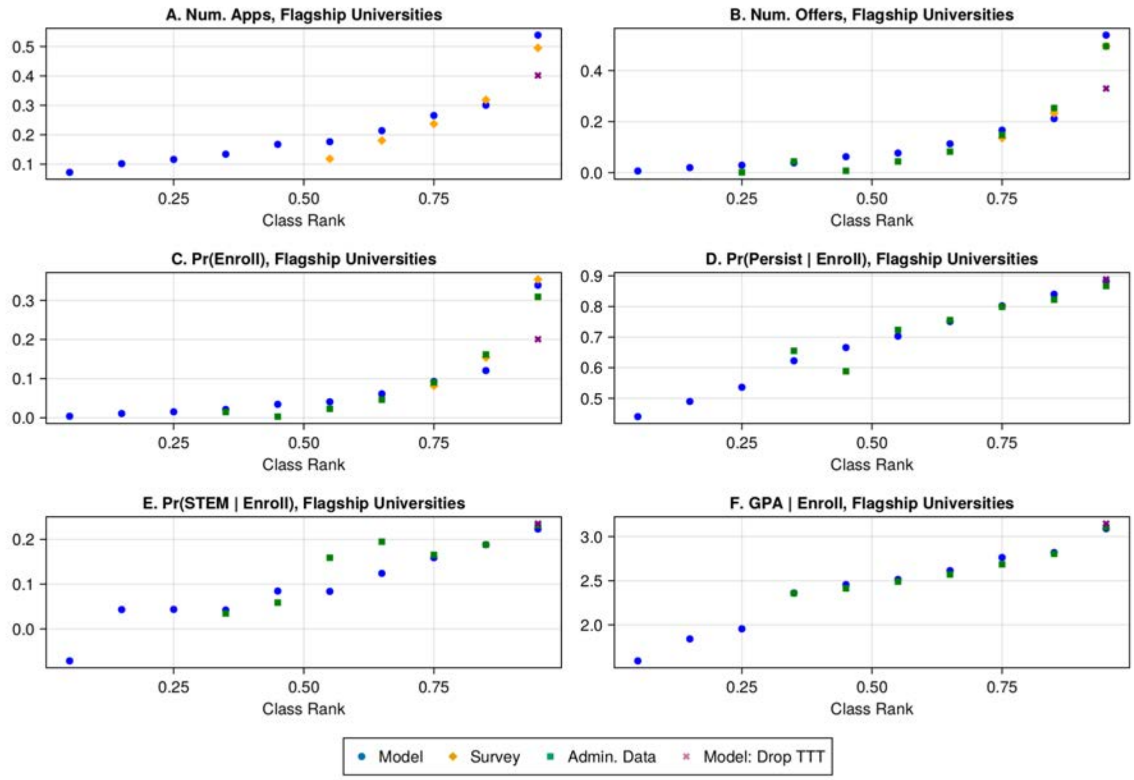
## 6.1 Descriptive Analyses, Total Impacts, and Model Fit

RDD estimates of the impact of crossing the 10% threshold, in the THEOP dataset used in this paper, show increases on flagship enrollment, especially for Latino students. Niu and Tienda [2010] obtain a point estimate of a nine percentage point increase in the probability of attending a flagship university for the average student near the tenth percentile of class rank. The results are stronger and significant for Latino students and students from majority-minority schools. Taken together, these results suggest that Texas Top Ten affected these students' application and matriculation decisions.

Of course, this threshold-crossing effect is not the total effect of the ten percent plan, which must include the impact of changes in admissions standards, a topic addressed in the next section. Using the model to simulate the impact of threshold crossing all else fixed, however, I find a 13.8 percentage point increase in the probability of flagship enrollment.[21] This estimate is within the 95% confidence bounds of Niu and Tienda's estimate.

[21] I simulate turning automatic admissions off, holding cutoffs fixed. I report the change in probability of flagship enrollment for top-decile students.

Figure 1: Model Fit, Flagship Universities




Note: This table shows expected number of applications to flagships, offers from flagships, and enrollment at Flagships, as well as outcome means conditional on enrollment, by decile of class rank. Model: as simulated at point estimates. Survey: population-weighted survey means. Model Drop TTT: turn off automatic admissions; hold cutoffs fixed.

I describe the data, decompose this threshold-crossing effect, and present model fit in Figure 1. I pool the flagship universities. Analogous figures for Texas A&M and UT Austin separately, as well figures for all public institutions and for all colleges, are provided in Online Appendix E, and other comparisons to the literature are given in Appendices E.3 and E.5.

Each panel of Figure 1 shows model-predicted and observed values of a given outcome by class-rank decile (in these figures only, ordered so that higher is better). Model outcomes are evaluated at the point estimates for each student. Model-predicted outcomes, and survey outcomes where available, are then averaged over students in the THEOP survey dataset, weighted by the survey's population weights. Administrative data, where available, are simple averages.

Panel A shows mean application probability by class-rank decile. This probability is rising in class rank for deciles below the top, then jumps approximately 17.7 percentage points in survey data (23.8 points according to model estimates) for top-decile applicants, with roughly 53.9% of top-decile students submitting applications.[22] Decomposing the threshold-crossing effect, the red "×" shows model-predicted top-decile application probabilities in the event there were no automatic admissions, holding admissions cutoffs fixed. Application probabilities would fall by roughly 13.7 points. This result is close to the high end of RD estimates from Daugherty et al. [2014] in an urban school district in Texas (IK optimal bandwidth specifications give point estimates of 13.6% and 11.0%) but is larger than the analogous impact on applications to UC campuses in California [Bleemer, 2024], where a similar "percent plan" guaranteed access to certain non-flagship UC campuses.[23]

Panel B of figure 1 shows the expected number of admission offers from flagships, and panel C shows the probability of enrolling in a flagship. The majority of applicants, and the vast majority of admitted students, come from the top four deciles. Cells with fewer than 10 students are not shown. The admissions model fits well. I slightly overestimate offers and matriculation probabilities of students with very low class rank relative to administrative data. This comparison to administrative data in panels B and C is out of sample, as administrative data on acceptances and enrollment decisions is not used in estimation.[24]

The final three panels show fit of post-enrollment outcomes conditional on enrollment. GPA and persistence fit well. Panel E considers the joint probability of majoring in STEM and persisting through 5 semesters. Here, the model fits the top three deciles reasonably well, but then underpredicts low-class-rank students' STEM participation. For transparency, this figure does not control for other covariates.

## 6.2 Selected Estimates

Table 2 presents admissions and outcome parameters. I estimate linear models for GPA, and for persistence and STEM indicators, which I use for counterfactuals. Admission chances are higher for students with high SAT scores (normalized to a 0-1 scale) and top class rank (the top decile has rank=.1). Admissions offices prefer students from high-poverty schools

[22] These numbers include an overestimate of the jump at Texas A&M relative to survey data. At UT Austin, the application probability jumps 14.1 percentage points in survey data (14.4 points according to model estimates) for top-decile applicants, with 29.9% of top-decile students submitting applications.

[23] See his Figure 3; Bleemer estimates a 6.5% increase in applications on a baseline of just over 60%.

[24] I observe administrative data only for students submitting an application. To construct the "administrative data" series in panels B and C, it is necessary to adjust for the proportion of students in each class-rank decile. While ex ante this share is 10% in each decile, I restrict the sample to students who have taken a college entrance exam. To construct the administrative series in panels B and C, I compute the proportion of applicants in each decile receiving an offer (enrolling), then multiply by the population-weighted measure of surveyed students in each decile.

but penalize students from schools with low SAT scores, as indicated by the coefficient on "SAT ratio". Weights on URM status are close to zero. Coefficients on scholarship availability and caliber are 0 and 1 by construction.

Turning to outcomes, SAT scores and class rank have the expected signs as well, but SAT ratio predicts majoring in STEM and is not significantly associated with lower persistence. "Caliber," as valued by admissions offices, predicts lower GPA and has smaller or zero effects on other outcomes. The availability of the Longhorn Opportunity Scholars program raises GPA and persistence for students from participating high schools who attend UT Austin, consistent with Andrews et al. [2020].

In Appendix table A4, I present probit specifications for binary outcomes. In Appendix E.7 I present estimates from an extension in which students' preferences, as well as their caliber, may correlate with potential outcomes.

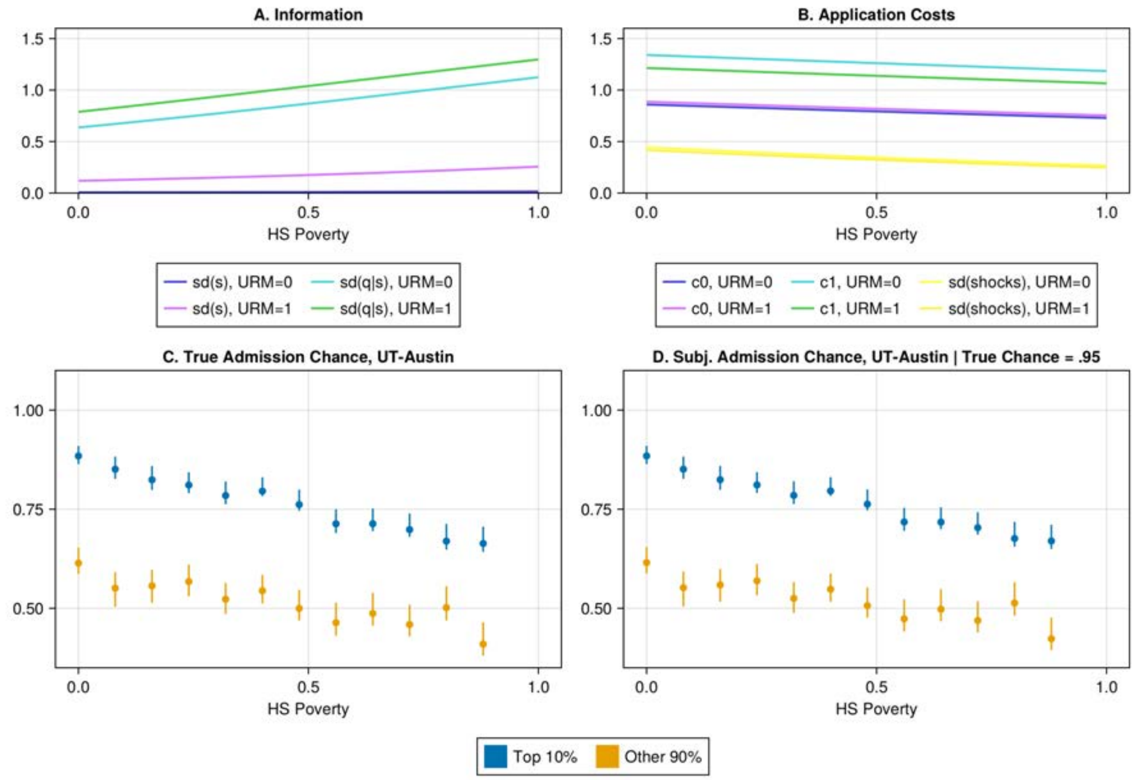
Table 2: Admissions and Outcome Parameters

| | $\gamma^{\text{admit}}$ | UTA GPA | LPM UTA Persist | LPM UTA STEM | TAMU GPA | LPM TAMU Persist | LPM TAMU STEM |
|---|---|---|---|---|---|---|---|
| Constant | 0.0 | 2.611 | 0.941 | -0.27 | 2.606 | 1.013 | -0.295 |
| | (0.0) | (0.077) | (0.064) | (0.043) | (0.105) | (0.052) | (0.069) |
| SAT | 5.624 | 4.737 | 0.188 | 0.229 | 3.724 | 0.129 | 0.399 |
| | (0.929) | (0.381) | (0.577) | (0.024) | (0.703) | (0.024) | (0.256) |
| Class Rank | -2.435 | -0.781 | -0.445 | -0.242 | -0.356 | -0.458 | -0.274 |
| | (0.193) | (0.299) | (0.133) | (0.167) | (0.295) | (0.205) | (0.194) |
| SAT Ratio | -1.2 | -2.692 | -0.104 | 0.266 | -2.122 | -0.09 | 0.384 |
| | (0.477) | (0.255) | (0.384) | (0.031) | (0.444) | (0.04) | (0.166) |
| Poverty | 0.278 | -0.235 | -0.241 | -0.163 | -0.278 | -0.232 | -0.32 |
| | (0.239) | (0.074) | (0.105) | (0.032) | (0.137) | (0.036) | (0.068) |
| URM | 0.011 | 0.055 | 0.008 | 0.055 | -0.006 | 0.001 | 0.082 |
| | (0.073) | (0.029) | (0.016) | (0.014) | (0.032) | (0.014) | (0.019) |
| Scholarship | 0.0 | 0.193 | 0.086 | 0.022 | 0.093 | 0.01 | 0.215 |
| | (0.0) | (0.043) | (0.026) | (0.029) | (0.12) | (0.078) | (0.089) |
| Caliber (q) | 1.0 | -0.988 | -0.002 | 0.152 | -1.018 | -0.1 | -0.118 |
| | (0.0) | (0.304) | (0.094) | (0.135) | (0.26) | (0.158) | (0.158) |

Note: This table shows admissions-index parameters $\gamma^{\text{admit}}$ (Equation (3)) and outcome indices (Equation (6)). By assumption the elements of $\gamma^{\text{admit}}$ multiplying Scholarship (LOS for UT-Austin, CS for Texas A&M) are zero and one, respectively.

Figure 2 describes information, application costs, and admission chances. Panels A and B of figure plot standard deviations of signals (panel A) and means and standard deviations

Figure 2: Costs, Information, and Admission Parameters

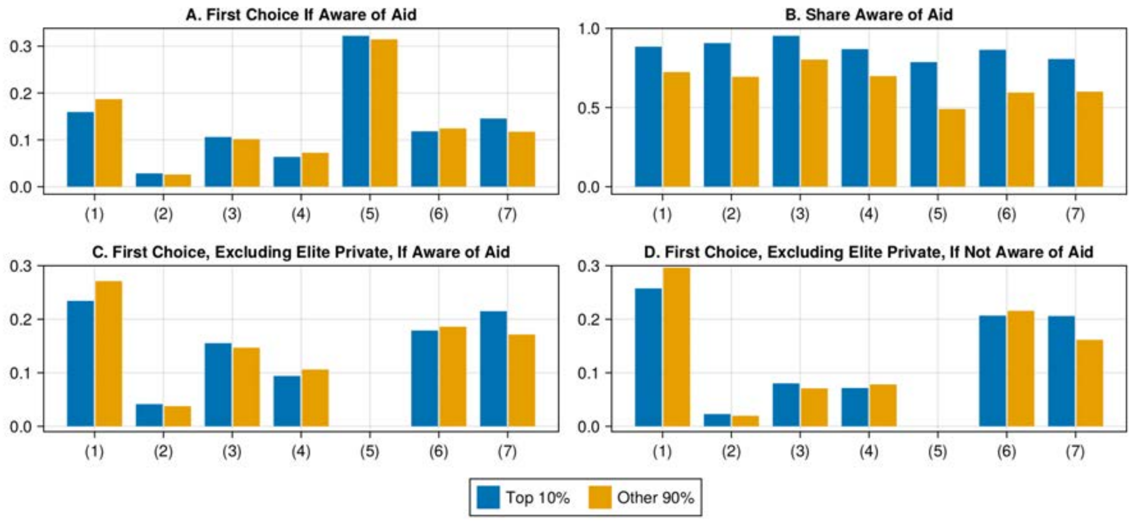



Note: Panels A and B show estimated standard deviations of signal and caliber, and application cost terms, by poverty rate and URM status. Panel C shows average true admissions chance at UT Austin by top-decile status and poverty rate (for top-decile students, model-predicted discretionary admissions chance is shown). Panel D: similar to panel C, but conditions on having the value of $q_i$ that makes admissions chance at UT Austin exactly equal to .95.

of costs (panel B) as functions of high school poverty rates. Panel A indicates that applicants face larger uncertainty in caliber $q$, conditional on their signal $s$, if they come from higher-poverty high schools. While the variance of the signal is also increasing at the point estimates, reflecting a greater extent of private information, the variance of $q$ conditional on $s$ is considerably larger. An interpretation is that students from poorer schools face greater residual uncertainty. Conditional on poverty rate, URM applicants' private signals $s$ are marginally informative, but they face greater residual uncertainty $q|s$. Hence these students may benefit more from an admissions guarantee.

Panel B of figure 2 shows that application costs, unlike uncertainty, do not seem to be increasing in poverty rate or URM status. While mean costs fall somewhat in poverty rate, so does the variance of shocks $\varepsilon^{\text{app}}$ that might explain application "mistakes". Underlying information and application-cost parameter estimates and standard errors are given in

Figure 3: Preferences and Awareness




Note: This figure shows average choice probabilities when $B_i$ contains one copy of each institution (Panel A), and when elite private schools are excluded (Panels C and D). Panel B shows the share of students aware of aid at each institution. Panels A and C: all students fully aware of aid. Panel D: aid awareness is set to zero for all students. Top-decile class rank and other students shown separately. Institutions: (1) Non-Flagship In-State Public; (2) Private Secular; (3) Religious; (4) Out-of-State Public; (5) Elite Private; (6) Texas A&M; (7) UT Austin.

Appendix Table A2.

The remainder of the figure shows that applicants from higher-poverty high schools face (or would face) lower admissions chances if not guaranteed admission, and that even if they are very likely to be admitted, they tend not to know it. Panel C shows mean admissions chances at UT Austin, by poverty-rate decile, separately for top-decile and non-top-decile applicants. For top-decile applicants, admissions chances via the discretionary channel (i.e. if Texas Top Ten were not available, but all else was held fixed) are shown. Points are means within a HS poverty decile; bars show 95% bootstrap confidence intervals. The average top-decile applicant from the most affluent high school would have a roughly 90% admissions chance; this figure falls to roughly 65% at the poorest schools. Chances for average non-top-decile students are lower, although not vanishing.

Panel D shows subjective admissions chances conditional on having a true admission chance of 95%, i.e. having caliber $q_i^*$ such that $Pr(z_i^{\text{admit}}\gamma + q_i^* - \pi_{\text{UTA}}) = .95$. For non-top-decile students $i$ from affluent schools, this value of $q_i^*$ is large, and students hold some private information. In contrast, at low-poverty schools students update very little. Students from the poorest high schools whose true admissions chances at UT Austin *without* Texas Top Ten would be 95% believe, on average, that their chances would be below 70%. This mean includes some students with even lower perceived chances. Because even students with very high admissions chances face subjective uncertainty, there may be room for an admissions guarantee to affect their application decisions.

Figure 3 summarizes preferences and awareness, plotting first-choice probabilities if $B_i$ consisted of one copy of each college (Panel A; Panels C and D if $j = 5$ is excluded) and awareness probabilities (Panel B). Estimates indicate that students would tend to prefer highly-selective private institutions (panel A), to which admissions are unlikely. Excluding these options from the choice set, the leading choices are UT Austin, non-flagship public institutions, and Texas A&M, each with roughly 20% share. (Preferences sum to less than 1, as some students prefer the outside option). Most top-decile students are aware of aid, while awareness rates are lower for non-top-decile students. If students were not aware of aid, the relative popularity of UT Austin would fall and the non-flagship share and outside-option share would grow.

Accounting for selection matters. Estimates based on final choices only, ignoring selection on applications and admissions, would have underestimated the share of students preferring elite private universities and non-flagship public institutions, and overestimated the share preferring flagships. I provide details in Appendix Figure A4.

## 6.3 Main Results

Having shown that roughly 40% of students prefer a flagship to other available options, and that students from high-poverty high schools with high objective chances of admission would face admissions uncertainty, we turn to the main results. I first decompose the percent plan as implemented in 2002. Second, I consider a hypothetical ideal financial-aid awareness intervention in which all students are made fully aware of aid at all institutions. Third, to isolate the role of transparency, I consider hypothetical "aligned" percent plans in which in-state public institutions commit to admit students on the basis of observables, but use the same weights on those observables as in their preferences.

In each case, I take as the baseline the (counterfactual) scenario in which there is no percent plan, holding college capacities and the population of high-school students in 2002 fixed. I decompose the effects of percent plans in the following order. First, I present mechanical effects, holding the baseline applicant pool fixed but admitting all "treated" applicants. Second, I add information effects, allowing the applications of treated students—those receiving an admissions guarantee—to respond. Third, I solve for equilibrium, allowing cutoffs and non-top-decile students' application decisions to adjust.

For the financial-aid awareness intervention, timing is as follows. I first make all students fully aware of financial aid, allowing applications to adjust but holding cutoffs fixed. This step can be interpreted as the informative effects of financial-aid information. Second, I solve for equilibrium, letting cutoffs adjust. Third, I introduce the percent plan in this scenario, and present mechanical, information, and equilibrium effects in order. This sequence makes it clear whether the effects of the percent plan are larger in the presence of better information about aid. However, other decompositions would be possible as well.

The "aligned" plans automatically admit the students with the highest values of $z^{\text{admit}}\gamma^{\text{admit}}$. Recall that student $i$ is admitted if $\pi_{ij} = z_i^{\text{admit}}\gamma^{\text{admit}} + q_i + \mu_{ij}^{\text{admit}}$ is greater than a cutoff. In effect, these plans commit to admit students with high observables in a way that is aligned with discretionary admissions, simply ignoring the components $(q_i, \mu_{ij}^{\text{admit}})$ that are not observed by students. I consider a sequence of plans of varying sizes, admitting the top 5%, 10%, 15%, and 20% in this way.

Table 3 presents the main results, decomposing changes in flagship-university enrollment of students receiving an admissions guarantee into mechanical and informative effects, and showing impacts on post-enrollment outcomes. Bootstrap standard errors are given in parentheses beneath point estimates.

**Isolating informative effects:** Row 1 of Table 3 shows that the top ten percent plan, as implemented in 2002, raised the share of top-decile students enrolling in flagships by about 9 percentage points, mainly via information, while improving academic outcomes at flagships. Columns 1 through 4 show probabilities of enrolling in a flagship university among the group targeted by the percent plan. The first column, "Base," indicates the share of such students who would enroll in a flagship university in equilibrium if there were no percent plan. I estimate that 24.7% of top-decile students would have done so. The second column, "Mech.", indicates that an additional 2.8 percent of the population of top-decile students would have enrolled in flagships via mechanical effects. Columns three and four respectively show that the informative effect would contribute a further 6.3 points, which would represent 69% of the total impact of the plan on top-decile enrollment.

**Academic outcomes:** Columns 5 through 7 of Table 3 show average differences in post-enrollment outcomes between the students who were "pulled in" and the non-top-decile students that they displaced from flagships in equilibrium. I report cumulative GPA, an indicator for persistence through two years, an indicator for persistence and majoring in

Table 3: Main Results

| | Pr(Enroll in Flagship), Treated Group | | | | Average Effect on Outcomes | | | |
|---|---|---|---|---|---|---|---|---|
| | Base | + Mech. | + Info | % Info | ΔGPA | ΔPersist | ΔSTEM | ΔPayoff |
| *A. Texas Top Ten* | | | | | | | | |
| (1): TTP | 24.72 | 2.81 | 6.33 | 69.23 | 0.71 | 15.66 | 10.46 | -0.56 |
| | (1.05) | (0.19) | (0.4) | (1.98) | (0.06) | (2.8) | (3.29) | (0.1) |
| (2) +Aware | 24.5 | 3.01 | 6.82 | 69.36 | 0.72 | 15.45 | 9.74 | -0.59 |
| | (1.03) | (0.25) | (0.57) | (1.97) | (0.07) | (2.94) | (3.51) | (0.08) |
| *B. Automatic Admission By* $z^{admit}\gamma$ | | | | | | | | |
| (3): Top 5% | 42.84 | 1.88 | 3.21 | 62.98 | 0.2 | 3.02 | 4.56 | 0.02 |
| | (2.41) | (0.25) | (0.41) | (2.38) | (0.11) | (7.3) | (12.19) | (1.54) |
| (4): Top 10% | 39.35 | 2.39 | 4.4 | 64.79 | 0.38 | 6.16 | 8.53 | 0.0 |
| | (1.97) | (0.25) | (0.42) | (2.3) | (0.06) | (5.05) | (7.43) | (0.88) |
| (5): Top 15% | 36.81 | 2.66 | 5.07 | 65.6 | 0.53 | 8.69 | 11.57 | -0.05 |
| | (1.73) | (0.24) | (0.42) | (2.23) | (0.07) | (3.96) | (5.16) | (0.56) |
| (6): Top 20% | 34.67 | 2.88 | 5.7 | 66.39 | 0.68 | 11.42 | 14.71 | -0.13 |
| | (1.49) | (0.24) | (0.42) | (2.14) | (0.07) | (2.74) | (3.21) | (0.3) |

Note: This table shows a decomposition of changes in flagship enrollment for treated households, and average impacts on academic outcomes and programs' payoffs. "Base": Share of "treated" group enrolling at baseline. "Mech": mechanical effect. "Info": information effect. "% Info": Share of total increase in treated-group enrollment due to information. Outcomes: total change in outcome (GPA, 1(Persist), 1(Persist and major in STEM), college payoff) / measure of top-decile students induced to enroll in flagships by mechanical and informative effects; this is equivalent to the average difference in outcomes between these students and the students they displaced.

STEM. Top-decile students induced to enroll by the mechanical and informative effects earn higher GPAs, are 16 percentage points more likely to persist, and 10 percentage points more likely to persist and choose a STEM major, than are the students that they displaced.[25]

Past studies have found that less-prepared students may select out of STEM majors [Arcidiacono et al., 2016], and that state merit aid programs' GPA eligibility programs may also cause students to select out of STEM [Sjoquist and Winters, 2015]. I do not find that this is the case for the percent plan. In practice, the University of Texas at Austin had a separate, selective admissions process for engineering majors. I abstract from this in my estimates. However, if lower demand would lead to lower admissions standards for those majors in a world without the percent plan, then my results might overstate the impacts of

[25]This calculation considers students enrolling in flagships. In Appendix E.4 I present evidence of overall gains in persistence across all institutions.

the percent plan on the number of STEM majors.

The final column of Table 3 considers the average payoff $E(\pi_{ij})$ that the program receives. Flagship program payoffs $\pi_{ij}$ would fall when top-decile students are automatically admitted, suggesting the presence of a trade-off.[26]

**Financial-Aid Information:** The second row of Table 3 considers an ideal financial-aid awareness intervention in which all students are made fully aware of aid at all institutions. In the absence of the percent plan, in equilibrium, providing this information would have no effect on the share of top-decile students who attend flagships.[27] However, the mechanical and informative effects of the percent plan would be larger, indicating that information about pricing and about admissions are complementary, with the introduction of the percent plan leading to a 0.47 percentage-point greater increase in the share of top-decile students than if aid information had not been provided.[28] Per-student impacts on outcomes are similar.

**Aligned "percent plans":** Next we turn to "aligned" percent plans which use observables exactly as in discretionary admissions, but simply commit to ignore unobservables for students with sufficiently high values of the observable index. Just as in the original percent plan, moderate-sized such plans would lead to enrollment changes roughly two-thirds due to information, and to improvements in academic outcomes. However, unlike the percent plan as implemented, they would not lead to significant decreases in programs' payoffs $\pi_{ij}$. Rows (3)-(6) of Table 3 show results as the size of the "aligned" percent plan increases from 5% to 20% of the population. In each case, the "treated group" is the set of students who will receive an admissions guarantee. For instance, row (6) shows that 35% of top-20% students would attend a flagship at baseline. This number would increase by $2.88 + 5.70 = 8.58$ percentage points under the percent plan, with 70% of this change due to information rather than mechanical effects. While the estimated impact on program's payoffs $E(\pi_{ij})$ is negative, one cannot reject zero change. Indeed, point estimates of impacts on program payoffs are positive (but noisy) for small aligned percent plans. This result implies that the impacts of the percent plan are not driven by the use of class rank alone, and that a small or moderate "percent plan" may improve academic outcomes at no

[26] The coefficient $\gamma_{\text{classrank}}$ is approximately -.24 per decile; hence program payoffs fall by the equivalent of admitting a top-40% rather than top-decile student all else equal.

[27] The point estimate, 24.50%, is not statistically distinguishable from 24.72% in the baseline scenario.

[28] A bootstrap 95% confidence interval for this difference is $(0.013, 2.192)$.

cost to programs.

**Demographics and Diversity:** Figure 4 shows impacts on top-decile enrollment and on measures of diversity under these counterfactuals, decomposing the effects into mechanical, information and equilibrium channels.

Panel A visually displays the changes in top-decile enrollment shares shown in Table 3. At baseline, a quarter of top-decile students enrolled in flagships. The direct effect of aid information would have been a 1.55% increase in this share. However, because aid information would have encouraged additional students to submit applications, universities' cutoffs would have to rise in equilibrium. This equilibrium response would lead to a 2.46% drop in the share of top-decile students, entirely reversing these gains, so that aid information alone, unaccompanied by admissions transparency, would provide zero additional access to flagships.

Panels B and C consider changes in the flagship enrollment share of URM (Black and Hispanic) students, and in students from high schools in the top quartile of poverty rates, as proxied by the share receiving free/reduced-price lunch.

As with top-decile students, financial-aid awareness alone would not raise these shares. At baseline, 8.4% of exam-taking URM students and 6.56% of exam-taking students from high-poverty high schools attend flagship universities. Providing information about aid without a percent plan, cutoffs fixed, would lead to small or negligible changes, 0.7% and roughly zero respectively, in the share attending flagships. Equilibrium responses would then more than undo these gains.

Turning to the impacts of percent plans, however, panel C shows that students from high-poverty high schools would experience large gains—12% and 29% respectively, from the mechanical and informative effects of the percent plan. About a third of this total effect would be clawed back in equilibrium as cutoffs for non-top-decile students rise.

Under the alternative percent plan that does not reserve seats by high school, in contrast, mechanical and direct effects would be smaller, and would be entirely undone in equilibrium. Thus, while the "aligned" percent plan leads to small losses in universities' payoffs, it does not raise enrollment shares from high-poverty schools.

Patterns for URM students (panel B) are similar, except that a larger share of the mechanical and direct effects would be reversed by equilibrium changes in admissions standards, suggesting that Texas Top Ten is an inefficient way to provide affirmative action.[29]

[29] Kain et al. [2005] show that most of the increases in URM enrollment at selective public universities post-1998 was

Panels D and E show impacts on affluent (bottom quartile of poverty rate) schools, and on a Theil index of high-school concentration. This index takes values between 0 and 1. If a fraction $p$ of all students attend flagships, then this index takes value 0 if each high school sends a proportion $p$ if its class, and equals 1 if a fraction $p$ of high schools send all students to flagships while the remaining high schools send none. Patterns for affluent high schools contrast with those of high-poverty high schools: direct and mechanical effects would be relatively small, and more than reversed by changes in equilibrium. The Theil index falls with each step: mechanical and informative effects bring in students from underrepresented high schools, while the displaced students in equilibrium come from schools that were overrepresented.

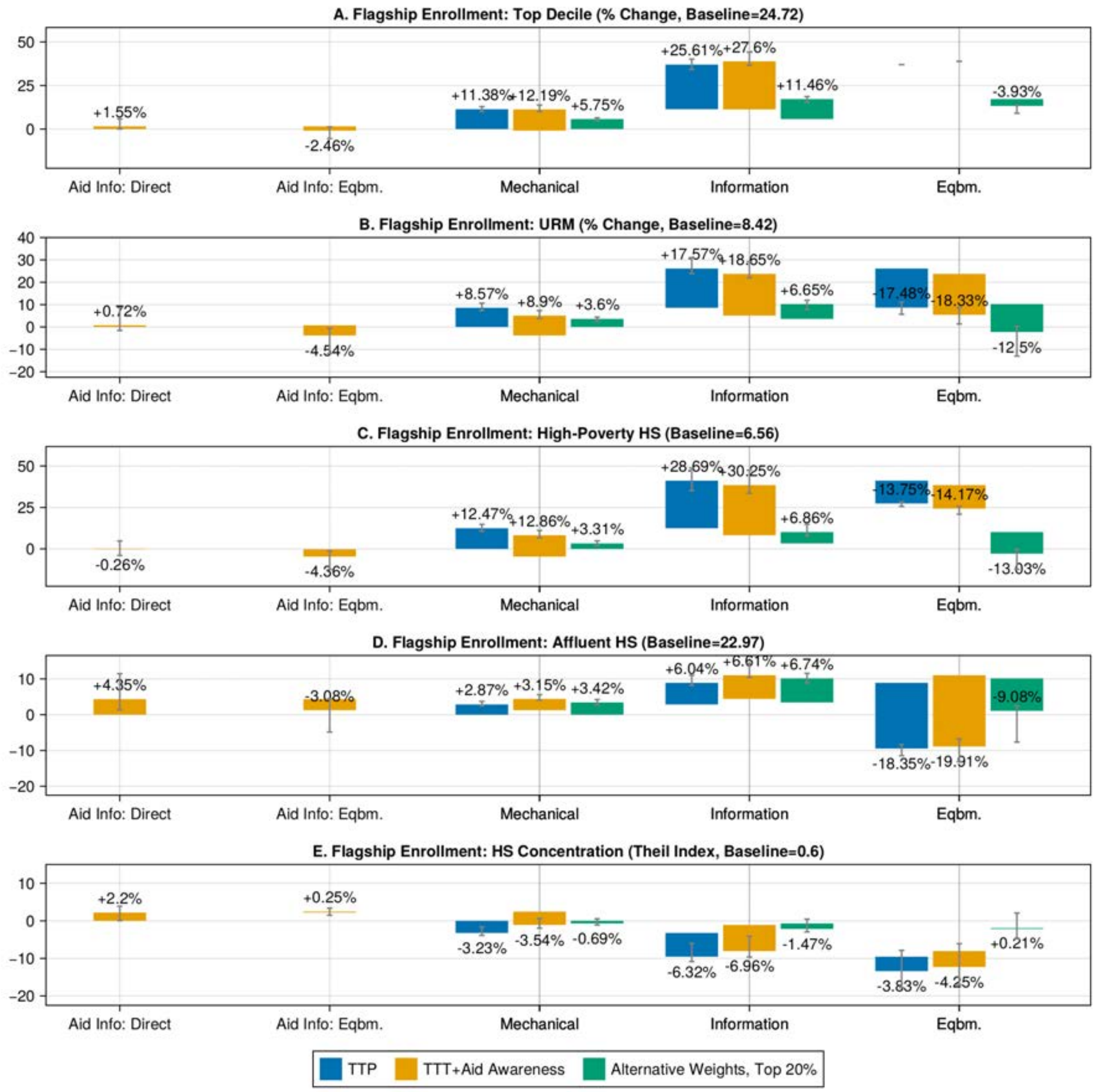
**Extensions:** In Appendix E.5 I consider the targeted scholarship programs, showing that estimated impacts are similar to results from the literature, and that, much like aid awareness, the LOS program and the percent plan have complementary effects.[30] In Appendix E.6 I consider robustness to a form of "class balancing" in which the weights on "personal" and "academic" characteristics change as the percent plan is introduced. If universities were to place more weight on academic characteristics in the absence of the percent plan, its effects on academic outcomes would shrink but not disappear, and impacts on diversity would be larger. In Appendix E.7 I extend the model of outcomes to allow for selection on students' private preferences as well as their caliber. Counterfactual results are essentially unchanged.

**Summary:** Taken together, the results indicate that there are many students who can be induced to enroll in flagship universities by resolving uncertainty about their admissions chances before they apply. The percent plan as implemented raised the share of top-decile students attending flagship universities by 9 percentage points, two thirds of which was due to information. these students earned higher grades, and were more likely to to persist and to major in STEM, than the students they displaced. Moreover, the percent plan increased economic diversity and attracted students from a broader range of high schools, primarily via information rather than mechanically, but at the cost of a lower payoff according to admissions offices' preferences.

driven by LOS schools. I provide more details on the impacts of scholarship programs, consistent with this finding, in Appendix E.5.

[30] This finding is consistent with evidence from the literature that multiple dimensions of programs intended to promote degree completion have complementary effects [Evans et al., 2019, Clotfelter et al., 2018].

Figure 4: Demographics and Diversity




Note: This figure decomposes impacts on enrollment and diversity measures, in percentage changes relative to baseline, due to mechanical, informative, and equilibrium effects. Blue bars ("TTP") show percent plan as implemented. Orange ("TTT+Aid Awareness") shows combined aid-information and percent plan counterfactual. Green ("Alternative Weights, Top 20%") shows aligned "top 20%" plan. "Aid Info: Direct": effect of aid intervention, cutoffs fixed (aid-awareness intervention only). "Aid Info: Eqbm": effect of subsequent change in cutoffs (aid-awareness intervention only.) "Mechanical," "Information," "Eqbm.": mechanical, informative, and equilibrium effects of percent plan, respectively.

An admissions guarantee that uses observables in the same way that they are used for discretionary admissions, however, but simply ignores student-unobserved components of applications, would also draw in students and achieve gains in academic outcomes, primarily via the information channel. Unlike the percent plan as implemented, this alternative plan would not draw more students from high-poverty schools, but it would also have small to zero impacts on colleges' payoffs. These results suggests that colleges may face a tradeoff between "caliber" and attracting students from a wider range of high schools. However, informative effects of transparency are sufficiently large that there is no such tradeoff between payoffs and academic outcomes: the payoff cost to Texas' flagship institutions to providing an admissions guarantee to a small to moderate number of top students is zero.

**Additional Analyses:** In Appendix E.3 I decompose academic outcomes and discuss selection. In Appendix E.4 I investigate outcomes at non-flagship universities and estimate student-level impacts. I find that students "pulled in" to flagships by informative effects are similar to, or slightly better-prepared than, "always-enrolling" students, and outperform the students they displace. Moreover, consistent with Black et al. [2023], top-decile students' fallback options are more likely to be no college or two-year programs, while non-top-decile students displaced from flagships are likely to attend non-flagship public institutions. In turn, these students displace other students. Nonetheless, accounting for these indirect effects, I find evidence that the percent plan led to overall gains in persistence.

# 7 Conclusions

How can Texas Top Ten improve outcomes if it acts as a constraint on colleges? One can imagine two reasons. First, admissions offices trade off personal and academic characteristics. They do not maximize the expected GPA of entering cohorts. Texas Top Ten forces colleges to admit students with high class rank, which predicts academic success. Second, the Texas Top Ten law is a credible promise of admission, which induced academically strong applicants with poor information to submit applications to flagship universities.

In order to define and then quantify the role of this promise, a model of imperfect information, and of selection on that information, is necessary. Using such a model, I show that this second effect drives the results. Roughly two thirds of the students drawn into flagship universities responded to the information provided by the percent plan. When they enrolled, they persisted and earned high grades, but in the plan's absence they would not have submitted applications. These findings are robust to alternative rules for calculating the "top X%" in which the same index is used for automatic and discretionary admissions. This result indicates that impacts on post-enrollment outcomes are not driven by misalignment between those outcomes and the observable indices used in discretionary admissions. Moreover, effects of the percent plan are larger if complemented by information about aid availability. These results suggest that it may be valuable for universities or university systems to commit to admissions policies which induce strong candidates to apply, and that transparency about admissions and pricing may be important elements of such policies.

Because capacity is limited at the flagship universities, one must be careful to account for the impacts of changes in universities' admissions standards. We may misstate the effects of large policy changes such as changes in the percent plan if we do not model or otherwise account for spillovers. For academic outcomes and high-school diversity, equilibrium effects reinforce the mechanical and informative effects of the policy. Pulled-in students come from underrepresented high schools, while the students displaced from flagships come from schools with a greater-than-average share of students attending flagships. In contrast, equilibrium responses undo the direct and informative effects of the percent plan on racial diversity. A caveat is that I hold colleges' capacities fixed in equilibrium, consistent with UT Austin having stopped expanding in the early 2000s. It would be an interesting extension to consider universities' choice to grow.

In evaluating the informative impacts of pre-commitment to admit certain students, I take the colleges' preferences and students' "caliber" as given. In practice, "caliber" is constructed from the inputs that these colleges choose to collect. I do not take a stand on what colleges ought to do, or consider what information they should collect, or how to optimally combine multiple signals to make admissions decisions. Collecting additional materials may impose costs on students, which may then interact with uncertainty to affect application decisions. I take applicants' costs, and the form of admissions uncertainty, as given as well. Under a substantial change in admissions policy, in equilibrium, students' might gather information differently. I leave these questions for future research.

Finally, this paper begins with SAT-taking and/or ACT-taking high school seniors, and ends after the first years of college. I take as exogenous the the decision to take a standardized test, the environment within high schools, and the high schools attended by students. While the Top Ten Percent plan ruled out some types manipulations by schools—in particular, at most 10% of students can be in the top decile— the use of observables such as scores, grades, and high-school characteristics can in general affect students' choice of schools and time allocation, and schools' incentives. I find that Texas Top Ten increases diversity in college and improves early academic outcomes at flagship universities, but a full assessment of the policy would consider the impact of information on students' human capital investment and on long-run outcomes.

# References


Atila Abdulkadiroglu, Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters. Do parents value school effectiveness? *American Economic Review*, 2020.

Rodney J. Andrews, Vimal Ranchhod, and Vijay Sathy. Estimating the responsiveness of college applications to the likelihood of acceptance and financial assistance: Evidence from texas. *Economics of Education Review*, (29):104–115, 2010.

Rodney J. Andrews, Jing Li, and Michael F Lovenheim. Heterogeneous paths through college: Detailed patterns and relationships with graduation and earnings. Working Paper, November 2011.

Rodney J. Andrews, Jing Li, and Michael F Lovenheim. Quantile treatment effects of college quality on earnings: Evidence from administrative data in texas. NBER Working Paper 18068, May 2012.

Rodney J Andrews, Scott A Imberman, and Michael F Lovenheim. Recruiting and supporting low-income, high-achieving students at flagship universities. *Economics of Education Review*, page 101923, 2020.

Joshua Angrist, Daniel Lang, and Philip Oreopoulos. Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, 1(1):136–63, 2009.

Kate Antonovics and Ben Backes. The effect of banning affirmative action on college admissions rules and student quality. April 2013a.

Kate Antonovics and Ben Backes. Were minority students discouraged from applying to university of california campuses after the affirmative action ban? *Education Finance and Policy*, 8(2):208–250, Spring 2013b.

Peter Arcidiacono. Affirmative action in higher education: How do admission and financial aid rules affect future earnings? *Econometrica*, 73(5):1477–1524, September 2005.

Peter Arcidiacono and Michael Lovenheim. Affirmative action and the quality-fit tradeoff. *Journal of Economic Literature*, 54(1), March 2016.

Peter Arcidiacono, Esteban Aucejo, Patrick Coate, and V Joseph Hotz. Affirmative action and university fit: Evidence from proposition 209. *IZA Journal of Labor Economics*, 2014. September.

Peter Arcidiacono, Esteban M Aucejo, and V Joseph Hotz. University differences in the graduation of minorities in stem fields: Evidence from california. *American Economic Review*, 106(3):525–562, 2016.

Christopher Avery and Caroline Hoxby. Do and should financial aid packages affect students' college choices? NBER working paper 9482, February 2003.

Christopher Avery and Jonathan Levin. Early admissions at selective colleges. *American Economic Review*, 100:1–34, December 2010.

Ben Backes. Do affirmative action bans lower minority college enrollment and attainment? evidence from statewide bans. *Journal of Human Resources*, 47(2):435–455, 2012.

Steven Berry, Amit Gandhi, and Philip Haile. Connected substitutes and invertibility of demand. *Econometrica*, 81(5):2087–2111, September 2013.

Marinho Bertanha and Guido W Imbens. External validity in fuzzy regression discontinuity designs. *Journal of Business & Economic Statistics*, 38(3):593–612, 2020.

Eric P. Bettinger, Bridget Terry Long, Philip Oreopolous, and Lisa Sanbonmatsu. The Role of Simplification and Information in College Decisions: Results From The H&R Block FAFSA Experiment. *Quarterly Journal of Economics*, 127(3):1–38, 2012.

Sandra E Black, Jeffrey T Denning, and Jesse Rothstein. Winners and losers? the effect of gaining and losing access to selective colleges on education and labor market outcomes. *American Economic Journal: Applied Economics*, 15(1):26–67, 2023.

Zachary Bleemer. Top Percent Policies and the Return to Postsecondary Selectivity. Technical report, January 2024. Working Paper.

H.D. Block and Jacob Marschak. Random orderings and stochastic theories of responses. In Ingram Olkin, editor, *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford University Press, 1960.

W. Bowen, M. Chingos, and M. McPherson. *Crossing the Finish Line: Completing College at America's Public Universities*. Princeton University Press, 2009.

David Card and Alan Krueger. Would the elimination of affirmative action affect highly qualified minority applicants? evidence from california and texas. *Industrial and Labor Relations Review*, 58(3), 2005.

Hector Chade and Lones Smith. Simultaneous search. *Econometrica*, 74(5):1293–1307, September 2006.

Hector Chade, Gregory Lewis, and Lones Smith. Student portfolios and the college admissions problem. *Review of Economic Studies*, February 2014.

Charles T Clotfelter, Steven W Hemelt, and Helen F Ladd. Multifaceted aid for low-income students and college outcomes: Evidence from north carolina. *Economic Inquiry*, 56(1): 278–303, 2018.

Kalena Cortes. Do bans on affirmative action hurt minority students? evidence from the texas top 10% plan. *Economics of Education Review*, 29:1110–1124, 2010.

Julie Berry Cullen, Mark Long, and Randall Reback. Jockeying for position: Strategic high-school choice under texas' top ten percent plan. *Journal of Public Economics*, 97, 2013.

Lindsay Daugherty, Paco Martorell, and Isaac McFarlin. Percent plans, automatic admissions, and college outcomes. *IZA Journal of Labor Economics*, 3:1–29, 2014.

Susan Dynarski, CJ Libassi, Katherine Michelmore, and Stephanie Owen. Closing the gap: The effect of a targeted, tuition-free promise on college choices of high-achieving, low-income students. Technical report, National Bureau of Economic Research, 2018.

Susan M. Dynarski and Judith E. Scott-Clayton. The cost of complexity in federal student aid: Lessons from optimal tax theory and behavioral economics. NBER Working Paper 12227, May 2006.

Glenn Ellison and Parag A Pathak. The efficiency of race-neutral alternatives to race-based affirmative action: Evidence from chicago's exam schools. Technical report, National Bureau of Economic Research, 2016.

Dennis Epple, Richard Romano, and Holger Sieg. Admission, tuition and financial aid policies in the market for higher education. *Econometrica*, 74(4):885–928, July 2006.

Dennis Epple, Richard Romano, and Holger Sieg. Diversity and affirmative action in higher education. *Journal of Public Economic Theory*, 4(10), 2008.

Dennis Epple, Richard Romano, Sinan Sarpca, and Holger Sieg. The u.s. market for higher education: A general equilibrium analysis of state and private colleges and public funding policies. NBER Working Paper 19298, 2014.

William N Evans, Melissa S Kearney, Brendan Perry, and James X Sullivan. A cautionary tale about emergency financial assistance without services: Evidence from a randomized controlled trial evaluation at a community college. In *AEA papers and proceedings*, volume 109, pages 218–222. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, 2019.

Jean-Claude Falmagne. A representation theorem for finite random scale systems. *Journal of Mathematical Psychology*, 18(1):52–72, 1978.

Ian Fillmore. Price discrimination and public policy in the u.s. college market. Working Paper, 2016.

Jason M. Fletcher and Adalbert Mayer. Tracing the effects of guaranteed admission through the college process: Evidence from a policy discontinuity in the texas 10% plan. *Contemporary Economic Policy*, 2013.

Stella M Flores and Catherine L Horn. Texas top ten percent plan: How it works, what are its limits, and recommendations to consider. 2016.

Roland G Fryer Jr, Glenn C Loury, and Tolga Yuret. An economic analysis of color-blind affirmative action. *The Journal of Law, Economics, & Organization*, 24(2):319–355, 2008.

Chao Fu. Equilibrium tuition, applications, admissions and enrollment in the college market. *Journal of Political Economy*, 122(2), 2014.

Michelle Sovinsky Goeree. Limited information and advertising in the us personal computer industry. *Econometrica*, 2008.

Philip Haile, Ali Hortacsu, and Grigory Kosenok. On the empirical content of quantal response equilibrium. *American Economic Review*, 98(1):180–200, 2008.

Brent R. Hickman and Aaron Bodoh-Creed. Pre-college human capital investment and affirmative action: A structural policy analysis of us college admissions. August 2018.

Jessica S. Howell. Assessing the impact of eliminating affirmative action in higher education. *Journal of Labor Economics*, 28(1):113–166, 2010.

Caroline Hoxby and Sarah Turner. Expanding college opportunities for high-achieving, low income students. SIEPR discussion paper 12-014, February 2014.

Caroline M. Hoxby and Christopher Avery. The missing "one-offs": The hidden supply of high-achieving, low income students. NBER Working Paper 15856, December 2012.

John F Kain, Daniel M O'Brien, and Paul A Jargowsky. Hopwood and the top 10 percent law: How they have affected the college enrollment decisions of texas high school graduates. Technical report, A Report to the Andrew Mellon Fpundation, 2005.

Jerome Karabel. *The chosen: The hidden history of admission and exclusion at Harvard, Yale, and Princeton*. Houghton Mifflin Harcourt, 2005.

Jack Mountjoy and Brent Hickman. The returns to college (s): Estimating value-added and match effects in higher education. Technical Report 2020-08, University of Chicago, Becker Friedman Institute for Economics Working Paper, 2020.

Sunny Xinchun Niu and Marta Tienda. The impact of the texas top 10 percent law on college enrollment: A regression discontinuity approach. *Journal of Policy Analysis and Management*, 29(1):84–110, 2010.

Sunny Xinchun Niu, Marta Tienda, and Kalena Cortes. College selectivity and the texas top 10% law. *Economics of Education Review*, 25:259–272, 2006.

Lindsay C Page, Stacy S Kehoe, Benjamin L Castleman, and Gumilang Aryo Sahadewo. More than dollars for scholars the impact of the dell scholars program on college access, persistence, and degree attainment. *Journal of Human Resources*, 54(3):683–725, 2019.

Amanda Pallais. Small differences that matter: Mistakes in applying to college. working paper, September 2009.

David L Sjoquist and John V Winters. The effect of Georgia's HOPE scholarship on college major: a focus on STEM. *IZA Journal of Labor Economics*, 4:1–29, 2015.

Juliette Thibaud. Aiming higher: Spillover effects of affirmative action in higher education. Working Paper, December 2019.

Christopher Walters. The demand for effective charter schools. *Journal of Political Economy*, 126(6), 2018.

# A Data Appendix [Online]

This section provides additional details on the construction of the dataset.

**ACT scores:** Some students take the ACT instead of the SAT; if a student took both exams, I keep that student's SAT scores. Otherwise, I convert ACT composite scores to SAT combined scores using the College board's 1999 ACT-SAT concordance tables.[31]

**Average SAT scores:** Raw SAT scores range from 0 (all questions answered incorrectly) to 1600. After converting ACT scores to SAT scores, I rescale SAT scores by dividing by 1600. The standard deviation of the rescaled measure is 0.119.

To construct the ratio of student *i*'s SAT scores to their high school's average score, I use school-level average scores reported by THEOP. I found and corrected one outlier in the school-level score data. Between 10 and 20 students at a single high school have a reported school-average score of 29 points out of 1600. Because students can obtain a score of 200 by leaving the exam blank, this value seems implausible. The next-lowest school average score is greater than 690 points out of 1600. I replace the (presumably erroneous) 29-point school average with the mean score of surveyed students at this high school, which is 933 points out of 1600. At other schools, I have confirmed that the means of surveyed students' scores are close to the reported school mean scores.[32]

**Income:** The data do not provide *i*'s household income; I do, however, observe the education and occupational category of each of *i*'s parents. I draw incomes from the March CPS. To do so, for each household I construct the household head's education and occupational category. For each simulation draw, I then draw from incomes in the 2002 and 2003 March CPS samples of Texas residents with the same occupational category and education. Because the THEOP dataset uses a different encoding of occupations, I convert the CPS sample occupation codes to 1990 CPS codes as described in the data appendix. If either parents' occupation category or education is missing in the THEOP data, but one variable is present, I draw from the March CPS conditional on the variable that is present.

[31] http://research.collegeboard.org/sites/default/files/publications/2012/7/researchnote-1999-7-concordance-sat-act-students.pdf

[32] The numbers may differ because of survey nonparticipation and because of the use of converted ACT scores for students who took only the ACT. The mean and variance of the difference between the within-school mean student score and the mean school score are 17 and 52 points out of 1600, respectively.

Table A1: Administrative dataset

| | Application Data | | College Transcript Data | |
|---|---|---|---|---|
| Institution | N | Years | N | Years |
| Texas A&M | 163,027 | 1992-2002 | 637,018 | 1992-2007 |
| Texas A&M Kingsville* | 18,872 | 1992-2002 | 91,106 | 1992-2004 |
| UT Arlington | 29,844 | 1994-2002 | 51,315 | 1994-2002 |
| UT Austin | 210,006 | 1991-2003 | 659,102 | 1991-2004 |
| UT Pan American** | 44,747 | 1995-2002 | 115,812 | 1995-2005 |
| UT San Antonio# | 61,221 | 1990-2004 | 160,604 | 1990-2004 |
| Texas Tech | 81,153 | 1995-2003 | 211,771 | 1995-2004 |
| Rice | 36,190 | 2000-2004 | 18,149 | 2000-2005 |
| SMU | 45,549 | 1998-2005 | 60,607 | 1998-2005 |

* Applicant data for enrollees only: 1992-1994
** Limited variables provided
# Applicant data for enrollees only, 1990-1997

http://www.texastop10.princeton.edu/admin_overview.html

**Expected family contribution:** In estimation, I use the family's expected family contribution (EFC) as a measure of the amount of financial aid the student is likely to receive. I use the federal government's "simplified EFC formula worksheet A" from the 2002-2003 FAFSA, using parents' income, the number of parents/guardians who live with the applicant, and the number who work full-time, if the applicant lives with two parents.

In estimation, I draw a large number of income draws for each student from the CPS. For each income draw, I calculate the EFC using the formula. I integrate over these draws when calculating the likelihood and moments. Details are provided in the computational appendix.

**Administrative dataset:** Table A1 provides a list of all institutions for which THEOP provided administrative data, together with the years covered and number of applications (in application data), or semester-individual observations (in transcript data).

The THEOP survey is based on a stratified sampling design that placed unequal weights on different types of high schools. Using the population weights, however, one obtains patterns very similar to those in the administrative data. A sequence of papers by Marta Tienda and coauthors describes the survey design and findings. I have confirmed that population-weighted mean SAT scores, class rank, and high school characteristics (poverty, SAT scores) of students in the survey who attend flagship universities closely match the actual numbers in the administrative data.

# B Model Appendix [Online]

## B.1 Optimality of cutoff rules

This section shows that cutoff rules are optimal for colleges.

As before, each college $j$ maximizes the expected quality of its entering class subject to the constraint that its expected enrollment is less than its capacity. The quality of a class $C(j) \subset I$, denoted $\Pi_j(C_j)$, is the sum of the quality of enrolled students:

$$\Pi_j(C(j)) = \sum_{i \in C(j)} \pi_{ij}.$$

Each college $j$ chooses an offer probability for each applicant, $B(j) \in [0,1]^{A(j)}$ where $A(j)$ denotes the set of applicants to $j$, in order to solve

$$\max_{B(j) \in [0,1]^{A(j)}} E\Pi_j(C(j)|A(j)) \text{ s.t. } E(C(j)|A(j)) \leq K_j.$$

college $j$ observes $(z_i, q_i, \mu_{ij})$ for all applicants before choosing who to admit. College $j$ does not observe $\mu_{ij'}$ for $j' \neq j$. Colleges cannot commit to an admission rule for non-automatically-admitted students that is not optimal given applications.

A college is said to be ***nonselective*** (given students' and other colleges' decisions) if its capacity constraint does not bind when it admits all students with $\pi_{ij} > 0$. Nonselective colleges admit all students for whom $\pi_{ij} > 0$.

Colleges whose capacity constraints bind are said to be *selective*. It is also optimal for selective colleges to employ cutoff rules, a result stated in the following Lemma:

**Lemma 1.** *Colleges' optimal admissions rules consist of cutoffs $\{\underline{\pi}_j\}_{j \in \mathscr{J}}$ such that an applicant is admitted to $j$ if and only if $\pi_{ij} > \underline{\pi}_j$, and either $\sum_{i:\pi_{ij} > \underline{\pi}_j, j \in A_i} Pr(C_i = j) = k_j$ or $\underline{\pi}_j = 0$.*

*Proof.* Consider a selective college. Let $B(j)$ be an admissions rule that satisfies the expected capacity constraint, and suppose it is not a cutoff rule. Then there are two applicants $i$ and $i'$ such that $\pi_{ij} < \pi_{i'j}$ but $B(j)(i) > 0$ and $B(j)(i') < 1$. If admitted, $i$ attends $j$ with probability $P_{ij}$ and $i'$ attends with probability $P_{i'j}$, for some $P_{ij}, P_{i'j} \in (0,1)$. It is feasible and profitable for $j$ to reduce $B(j)(i)$ by $\frac{\varepsilon}{P_{ij}}$ and increase $B(j)(i')$ by $\frac{\varepsilon}{P_{i'j}}$ for some $\varepsilon$. $\square$

## B.2 Rational-Expectations Equilibrium

A rational-expectations equilibrium is a tuple

$$\left\{\{A_i\}_{i\in I}, \{Admit(j)\}_{j\in J}\right\}$$

that satisfies the following properties:

1. $A_i \in \mathscr{A}$ solves $i$'s application problem

$$\max_{A\in\mathscr{A}} V_i(A)$$

   given admissions rules $Admit(j)$and $i$'s characteristics, where $\mathscr{A}$ is the set of feasible portfolios.

2. For each college $j$, $Admit(j) \subset I$ maximizes

$$\mathbb{E}\left(\Pi_j | x_i, z_i, q_i, \mu_i, Admit_{-j}\right)$$

   subject to

$$\sum_{i\in Admit(j)} Pr(C_i = j | x_i, z_i, q_i, Admit_{-j}) \leq k_j,$$

   where $Pr(C_i = j | x_i, z_i, q_i, \mu_i; Admit_{-j})$ are the matriculation probabilities induced by $A$, integrating over the distribution of the student's other applications and admissions offers given the characteristics $x_i, z_i, q_i, \mu_i$ observed by college $j$ and the other colleges' policies $Admit_{-j}$.

3. $i \in Admit_j$ if and only $j \in B_i$.

Because optimal admissions rules are cutoffs, I identify admissions rules with the corresponding cutoffs and consider equilibria of the form:

$$\left\{\{A_i\}_{i\in I}, \{\underline{\pi}_j\}_{j\in J}\right\}.$$

## B.3 Existence of Equilibrium

Suppose there is a continuum of students with type $\omega \in \Omega \subseteq R^N$ and measure $F(\omega)$ with density $f$ with respect to Lebesgue measure. (In the model, an individual is defined by his observables $z_i$, his preference terms and random coefficients, and his caliber and signal $(q,s)$.) Each cutoff vector $\underline{\pi}$ induces a joint distribution of application portfolios $A \in \mathscr{A}$ and types $\omega$. In particular, for almost all $\omega$ there is a unique portfolio $A(\underline{\pi},\omega) \in \mathscr{A}$ that is optimal.

To prove existence, we need to show that there is a cutoff $\underline{\pi}$ such that $A_i$ is a best response to $\underline{\pi}$ for all $i$, and that $\underline{\pi}$ maximizes quality subject to capacity constraints given applications $A(\underline{\pi},\omega)_{\omega\in\Omega}$.

Let $\tilde{A}(\underline{\pi})$ denote the joint distribution of $(A \in \mathscr{A}, \omega \in \Omega)$ induced by students' best responses to $\underline{\pi}$. Because $\omega$ admits a density and the indirect utility $U(\mathscr{A};\underline{\pi})$ is continuous in $\underline{\pi}$, the share of applicants to each portfolio $A \in \mathscr{A}$ changes continuously in $\underline{\pi}$. Moreover, the share of students with admission set $B$ given application set $A$ changes continuously in $\underline{\pi}$, for all $A, B \in \mathscr{A}$. Conditional on admission set $B$, application set $A$, and $j$'s information about $\omega$, the probability of attending $j \in B$ is continuous in $\underline{\pi}$. As a result the best-response cutoff function $h(\underline{\pi}) \equiv \underline{\pi}^*(\tilde{A}(\underline{\pi})) : R^J \to R^J$ is continuous in $\underline{\pi}$.

We now show that $h(\cdot)$ is bounded, i.e. we can restrict attention to the map $h(\cdot)$ on a box in $R^J$. For each $j$, taking the distribution of applications $A$ that has every student apply to $\{j\}$ with probability 1 gives an upper bound $\underline{\pi}_j^{high} \leq \underline{\pi}^*(E(A))$. There exists a cutoff $\underline{\pi}_j^{low}$ such that $j$ strictly prefers empty seats to students with caliber less than $\underline{\pi}_j^{low}$.

From Brouwer's fixed point theorem, $h$ has a fixed point $\underline{\pi}^*$. By construction $\underline{\pi}^*$ is a set of market-clearing cutoffs given applications $\tilde{A}(\underline{\pi})$, and applications $\tilde{A}$ are optimal given $\underline{\pi}^*$.

## B.4 Limited multiplicity of equilibria

In general, there is no guarantee that equilibrium is unique.[33] In this section, I show that given fixed distribution of application portfolio choices, there is a unique mutual best response by colleges. That is, there is precisely one vector of cutoffs $\underline{\pi}$ such that each $\underline{\pi}_j$ is optimal given $\underline{\pi}_{-j}$ and applications.

To show these results, I show that holding applications fixed, students' matriculation probabilities are isotone in $\underline{\pi}$. As a result, market-clearing cutoffs $\underline{\pi}$ form a complete lattice. As a corollary, for any set of applications, cutoffs exist that satisfy the colleges' problems. Letting $\underline{\pi}_L$ denote the lowest cutoff and $\underline{\pi}_H$ denote the highest, When students' choice probabilities satisfy a substitutes property, I show that $\underline{\pi}_L \neq \underline{\pi}_H$ leads to a contradiction,

[33] See Chade, Lewis, and Smith (2014).

as the share of students attending the outside option and/or nonselective colleges must be strictly higher under $\underline{\pi}_H$ than under $\underline{\pi}_L$, but the share at each selective college cannot decrease.

Let $Pr(C_i = j|B)$ denote the probability that student $i$ enrolls in college $j$ given admission to a set of colleges $B \subseteq \mathscr{A}$. Let $Pr(C_i = 0|B)$ denote the probability that $i$ chooses the outside option.

**Condition** (Substitutes)**.** For each individual, the following condition holds:

$$B \subseteq B' \implies Pr(C_i = 0|B') < Pr(C_i = 0|B).$$

This condition holds generally whenever the outside option is a possible substitute for each college.[34] In particular it is satisfied for the nested logit specification of utility in this model conditional on random coefficients, and hence for the mixed nested logit specification.

**Proposition 2.** *If the substitutes condition holds, then conditional on applications there is a unique cutoff vector $\underline{\pi}$ that satisfies equilibrium condition (3).*

*Proof.* For each set $B \subseteq \mathscr{A}$ and and $B' \subseteq A$ with $B \subseteq B'$, each agent's choice probabilities must satisfy[35]

$$Pr(C_i = j|B') \leq Pr(C_i = j|B). \qquad (*)$$

Define $\underline{\pi}^*(\underline{\pi}) : R^{|J|} \to R^{|J|}$ by $\underline{\pi}^*_j(\underline{\pi}) = \underline{\pi}^*_j(\underline{\pi}_{-j})$. That is, the $j$th component of $\underline{\pi}^*$ is the cutoff that college $j$ optimally chooses given (fixed) applications and student preferences, and the cutoffs of other colleges. By (*), the function $\underline{\pi}^*$ is isotone in $R^{|J|}$. Its fixed points therefore form a complete lattice.

Let $\underline{\pi}^*_L$ be the lowest equilibrium cutoffs, and $\underline{\pi}^*_H$ be the highest. Let $\mathscr{J}_u \subseteq \mathscr{J}$ denote the set of schools that are not selective at $\underline{\pi}_L$, i.e. the union of the outside option and the set of schools for which $\underline{\pi}^*_{Lj} = \pi^0_j$. For a contradiction, suppose $B_i(\underline{\pi}^*_H) \neq B_i(\underline{\pi}^*_L)$ for some student $i$.

The share of students attending the outside option must strictly increase. Moreover, in any random utility model the the share of students attending other colleges in $\mathscr{J}_u$ must also

[34] The condition is related to the "connected substitutes" condition of Berry, Gandhi and Haile with continuous prices. In my setting, what is needed is that when a capacity-constrained college is removed from the choice set, the probability of not attending a capacity-constrained college strictly increases.

[35] This result is due to Block and Marshak (1960) and Falmagne (1978). See also Haile, Hortacsu and Kosenok (2008). In particular, a direct application of Theorem 1 of Falmagne (1978) gives that for any sets $B_0, B_1 \subseteq \mathscr{A}$, we have $Pr(C_i = j|B_0) - Pr(C_i = j|B_0 \cup B_1) \geq 0$.

weakly increase.[36] Therefore the share of students attending colleges in $\mathscr{J}_u$ must strictly increase:

$$\sum_i Pr(C_i \in \mathscr{J}_u | \underline{\pi}_H^*) > \sum_i Pr(C_i \in \mathscr{J}_u | \underline{\pi}_L^*).$$

Cutoffs of college in $\mathscr{J} \setminus \mathscr{J}_u$ are weakly higher under $\underline{\pi}_H^*$, however, implying that they are at full capacity in both $\underline{\pi}_L^*$ and $\underline{\pi}_H^*$, i.e.

$$\sum_i Pr(C_i \notin \mathscr{J}_u | \underline{\pi}_H^*) = \sum_i Pr(C_i \notin \mathscr{J}_u | \underline{\pi}_L^*),$$

which is a contradiction. □

# C Estimation Appendix [Online]

## C.1 Likelihood

Let $\theta \in \Theta \subset R^N$ be a vector of parameters, where $\Theta$ is the set of allowed parameter values. Recall that $A_{ij}, B_{ij}, C_{ij}$ denote applications, admission, and matriculation respectively for student $i$ and college $j$. $Aware_{ij}^{obs} \in \{0,1\}$ is an indicator for completing an application for financial aid at college $j$ and is observed only if $A_{ij} > 0$. $Aware_{ij}^{obs} = 1(\text{i applied for financial aid at j})$ is generally not the full vector of financial-aid awareness, because we do not observe whether $i$ would have completed an application for aid at colleges outside his application set.

Let $\omega_i$ denote the vector of random coefficients $\beta_i$ of individual $i$ as well as income $y_i$, expected family contribution $EFC_i$, random terms in aid, and shocks $\nu_i$.

The likelihood of applications $\ell_i^{Ap}$ is the measure of the set of preference and signal draws such that the observed application portfolio maximizes expected utility among all application portfolios net of costs:

$$P(A_i) = \frac{\exp \lambda^{\text{app}} \left( V_{iA} - c_{z_i^{\text{app}}}^{fixed} - c_{z_i^{\text{app}}}^{var} |A| \right)}{\sum_{A \in \mathscr{A}} \exp \lambda^{\text{app}} \left( V_{iA} - c_{z_i^{\text{app}}}^{fixed} - c_{z_i^{\text{app}}}^{var} |A| \right)}.$$

[36] For $j \in \mathscr{J}_u$, if $\underline{\pi}_{Hj}^* = \underline{\pi}_{Lj}^*$ then the share must weakly increase by (*). Alternatively, $\underline{\pi}_{Hj}^* > \underline{\pi}_{Lj}^*$ if and only if college $j$ has filled its capacity under $\underline{\pi}_H^*$.

The following expression is the likelihood of admissions conditional on a signal $s$:

$$\ell_i^{B|A}(\theta, s) = \int \prod_{j\in B}\Big(\Phi(z_{ij}\gamma_j + q_i - \underline{\pi}_j)\Big) \prod_{j\in A\setminus B}\Big(1-\Phi(z_{ij'}\gamma_{j'} + q_i - \underline{\pi}_{j'})\Big)\, dF_i(q|s;\sigma_{q|s}).$$

The likelihood of financial-aid awareness $\ell_i^{\text{Aware}|A}$ is analogous to the term above, except that a double integral is taken over the distribution of income $y$ and random coefficient $\alpha_i^{\text{aware}}$ rather than an integral over $q$, and coefficients on $y$, $x$, $z_i^{\text{admit}}$, and program indicators are estimated.

The likelihood of matriculation is the nested-logit choice probability of school $j$ conditional on $i$'s characteristics:

$$\ell_i^{C|B}(\theta, \omega_i) = \frac{\exp(u_{ij}(\omega_i)/\lambda)\left(\sum_{j'\in B}\exp(u_{ij'}(\omega_i)/\lambda)\right)^{\lambda-1}}{1+\left(\sum_{j'\in B}\exp(u_{ij'}(\omega_i)/\lambda)\right)^{\lambda}}.$$

The dependence on aid awareness is implicit, via the impact of awareness at $j$ on $u_{ij}$. For students who appear in wave 1 of the survey, I set $\ell_i^{C|B} = 1$, as matriculation decisions are not observed. The likelihood of all observables in the data is given by

$$\ell_i(\theta) = \int_s\int_{\omega_i} \ell_i^A(\theta,\omega_i,s)\ell_i^{Aware|A}(\theta,\omega_i)\ell_i^{B|A}(\theta,s)\ell_i^{C|B}(\theta,\omega_i)dF_i(s;\theta)dG_i(\omega_i;\theta).$$

## C.2 Calculating the objective

Consider the optimization problem

$$\max_\theta \frac{1}{N}\sum_{i=1}^N \log \ell_i(\theta) + g(\theta)W_g g(\theta)^T, \tag{8}$$

where

$$g(\theta) = \frac{1}{N}\sum_i g_i(\theta)$$

are non-likelihood moments, $\log \ell_i(\cdot)$ is the log-likelihood, and $W_g$ is a weighting matrix.

Suppose that the objective in 8 is maximized at $\theta_0$. Then the FOC of objective (8) is given by

$$\frac{1}{N}\sum_i \nabla_\theta \log \ell_i(\theta_0) + 2g(\theta)W_g(\nabla_\theta g(\theta_0))^T = 0. \tag{FOC1}$$

Define the matrix $W$ by

$$W = \begin{pmatrix} \frac{1}{2}\left[\frac{1}{N}\sum_i \left((\nabla_\theta \log \ell_i(\theta_0))^T * (\nabla_\theta \log \ell_i(\theta_0))\right)\right]^{-1} & 0 \\ 0 & W_g \end{pmatrix}$$
$$= \begin{pmatrix} \frac{1}{2}\cdot \mathscr{I}_{\theta_0}^{-1} & 0 \\ 0 & W_g \end{pmatrix}$$
$$= \begin{pmatrix} -\frac{1}{2}\left[\frac{1}{N}\sum_i \frac{\partial^2}{\partial\theta\partial\theta'} log\ell_i(\theta_0)\right]^{-1} & 0 \\ 0 & W_g \end{pmatrix},$$

where $\mathscr{I}_{\theta_0}$ is the Fisher information matrix evaluated at $\theta_0$. Define $loglik(\theta) = \sum_i \log \ell_i(\theta)$. Then $FOC1$ is also the first-order condition of the following GMM objective function:

$$\min_\theta [\frac{1}{N}\nabla_\theta \log \ell(\theta), g(\theta)] W [\frac{1}{N}\nabla_\theta \log \ell(\theta), g(\theta)]^T, \tag{9}$$

Hence the solution to 8 will also satisfy 9. The advantage of 8 for my purposes is that it is possible to use gradient-based optimization methods without needing to calculate second derivatives of $\log \ell(\theta)$. Therefore, 8 is a relatively computationally inexpensive way to optimize the GMM objective 9.

The weight matrix is not known in advance; however it is a fixed function of the parameters $\theta_0$ that optimize the penalized-likelihood objective.

## C.3 Standard errors

I estimate standard errors via the bootstrap. I resample applicants with replacement from the survey data, and draw new draws of financial-aid awareness, random coefficients, shocks, and other latent variables in order to allow my procedure to account for simulation error as well as sampling error. Using this new resampled dataset, and taking the point estimates as starting values, I re-estimate the “first step” parameters $\theta$.

Population weights are needed for second-step estimation and counterfactuals. I construct new population weights proportional to the weights of each resampled individual, but summing to one.

Next, I redraw auxiliary-model parameters using their asymptotic mean and covariance matrix as estimated on the administrative datasets, then re-estimate the indirect-inference outcome specifications. Finally, I re-solve the model in counterfactuals and compute outcomes.

# D Computation [Online]

## D.1 Computing the value of application sets via the inclusion-exclusion principle

Computing the value of an application set $A \subset \mathscr{A}$ requires integrating over all possible outcomes $B \subseteq A$, as the value depends on the probabilities and utilities of each admissions set $B$ that is possible given application to $A$.

In principle, and dropping $i$ subscripts, computing $V(A)$ for all $A$ requires computing $|\mathscr{A}|$ utility terms $\{U(B)\}_{B \in \mathscr{A}}$, and $\mathscr{O}(|\mathscr{A}|^2)$ multivariate normal CDF evaluations $\{P(B|A)\}_{B \subseteq A, A \in \mathscr{A}}$. It would be expensive to compute all of these probabilities directly for each draw for each individual at each trial parameter value. In what follows, I show that one needs only evaluate $|\mathscr{A}|$ multivariate normal CDFs and perform some matrix multiplication.

The argument does not rely on the presence of a single common correlating factor, $q$. Therefore, this method may be useful in cases beyond this paper where it is necessary to compute all conditional admissions probabilities under a complicated correlation structure.

Define

$$P_B = \int_q \prod_{j \in B} \Phi(q_i + z_{ij}^{\text{admit}} \gamma_j - \underline{\pi}_j) dF(q|s)$$

as the probability of admission to every school in $B$ given application set $B$ and a realization of the applicant's information $q^s$. Let $X_B$ be the event that $i$ is admitted to all schools in $B$. Let $X_{A;B}$ denote the event that $i$ is admitted to all schools in $B$ and rejected from all schools in $A \backslash B$, and $P_{A;B}$ the probability of this event. If $B \nsubseteq A$ then let $X_{A;B}$ be empty and $P_{A;B} = 0$. Let $\mathbb{P}$ be the probability measure associated with $i$'s characteristics and signal $s$, so that $P_{A;B} = \mathbb{P}(X_{A;B})$.

**Proposition.** *(inclusion-exclusion formula) Given the above definitions, the following result holds:*

$$P_{A;B} = \sum_{B' : B \subseteq B' \subseteq A} P_{B'} \cdot (-1)^{|A| - |B'|} \textit{ for all sets } A, B \in \mathscr{A} \textit{ with } B \subseteq A.$$

*Proof.*

$$\begin{aligned}
P_{A;B} &= \mathbb{P}\left(X_B \setminus \left(\cup_{j\in A\setminus B}\left(X_{B\cup\{j\}}\right)\right)\right) \\
&= \mathbb{P}(X_B) - \mathbb{P}\left(\cup_{j\in A\setminus B}\left(X_{B\cup\{j\}}\right)\right) \\
&= P_B - \sum_{j\in B\setminus A} \mathbb{P}(X_{B\cup\{j\}}) + \sum_{j_1,j_2\in B\setminus A} \mathbb{P}(X_{B\cup\{j_1\}} \cap X_{B\cup\{j_2\}}) - \ldots + \mathbb{P}(\cap_{j\in A\setminus B} X_{B\cup\{j\}}) \\
&= P_B - \sum_{B':B\subseteq B'\subseteq A} P_{B'} \cdot (-1)^{|A|-|B'|+1}.
\end{aligned}$$

The second line follows because $X_{B\cup\{j\}} \subseteq X_B$ and the third line is the standard inclusion-exclusion formula. □

List all the portfolios $A_0, A_{1,}\ldots, A_{|\mathscr{A}|}$ in some order, and define the matrix $T$ by

$$T_{kl} = 1_{A_k\subseteq A_l} \prod_{j\in A_k} \binom{A_{lj}}{A_{kj}},$$

where $A_{lj}$ is the number of applications to $j$ in portfolio $A_l$. Similarly, define the matrix $S = T^{-1}$. We have

$$S_{kl} = (-1)^{|A_l|-|A_k|} T_{kl}.$$

It follows that $P_{A_k;A_l} = \sum_{A_k\subseteq A_r\subseteq A_l} T_{kr} P_r$

**Corollary.** *Let $P$ be the diagonal matrix with kth entry $P_{A_k}$. The vector $V = \{V_A\}_{A\in\mathscr{A}}$ is given by*

$$V = T * Diag(P) * S * U.$$

Computing only the admissions probabilities $P_B$ rather than all $P_{A;B}$ simplifies computation. As an example, if $\mathscr{A} = \{\{1\},\{2\},\{1,2\}\}$ we have the following calculation:

$$\begin{pmatrix} V_{\{1\}} \\ V_{\{2\}} \\ V_{\{1,2\}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} * \begin{pmatrix} P_{\{1\}} & 0 & 0 \\ 0 & P_{\{2\}} & 0 \\ 0 & 0 & P_{\{1,2\}} \end{pmatrix} * \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} * \begin{pmatrix} U_{\{1\}} \\ U_{\{2\}} \\ U_{\{1,2\}} \end{pmatrix}.$$

By using the inclusion-exclusion principle, we avoid having to compute $P_{\{1,2\};\{1\}}$ or $P_{\{1,2\};\{2\}}$ by integrating over multivariate normal densities.

## D.2 Discretization

To compute integrals over $s$, and over the residual uncertainty $q|s$, I use Gauss-Hermite quadrature with 9 nodes.

I draw $M = 384$ draws $\omega_{i,m}$ for each individual $i$, which contain random coefficients, income $y_{im}$, EFC $efc_{im}$, and preference shocks $\nu$. The construction of income and EFC draws is described in the data appendix. To draw random coefficients and shocks $\nu$, I fix a set of iid standard Normal draws. I use an importance-sampling procedure, which I describe below to integrate over financial-aid awareness draws.

## D.3 Simulation of financial-aid awareness

I observe financial-aid application or failure to apply for aid only at schools to which $i$ submits an application. In order to compute the likelihood that $i$ chooses his observed application portfolio, however, we need to calculate the value of all application portfolios he could have chosen. Therefore calculating the likelihood requires integrating over financial-aid awareness at colleges to which $i$ did not apply.

In order to obtain an estimator that is smooth in $\theta$, I draw financial-aid awareness for student $i$ using an importance sampling procedure. Before searching over parameter values, I draw a vector of outcomes $aware_{ijm} \in \{0,1\}$ once for each simulation draw, for all colleges $j$ to which $i$ did not submit an application. I estimate weights on each vector for each parameter vector $\theta$. These weights are then smooth functions of $\theta$.

Initially, I estimate the parameters of equation (4), taking the set of observed applications as given, ignoring selection, and integrating over the distribution of the random effect $\alpha_i^{\text{aware}}$.

for each individual $i$ and simulation draw $m$, I then draw a vector of financial-aid awareness outcomes for the schools to which $i$ did not apply, according to their distribution given the parameters estimated above, and the agent's observed financial-aid awareness, subject to the constraint that, at colleges to which $i$ applied, awareness must take its observed value.

Let $P_{0i}^{aware}$ denote the probability of the vector of financial aid awareness draws under $\left\{Pr_{0ijm}\right\}_{j\in\mathscr{J}}$ for simulation draws $m = 1,\ldots,M$.

For each trial parameter value $\theta$ in estimation, I calculate the probability of financial-aid awareness,

$$P_{im}^{aware}(\theta) = \int_\alpha \prod_{j\in\mathscr{J}} \left[\left(Pr(Aware_{ij} = 1|\theta, \alpha_i^{\text{aware}})\right)^{aware_{ijm}} \left(Pr(Aware_{ij} = 0|\theta, \alpha_i^{\text{aware}})\right)^{1-aware_{ijm}} dF(\alpha_i^{\text{aware}}|\theta)\right].$$

Note that the probability of the financial aid awareness draw under parameters $\theta$ includes the probability of $aware_{ijm}$ for $j \in A_i$.

The likelihood of financial aid awareness is then approximated by

$$\hat{\ell}_i^{F^{aware}|A,B}(\theta,m) \approx \frac{P_{im}^{aware}(\theta)}{P_{0im}^{aware}}.$$

In estimation, I obtain a starting guess for financial-aid awareness parameters as described above, which forms part of the starting values $\theta^0$. I draw financial-aid awareness according to these parameters. I then run the estimation procedure, arriving at estimates $\theta^1$, and use the estimated values as new initial parameters for constructing $p_{0im}^{aware}$. I then run the estimation procedure again to obtain estimates $\theta^2$. Finally, I redraw financial-aid awareness again, and rerun the estimation procedure, to obtain final point estimate $\theta^*$.

### D.4 Implementation

The first-stage estimation procedure uses an implementation of the L-BFGS algorithm provided by the NLopt solver package. I provide efficient analytic first derivatives to the solver for first-stage estimation.

The second-stage uses the L-BFGS algorithm as well. I use automatic differentiation to obtain gradients for the outcome equations in the second stage.

Counterfactuals require solving for a vector of new cutoffs at three programs. I use the baseline cutoffs, recovered during estimation, as starting values. I use the L-BFGS algorithm to minimize the sum of squared differences between total enrollments in the counterfactual and at baseline, using automatic differentiation to obtain the gradient. In practice this procedure converges rapidly.

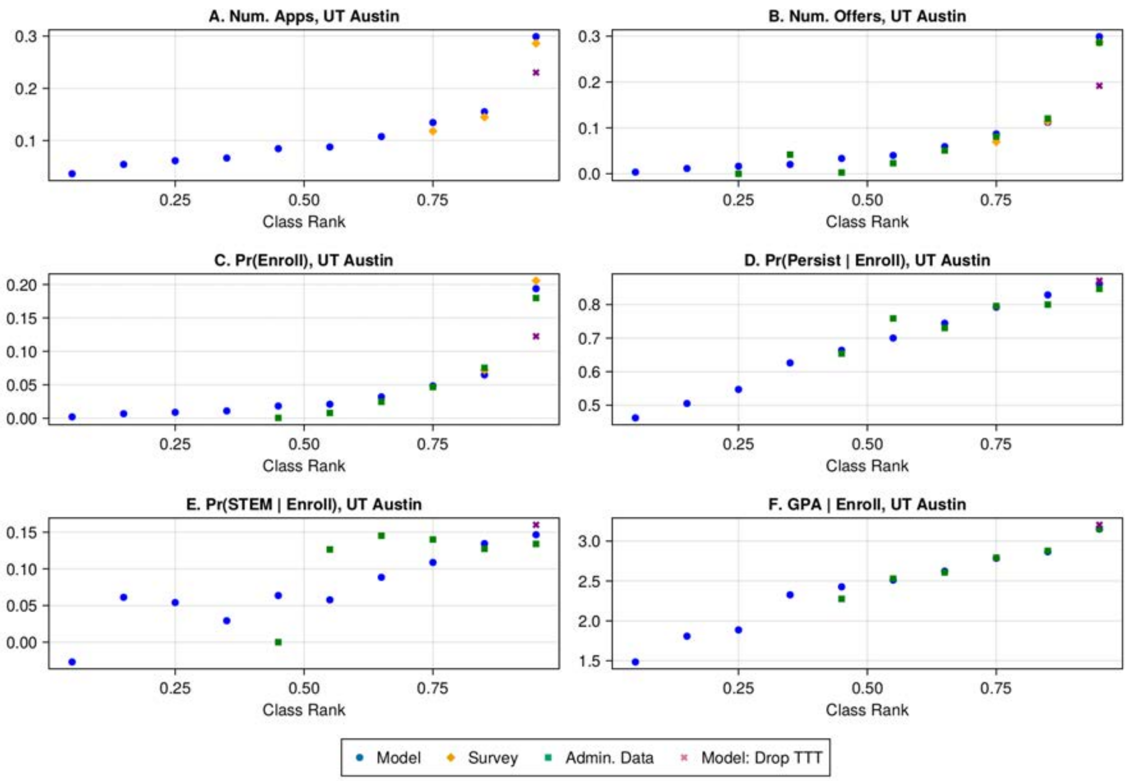
## E Additional Estimates and Results [Online]

This section provides additional model-fit results, results on the importance of accounting for selection, and additional counterfactuals and extensions.

### E.1 Additional Model-Fit Figures

Figures A1 and A2 show the analogues of Figure 1, separately for each flagship institution. Figure A3 shows fit of applications, admissions, and matriculation by class rank decile for all public institutions, and for all four-year institutions. Results are similar to those in the main text.

Figure A1: Model Fit: UT Austin




## E.2 Parameter Estimates

This section provides additional information on parameters. A full set of estimates is given in the supplementary material. Table A2 shows cost and information parameters. While estimates of information parameters $\gamma^s$ are noisy, we can reject large values of $\sigma^s$. Table A3 shows the estimated change in cutoffs across counterfactual scenarios. Table A4 shows probit versions of the linear specifications reported in Table 2 for binary outcomes.

## E.3 Additional Results

This section provides additional results and context. I first consider the importance of accounting for selection. I then decompose post-enrollment outcomes at flagship universities.

Figure A2: Model Fit: Texas A&M

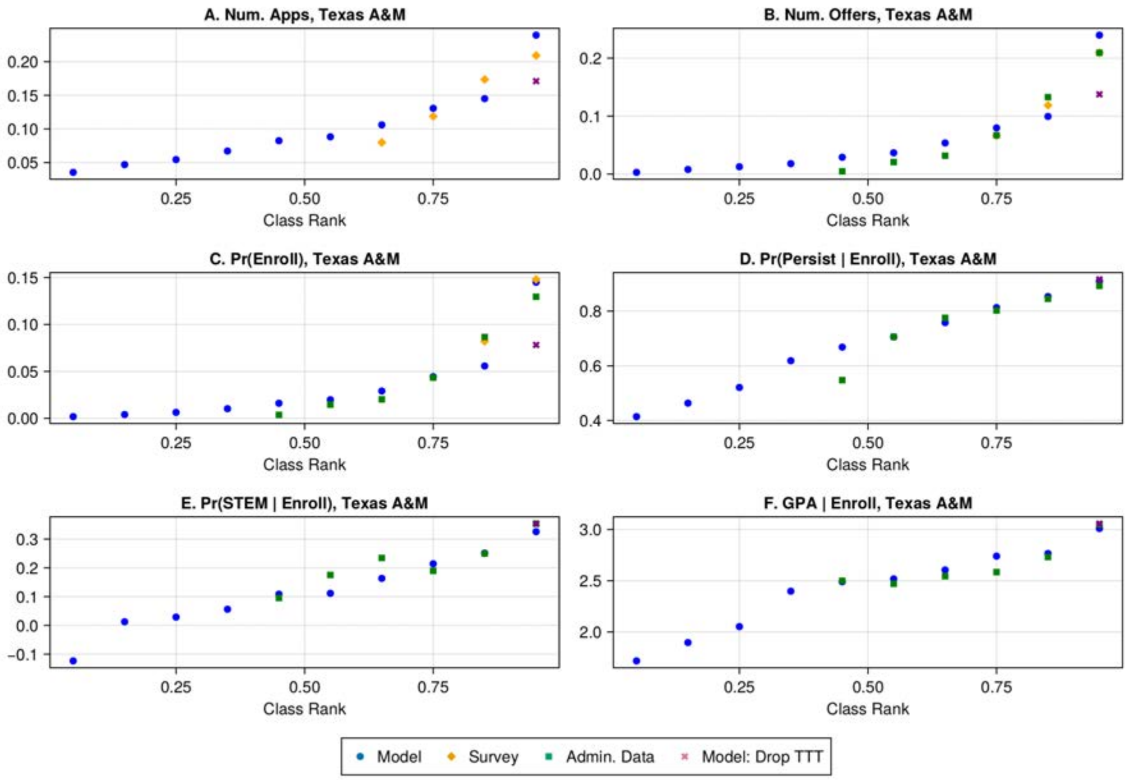


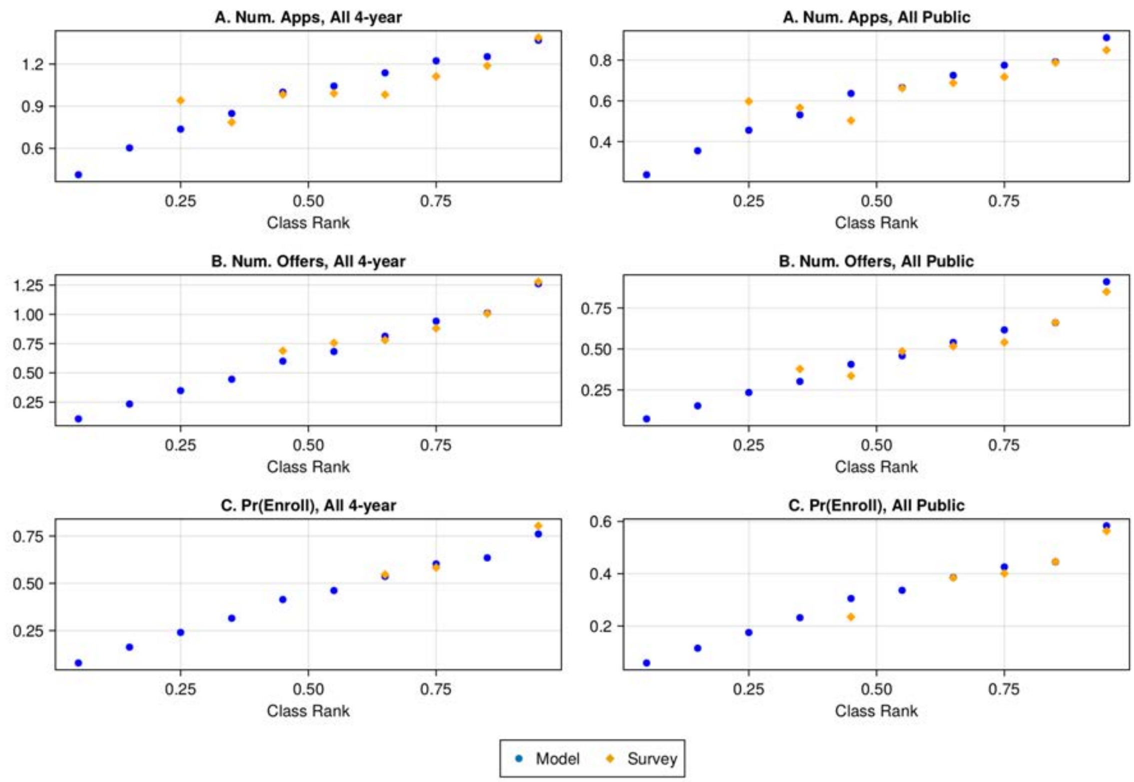

Table A2: Information Parameters

| | Student Info. $\gamma_s$ | Uncertainty $\gamma_{q\|s}$ | Fixed Cost $\gamma_0^c$ | Var. Cost $\gamma_1^c$ | Var. Shocks $\gamma^{\varepsilon_{\text{app}}}$ |
|---|---|---|---|---|---|
| Constant | -10.0 | -0.692 | 0.31 | 1.038 | -0.88 |
| | (0.556) | (0.353) | (0.091) | (0.053) | (0.094) |
| Poverty | 1.563 | 1.625 | -0.239 | -0.219 | -0.592 |
| | (8.376) | (0.763) | (0.111) | (0.119) | (0.21) |
| URM | 5.748 | 0.545 | 0.047 | -0.176 | -0.075 |
| | (6.26) | (0.305) | (0.06) | (0.045) | (0.079) |

Note: this table shows information and cost parameter estimates.

Figure A4 shows the impact of accounting for selection on applications and admissions. Estimates using matriculation choices only, taking choice sets as given, would have underestimated demand for non-flagship institutions and overstated the popularity of flagships.

Figure A3: Model Fit: Public Institutions and All Four-Year Institutions




While I estimate that top-decile applicants are generally well-informed about aid availability, Panel B of A4 shows that naive estimates ignoring selection would have slightly overestimated aid awareness among non-top-decile students at elite private institutions, at out-of-state institutions, and at Texas A&M.

In Figure A5 I show means of students enrolled at baseline, of students "pulled in" [Black et al., 2023] by mechanical and informative effects, and of students "pushed out" by equilibrium responses. I show decompositions of the percent plan as implemented (blue bars), of the percent plan combined with aid information (orange bars), and of the "top 20%" aligned percent plan (green bars). The leftmost column, "baseline," shows means of students enrolled in flagships at baseline. "Aid info" shows means among students pulled in by aid information without the percent plan (e.g. 3.051 GPA, top panel) and means among students "pushed out" by equililbrium cutoff changes (2.926 GPA, top panel). The next columns, "Mechanical" and "informative," show means of students pulled in by mechanical and informative effects respectively. Finally, the last column, "Eqbm," shows means of

Table A3: Change in Cutoffs Relative to Baseline

| Policy | UT Austin | Texas A&M | Other Public |
|---|---|---|---|
| TTP | 0.404 | 0.382 | 0.124 |
| | (0.03) | (0.03) | (0.01) |
| Aid Awareness Only | 0.092 | 0.021 | 0.062 |
| | (0.03) | (0.05) | (0.06) |
| Aid Awareness + TTP | 0.537 | 0.387 | 0.185 |
| | (0.06) | (0.09) | (0.06) |
| Top 5% zg | 0.029 | 0.019 | 0.007 |
| | (0.12) | (0.12) | (0.04) |
| Top 10% zg | 0.08 | 0.061 | 0.018 |
| | (0.11) | (0.11) | (0.04) |
| Top 15% zg | 0.138 | 0.109 | 0.03 |
| | (0.09) | (0.09) | (0.03) |
| Top 20% zg | 0.221 | 0.181 | 0.046 |
| | (0.07) | (0.06) | (0.03) |

Note: This table shows estimates of changes in program cutoffs relative to baseline. Lower cutoffs denote lower admissions standards.

Table A4: Outcome Probit Models

| | Probit UTA Persist | Probit UTA STEM | Probit TAMU Persist | Probit TAMU STEM |
|---|---|---|---|---|
| Constant | 1.403 | -3.098 | 1.755 | -2.542 |
| | (0.218) | (0.313) | (0.34) | (0.305) |
| SAT | 0.314 | 1.088 | 0.214 | 1.155 |
| | (0.112) | (0.177) | (0.108) | (0.492) |
| Class Rank | -0.714 | -1.661 | -0.845 | -0.923 |
| | (0.518) | (1.188) | (0.569) | (0.962) |
| SAT Ratio | -0.208 | 1.35 | -0.25 | 1.41 |
| | (0.141) | (0.207) | (0.212) | (0.369) |
| Poverty | -0.953 | -0.855 | -0.908 | -1.118 |
| | (0.109) | (0.229) | (0.179) | (0.173) |
| URM | 0.067 | 0.252 | 0.095 | 0.238 |
| | (0.065) | (0.073) | (0.08) | (0.061) |
| Scholarship | 0.333 | 0.104 | -0.227 | 0.703 |
| | (0.101) | (0.228) | (0.236) | (0.317) |
| Caliber (q) | -0.773 | 1.137 | -1.314 | -0.383 |
| | (0.346) | (1.002) | (0.581) | (0.593) |

Note: This table shows probit specifications for outcome equations (Equation (6)).

Figure A4: Preferences and Awareness, Estimates vs. Naive Models


Note: This figure shows average choice probabilities when $B_i$ contains one copy of each institution (Panel A), and when elite private schools are excluded (Panels C and D). Panel B shows the share of students aware of aid at each institution. Panels A and C: all students fully aware of aid. Panel D: aid awareness is set to zero for all students. “Ignore selection”: In panels A, C and D, use estimates from a random-coefficients nested-logit specification estimated on final enrollment decisions only, taking choice sets as given; in Panel B, use observed aid awareness only, ignoring selection into applications. Top-decile class rank and other students shown separately. Institutions: (1) Non-Flagship In-State Public; (2) Private Secular; (3) Religious; (4) Out-of-State Public; (5) Elite Private; (6) Texas A&M; (7) UT Austin.

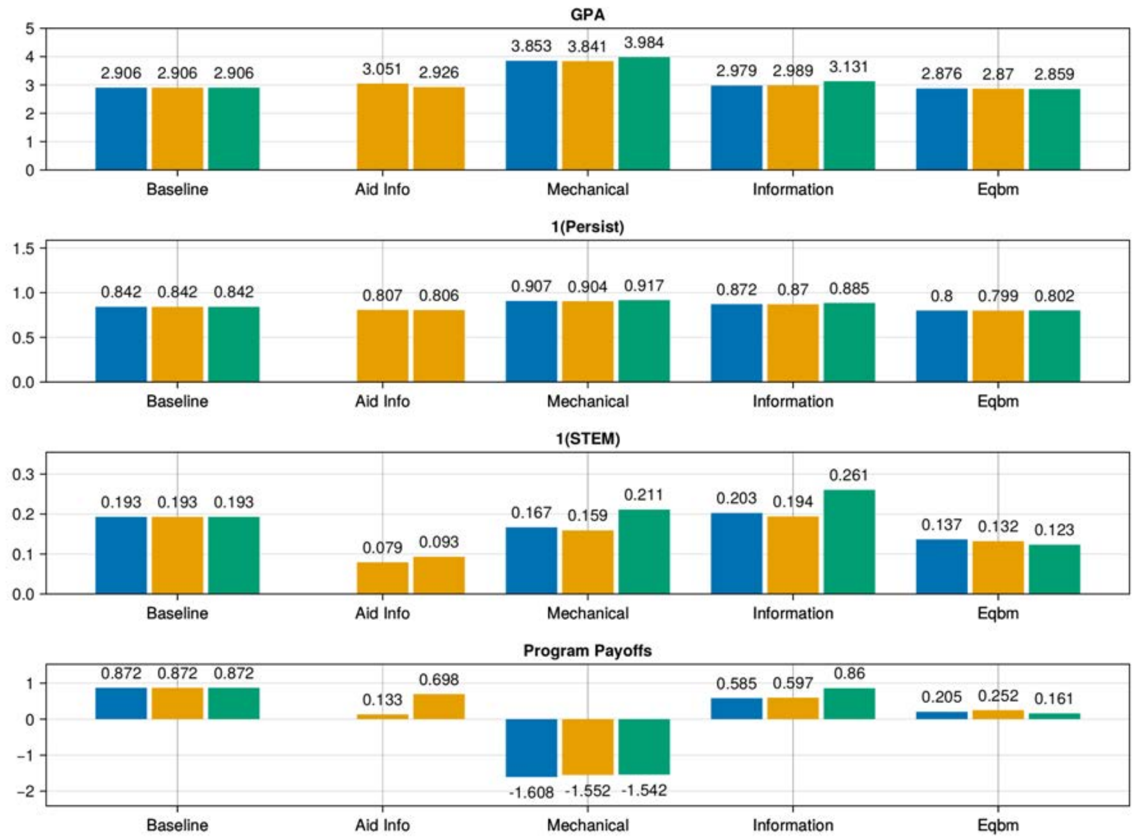
students “pushed out” by cutoff changes due to the percent plan.

In all academic measures, students pulled in by either the mechanical or informative channels outperform those pushed out.

Those induced to enroll by information are comparable to those enrolled at baseline in propensity to enroll in STEM, and are slightly more likely to persist and to earn high grades. Those pulled in mechanically earn even higher GPAs, but are less likely to major in STEM. Unsurprisingly, given that they were below the bar, those pulled in by mechanical effects provide very low payoffs to colleges.

Interestingly, when only aid information is provided, students “pulled in” by this information earn slightly higher GPAs but are less likely to persist and to major in STEM than the average student or the students they displace.

Figure A5: Decomposition of Post-Enrollment Outcomes




Note: This table shows mean post-enrollment outcomes of students enrolled in flagships, and of students shifted into or out of flagships by mechanical, informative, and equilibrium effects. "Baseline": means of students enrolled in flagships at baseline. "Aid info": left column shows means among students pulled in by aid information without the percent plan; right column shows means among students in turn "pushed out" by equililbrium cutoff changes. "Mechanical" and "informative": means of students pulled in by mechanical and informative effects respectively. "Eqbm": means of students "pushed out" by cutoff changes due to the percent plan.

## E.4 Enrollment, Match Effects, and Student-Level Impacts

So far, I have focused on outcomes at flagship institutions. These are of interest in evaluating transparency. For instance, if the students shifted into flagships by the percent plan were to drop out at high rates, this would provide an argument against the percent plan. To quantify impacts on students, however, one needs to understand what would have happened had they not attended flagship institutions. This in turn depends on two factors: (1) where students would counterfactually enroll, and (2) the presence or absence of "match effects". In this section I first show impacts on enrollment, then discuss the possibility of match effects. Finally, I present estimated impacts of policy changes on the number of students persisting two years across all institutions, not just flagships. I find that the gains for top-decile students were larger than the losses for those displaced, leading to an overall increase in persistence.

**Enrollment impacts:** Figure A6 describes counterfactual enrollment, comparing equilibrium under the percent plan to the baseline scenario, in equilibrium, in which there is no such plan. Blue bars show the gains in enrollment at each institution in equilibrium under the percent plan, relative to baseline, among top-decile students, as a percentage of the total measure of top-decile students. Orange bars show *decreases* in enrollment at each institution in equilibrium under the percent plan, relative to baseline, among non-top-decile students. To make the scale comparable, changes are divided by the measure of top-decile students. This figure shows that the probability of enrollment at each flagship institution among top-decile students increases by roughly 4.5 points. At UT Austin, this is comparable to the 5.3-point increase among "pulled in" students estimated in Black et al. [2023]. In contrast, top-decile enrollment at other institutions falls. In particular, despite the presence of an admissions guarantee at non-flagship in-state public institutions, in fact the probability of enrollment there among top-decile students falls slightly with the percent plan, as some of these students substitute to flagships.[37]

Importantly, Under my counterfactuals, total four-year college enrollment essentially does not change. Because capacities are held fixed, necessarily an equal measure of non-top-decile students is displaced from flagships and from public non-flagship institutions. Moreover, while I do not hold capacities at the remaining four options fixed, aggregate changes in enrollment there are negligible. The total increase in four-year college enrollment among top-decile students is equal to 5.26% of top-decile students. The total decrease in four-year college enrollment among non-decile students is equal to 2.38% of non-top-decile students. In total these changes essentially offset, leading to a 0.01% increase in total four-year enrollment.

My results here differ from those of Black et al. [2023] because, in addition to the direct effects of admissions changes at flagships, I am including indirect effects in which students who are displaced from flagships in turn displace other students, who then displace others, and so on. By solving for equilibrium, I am accounting for the possibility that non-top-

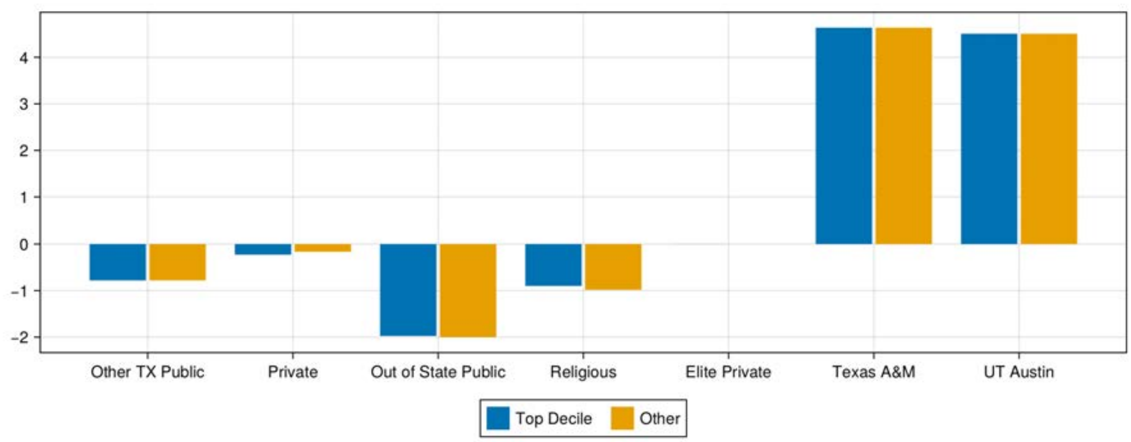
[37] I estimate a 5.26 point increase in overall four-year college enrollment for this group, comprising decreases in out-of-state and private institution enrollment and a 8.36 point increase in in-state public university enrollment. This latter number is larger than the 6.6 point increase in public four-year college enrollment estimated by Black et al. [2023] (p.49), and is driven by larger estimated enrollment increases at Texas A&M University, which may be due to differences in the population being considered.

decile students whom the percent plan pulled out of flagships set off chains of displacement.

Alternatively, one may wish to focus on specific groups of students defined by their potential assignments, as in Black et al. [2023]. There, "Pulled In" students are those who attend a flagship if the percent plan is in place and otherwise do not. Conversely, "Pushed Out" students attend a flagship if and only if the percent plan is not in place. Most students who will in turn be displaced from non-flagship institutions by the "Pushed Out" group would not have attended flagships under either scenario. They are therefore not part of either the "pulled in" or "pushed out" group.

To make my analysis comparable to Black et al. [2023], I compute a counterfactual in which flagships (only) provide automatic admission, and flagship programs' cutoffs adjust to their equilibrium values, but there is no automatic admission at non-flagship institutions, and cutoffs at non-flagship institutions are held fixed at their baseline values. This counterfactual captures the direct effect of the guarantee at flagships, holding the probability of having an option to attend a non-flagship institution fixed at its baseline value for every student.

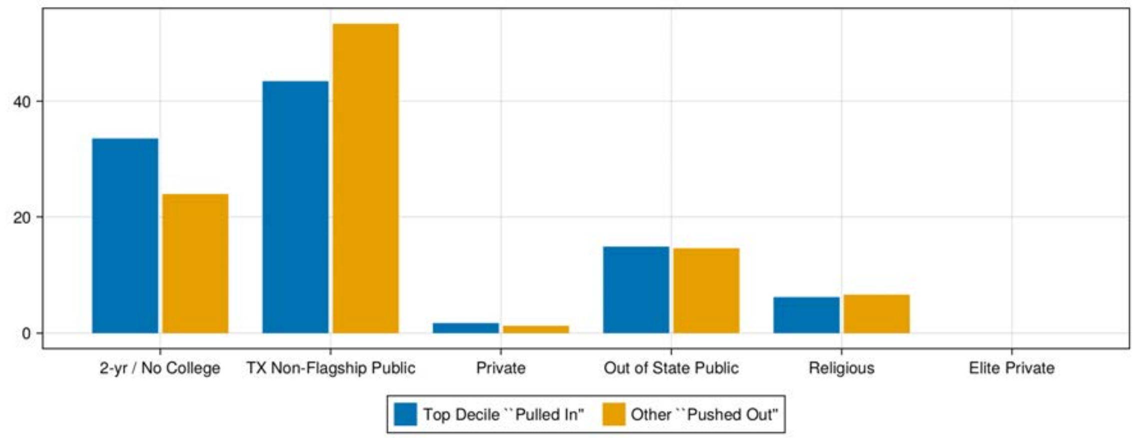
Figure A6: Equilibrium Changes in Enrollment




Note: This figure shows gains in enrollment shares in equilibrium under the percent plan, relative to baseline, for top-decile students (blue bars), and decreases in enrollment shares in equilibrium, relative to baseline, for non-top-decile students (orange bars).

Figure A7 provides results, showing the institutions that top-decile students who are "pulled in" to flagship institutions come from (blue bars), and the institutions that non-top-decile students "pushed out" of flagships attend (orange bars). I find that 34 percent of students "pulled in" substitute into flagships from the outside option, while only 24% of students "pushed out" substitute into the outside option (leftmost column). Instead, non-top-decile students displaced by the percent plan are more likely to substitute to non-flagship in-state institutions. These substitution patterns are broadly consistent with Black et al. [2023]. They find that about two thirds of "pushed out" students substitute into non-flagship institutions, while one third substitute into community colleges, while "pulled in" students are more likely to have been pulled into the public four-year sector from outside. Differences between Figures A7 and A6 are due to the subsequent adjustments in programs' cutoffs, and chains of rejections, set off by non-top-decile students' moves.

Figure A7: Where do Students Substitute From/To?




Note: This figure considers a counterfactual in which automatic enrollment is provided to top-decile students at flagships only. It shows enrollment shares, under this counterfactual, for top-decile students who attend flagship universities under this counterfactual but not at baseline (blue bars), and enrollment shares under the percent plan of students who attend flagships at baseline but are displaced when flagships' cutoffs adjust after the percent plan is introduced (orange bars).

**Match Effects:** The previous results suggest that, if there are gains or losses in aggregate human capital production, they will have to be driven by match effects. While students "pushed out" of flagships are more likely to attend other four-year institutions than those "pulled in" in the event they do not attend a flagship, the "pushed out" students in turn displace other students, so that overall effects on enrollment at any institution type are close to zero in equilibrium.

Estimating these effects requires estimating outcomes at non-flagship universities. To do so, I use THEOP public data from the non-flagship universities listed in Table A1. I restrict to Texas-resident students from public schools entering in Fall 2002, and drop institutions that do not have transcript data through at least 2004. As an outcome, I focus on two-year persistence.

I pool Rice University and SMU, and ascribe potential outcomes estimated at those universities to students enrolling in options 2 ("Secular Private"), 4 ("Religious"), and 5 ("Elite Private"). The remaining institutions, which are public, are used to estimate persistence at college 1 ("In-State Non-Flagship Public"). In addition, as I lack data on out-of-state institutions, I use the outcome equation estimated there to predict persistence for students matched to out-of-state public institutions.

Estimation is as described in Section 5.4. However, the auxiliary-model covariates $w_{ij}^{\text{outcome}}$ in Equation (7) differ. I observe high schools' poverty quartile, but not exact FRPL rate, mean SAT scores, or the availability of the LOS or CS programs. Hence, for non-flagship options $j$, the auxiliary-model covariates include indicators for poverty quartiles, and do not include "SAT ratio":

$$w_{ij}^{\text{outcome}} = \left(1, \text{SAT}_i, \text{classrank}_i, 1((\text{poverty quartile})_{h(i)}), \text{classrank}_i^2, \text{classrank}_i^3, \text{topdecile}_i, \text{urm}_i\right).$$

A second difference is that I do not include scholarships among the outcome covariates, as there were no targeted scholarship programs for these institutions comparable to the LOS or CS programs.

Table A5: Persistence Parameters, Non-Flagship Institutions

| | Non-Flagship Public | Private |
|---|---|---|
| Constant | 0.22 | 0.578 |
| SAT | -2.754 | 1.433 |
| Class Rank | 0.14 | -0.311 |
| SAT Ratio | 2.193 | -0.681 |
| Poverty | -0.142 | 0.015 |
| URM | 0.152 | 0.005 |
| Caliber (q) | -0.89 | 0.032 |

Note: This table shows outcome indices (Equation (6)) for non-flagship institutions. As these institutions do not have targeted (LOS or Century Scholars) scholarships, "Scholarship" is omitted.

Table A5 gives estimates of persistence parameters. While the coefficient on SAT is negative at public institutions, the coefficient on SAT ratio is also highly positive, leading to positive effects of increases in SAT scores all else equal (for the mean student enrolling in non-flagship public institutions, the peer-average SAT score at baseline is equal to 0.68.) Caliber is negatively associated with persistence at public institutions but not relevant at private institutions.

Figure A8 shows the share of students who enroll and persist at least two years in some four-year institution under my counterfactuals of interest overall (Panel A) and by subgroups (Panels B-F).

Overall persistence is roughly half a point higher under the percent plan than at baseline, and would be a further half-point higher if accompanied by aid information. Persistence under the "aligned" percent plan is similar to baseline, however.

Turning to subgroups, as the percent plan raises top-decile enrollment, the fraction of top-decile students enrolling and persisting grows, from 50% at baseline to 57.5% with the percent plan and aid-information intervention. The percent plan, with or without the aid intervention, would also raise the share of URM students (Panel C) and students from lowest-quartile schools (Panel D) who enroll and persist.

In sum, this evidence suggests that the percent plan raised total persistence, and improved outcomes for disadvantaged subgroups, by assigning students to colleges who were more likely to benefit. However, in equilibrium, tracing chains of rejection, some non-top-decile students were crowded out.

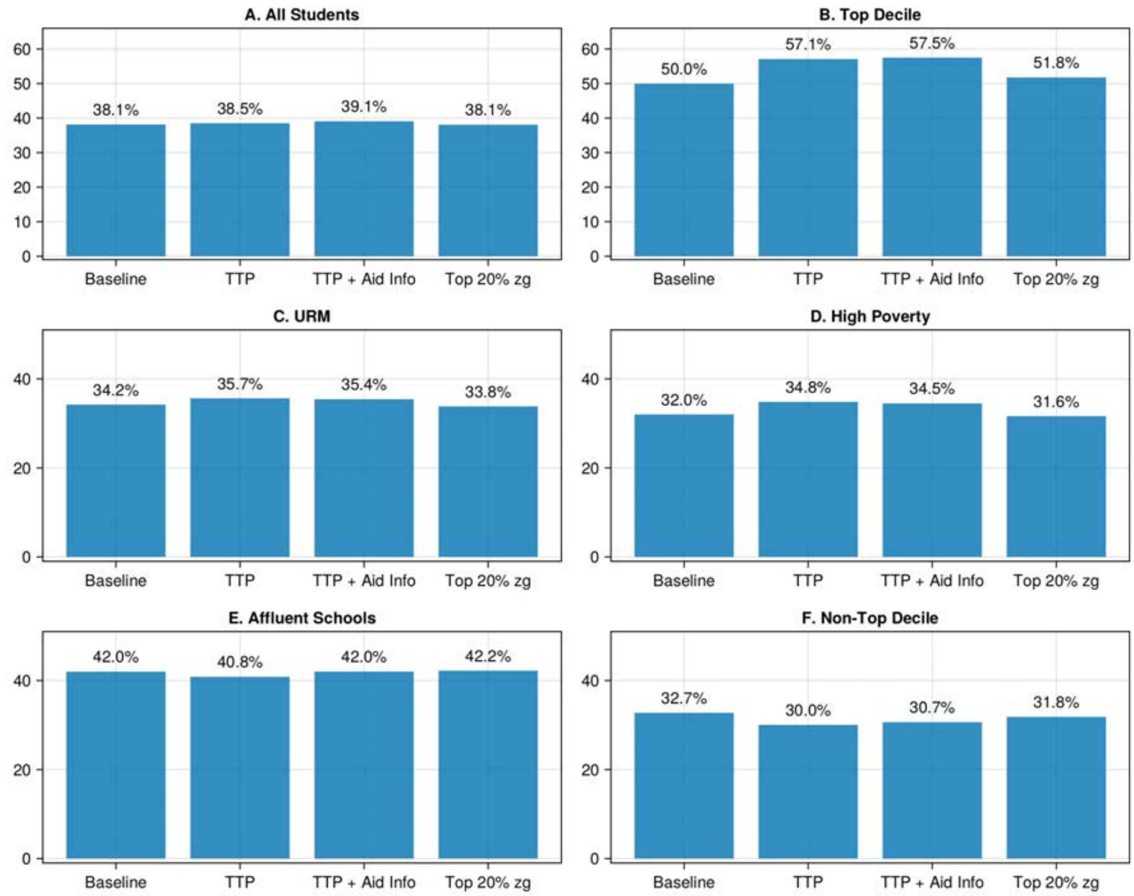
A caveat is that I am holding capacity at non-flagship in-state institutions fixed. If there is slack at some of those institutions, then total enrollment could increase with the percent plan (See Figure A7). This in turn would mechanically lead to further increases in the number of students persisting.

Black et al. [2023] conjecture that "Pushed Out students are likely to come from families with more support for college success, so may be less dependent on inputs received from the college itself."(p.30) My results are consistent with this conjecture. Conversely, Bleemer [2024] argues that gains from selectivity may be larger for students who are less well prepared. In contrast to the setting of Bleemer [2024], "Pulled In" students in my setting earn higher grades, and are more likely to persist, than the average student at the institutions they are pulled into.

## E.5 Effects of the Longhorn Opportunity Scholars (LOS) Program

This section provides an analysis of the effects of the LOS program, under the assumption that the assignment of the program to high schools was "as good as random" conditional on schools' poverty rates, average SAT scores, and other observables. I first show that estimated effects are consistent with results from the literature. I then provide further evidence of complementarities between admissions transparency and other forms of support—in this case, the availability of scholarship and mentorship programs.

Figure A8: Two-Year Persistence, All Institutions

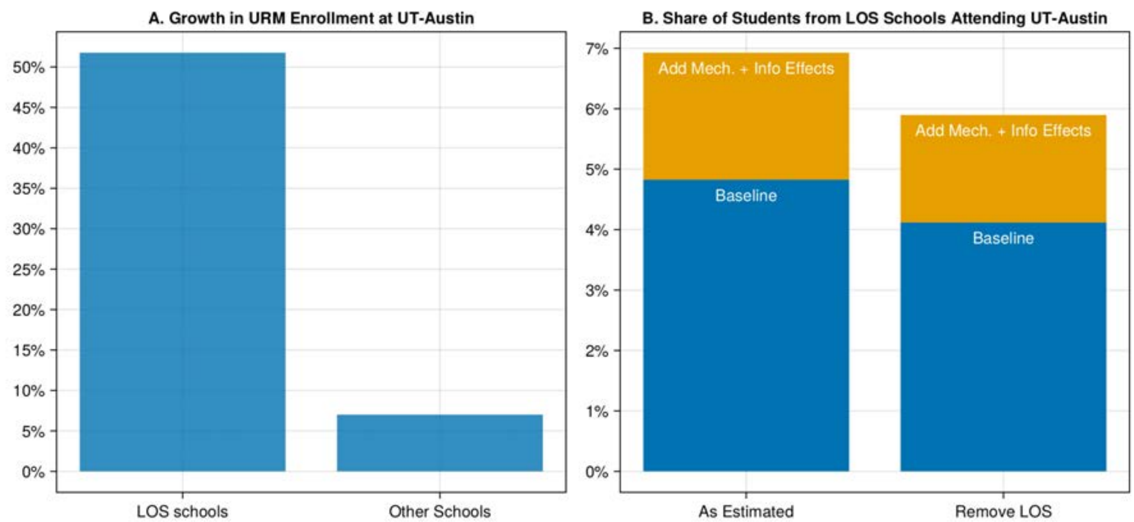



Panel A: Percentage of all students (population-weighted) who enroll in a four-year college and persist at least two years. Panels B through F: share of students in given subgroup who enroll in a four-year college and persist at least two years. "Top Decile": top 10% by class rank; "URM": Black or Hispanic; "High Poverty": top quartile of HS ever-received-FRPL; "Affluent Schools": bottom quartile of HS ever-received-FRPL; "Non-Top Decile": class rank outside of top decile.

Panel A of Figure A9 decomposes differences between my equilibrium and baseline counterfactuals, showing percentage growth in URM enrollment at UT Austin, relative to baseline enrollment numbers, separately by LOS schools—defined as those schools that participated in the LOS program in 2002—and other schools. At the baseline admissions cutoffs that I estimate, if there were no LOS program, 13.0% of URM students at UT Austin would have come from LOS schools. I estimate that the percent plan and the LOS program combined caused this share to grow by 51.8%. In contrast, the share of minority students from other high schools would have grown by 7.0%. As such, most of the growth in URM enrollment at UT Austin was driven by LOS schools.

This result is consistent with the argument and descriptive evidence in Kain et al. [2005] that effects were driven by LOS schools. Those authors argue that apparent increases in minority enrollment at UT Austin elsewhere may be due to other demographic changes, such as increases in the share of minority students at suburban high schools over the years 1998-2002, which were not caused by the percent plan. An advantage of my model-based approach is that it is able to evaluate policy changes while holding demographics fixed.

Figure A9: Effects of LOS Program




Note: Panel A shows growth in URM enrollment, relative to baseline total among students from LOS schools (left column) and other schools (right column), when the percent plan is introduced. Results come from model simulations, comparing "baseline" and "equilibrium" (with TTT) scenarios. In both cases, the LOS program is present at LOS schools. Panel B shows the impacts of the LOS program. Left column: baseline (no TTP) and mechanical+information effects of percent plan, taking LOS as given, among students from LOS schools. Right column: baseline (no TTP) and mechanical+information effects of percent plan on UT Austin enrollment among students from LOS schools, if there were no LOS program.

Panel B of Figure A9 shows baseline shares of students from LOS schools attending UT Austin (blue), and baseline shares + mechanical and informative effects (orange). The left column, "as estimated," reports mean estimated propensities to enroll at UT Austin from the main specification, restricting to students from LOS schools. In the right column, "Remove LOS", I consider the same set of students, but counterfactually remove the LOS program. I find that the LOS program would have increases UT Austin enrollment even without the percent plan, with 4.8% of students from LOS schools attending UT Austin, as opposed to 4.1% if there were no LOS program. Moreover, enrollment effects of the percent plan and the LOS program are complementary. UT-Austin enrollment would have increased by 2.1 points, in this population, if LOS and the percent plan were both introduced, as they were in the data (left column). If only the percent plan had been introduced, the increase in enrollment would have been 1.8 points, which is smaller. Hence, in the same way that admissions transparency and aid information are complements, admissions transparency and targeted scholarship/mentorship programs are complementary as well.

There are two important caveats to this analysis. First, for simplicity, and to isolate complementarities holding all else fixed, I am using previously-estimated baseline and equilibrium cutoffs which held the presence of the LOS program fixed. It would be possible to solve for equilibrium without the LOS program. Second, I am not modeling school-level unobservables with which LOS participation might be correlated. If LOS schools were negatively selected on baseline share of students sent to UT Austin, then my estimates of these complementarities may be lower bounds.

## E.6 Extension: Class balancing

In this section I consider robustness to a form of class balancing. Indices for GPA and persistence load heavily on class rank and exam scores. Because the Top Ten Percent plan forces colleges to admit students with high class rank, a college that wants some students with strong personal scores might place additional weight on personal characteristics for the remaining seats when the percent plan is in effect. In the absence of the percent plan, it might stop doing so [Antonovics and Backes, 2013b].

Accordingly, I allow admissions offices to place additional weight on academic characteristics in the baseline scenario in which there is no percent plan. To operationalize this idea, I take $\gamma_{UTA}^{\text{persist}}$, the coefficients on $(z, q)$ in equation (6), as the academic index (results are similar for Texas A&M's indices, and for GPA indices). I define $\gamma^{\text{personal}} \equiv (\gamma^{\text{admit}}, \gamma_q^{\text{admit}}) - (\gamma_{UTA,z}^{\text{persist}}, \gamma_{UTA,q}^{\text{persist}})$, where $\gamma_q^{\text{admit}} = 1$. Table A6 provides results. In the first row, at baseline, the admissions index is given by $\gamma^{\text{personal}} + 4\gamma_{UTA}^{\text{persist}}$. (I also consider a $2\times$ weight). Once top-decile students are automatically admitted, colleges return to their estimated preferences for the remaining applicants.

In this case, mechanical effects on top decile students' enrollment are not distinguishable from zero, and (unsurprisingly) impacts on persistence and other outcomes are attenuated. However, informative effects remain large and would raise enrollment of top-decile students by up to eight percentage points.

Table A6: Effects of TTP with Class Balancing

| | Pr(Enroll in Flagship), Treated Group | | | | Average Effect on Outcomes | | |
|---|---|---|---|---|---|---|---|
| | Base | + Mech. | + Info | % Info | ΔGPA | ΔPersist | ΔSTEM |
| *Class Balance: Additional Weight on Acad. Index* | | | | | | | |
| (7): 4× | 26.13 | 1.41 | 6.33 | 81.78 | 0.52 | 7.51 | 1.88 |
| | (2.74) | (2.73) | (0.4) | (13.75) | (0.33) | (6.5) | (8.43) |
| (8): 2× | 25.22 | 2.32 | 6.33 | 73.21 | 0.66 | 13.08 | 7.68 |
| | (2.04) | (1.93) | (0.4) | (9.87) | (0.12) | (5.72) | (7.32) |

Note: This table shows average impacts on academic outcomes, allowing for increased weight on an academic index in the baseline scenario. "Base": Share of "treated" group enrolling at baseline. "Mech": mechanical effect. "Info": information effect. "% Info": Share of total increase in treated-group enrollment due to information. Outcomes: total change in outcome (GPA, 1(Persist), 1(Persist and major in STEM), college payoff) / measure of top-decile students induced to enroll in flagships by mechanical and informative effects; this is equivalent to the average difference in outcomes between these students and the students they displaced.

## E.7 Extension: Selection on student preferences and caliber

In this section I extend the model of outcomes to include selection on students' preferences. Let $\tilde{v}_{ij} = v_{ij} + w_j \beta_i^w$ denote the value of preference unobservables—application-time shocks plus terms arising from random coefficients—that $i$ will receive if she enrolls in $j$. I assume:

$$\text{outcome}_{ij} = z_i^{\text{outcome}} \gamma_{z,j}^{\text{outcome}} + q_i \gamma_{q,j}^{\text{outcome}} + \tilde{v}_{ij} \gamma_{v,j}^{\text{outcome}} + \mu_{ij}^{\text{outcome}}. \tag{10}$$

Thus outcomes may be selected both on "caliber," which is observed by colleges, and on students' preferences.[38] For instance, the model allows students with a stronger preference for Texas A&M University to be more likely to persist if enrolled there, conditional on caliber. Estimation is as before. Implicitly, the indirect-inference procedure is making use of the "opposite" scholarship, e.g. the availability of the Longhorn Scholarship for students who enroll at Texas A&M. Intuitively, because the LOS program makes UT Austin observably more appealing, if a student from a LOS school attends Texas A&M this student must have had a higher preference shock for Texas A&M, in expectation, all else equal.

I hold the application-admissions-matriculation model parameters fixed, so that patterns of enrollment are as in the main specification. Table A7 presents parameters of the

[38] I have also estimated specifications of the form

$$\text{outcome}_{ij} = z_i^{\text{outcome}} \gamma_{z,j}^{\text{outcome}} + q_i \gamma_{q,j}^{\text{outcome}} + v_{ij} \gamma_{v,j}^{\text{outcome}} + \mu_{ij}^{\text{outcome}}$$

in which selection on preferences occurs only through the match-level shock $v$. Results are similar.

outcome equations, and Table A8 presents post-enrollment outcomes under counterfactuals, analogous to the main results in Table 3.

Table A7: Outcome Parameters Allowing Selection on Preferences

| | UTA GPA | LPM UTA Persist | LPM UTA STEM | Probit UTA Persist | Probit UTA STEM | TAMU GPA | LPM TAMU Persist | LPM TAMU STEM | Probit TAMU Persist | Probit TAMU STEM |
|---|---|---|---|---|---|---|---|---|---|---|
| Constant | 2.646 | 0.598 | 0.964 | -0.095 | -1.586 | 2.718 | 0.962 | 1.635 | -0.299 | -2.327 |
| | (0.382) | (0.194) | (0.071) | (0.263) | (0.739) | (0.492) | (0.098) | (0.146) | (0.372) | (0.535) |
| SAT | 4.748 | 0.251 | 0.541 | 0.191 | 0.522 | 3.866 | 0.132 | 0.215 | 0.34 | 1.586 |
| | (0.357) | (0.536) | (0.03) | (0.117) | (0.205) | (0.856) | (0.026) | (0.248) | (0.307) | (0.822) |
| Class Rank | -0.787 | -0.394 | -1.202 | -0.267 | -0.424 | -0.38 | -0.454 | -0.805 | -0.18 | -1.799 |
| | (0.325) | (0.152) | (0.164) | (0.502) | (1.809) | (0.493) | (0.178) | (0.208) | (0.536) | (0.797) |
| SAT Ratio | -2.714 | 0.005 | -0.121 | 0.215 | 0.951 | -2.223 | -0.088 | -0.244 | 0.436 | 1.161 |
| | (0.239) | (0.393) | (0.032) | (0.135) | (0.3) | (0.459) | (0.044) | (0.229) | (0.317) | (0.501) |
| Poverty | -0.22 | -0.342 | -1.124 | -0.111 | -0.346 | -0.176 | -0.262 | -0.974 | -0.296 | -0.961 |
| | (0.157) | (0.145) | (0.031) | (0.148) | (0.264) | (0.511) | (0.085) | (0.105) | (0.224) | (0.273) |
| URM | 0.053 | 0.028 | 0.08 | 0.045 | 0.184 | -0.009 | 0.002 | 0.098 | 0.085 | 0.227 |
| | (0.038) | (0.023) | (0.014) | (0.064) | (0.087) | (0.071) | (0.022) | (0.024) | (0.087) | (0.094) |
| Scholarship | 0.192 | 0.094 | 0.345 | 0.018 | 0.073 | 0.077 | 0.016 | -0.217 | 0.15 | 0.586 |
| | (0.054) | (0.03) | (0.029) | (0.1) | (0.197) | (0.137) | (0.072) | (0.1) | (2.207) | (0.275) |
| Caliber (q) | -0.991 | 0.042 | -0.298 | 0.129 | -0.129 | -1.097 | -0.062 | -1.25 | -0.199 | 0.128 |
| | (0.349) | (0.124) | (0.126) | (0.367) | (1.113) | (0.479) | (0.141) | (0.211) | (0.493) | (0.565) |
| Pref. Shock | -0.043 | 0.389 | 0.543 | -0.202 | -1.837 | -0.51 | 0.21 | 0.489 | -0.096 | -0.827 |
| | (0.71) | (0.253) | (0.059) | (0.243) | (0.579) | (1.499) | (0.149) | (0.489) | (0.326) | (0.835) |

Note: This table shows outcome indices (Equation (10)), extending the model to allow selection on unobserved components of utility.

GPA parameters are nearly unchanged. Persistence and STEM parameters are similar to the main estimates, but there is evidence of selection on match quality, although the preference coefficients (final row) are noisily estimated. However, even though preferences appear to matter for persistence in this specification—i.e. students who like their matched programs more indeed are more likely to persist—the counterfactual results are statistically indistinguishable from those of the main specification.

Turning to counterfactuals in Table A8, estimated impacts on GPA are nearly identical to those of the main specification, reported in Table 3. Point estimates of impacts on persistence and STEM major are slightly smaller and slightly larger, respectively, than in the main specification, but estimates here are not statistically distinguishable from the main estimates.

An implication of outcomes' dependence on preference shocks is that, to the extent that colleges value those outcomes, they would now face a selection problem in their admissions decisions. If colleges were to maximize persistence, and this in turn depended on $\nu_{ij}$, then the conditional expectation of outcome-relevant unobservables, conditional on

Table A8: Changes in Outcomes, Allowing Selection on Preferences

| | ΔGPA | ΔPersist | ΔSTEM |
|---|---|---|---|
| *A. Texas Top Ten* | | | |
| (1): TTP | 0.72 | 13.1 | 11.96 |
| | (0.08) | (2.77) | (3.25) |
| (2) +Aware | 0.73 | 13.32 | 11.01 |
| | (0.08) | (2.81) | (3.43) |
| *B. Automatic Admission By $z^{admit}\gamma$* | | | |
| (3): Top 5% | 0.21 | 2.02 | 5.15 |
| | (0.12) | (6.98) | (13.04) |
| (4): Top 10% | 0.39 | 4.36 | 9.58 |
| | (0.05) | (4.81) | (8.08) |
| (5): Top 15% | 0.54 | 6.34 | 12.97 |
| | (0.06) | (3.82) | (5.62) |
| (6): Top 20% | 0.69 | 8.5 | 16.51 |
| | (0.07) | (2.75) | (3.49) |

Note: This table shows a decomposition of changes in flagship enrollment for treated households, and average impacts on academic outcomes and programs' payoffs, allowing academic outcomes to be selected on preference shocks. "Base": Share of "treated" group enrolling at baseline. "Mech": mechanical effect. "Info": information effect. "% Info": Share of total increase in treated-group enrollment due to information. Outcomes: total change in outcome (GPA, 1(Persist), 1(Persist and major in STEM), college payoff) / measure of top-decile students induced to enroll in flagships by mechanical and informative effects; this is equivalent to the average difference in outcomes between these students and the students they displaced.

students' observables and accepting the offer, could differ across counterfactuals or policy changes. For instance, top-decile students might be more positively selected on preference shocks if there were no top-ten-percent plan. Put differently, if colleges did not also observe preference shocks, they would no longer have private values. For this reason, in counterfactuals the score $\pi_{ij}$ or relative ranking of students could change in principle, in addition to changes in "cutoffs".

As one obtains essentially the same impacts on counterfactual persistence, grades, and STEM majors abstracting from this source of selection, I abstract from it in the main specification.

# F Supplementary Materials [Not For Publication]

## F.1 Additional Descriptives

Table S1 reports THEOP sample means, using the population weights provided by the survey. According to these weights, about 7% of students attend a LOS school, and 15% attend a school with a poverty rate greater than 60%. Table S2 shows sample means at the student-by-college level. Tables S3 through S5 study how students' demographic covariates and application behavior differ by survey wave 2 participation, top-decile status, and participation in the Longhorn Opportunity Scholars program, respectively. Wave-2 students are disproportionately Black and Asian. These differences reflect an intentional decision by survey designers to oversample Black and Asian students. They are similar to wave-1 students in SAT scores and high school poverty rates. Top-decile applicants are about 6 percentage points more female than non-top-decile students. They are more likely to be Asian, and less likely to be Black, than non-top-decile students, but are equally likely to be White and/or Hispanic. They are much more likely to apply to UT-Austin, to apply for aid there, and to be admitted there. LOS participants are more likely to be Black, equally likely to be Hispanic, and less likely to be White or Asian, than the general population. They come from poorer schools than average and have lower SAT scores.

Table S6 provides statistics on admissions and matriculations by in-state and out-of-state applicants at UT Austin over the years 1991 through 2003 as given in the administrative data. The fraction of the enrolled students whom the university classified as in-state applicants remained slightly above 90% in each year with the exception of 1997, where 89.9% of first-time freshmen were in-state. The share of students from in-state public high schools is also nearly constant.

## F.2 Parameter Estimates

Tables S7 through S9 provide estimates from the main specification.

Table S1: Sample Means: Students (Pop. Weighted)

| **Variable** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
|---|---|---|---|---|
| Participates in Century Scholars Program | 0.025 | 0.157 | 0 | 1 |
| Participates in Longhorn Opp. Scholars Program | 0.069 | 0.253 | 0 | 1 |
| White | 0.626 | 0.484 | 0 | 1 |
| Black | 0.077 | 0.266 | 0 | 1 |
| Latino | 0.192 | 0.394 | 0 | 1 |
| Asian | 0.049 | 0.215 | 0 | 1 |
| Female | 0.567 | 0.496 | 0 | 1 |
| Num. Guardians | 1.714 | 0.546 | 0 | 2 |
| Class Rank | 0.268 | 0.181 | 0.1 | 1 |
| SAT* | 0.67 | 0.117 | 0.006 | 1 |
| HS mean SAT | 0.682 | 0.055 | 0.431 | 0.806 |
| SAT / HS mean SAT | 0.984 | 0.161 | 0.008 | 1.553 |
| HS Poverty (share ever FRPL) | 0.323 | 0.244 | 0.023 | 1 |
| High Poverty HS | 0.147 | 0.354 | 0 | 1 |
| Affluent HS | 0.283 | 0.45 | 0 | 1 |
| Applied Anywhere | 0.861 | 0.346 | 0 | 1 |
| Applications | 1.417 | 1.142 | 0 | 5 |
| Admissions Offers | 1.085 | 0.998 | 0 | 5 |
| Aid Applications | 1.041 | 1.067 | 0 | 5 |
| Enrolled | 0.326 | 0.469 | 0 | 1 |
| Did Not Enroll in 4-year Institution | 0.14 | 0.347 | 0 | 1 |
| Not in Wave 2 | 0.534 | 0.499 | 0 | 1 |
| N | | 4143 | | |

Table S2: Sample Means: Applications

| | All Potential Applications | Applications | Offers | Matriculations |
|---|---|---|---|---|
| Distance (100mi) | 188.32 | 104.07 | 97.79 | 86.39 |
| Cost of Attendance ($10000) | 1.83 | 1.53 | 1.49 | 1.39 |
| Frac. minority | 0.20 | 0.28 | 0.28 | 0.29 |
| Frac. own race | 0.55 | 0.69 | 0.69 | 0.70 |
| 1(Top Ten) | 0.12 | 0.23 | 0.28 | 0.30 |
| Applied for aid | 0.13 | 0.86 | 0.91 | 0.93 |
| Applications | 0.18 | 1.20 | 1.22 | 1.20 |
| Admitted | 0.26 | 0.89 | 1.16 | 1.01 |
| In Wave 2 | 0.48 | 0.49 | 0.49 | 1.00 |
| Enrolled | 0.06 | 0.27 | 0.31 | 1.00 |
| *N* | 33,144 | 4,945 | 3,780 | 1,382 |

Table S3: Sample Balance, Wave 2 Participation

| Variable | Mean, Not wv2 | Increase if wv2 | P-value | N |
|---|---|---|---|---|
| Female | 0.57 | 0.01 | 0.54 | 4,143 |
| Black | 0.04 | 0.09*** | 0.00 | 4,143 |
| Hispanic | 0.20 | -0.02 | 0.15 | 4,143 |
| White | 0.67 | -0.12*** | 0.00 | 4,143 |
| Asian | 0.02 | 0.07*** | 0.00 | 4,143 |
| SAT | 0.67 | 0.01** | 0.05 | 4,143 |
| SAT / HS mean SAT | 0.98 | 0.01*** | 0.01 | 4,143 |
| HS poverty rate | 0.32 | 0.01* | 0.06 | 4,143 |
| Apply Anywhere | 0.85 | 0.03*** | 0.00 | 4,143 |
| Num. Apps | 2.37 | 0.20*** | 0.00 | 4,143 |
| Apply to UT-A | 0.14 | 0.04*** | 0.00 | 4,143 |
| Apply for aid at UT-A | 0.09 | 0.04*** | 0.00 | 4,143 |
| Admitted to UT-A | 0.12 | 0.04*** | 0.00 | 4,143 |

Note: this table shows means among surveyed students not participating in Survey Wave 2, and differences in means between Wave 2 and non-Wave 2 students. Wave 1 population weights are used. Oversampling of Black students in Wave 2 reflects an intentional choice on the part of the survey designers. *** indicates $p < 0.01$. ** indicates $p < 0.05$. * indicates $p < 0.1$.

Table S4: Sample Balance, Top-Decile Status

| Variable | Mean, Not ttt | Increase if ttt | P-value | N |
|---|---|---|---|---|
| Female | 0.54 | 0.06*** | 0.00 | 4,143 |
| Black | 0.09 | -0.05*** | 0.00 | 4,143 |
| Hispanic | 0.19 | 0.01 | 0.43 | 4,143 |
| White | 0.63 | -0.01 | 0.45 | 4,143 |
| Asian | 0.03 | 0.06*** | 0.00 | 4,143 |
| SAT | 0.64 | 0.10*** | 0.00 | 4,143 |
| SAT / HS mean SAT | 0.94 | 0.16*** | 0.00 | 4,143 |
| HS poverty rate | 0.31 | 0.05*** | 0.00 | 4,143 |
| Apply Anywhere | 0.82 | 0.11*** | 0.00 | 4,143 |
| Num. Apps | 2.34 | 0.32*** | 0.00 | 4,143 |
| Apply to UT-A | 0.11 | 0.19*** | 0.00 | 4,143 |
| Apply for aid at UT-A | 0.06 | 0.17*** | 0.00 | 4,143 |
| Admitted to UT-A | 0.07 | 0.22*** | 0.00 | 4,143 |

Note: this table shows means among surveyed students not placing in the top decile of their class, and differences in means between top-decile and non-top-decile students. Wave 1 population weights are used. *** indicates $p < 0.01$. ** indicates $p < 0.05$. * indicates $p < 0.1$.

## F.3 Data Supplement

### Calculating EFC

**Matching THEOP and CPS datasets** In order to draw incomes from the CPS, I recode parents' education in the THEOP survey to match the CPS dataset. I then draw incomes for each student in THEOP from the distribution of income given parents' occupation and the education of the most-educated parent.

1. THEOP senior survey

    - Mother's education: q71
    - Father's education: q67
    - `label list q71`
      `q71:`
      `1 No Schooling`
      `2 Elementary School`

Table S5: Sample Balance, LOS Participation

| Variable | Mean, Not longhorn | Increase if longhorn | P-value | N |
|---|---|---|---|---|
| Female | 0.56 | -0.01 | 0.63 | 4,143 |
| Black | 0.06 | 0.28*** | 0.00 | 4,143 |
| Hispanic | 0.17 | 0.29*** | 0.00 | 4,143 |
| White | 0.67 | -0.55*** | 0.00 | 4,143 |
| Asian | 0.05 | -0.05*** | 0.00 | 4,143 |
| SAT | 0.68 | -0.09*** | 0.00 | 4,143 |
| SAT / HS mean SAT | 0.98 | 0.02** | 0.03 | 4,143 |
| HS poverty rate | 0.29 | 0.51*** | 0.00 | 4,143 |
| Apply Anywhere | 0.86 | -0.07*** | 0.00 | 4,143 |
| Num. Apps | 2.45 | -0.11 | 0.34 | 4,143 |
| Apply to UT-A | 0.16 | -0.06*** | 0.00 | 4,143 |
| Apply for aid at UT-A | 0.11 | -0.03 | 0.12 | 4,143 |
| Admitted to UT-A | 0.13 | -0.05*** | 0.00 | 4,143 |

Note: this table shows means among surveyed students whose high schools do not participate in the Longhorn Opportunity Scholars program, and differences in means between students whose schools participate and those who don't. Wave 1 population weights are used. *** indicates $p < 0.01$. ** indicates $p < 0.05$. * indicates $p < 0.1$.

Table S6: Fraction of admissions and matriculations from in-state applicants, UT Austin

| year | admitted (total) | enrolled (total) | % admitted (in-state) | (TX public) | % enrolled (in-state) | (TX public) |
|---|---|---|---|---|---|---|
| 1991 | 10403 | 5818 | 85.0 | 71.8 | 90.4 | 84.1 |
| 1992 | 9726 | 5613 | 85.0 | 71.5 | 90.9 | 84.2 |
| 1993 | 10085 | 5861 | 86.0 | 72.3 | 91.6 | 84.3 |
| 1994 | 10278 | 6156 | 84.0 | 74.3 | 89.6 | 81.7 |
| 1995 | 10506 | 6346 | 84.3 | 74.3 | 90.9 | 82.8 |
| 1996 | 11041 | 6381 | 81.7 | 71.7 | 90.1 | 81.7 |
| 1997 | 11352 | 7138 | 82.2 | 72.8 | 89.9 | 82.2 |
| 1998 | 10693 | 6833 | 84.9 | 75.7 | 92.0 | 83.2 |
| 1999 | 10990 | 7288 | 89.1 | 79.4 | 93.5 | 84.6 |
| 2000 | 13061 | 8118 | 87.5 | 77.8 | 93.1 | 84.7 |
| 2001 | 12564 | 7243 | 85.0 | 76.7 | 91.1 | 83.8 |
| 2002 | 14138 | 7868 | 86.3 | 77.7 | 90.9 | 83.6 |
| 2003 | 10820 | 6493 | 87.1 | 80.0 | 92.9 | 86.9 |

Note: this table shows the fraction of admissions and matriculations from in-state applicants and from in-state applicants who attended public schools at the University of Texas at Austin.

```
3 Some High School
4 High School Graduate
```

Table S7: Parameters 1/3

| | Parameter | Estimate | SE | p5 | p95 |
|---|---|---|---|---|---|
| 1 | $\underline{\pi}$ Eqbm. (In-State Public) | 0.373 | 0.354 | -0.039 | 1.104 |
| 2 | $\underline{\pi}$ Eqbm. (Private) | 1.626 | 0.44 | 1.182 | 2.405 |
| 3 | $\underline{\pi}$ Eqbm. (Out-of-State Public) | 0.717 | 0.363 | 0.328 | 1.54 |
| 4 | $\underline{\pi}$ Eqbm. (Relig.) | 0.901 | 0.377 | 0.433 | 1.654 |
| 5 | $\underline{\pi}$ Eqbm. (Elite Private) | 4.529 | 0.418 | 3.987 | 5.354 |
| 6 | $\underline{\pi}$ Eqbm. (Texas A&M) | 1.669 | 0.372 | 1.243 | 2.526 |
| 7 | $\underline{\pi}$ Eqbm. (UT Austin) | 1.573 | 0.38 | 1.153 | 2.423 |
| 8 | $\gamma$ (SAT) | 5.624 | 0.948 | 4.318 | 7.43 |
| 9 | $\gamma$ (Class Rank) | -2.435 | 0.197 | -2.73 | -2.11 |
| 10 | $\gamma$ (SAT/HS mean SAT) | -1.2 | 0.486 | -1.929 | -0.318 |
| 11 | $\gamma$ (Poverty) | 0.278 | 0.244 | -0.123 | 0.683 |
| 12 | $\gamma$ (URM) | 0.011 | 0.074 | -0.108 | 0.104 |
| 13 | $\gamma^s$ (Const) | -10.0 | 0.556 | -10.0 | -8.715 |
| 14 | $\gamma^s$ (Poverty) | 1.563 | 8.376 | -10.0 | 10.0 |
| 15 | $\gamma^s$ (URM) | 5.748 | 6.26 | -10.0 | 6.596 |
| 16 | $\gamma^{a\|s}$ (Const) | -0.692 | 0.353 | -1.217 | -0.31 |
| 17 | $\gamma^{a\|s}$ (Poverty) | 1.625 | 0.763 | 0.442 | 2.705 |
| 18 | $\gamma^{a\|s}$ (URM) | 0.545 | 0.305 | 0.145 | 1.045 |
| 19 | $\lambda$ (Matric. shock scale) | 10.0 | 0.0 | 10.0 | 10.0 |
| 20 | $\gamma^{fixed}$ (Const) | 0.31 | 0.091 | 0.19 | 0.44 |
| 21 | $\gamma^{fixed}$ (Poverty) | -0.239 | 0.111 | -0.361 | -0.062 |
| 22 | $\gamma^{fixed}$ (URM) | 0.047 | 0.06 | -0.077 | 0.122 |
| 23 | $\gamma^{var}$ (Const) | 1.038 | 0.053 | 0.95 | 1.099 |
| 24 | $\gamma^{var}$ (Poverty) | -0.219 | 0.119 | -0.46 | -0.064 |
| 25 | $\gamma^{var}$ (URM) | -0.176 | 0.045 | -0.225 | -0.098 |
| 26 | $\gamma^{shock}$ (Const) | -0.88 | 0.094 | -1.044 | -0.795 |
| 27 | $\gamma^{shock}$ (Poverty) | -0.592 | 0.21 | -0.965 | -0.272 |
| 28 | $\gamma^{shock}$ (URM) | -0.075 | 0.079 | -0.168 | 0.088 |
| 29 | $\beta^p$ (Const) | 2.808 | 0.555 | 2.57 | 4.309 |
| 30 | $\beta^p$ (income) | -0.422 | 0.055 | -0.534 | -0.369 |
| 31 | $\beta^{(w,x,z)}$ (In-State Public) | 3.058 | 0.398 | 2.876 | 4.034 |
| 32 | $\beta^{(w,x,z)}$ (Private) | -0.898 | 1.03 | -1.442 | 0.798 |

Note: This table shows all estimated parameters (Part 1 of 3).

```
5 Some College
6 Two-Year College
7 Four-Year College
8 Master's Degree
9 Professional Degree
```

2. 2002 March CPS

- Education of household head
- `label list EDUC_HEAD`

Table S8: Parameters 2/3

| | Parameter | Estimate | SE | p5 | p95 |
|---|---|---|---|---|---|
| 33 | $\beta^{(w,x,z)}$ (Out-of-State Public) | 0.937 | 0.583 | 0.493 | 2.112 |
| 34 | $\beta^{(w,x,z)}$ (Relig.) | 2.822 | 0.613 | 2.076 | 4.049 |
| 35 | $\beta^{(w,x,z)}$ (Elite Private) | 1.798 | 2.688 | -0.611 | 5.286 |
| 36 | $\beta^{(w,x,z)}$ (Texas A&M) | 2.888 | 0.605 | 2.362 | 4.175 |
| 37 | $\beta^{(w,x,z)}$ (UT Austin) | 0.208 | 0.707 | -0.422 | 1.898 |
| 38 | $\beta^{(w,x,z)}$ (Dist < 25) | 0.17 | 0.07 | 0.053 | 0.27 |
| 39 | $\beta^{(w,x,z)}$ (Distance) | -0.184 | 0.028 | -0.229 | -0.14 |
| 40 | $\beta^{(w,x,z)}$ (URM) | -0.108 | 0.133 | -0.346 | 0.101 |
| 41 | $\beta^{(w,x,z)}$ (URM X UTA) | 0.127 | 0.095 | -0.011 | 0.268 |
| 42 | $\beta^{(w,x,z)}$ (URM X TAMU) | -0.367 | 0.19 | -0.732 | -0.148 |
| 43 | $\beta^{(w,x,z)}$ (Poverty) | -0.7 | 0.397 | -1.374 | -0.147 |
| 44 | $\beta^{(w,x,z)}$ (Poverty X UTA) | -0.354 | 0.188 | -0.745 | -0.174 |
| 45 | $\beta^{(w,x,z)}$ (Poverty X TAMU) | -0.708 | 0.353 | -1.401 | -0.423 |
| 46 | $\beta^{(w,x,z)}$ (LOS X UTA) | 0.099 | 0.175 | -0.217 | 0.289 |
| 47 | $\beta^{(w,x,z)}$ (Century X TAMU) | -0.203 | 0.3 | -0.763 | 0.1 |
| 48 | $\beta^{(w,x,z)}$ (SAT) | -1.141 | 0.647 | -2.047 | -0.277 |
| 49 | $\beta^{(w,x,z)}$ (Class Rank) | -0.098 | 0.545 | -0.926 | 0.683 |
| 50 | $\beta^{(w,x,z)}$ (SAT / HS Mean SAT) | -0.222 | 0.377 | -0.702 | 0.338 |
| 51 | $\beta^{(w,x,z)}$ (SAT X TAMU) | 1.296 | 0.532 | 0.094 | 1.782 |
| 52 | $\beta^{(w,x,z)}$ (SAT X UTA) | 4.485 | 0.637 | 3.186 | 5.174 |
| 53 | $\beta^{(w,x,z)}$ (SAT X Private) | 2.502 | 0.533 | 1.617 | 3.251 |
| 54 | $\log(\sigma^{rc})$ (Distance) | -0.61 | 3.104 | -10.0 | -0.249 |
| 55 | $\log(\sigma^{rc})$ (S/F Ratio) | -10.0 | 0.005 | -10.0 | -10.0 |
| 56 | $\log(\sigma^{rc})$ (UTA vs TAMU) | -10.0 | 1.529 | -10.0 | -10.0 |
| 57 | $\beta^{aware}$ (In-State Public) | 4.014 | 2.686 | -1.403 | 5.895 |
| 58 | $\beta^{aware}$ (Private) | -0.261 | 3.455 | -7.282 | 0.251 |
| 59 | $\beta^{aware}$ (Out-of-State Public) | 0.832 | 2.873 | -5.287 | 2.077 |
| 60 | $\beta^{aware}$ (Relig.) | 3.542 | 2.549 | -1.723 | 5.399 |
| 61 | $\beta^{aware}$ (Elite Private) | -1.675 | 4.038 | -10.0 | 0.208 |
| 62 | $\beta^{aware}$ (Texas A&M) | -3.733 | 3.42 | -10.0 | -0.113 |
| 63 | $\beta^{aware}$ (UT Austin) | 1.023 | 3.602 | -9.453 | 0.791 |
| 64 | $\beta^{aware}$ (Dist < 25) | -0.421 | 0.706 | -0.884 | 0.886 |

Note: This table shows all estimated parameters (Part 2 of 3).

```
EDUC_HEAD:
0 NIU or no schooling
1 niu
2 None or preschool
10 Grades 1, 2, 3, or 4
11 Grade 1
12 Grade 2
13 Grade 3
14 Grade 4
```

## Table S9: Parameters 3/3

| | Parameter | Estimate | SE | p5 | p95 |
|---|---|---|---|---|---|
| 65 | $\beta^{aware}$ (Distance) | -0.05 | 0.194 | -0.334 | 0.157 |
| 66 | $\beta^{aware}$ (URM) | 0.779 | 0.55 | 0.327 | 1.797 |
| 67 | $\beta^{aware}$ (URM X UTA) | 1.168 | 1.15 | -0.342 | 3.156 |
| 68 | $\beta^{aware}$ (URM X TAMU) | -0.404 | 1.806 | -1.535 | 3.772 |
| 69 | $\beta^{aware}$ (Poverty) | 2.319 | 1.111 | 1.683 | 5.004 |
| 70 | $\beta^{aware}$ (Poverty X UTA) | 0.749 | 3.027 | -1.792 | 7.566 |
| 71 | $\beta^{aware}$ (Poverty X TAMU) | 7.591 | 2.649 | 3.327 | 10.0 |
| 72 | $\beta^{aware}$ (LOS X UTA) | 10.0 | 1.853 | 5.345 | 10.0 |
| 73 | $\beta^{aware}$ (Century X TAMU) | -3.785 | 3.652 | -9.571 | 1.994 |
| 74 | $\beta^{aware}$ (SAT) | 3.394 | 3.351 | 0.643 | 10.0 |
| 75 | $\beta^{aware}$ (Class Rank) | -10.0 | 0.0 | -10.0 | -10.0 |
| 76 | $\beta^{aware}$ (SAT / HS Mean SAT) | -2.208 | 1.664 | -3.99 | 0.722 |
| 77 | $\beta^{aware}$ (SAT X TAMU) | 7.162 | 3.461 | 0.963 | 10.0 |
| 78 | $\beta^{aware}$ (SAT X UTA) | 1.775 | 4.757 | -2.407 | 10.0 |
| 79 | $\beta^{aware}$ (SAT X Private) | 6.043 | 2.662 | 2.119 | 10.0 |
| 80 | $\beta^{aware}$ (income) | -0.066 | 0.112 | -0.155 | 0.233 |
| 81 | $\log(\sigma)(\beta_i^0)$ | 0.667 | 0.219 | 0.711 | 1.492 |
| 82 | $\alpha^{aid}$ (In-State Public) | -0.505 | 0.128 | -0.604 | -0.207 |
| 83 | $\alpha^{aid}$ (Private) | -0.332 | 0.077 | -0.423 | -0.187 |
| 84 | $\alpha^{aid}$ (Out-of-State Public) | -0.175 | 0.039 | -0.181 | -0.079 |
| 85 | $\alpha^{aid}$ (Relig.) | -0.461 | 0.084 | -0.497 | -0.252 |
| 86 | $\alpha^{aid}$ (Elite Private) | -0.55 | 0.195 | -0.817 | -0.297 |
| 87 | $\alpha^{aid}$ (Texas A&M) | -1.044 | 0.245 | -1.246 | -0.514 |
| 88 | $\alpha^{aid}$ (UT Austin) | -0.405 | 0.123 | -0.437 | -0.094 |
| 89 | $\log(\sigma^e)$ (In-State Public) | -10.0 | 1.646 | -10.0 | -10.0 |
| 90 | $\log(\sigma^e)$ (Private) | 0.732 | 0.242 | 0.274 | 1.087 |
| 91 | $\log(\sigma^e)$ (Out-of-State Public) | -0.309 | 2.665 | -8.246 | -0.046 |
| 92 | $\log(\sigma^e)$ (Relig.) | -0.262 | 2.763 | -8.744 | 0.288 |
| 93 | $\log(\sigma^e)$ (Elite Private) | 1.699 | 0.503 | 0.864 | 2.321 |
| 94 | $\log(\sigma^e)$ (Texas A&M) | -1.075 | 4.125 | -10.0 | -0.13 |
| 95 | $\log(\sigma^e)$ (UT Austin) | -0.578 | 3.074 | -10.0 | -0.187 |

Note: This table shows all estimated parameters (Part 3 of 3).

```
20 Grades 5 or 6
21 Grade 5
22 Grade 6
30 Grades 7 or 8
31 Grade 7
32 Grade 8
40 Grade 9
50 Grade 10
60 Grade 11
70 Grade 12
```

```
71 12th grade, no diploma
72 12th grade, diploma unclear
73 High school diploma or equivalent
80 1 year of college
81 Some college but no degree
90 2 years of college
91 Associate's degree, occupational/vocational program
92 Associate's degree, academic program
100 3 years of college
110 4 years of college
111 Bachelor's degree
120 5+ years of college
121 5 years of college
122 6+ years of college
123 Master's degree
124 Professional school degree
125 Doctorate degree
999 Missing/Unknown
```

3. Matching procedure:

   - EDUC_HEAD $\in \{0,1\}$: $Ed = 1$
   - EDUC_HEAD $\in \{10,...,32\}$: $Ed = 2$
   - EDUC_HEAD $\in \{40,...,71\} \mapsto Ed = 3$.
     - Note that in the 2002 March CPS, TX subsample there are no household heads with unclear graduation status EDUC_HEAD=72.

   EDUC_HEAD $\in \{73\} \mapsto Ed = 4$.

   - EDUC_HEAD $\in \{81\} \mapsto Ed = 5$.
   - EDUC_HEAD $\in \{91,92\} \mapsto Ed = 6$.
   - EDUC_HEAD $\in \{111\} \mapsto Ed = 7$.
   - EDUC_HEAD $\in \{123\} \mapsto Ed = 8$.
   - EDUC_HEAD $\in \{124\} \mapsto Ed = 9$.

4. Administrative data: UT Austin codes parents' education similarly to the THEOP survey.

### Aggregate colleges

In this section I provide lists of the colleges and universities that make up each aggregate college. The data use agreement requires me to aggregate each college that has fewer than ten applicants in the data.

| Institution | Sum | |
|---|---|---|
| | apply | admit |
| ANGELO STATE UNIVERSITY | 148 | 119 |
| Institutions with ten or fewer apps | 10 | 9 |
| LAMAR UNIVERSITY-BEAUMONT | 80 | 61 |
| MIDWESTERN STATE UNIVERSITY | 30 | 24 |
| PRAIRIE VIEW A & M UNIVERSITY | 95 | 49 |
| SAM HOUSTON STATE UNIVERSITY | 187 | 143 |
| SOUTHWEST TEXAS STATE UNIVERSITY | 329 | 266 |
| STEPHEN F AUSTIN STATE UNIVERSITY | 282 | 227 |
| SUL ROSS STATE UNIVERSITY | 27 | 15 |
| TARLETON STATE UNIVERSITY | 71 | 49 |
| TEXAS A & M INTERNATIONAL UNIVERSITY | 56 | 37 |
| TEXAS A & M UNIVERSITY-CORPUS CHRISTI | 55 | 43 |
| TEXAS A & M UNIVERSITY-GALVESTON | 26 | 21 |
| TEXAS A & M UNIVERSITY-KINGSVILLE | 77 | 61 |
| TEXAS A&M UNIVERSITY-COMMERCE | 47 | 30 |
| TEXAS SOUTHERN UNIVERSITY | 125 | 79 |
| TEXAS TECH UNIVERSITY | 425 | 360 |
| TEXAS WOMAN'S UNIVERSITY | 38 | 27 |
| THE UNIVERSITY OF TEXAS AT ARLINGTON | 170 | 120 |
| THE UNIVERSITY OF TEXAS AT BROWNSVILLE | 62 | 40 |
| THE UNIVERSITY OF TEXAS AT DALLAS | 86 | 61 |
| THE UNIVERSITY OF TEXAS AT EL PASO | 275 | 183 |
| THE UNIVERSITY OF TEXAS AT SAN ANTONIO | 166 | 121 |
| THE UNIVERSITY OF TEXAS AT TYLER | 27 | 17 |
| THE UNIVERSITY OF TEXAS OF THE PERMIAN BASIN | 30 | 22 |
| THE UNIVERSITY OF TEXAS-PAN AMERICAN | 194 | 165 |
| UNIVERSITY OF HOUSTON-DOWNTOWN | 37 | 18 |
| UNIVERSITY OF HOUSTON-UNIVERSITY PARK | 470 | 305 |
| UNIVERSITY OF NORTH TEXAS | 317 | 231 |
| WEST TEXAS A & M UNIVERSITY | 42 | 37 |
| Total | 3,984 | 2,940 |

Colleges comprising OTHER TX PUBLIC 4-YEAR

| Institution | Sum | |
|---|---|---|
| | apply | admit |
| EMBRY-RIDDLE AERONAUTICAL UNIVERSITY | 12 | 10 |
| Institutions with ten or fewer apps | 235 | 190 |
| NEW YORK UNIVERSITY | 29 | 21 |
| NORTHWOOD UNIVERSITY | 21 | 17 |
| TULANE UNIVERSITY OF LOUISIANA | 17 | 16 |
| UNIVERSITY OF SOUTHERN CALIFORNIA | 23 | 22 |
| VANDERBILT UNIVERSITY | 12 | 11 |
| Total | 349 | 287 |

Colleges comprising PRIVATE NONRELIGIOUS

| Institution | Sum | |
| --- | --- | --- |
| | apply | admit |
| ABILENE CHRISTIAN UNIVERSITY | 60 | 53 |
| AUSTIN COLLEGE | 16 | 15 |
| BAYLOR UNIVERSITY | 242 | 213 |
| BRIGHAM YOUNG UNIVERSITY | 28 | 24 |
| DILLARD UNIVERSITY | 18 | 14 |
| EAST TEXAS BAPTIST UNIVERSITY | 29 | 21 |
| HARDIN-SIMMONS UNIVERSITY | 21 | 19 |
| HOUSTON BAPTIST UNIVERSITY | 41 | 29 |
| HOWARD PAYNE UNIVERSITY | 27 | 19 |
| Institutions with 15 or fewer apps | 383 | 310 |
| LUBBOCK CHRISTIAN UNIVERSITY | 17 | 14 |
| MCMURRY UNIVERSITY | 19 | 17 |
| OUR LADY OF THE LAKE UNIVERSITY-SAN ANTONIO | 53 | 42 |
| SAINT EDWARDS UNIVERSITY | 35 | 29 |
| SOUTHERN METHODIST UNIVERSITY | 58 | 50 |
| SOUTHWESTERN UNIVERSITY | 33 | 32 |
| ST MARYS UNIVERSITY | 59 | 47 |
| TEXAS CHRISTIAN UNIVERSITY | 126 | 106 |
| TRINITY UNIVERSITY | 40 | 39 |
| UNIVERSITY OF MARY HARDIN BAYLOR | 19 | 13 |
| UNIVERSITY OF SAINT THOMAS | 28 | 23 |
| UNIVERSITY OF THE INCARNATE WORD | 25 | 23 |
| XAVIER UNIVERSITY OF LOUISIANA | 20 | 16 |
| Total | 1,397 | 1,168 |

Colleges comprising RELIGIOUS

| Institution | Sum | |
| --- | --- | --- |
| | apply | admit |
| ARIZONA STATE UNIVERSITY-MAIN CAMPUS | 30 | 27 |
| COLORADO STATE UNIVERSITY | 14 | 12 |
| FLORIDA AGRICULTURAL AND MECHANICAL UNIVERSITY | 11 | 10 |
| FLORIDA STATE UNIVERSITY | 14 | 12 |
| GRAMBLING STATE UNIVERSITY | 11 | 6 |
| Institutions with ten or fewer apps | 421 | 346 |
| KANSAS STATE UNIVERSITY OF AGRICULTURE AND APP SCI | 11 | 9 |
| LOUISIANA ST UNIV & AGRL & MECH & HEBERT LAWS CTR | 64 | 51 |
| NEW MEXICO STATE UNIVERSITY-MAIN CAMPUS | 102 | 76 |
| OKLAHOMA STATE UNIVERSITY-MAIN CAMPUS | 45 | 42 |
| PURDUE UNIVERSITY-MAIN CAMPUS | 12 | 12 |
| SOUTHERN UNIVERSITY-NEW ORLEANS | 10 | 3 |
| THE UNIVERSITY OF ALABAMA | 13 | 12 |
| UNITED STATES AIR FORCE ACADEMY | 13 | 12 |
| UNITED STATES NAVAL ACADEMY | 11 | 8 |
| UNIVERSITY OF CALIFORNIA-BERKELEY | 16 | 14 |
| UNIVERSITY OF COLORADO AT BOULDER | 15 | 15 |
| UNIVERSITY OF GEORGIA | 12 | 10 |
| UNIVERSITY OF MISSISSIPPI MAIN CAMPUS | 11 | 10 |
| UNIVERSITY OF NEW MEXICO-MAIN CAMPUS | 17 | 9 |
| UNIVERSITY OF OKLAHOMA NORMAN CAMPUS | 36 | 32 |
| Total | 889 | 728 |

Colleges comprising NON-TX PUBLIC

| Institution | Sum | |
|---|---|---|
| | apply | admit |
| HARVARD UNIVERSITY | 16 | 12 |
| Institutions with ten or fewer apps | 55 | 41 |
| MASSACHUSETTS INSTITUTE OF TECHNOLOGY | 12 | 11 |
| NORTHWESTERN UNIVERSITY | 13 | 12 |
| PRINCETON UNIVERSITY | 14 | 10 |
| RICE UNIVERSITY | 52 | 44 |
| STANFORD UNIVERSITY | 19 | 17 |
| WASHINGTON UNIVERSITY | 19 | 17 |
| YALE UNIVERSITY | 11 | 7 |
| Total | 211 | 171 |

Colleges comprising SELECTIVE PRIVATE