Inequity in Centralized College Admissions with Public and Private Universities: Evidence from Albania*

Iris Vrioni†

Job Market Paper

November 8, 2023

Updated frequently, please click here for latest version.

Abstract

Centralized assignment systems are a popular policy tool to improve fairness and efficiency in allocating students to public college seats. In most implementations, however, private college admissions remain decentralized, which may give high socio-economic status (SES) students a strategic advantage in the centralized public match because high-SES students derive higher value from expensive private alternatives. I empirically study application behavior and the allocation of students in markets where only public college seats are centrally assigned with new data from the college match in Albania. Using a policy change that incorporated all private colleges in the centralized platform, which differentially shifted outside alternatives by SES, I find that when private colleges operate outside the match, high-SES students apply to more selective portfolios and enroll in more selective public programs, but the selectivity gap in applications shrinks after the policy change. I build and estimate a model of applications and matriculation that uses the unique institutional features of the Albanian college admissions to disentangle the effects of heterogeneous beliefs, preferences, and outside options on choice, and evaluate the distributional consequences of counterfactual admissions design. I find that removing outside options reverses the welfare gap in favor of lower-SES students, but at the expense of overall market efficiency. This is driven by the fact that outside options dampen the distortionary effects of list size restrictions and incorrect beliefs on choice.

*I am incredibly grateful to my advisors Basit Zafar, Ying Fan, Zach Brown, and Kevin Stange for their invaluable guidance and support. This paper has greatly benefited from conversations with John Bound, Adam Kapor, Victoria Barone, Emir Murathanoglu, Hanna Onyshchenko, Nikhil Rao, Brock Rowberry, Carolina Santos, Agostina Brinatti and many participants in the University of Michigan Labor Seminar, IO and Labor Lunches, UM-MSU-Western Labour Day Conference, and Warwick PhD Conference. I am immensely grateful to Arjan Xhelaj and RASH for providing me with the data and institutional insight. All errors are my own.

†Department of Economics, University of Michigan. Email: ivrioni@umich.edu.

1 Introduction

The importance of fairness and efficiency in assigning students to scarce public school or university seats has led to an increase in the practice of centralized admissions at all levels of education across the world. In higher education alone, every year, over 20 million students are matched to colleges and majors through centralized mechanisms.[1]

Two prominent features of most implementations of centralized mechanisms are: (1) list size restrictions and (2) the exclusion of some institutions, generally private ones, from the match. List size restrictions limit the number of programs one can apply to through the mechanism and force students to weigh their preferences and ambitions against the chances of admission, inducing them to strategize on what programs to include in their application. At the same time, private universities manage their admissions outside of the mechanism and can serve as much as 75% of the market.[2] In this context, for applications within the match, it may matter strategically what outside options a student has. Those with better outside options can apply more ambitiously within the mechanism and may ultimately be assigned to more preferred programs than students with worse outside options. This strategic response to market structure may have both efficiency and equity consequences. If outside alternatives, which are often expensive, are more desirable for high-SES students, it gives them not only higher direct value from choosing these options, but also the ability to take more risk within the match.

This paper assesses the importance of market structure, in particular the extent of centralization, on strategic applications in centralized assignment systems, focusing on Deferred Acceptance (DA) mechanisms with list size restrictions. With data from a market that changed its structure from partial centralization with a public match and private decentralized admissions to a fully centralized system, I build a model of student applications and enrollment decisions to college-major pairs ("programs") on the platform. The model and estimation take advantage of the unique features of the setting, which allow separate identification of student beliefs about chances of admission and their preferences for programs. The level of uncertainty and bias about the expected selectivity of programs are central in determining how constraining the list size restrictions are. In addition, preferences for inside

[1] Around 10 million students participate in the college match in China (Chen and Kesten 2017), more than 2.5 million in Brazil (Otero et al. 2021), and millions more in Chile, Germany, India, Kenya, and Turkey, among others.

[2] In Brazil, more than 75% of students enrolled in a college degree attend a private institution. In my empirical setting, 27% of students are enrolled in private universities.

and outside options determine the extent to which it is important to students to make sure they are admitted to an acceptable program in the match rather than take a chance with their favorite programs. I quantify these elements and assess choice, resulting allocations and welfare in commonly observed market structures.

My empirical setting is the centralized admissions system in Albania, which underwent a unique change that incorporated all private colleges into the centralized admissions process. Prior to 2016, the Albanian college admissions were representative of common partially centralized systems with a market that was partitioned into a public match with decentralized private college admissions that happened roughly at the same time. Assignments to public universities were mediated by a clearinghouse, which took in stated student preferences for programs as well as program priorities.[3] Student preferences were reported through rank order lists restricted to 10 or fewer programs, and the clearinghouse produced matches through a standard DA algorithm. In 2016, a major higher education reform changed admissions in two significant ways. First, all private colleges got incorporated into the centralized match with no programs allowed to conduct admissions outside of the national match. This expanded options on the national platform by over 50% with list size restrictions remaining the same as before the reform, and any possibility of enrolling in college outside of the match was eliminated. Second, the reform changed the assignment procedure to a live multi-offer mechanism with 7 phases in its main round. Students submit their application lists to the clearinghouse and programs submit their priorities. Then, as a program-proposing Gale-Shapley algorithm would begin, in the first phase, programs make initial proposals to students ranked at the top. Under the new procedures, students observe all offers and decide within 48 hours whether to enroll in any of the proposing programs or forgo all first phase offers and wait for a better offer in the next phase. The multi-offer phases continue until the last phase or until each school has filled its seats.

This setting is attractive because it allows me to overcome two key challenges that have so far prevented a clear answer on the extent to which outside options matter for strategic applications in centralized mechanisms. The first is that of observing choice over in- and out-of-match programs in order to infer preferences for private programs. Data on applications and enrollments are generally only available for programs on the match, which makes it impossible to understand the value of outside options to students. I overcome this challenge with data from the post-reform period in which choices to apply and enroll in all programs

[3] Program priorities come in the from of weights for GPA and end-of-high school exams that the clearinghouse uses to produce weighted average scores for each student applying to each program.

are made on the platform.

The second and more complex challenge is one of identification. In a setting with constrained applications, both preferences for programs and perceived probabilities of admission play a crucial role in determining the extent to which market structure affects on-platform applications. It is necessary to separate preferences from student beliefs about probabilities of admission. If students have perfect information about which schools on the match they would be admitted to, there is no scope for strategic applications in general and no role for outside options to affect applications through the strategic channel in particular. The importance of private outside options is tied to the level of uncertainty and bias that students have about their chances of admission to centralized programs. In most centralized settings the only observed choice is the selection of application lists, which are insufficient to separately identify preferences from beliefs about admission chances (Agarwal and Somaini 2020). The school choice literature often makes strict assumptions about belief formation where agents have rational expectations about program cutoffs (Agarwal and Somaini 2018; Idoux 2022). This assumption is inadequate to evaluate the question because it drastically limits the ability of the model to explain choices as arising from a strategic channel.[4] I instead allow beliefs to depart form rational expectations and capture the uncertainty and bias in the market in a reduced form. I overcome the identification issue by taking advantage of the post-reform mechanism, which allowed multiple offers at the admissions phase. Students observe which programs have admitted them, and many are admitted to 2 or more programs in their choice set. The choices to enroll and which program students enroll in pins down preferences, and the application portfolio choice and decisions to wait for future phases can be exploited to identify beliefs about probabilities of admission.

Using rich data on applications and enrollments in years 2013-2019, I first provide descriptive evidence that when private options are available outside the centralized match, high-SES students apply to and enroll in more selective public programs than their lower-SES peers with the same high school and exam performance. High-SES students are also more likely to end up without an assignment in the public match. These are striking facts because these differences cannot be explained by geographic access to more selective programs for high-SES students or lack of affordability of selective public programs given that public institutions are tuition-free.

[4] Non-degenerate beliefs about distribution of cutoffs from which perceived probabilities of admission arise exist only because of sampling variation and hence probabilities of admission depend only on student preferences and the size of the market, holding the mechanism constant.

I then analyze the policy change that enforced participation by all colleges in the platform. I use an event study design to measure the effects of centralizing all available alternatives on application behavior. Comparing applications of high-SES students to lower-SES students before and after the reform, I find that high-SES students change their applications more: they reduce the number of public programs they apply to by 1.2 more than lower-SES students. In addition, the selectivity gap in applications between the two groups shrinks after the policy change, both for the overall portfolio and for the reach programs. The shrinkage is driven by high-SES students decreasing the selectivity of their applications by more than lower-SES students do. Finally, the variance in the selectivity of public programs in the applications declines, indicating that students are giving up more selective programs rather than shifting the entire application toward less selective programs.

Motivated by the reduced-form results, I quantify the welfare and distributional impacts of a partitioned market structure. Based on the features of the post-reform period, I build a structural model of student decision to apply to college, application portfolio selection, and enrollment and waiting decisions on the waitlist. In the model, students take national exams and after observing their score, decide whether and where to apply to college. Each graduating high-schooler applies through the match if there is at least one program on the platform that they prefer to their outside option. Crucially, applications are allowed to be strategic: students are allowed to prefer more than ten programs to their outside options, but are only allowed to apply to ten, which induces them to exclude certain programs in such a way that the resulting portfolio maximizes their expected utility from the lottery over outcomes induced by the application portfolio. Portfolios are constructed as in Chade and Smith (2006)—students understand that the marginal value of each option included in the portfolio depends not only on the probability of admission to this program and the value of attending it, but also on the admission probabilities and value from all the other choices on the portfolio. Therefore portfolios are chosen as an optimization problem over all possible lotteries induced by portfolios of size ten.

Students form beliefs over probabilities of admission that depend only on the final cutoff at the last phase of admission and I assume they disregard the distribution of possible cutoffs in intermediate stages of the mechanism. This assumption is consistent with the information available to students at the time of application and alleviates the intractability problem that arises from the fact that each portfolio choice induces a distribution of waitlist states in each of the 7 rounds of these dynamic admissions.[5] I model beliefs over cutoffs as based on the previous year's cutoff for that program. Students expect a mean shift from the previous year's cutoff and have uncertainty over the realization of cutoffs, which I model as a distribution in possible cutoffs that is centered around the shifted previous year's threshold. The scale of distribution is allowed to be heterogeneous by SES group and by program selectivity. By parameterizing beliefs as a normal distribution over program cutoffs with a shift and a scale parameter, I capture a very complicated multi-dimensional object with data and a small set of parameters.

I assume that choices on the first phase of admissions are made with the same information and preferences as those in the application stage. While it is possible to model learning in this context, I find that differences in the first-phase cutoffs from the previous year do not predict the likelihood of students accepting first-phase offers or the likelihood of waiting for the next phase. At this stage, applicants observe offers and choose whether to accept a given offer, wait for the next phase, or exit the mechanism unmatched. Choices to enroll in programs in this phase offer the main source of identification for student preferences for programs.

I estimate the model by Simulated Maximum Likelihood. The model does not generate a closed form expression for the likelihood of the sequence of observed student actions. In particular, a major challenge in estimation is computing the conditional likelihood of the chosen portfolio being optimal. Given the large number of available options, it is infeasible to compute the probability of an observed portfolio being optimal among those in the vast choice set. I use a method developed in Larroucau and Rios (2020) which derives, for the portfolio problem of the Chade and Smith (2006) type, a small sufficient set of deviations from the observed portfolio that need to be checked for optimality. This allows for both the tractable simulation of choice sets at the application stage and an estimation routine that maximizes a likelihood function that is not prohibitively flat.

[5] The only paper to model a multi-stage mechanism that resembles the one in my setting is Waldinger (2021). In that problem, a two-stage mechanism of housing development choice is modeled in which the first stage consists of applications of size up to 3 among 18 possible choices and each portfolio generates a distribution of possible waitlist positions in the second stage, each of which generates a distribution waiting times for chosen developments. By contrast, students in my setting choose 10 options among over 500 in the first stage which induces a distribution of waitlist positions in the first admission stage, and from there, each preceding vector of waitlist states induces a distribution of waitlist states in the following phase. Accounting for these dynamic considerations becomes quickly intractable. The platform authority itself avoids distributing information about intermediate program cutoffs at the time of application with the goal of discouraging students from considering intermediate stages.

Estimates of the model imply that students face significant uncertainty about their admission probabilities in this market. Both high and lower-SES students perceive on average a lower probability of being accepted to each program than they would have with perceived distributions centered at the previous year's cutoffs. The mean of the belief distribution for high-SES students is shifted further up the range of cutoffs implying they are slightly more optimistic for lower-cutoff programs, but the slope of the mean is less steep than for the lower-SES group. These estimates suggest that student information is far from perfect. It is crucial then to evaluate both the role of application constraints and how students' outside options interact with the platform application constraints.

With estimated taste and belief parameters, I conduct counterfactual analyses that evaluate the role of market structure on applicant behavior, allocations of students to programs and welfare. I assess a centralizing policy change (an "all-in" structure) from a market where all outside alternatives are private (a "partitioned" structure). In the partitioned structure, both the centralized and decentralized admissions operate simultaneously but separately in a single-stage application with DA assignment. In reality, decentralized markets suffer from congestion and matching frictions which can affect assignments in both the public match and the private market (Abdulkadiroğlu et al. 2017; Kapor et al. 2022), but the scope of this paper is not to assess such frictions.[6] The assumption of an unrestricted-list DA mechanism for the private market serves to abstract away from these frictions to focus on the effect of the strategic channel for platform applications.

I find that centralization with list-size restrictions reduces enrollment by 4.2pp (1.7pp) for lower (high)-SES students relative to partial centralization. The allocation into private and public programs changes too. For the students that go to college, assignments worsen for some and improve for others with net losses for 2.9% of high-SES students and net wins for 1.1% of lower-SES students. Welfare calculations also imply a net loss in the market, which accumulates mostly in the high-SES group, leading to slightly improved equity at a high cost of efficiency. I measure welfare relative to the gains possible under an unrestricted DA mechanism and find that an all-in policy increases the welfare gap relative to the unrestricted DA by €160 for high-SES students and €98 for lower-SES students.

I investigate the channels next. First, the direct response from a less valuable outside

[6] Other work has studied these frictions. Most recently, Kapor et al. (2022) find frictions that come from chains of on-platform offer rejections in favor of off-platform options that lead to vacancies in platform programs or mismatches as platform programs try to contact individuals that were initially rejected.

option induces some students to exclude a public program they would have been marginally admitted to and would have preferred relative to the assignment they get. This accounts for a minority of those with worse outcomes. Second, with outside options incorporated in the mechanism, there are more programs to choose from in addition to a lower-valued outside option for many, which induces some students to reduce the number of public programs in their list in order to accommodate private programs. This channel accounts for the majority of the losers. Third, programs' capacity constraints generate spillovers from those who change their application through the first two channels. Some students do not change their applications and get pushed to less preferred programs because more have applied to their otherwise feasible program. The first two channels are much stronger for high-SES students, while the third affects everyone. The lessons from this decomposition imply that the constraint channel is far stronger than the outside-option channel in determining the effects of the market structure on applications and assignments of students. The inequity in observed outcomes with the partitioned structure is due to the fact that high-SES students behave almost as though unconstrained while constraints are more binding for lower-SES students. The all-in market structure forces binding list-size constraints on high-SES students, reducing inequity, but also efficiency.

Since the effect of market structure through the strategic channel is largely determined by constraints rather than outside options, I consider market designs that keep the all-in structure, but alleviate list-size restrictions. I find that an increase in list size of just 4 additional slots recovers more than half the losses from market structure change for lower-SES.

This paper contributes to the empirical literature on implementations of centralized assignment. A small body of work documents welfare implications of different aspects of common implementations of matching (Abdulkadiroğlu et al. 2017; Calsamiglia et al. 2020; Fack et al. 2019). Luflade (2017) uses a setting in Tunisia to show that imperfect information on probabilities of admission affects strategic applications. In the same vein, Ajayi and Sidibe (2020) use a setting in Ghana to estimate welfare effects of changing the allowed number of applications students can submit to the centralized mechanism. Most similarly to my setting, Kapor et al. (2022) use a centralized platform expansion in Chile to evaluate the welfare consequences of matching aftermarkets. My paper provides the first empirical evidence of the interaction between market structure and strategic applications.

My paper also contributes to the theoretical market design literature studying incentives generated by outside options in manipulable mechanisms as well as market structure. Akbarpour et al. (2021) formalize theoretical predictions for the effect of outside options on manipulable centralized mechanisms with an application to elementary school matching. Andersson et al. (2019) study the implications of a sequential public and private school matching mechanism. I add empirics to this largely theoretical literature to assess the practical importance of outside options in partially centralized settings.

Finally, I add to the literature studying educational decisions under imperfect information. Kapor et al. (2020) use surveys to elicit beliefs about chances of admission to schools and find that students and parents have heterogeneous beliefs that depart from rational expectations. Luflade (2017) estimates beliefs that rationalize untruthful applications in a DA mechanism while extrapolating preferences from a small subset of plausibly truthful applicants. My paper is the first to jointly estimate preferences and beliefs in a centralized mechanism setting.

This paper is organized as follows: Section 2 describes the Albanian college admissions, policy variation, and data. Section 3 provides descriptive facts on application patterns that are consistent with predictions from a simple model of length-restricted college applications in which higher-SES students have better outside options. Section 4 analyzes the effects of a centralizing policy. Section 5 then provides a full model of strategic application behavior and estimation details of its primitives, and Section 6 presents model results. With estimated model parameters, Section 7 analyzes the effect of market structure on application behavior, on the allocation of students to colleges and majors and welfare and assesses the relevance of strategy in determining outcomes for students under different market configurations. Finally, Section 8 concludes.

2 Background and Data

2.1 Higher Education in Albania

Higher education in Albania is delivered by 12 public and 26 private universities.[7] While public universities have always served the majority of students (73% of college enrollees in 2019), private universities have enrolled an increasing share of students over the past two

[7] These counts reflect the market in the period 2016-2019. More detail on public and private programs in the system can be found in Table 1.

decades.[8] The quality of the degrees varies for both types of universities with significant overlap. Programs offered by public universities in the capital are on average the most competitive and most likely to be oversubscribed. Appendix Figure A-1 displays the distribution of average scores for enrollees in public and private institutions. Tirana campuses of public universities enroll better performing students than private programs on average and several of the most competitive programs in the country are public. Public regional universities on the other hand tend to offer programs that are less competitive than private universities. Geographically, all private university campuses are located in the capital, whereas public college campuses are spread across most regions of the country. Public colleges are generally tuition-free, and impose a small fee for students to attend, whereas tuition to private colleges varies by college and major and ranges between \$500 and \$5000 per year with little need-based financial aid in the private system. While scholarships are offered in private colleges, they are merit-based and generally subject to strict score cutoffs for qualification.[9] Finally, as in many other countries, higher education is immediately specialized, with students making application and enrollment decisions to college-major pairs rather than just institutions.

2.2 Admissions procedures in the partitioned pre-2016 market

Before 2016, a national clearinghouse managed public college admissions alone. Admissions proceeded as follows. At the end of high school in June of each year, students took national exams (called Matura Exams) that included mandatory tests in math and Albanian language and elective subject tests in two subjects among those covered in the high school curriculum. After results of the national exams were announced, students began their application process to public programs through the national clearinghouse. Applicants submitted a rank-ordered list of up to ten public college-major pairs to the mechanism. On the colleges' side, programs ranked students through predetermined formulas that computed a weighted average of the high school GPA and scores in the national exams. These weighted averages

[8] Private provision of higher education only became possible after the end of the communist regime in 1991. In particular since the early 2000's there was a proliferation of private for-profit higher education institutions. Concerns over the quality of these institutions led to a government crackdown on private for-profit colleges and the closure of 18 private universities. With more quality oversight, there have been 26 private universities that have operated between 2014 and today. See the closure order here: https://arsimi.gov.al/wp-content/uploads/2018/07/VKM_per_heqjen_e_licences.pdf.

[9] Scholarships are also offered for four additional categories independently of their academic performance: children of policemen killed in the line of duty, athletes with high achievement at the national level in their respective Olympic sport, members of the Roma/Egyptian community, and orphaned children from low-income families. These, however, are a very small number of scholarships for each program.

would weight each component depending on the course content of the program.[10] Because the majority of public programs, in particular those in the capital, were oversubscribed, a Deferred Acceptance algorithm was run by the clearinghouse to allocate students to a single program.

Admissions to private programs, on the other hand, were decentralized and spanned a period beginning before the centralized public match and ending after it. Students were able to apply and enroll in a private program at any point during the public match. Since a significant portion of the students assigned to public university seats rejected their assignment for private off-match programs, the main phase of the match left universities with vacant seats that could be filled by students that would prefer those seats relative to their assignments.[11] Therefore, a supplemental assignment round was conducted in which all participants still present in the public match would be reallocated to a choice at least as highly ranked in their initial list as the one they were assigned to in the first round. Appendix B.1 describes in detail the timeline of admissions and allocation mechanism for the public match before 2016.

2.3 Admissions policy change

A higher education reform, which was signed into law in 2015, changed the configuration and operation of the college admissions market for the graduating high school class of 2016 and all following cohorts. Below I describe the features of the policy change:

(1) All private universities joined the centralized platform. The policy change incorporated all private programs into the centralized application platform such that no students could gain admission to any program in the country without going through the centralized admissions process. Private colleges were required to produce their own criteria and formulas for admission, which would be made public on the clearinghouse website.[12] Figure 1 shows the expansion of the platform in 2016 from around 300 programs to over 500 programs.

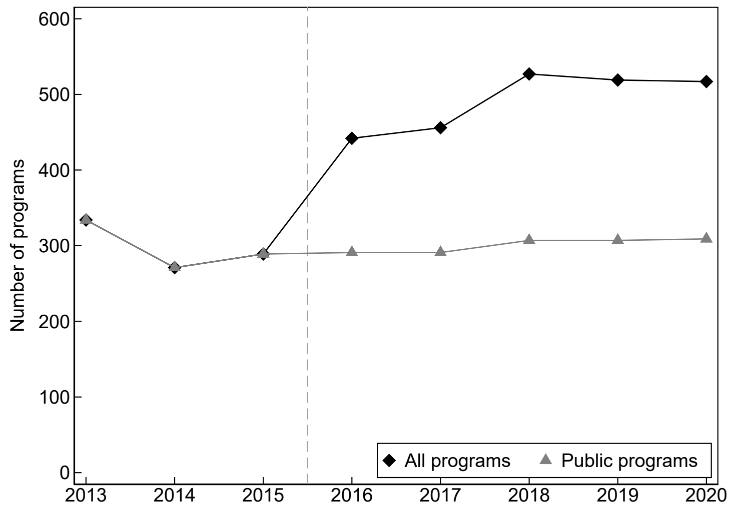
The incorporation of all private programs into the platform was immediate and complete

[10] For example, the Mathematics degree at the University of Tirana would give a weight of 1.4 to a Math subject score in the national exam, and only a weight of 1 to the History subject score.

[11] This externality in the match generated by the partitioned nature of the market and private off-platform options is empirically evaluated in Chile by Kapor et al. (2022). I instead focus on externalities from off-platform options at the application stage rather than during the match process.

[12] This website is also where college applications would be submitted and information about previous years' program cutoffs would be posted (www.ualbania.al).

Figure 1
Number of Programs on the Centralized System



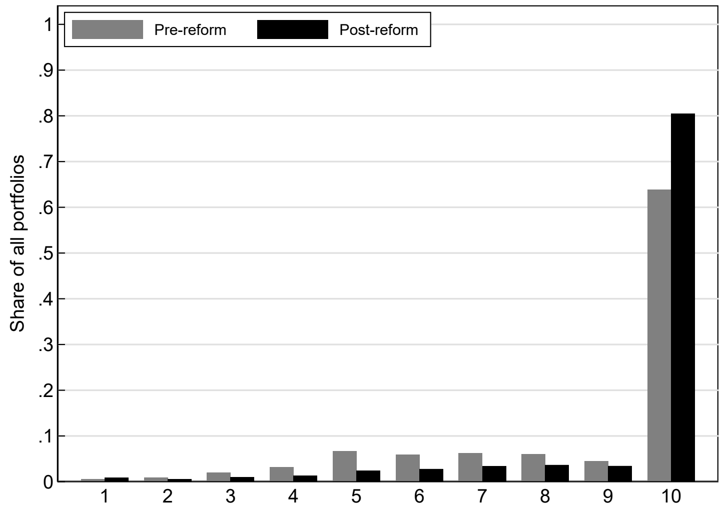
Note: Chart shows the number of participating programs in the centralized system. In 2016, all of the private universities in the country joined the centralized system. The increasing number of programs after 2016 reflects private universities in the system increasing their program offerings.

and the jump to 430 available platform programs in 2016 reflects this implementation. In the years that followed the reform, the number of on-platform programs continued to increase as universities introduced new programs, but there was no movement of programs into or out of the platform.

(2) List size restrictions remained the same. Despite the expansion of the platform, the number of programs students were allowed to apply to remained restricted to 10, the same as in the pre-reform period when the platform only had about 60% of the post-reform programs on it. Figure 2 shows a histogram of submitted application sizes before and after the reform. The list size restriction became more binding after the policy change with over 80% of students filling their lists relative to the 62% before the reform. Not only did students apply to more programs after the reform, but a larger share of national exam takers applied to college through the centralized platform among both high and lower-SES students (Appendix Figure A-2).

(3) The assignment procedure changed from DA to a multi-offer dynamic procedure. The reform was accompanied with a change in the mechanism that allocated students to programs. Instead of DA, the new mechanism is a multi-offer dynamic mechanism with 7 phases in its main round. As before, students submit their application lists of at most 10 programs to the clearinghouse and programs submit their priorities. Then, as a program-proposing

Figure 2
Application Sizes Before and After the Reform



Note: Chart shows the number of programs submitted in an application portfolio before and after the reform.

Gale-Shapley algorithm would begin, in the first phase, programs make initial proposals to students ranked at the top. In a college-proposing DA, the first step would produce matches between proposing programs and students that rank these programs first. Any offer from programs that students don't rank first would be rejected. Then in following steps any programs with empty seats would propose in order of priority to students without a match and matches would be produced in the same way as the first step until either all students have been assigned, or there are no seats available in programs with unmatched applicants. Under the new procedures, in the first phase, students observe all initial offers and decide within 48 hours whether to enroll in any of the proposing programs or forgo all first phase offers and wait for a better offer in the next phase. Matches are then produced with active participation by the students. Differently from a college-proposing DA, the first stage matches include not only matches between programs and qualified students that ranked them first, but also matches between qualified students and programs preferred relative to expected payoff from the remainder of the portfolio. All unmatched students and empty seats are then carried to the next phase of the mechanism and colleges propose in the same way as in the first phase. The multi-offer phases continue until the 7th phase of the main round or until each school has filled its seats. In Appendix B.2 I describe in detail the post-reform admissions procedures.

2.4 Data

I bring together application, assignment, offer, and enrollment data from several sources for three years just before the reform (2013-2015) and four years after the reform (2016-2019).

Pre-reform applications: For applications on the centralized system for the three years before the reform, I use applicant-level data from the Center for Educational Services of Albania (*Qendra e Shërbimeve Arsimore, QSHA*), which was the agency that administered the national high school exams and managed college applications before the reform. The data contain the rank order lists of programs in each application, as well as information on each applicant's district, high school, GPA, national exams taken and exam scores. The data do not include any information on applications to private programs, a limitation common to all application data in centralized assignment systems that do not include all available options. Despite this limitation, the data allow me to compare application behavior within the centralized system, in particular for applications to public programs before and after the reform.

Pre-reform assignments: I have data on the assignments of students to public programs for the years 2013-2015, which can be linked uniquely to application data through the Matura ID, the unique ID assigned to each student at the time of national exams. Assignment data are available from QSHA for both the initial placement and the final student placement after the reassignment round once those with better off-platform options have rejected their public program assignment.

Post-reform applications: For applications in the post-reform period, I combine data from a number of sources. College application portfolios from the early post-reform years 2016-2017 are publicly available as part of a transparency effort. For the year 2018, I obtain individual application and admissions datasets directly from Albanian universities. While I was not able to obtain application data for all private universities in 2018, I collected applications to all public universities, which is sufficient to analyze strategic behavior in applying to public programs. For the year 2019, I obtained all applications from the Academic Network of Albania (*Rrjeti Akademik Shqiptar, RASH*), the agency formed after the reform to manage the college application process. Unlike application data before the reform, these data do not contain any details on applicant district or high school. I extract applicant district and high school from applicant IDs.[13] I then supplement applicant data with public

[13] I observe a regularity in the structure of IDs in all years of my data. IDs in 2019 have the following structure: 19ABcdeVWXYZ. The first two digits of the ID correspond to the cohort of the student. In

information on exam scores from the national exams.

Dynamic mechanism offers and enrollments: Initial (phase I) priority rankings from each program are published on the clearinghouse website every year. In addition, the clearinghouse provides initial seat counts available for each program. I obtain enrollment decisions in each of the phases of the multi-offer dynamic mechanism from RASH and combine them with priority rankings of applicants and seat counts for each program to generate offers the students receive in each of the phases of the main round.

I present summary statistics on applications and market characteristics in Table 1. The sample contains over 84,000 applicants in the three years before the reform and around 98,000 in the four years after the reform. Public high schools educate the majority of the applicants, about 85% in both periods, and they perform slightly worse on average than private schools in the national exams. The consolidation of private universities into the centralized application system increased the choices available through the same application from the 12 public universities to all 38 higher education institutions and increased the available number of programs through the platform from 289 to 517 over a period of five years. After the reform, students from private high schools have slightly longer application lists than students from public high schools and 21% of their application lists are private programs. In contrast, only 9% of public high school students' post-reform application lists are private programs.

3 Application patterns with a partitioned market

3.1 Do students of different SES groups have different applications and outcomes in the public-only match?

In this section I show descriptive evidence of differences in applications and assignment outcomes between high and lower-SES students in the centralized system when only public university seats are assigned centrally. This evidence is consistent with high-SES students applying to more selective colleges conditional on exam and high school performance. This behavior is rewarded for some through enrollment in more selective degrees. I formally test

the example above, the student is applying to college in 2019. I infer that the second two digits reflect the district where the student went to high school, and the next three digits correspond to the high school from which the student is graduating. Finally, the last 5 digits are unique to the individual. I then map the district code and high school code to the corresponding district and high school using data from 2013-2015, in which I have district and high school information.

Table 1
Summary statistics on applicants and programs

	Pre-Reform (2013-2015)			Post-Reform (2016-2019)		
	All	Public HS	Private HS	All	Public HS	Private HS
a. Applications on the platform						
Number of applicants	84,931	72,766	12,165	98,459	82,499	15,527
Share of applicants		0.86	0.14		0.84	0.16
Share from capital	0.20	0.20	0.25	0.23	0.22	0.28
Exam score average	6.99	6.92	7.41	7.30	7.22	7.70
Application portfolio size	8.60	8.61	8.60	9.19	9.18	9.26
Public share of portfolio	1.00	1.00	1.00	0.89	0.91	0.79
	All	Public Uni.	Private Uni.	All	Public Uni.	Private Uni.
b. Programs on the platform						
Number of colleges	12	12	0	38	12	26
Number of programs	289	289	0	517	309	208

Notes: Application data come from the Center for Educational Services for 2013-2015, publications of the Ministry of Education for 2016-2017, the Academic Network of Albania for 2019, and individual colleges for 2018. Exam score data come from publications of the Center for Educational Services. The averages of exam scores and portfolio sizes exclude 2018 as application data are missing for some colleges for that year.

differences in application and assignments by regressing measures of selectivity of application lists and assignment outcomes (y_{idt}) on an indicator for SES, and exam and high school performance for the sample of applicants in the period with a public-only match:

$$y_{idt} = \delta_1 \text{lowSES}_{it} + \delta_2 score_{it} + \delta_{dt} + \varepsilon_{idt} \tag{1}$$

Score is the average of high school GPA and end-of-high school exam scores. The regression includes district-by-year (δ_{dt}) fixed effects such that the comparison is between high and lower-SES students within the same district and year. Outcomes include the assigned program's rank in the student's rank order list, the assigned program's selectivity and its selectivity rank among all programs.

First, I show that high-SES students apply to more selective programs and conditional on enrollment, also enroll in more selective programs. Table 2 shows the coefficient estimates $\hat{\delta}_1$ for the selectivity of the top three listed programs in students' applications. For the selectivity of a program is measured as its score cutoff, the lowest average score for a student who was assigned a seat in the program. Columns 1-3 show that each of the top three choices in the applications of public high school students is less selective than each of the top three choices of private high school applicants by 0.2 grade points. This difference is statistically significant and equivalent to 0.23 standard deviations of the distribution of top choices across the full set of applicants.[14] Columns 4 and 5 in table Table 2 show that the differences in selectivity are not limited to programs to which students applied, but also to those where they enrolled. While the DA mechanism narrows the selectivity difference between schools to which students are assigned relative to the schools to which they apply, there remains a difference of 0.054 grade points and 4 ranks in the programs that high and lower-SES students enroll in conditioning on their scores.

Table 2
SES differences in selectivity of application and enrollment

	Application Selectivity			Enrollment Selectivity	
	Top ranked	Second ranked	Third ranked	Cutoff	Rank
Public HS	-0.210***	-0.192***	-0.196***	-0.054***	3.896***
	(0.011)	(0.010)	(0.010)	(0.010)	(1.040)
Priv. HS Mean	9.118	9.078	9.056	8.682	121.993
Adjusted R^2	0.243	0.282	0.289	0.406	0.445
Obs.	50,947	50,753	50,456	45,595	45,595
Condition on Score	✓	✓	✓	✓	✓
District-by-year FEs	✓	✓	✓	✓	✓

Notes: Sample includes years 2013-2015. Standard errors are clustered at the high school level.

Second, the rate of assignment to a program on the platform overall, and to top-listed programs in particular differs by SES. Table 3 shows the coefficient estimates $\hat{\delta_1}$ for the rate of assignment to the top-one and one of the top-three listed choices as well as the rate of remaining unassigned. For the same average score, lower-SES students are more likely than high-SES students to be assigned to a higher listed choice. They are roughly 20% (6pp) more likely than high-SES students in the same district to be assigned to their first listed choice and 18% (9pp) more likely to be assigned to one of their top three choices. These results show that lower-SES students get assigned more frequently to programs higher in their lists, which implies, consistently with the theoretical prediction, that they apply to relatively less risky choices at the top of their lists. Lower-SES students are also about 12% (2pp) less likely to remain unassigned after conditioning on scores, which indicates that the difference in aggressiveness of applications is not only present among the top choices, but also for the entire portfolio.

Specifications with additional controls confirm that the differences in assignments are robust

[14]This magnitude is also four times the size of the difference in average selectivity between the top and the third listed option for high-SES students, as shown in the third row of Table 2.

to comparing applications within high school track and type of high school. In addition, not only do the differences appear once reassignment requests are accommodated, but also for initial offers (Appendix Table A-1) and assignment of only students who did not reject their platform offers (Appendix Table A-4).

Table 3
Relationship Between Attending a Public HS and Final Assignment Outcomes

	National Outcomes			Outcomes in Capital		
	Assigned to Top Choice	Assigned to a Top-Three Choice	Unassigned	Assigned to Top Choice	Assigned to a Top-Three Choice	Unassigned
Public HS	0.056***	0.091***	-0.019**	0.062***	0.111***	-0.009
	(0.012)	(0.014)	(0.009)	(0.019)	(0.025)	(0.018)
Private HS Mean	0.288	0.498	0.134	0.173	0.346	0.212
Adjusted R^2	0.109	0.154	0.221	0.123	0.209	0.268
Observations	84,931	84,931	84,931	17,336	17,336	17,336
Condition on Score	✓	✓	✓	✓	✓	✓
District-by-year FEs	✓	✓	✓			
Year FEs	✓	✓	✓	✓	✓	✓

Notes: Sample includes years 2013-2015. Private HS Mean is the unconditional mean of the outcome variable for students attending private high schools. Standard errors are clustered at the district level for the national sample and are robust for the capital-only sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The evidence from both differences in the rate of assignment to top choices and any choices, and the differences in the measures of of selectivity for programs in the rank order lists and those in which students enrolled are consistent with differential risk-taking behavior that may arise when the appeal of outside options is higher for the high-SES group.

Alternative explanations: There are several possible alternative explanations for these observed patterns that may not be related to outside options.

Unequal choice sets: Higher SES students may have different choice sets even within the public match that may explain the patterns. For example, consider two students from different SES backgrounds applying to college from the same district. The lower SES student may not be able to afford relocating to a different city to attend college, so her choice set may be limited to programs available at the local university, which may be less selective than programs further away. It may appear then as though this student is exhibiting less risky application behavior than the high-SES student, though her choices are not reflective of strategic misrepresentation due to preferences for outside option, but are a result of her having a different choice set than the high-SES student even within the set of available public programs. When focusing on applicants with similar choice sets within the options offered on the centralized system, the differences in assignments and applications remain, suggesting that the disparities are not explained by differential access to public programs. Columns 4-6 of Table 3 show final assignment outcome comparisons by type of school for applicants from the capital. The restriction to students in the capital allows comparisons of students with very similar choice sets. The differences in applications and assignments are even more stark here than at the national level. Public HS students are assigned to their first choice 6pp more frequently than private HS students, who get assigned to their first choice an average of 17% of the time. The difference is 11pp for likelihood of assignment to the top three listed choices.

Differential grade inflation in high schools: Higher grade inflation in public high schools relative to private high schools may explain application patterns. The formulas for admission to public programs are a function of both high school GPA and national exam scores. If private high schools exhibit lower grade inflation than public schools, it may be that students of the same ability level have different high school GPA components in their score, but they apply according to their privately perceived ability rather than the observed scores. To alleviate this concern, I regress national exam scores on indicators for public status of high school, high school GPA, and an interaction term between high school grades and public status. I find evidence of the opposite: while distributions of within-high school performance differences overlap across private and public high schools, private high schools on average have higher high school grades in math and language than their exams, so that their high school GPA on average overstates their ability to the extent that ability is correctly reflected in their exam grades.

Informational differences: Lower-SES students may have better information about their admission chances to programs or alternatively high-SES students may be overconfident about their chances. This cannot be ruled out at this stage, but the policy analysis in Section 4 shows that systematic differences in beliefs cannot explain the full difference in applications in the pre-reform period. In addition, the model I specify in section Section 5 will allow for students from different backgrounds to have different beliefs about their chances of admission in order to assess the importance of informational gaps in explaining different applications by SES groups.

Different preferences for selectivity: One may be concerned that the observed data patterns are due to high-SES applicants having stronger preferences for more selective programs. While this cannot be ruled out, the model presented in section Section 5 will allow the separate identification of preferences for selectivity for high and lower-SES students in order to assess the relative importance of preferences for selectivity in students' applications.

As a whole, this descriptive evidence documents disparities along SES lines in assignments through a centralized process to free public colleges. The disparities arise from differences in application behavior that is consistent with more aggressive applications for high-SES students that is detectable both directly, and through assignment patterns. This does not imply that better outside options are responsible for more selective applications by high-SES students. Importantly, differences in beliefs about admission chances or strength of preferences for selectivity could explain this behavior. In the next section, I provide evidence from a policy that incorporated outside options into the centralized application to show additional evidence that the availability of private programs in the choice set, rather than preferences or information, is an important factor contributing to applications in the public program match.

4 Effects of the Reform on Applications

In this section, I turn to providing evidence that a shock that differentially impacted the outside and on-platform options of high-SES students relative to lower-SES students closed much of the gap in the selectivity of applications for the two groups. I will analyze the 2016 reform and show that the collapse of the off-platform choice set of private programs and the expansion of the within-platform choices decreases the selectivity gap between high and lower-SES students observed in the pre-reform period. This provides additional evidence that the presence of private programs outside of the centralized system offers high SES students an advantage in public program applications.

4.1 Event Study Specification

Difference-in-differences event study: In an ideal experiment, to estimate the effect of outside options on strategic misreporting in applications, I compare two groups that are identical, except one has access to outside options and the other does not, such that when outside options are completely removed, the changes in the affected group's applications relative to the unaffected group reflect the effect of the shock to outside options. Instead, the Albanian policy environment calls for a comparison between the application behavior of high-SES students to lower-SES students before and after the reform.. This is an imperfect comparison. In the Albanian setting, private high schools are a good proxy for high-SES students, whereas public high school students are a worse proxy for lower-SES students as there are many high-SES students among those that attend public high schools. The comparison between these two groups in the data relies on the fact that all private high school (high-SES) students will be treated by the policy (the value for the outside option will shift for all of them), whereas public high school students (lower-SES) have a lower treatment rate. The event study specification is:

$$y_{idt} = \sum_{k=2013}^{2019} \beta_k \mathbb{1}\left[t = k, t \neq 2015\right] + \sum_{k=2013}^{2019} \beta_k^{\text{LowSES}} \mathbb{1}\left[t = k\right] \times \mathbb{1}\left[\text{LowSES}\right] + \beta_s \text{score}_i + +\delta_d + \varepsilon_{it} \tag{2}$$

where the omitted group is high SES students in the year 2015 and controls include district fixed effects, and scores. Because the policy change will impact both the outside options of high and lower-SES students, the effect of interest would be the net effect of the reform, the extent to which the reform differentially changed the applications of high SES students compared to lower-SES students. This is captured by the difference $(\beta_k - \beta_k^{\text{LowSES}})$ for each $k \in \{2013, ..., 2019\}$.

Even though this event study specification will allow me to trace out the time path of the estimated effects of the reform, it is susceptible to time-varying confounds. For example, the reform of 2016 also changed the mechanism that allocated students to programs from an algorithmic DA to a live-DA-type assignment system with exploding offers[15]. This may have changed students' beliefs about their probability of admission differentially for high and lower-SES students, which would confound the estimates and would not appear in the pre-trends.

Triple differences event study: To address any time-varying effect of the broader changes to the system that would have systematic effects for all students of the same background but would differ across backgrounds, I use an additional source of cross-sectional variation that would absorb such time-varying effects. This source of variation is merit eligibility for scholarships to private colleges. As discussed in Section 2.1, private schools do not generally offer scholarships based on need, but they do offer scholarships based on merit. A common rule of thumb private colleges use to offer scholarships, which is advertised widely, is to give full rides or scholarships for the majority of the tuition to students above a grade

[15]Details of this mechanism change are discussed in Section 2.3 extensively Appendix B.2.

threshold.[16] These scholarship policies imply identical, or near identical access to private university options for top students in both private and public high schools. However, if there are any changes in beliefs about chances of admission that are systematically different across backgrounds after the reform, lower-SES merit-eligible students would be differentially affected by these changes, which would isolate the effect of these changes on lower-SES students relative to high-SES students.

To formalize this empirical strategy, the event study specification is the following:

$$
\begin{aligned}
y_{idt} = & \sum_{k=2013}^{2019} \beta_k \ \mathbb{1}\big(t = k, t \neq 2015\big) + \sum_{k=2013}^{2019} \beta_k^{\text{LowSES}} \ \mathbb{1}\big(t = k\big) \times \mathbb{1}\big(\text{LowSES}\big) \\
& + \sum_{k=2013}^{2019} \beta_k^{\text{NM}} \ \mathbb{1}\big(t = k\big) \times \mathbb{1}\big(\text{NM}\big) \ + \sum_{k=2013}^{2019} \beta_k^{\text{LowSES,NM}} \ \mathbb{1}\big(t = k\big) \times \mathbb{1}\big(\text{NM}\big) \times \mathbb{1}\big(\text{LowSES}\big) \\
& + \gamma_t \text{score}_i + \delta_d + \varepsilon_{idt}
\end{aligned}
$$

where the omitted category is applicants from private high schools in the year just before the reform. In this specification, β_k are the coefficients for merit-elegible high-SES students, β_k^{LowSES} are coefficients for merit-eligible lower-SES students, β_k^{NM} are coefficients for high-SES non-merit eligible students, and $\beta_k^{\text{LowSES, NM}}$ are the coefficients for lower-SES non-merit students. The net effect of the reform is then the difference $(\beta_k - \beta_k^{\text{LowSES}}) - (\beta_k^{\text{NM}} - \beta_k^{\text{LowSES, NM}})$ which reflects the outcome changes for high-SES students relative to lower-SES students after differencing out other time-varying confounds that affect students differently across backgrounds but identically within backgrounds.

In addition, the effect of the reform will be the combined effect of a shock that differentially affects outside options *and* expands choice within the platform.

4.2 Triple Differences

Finally, the triple difference specification is:

$$
y_{iht} = \beta_1 LowSES_{iht} \times NM_{iht} \times Post_{iht}
$$

[16] For many years, that threshold has been 9 for most schools. This corresponds to roughly the top decile of applicants, or about 2,000 students. Further details can be found in Appendix C.1.

$$+ \beta_2 LowSES_{iht} \times NM_{iht} + \beta_3 LowSES_{iht} \times Post_{iht} + \beta_4 NM_{iht} \times Post_{iht}$$
$$+ \beta_5 LowSES_{iht} + \beta_6 LowSES_{iht} + \beta_7 LowSES_{iht} + \gamma score_{iht} + \delta_d + \varepsilon_{iht}.$$

For student i, attending a high school with status $h \in \{\text{private}, \text{public}\}$ and applying in year t, outcome variables y_{iht} capture application characteristics such as the selectivity of the most selective program or the top two and three most selective public programs in the portfolio. $LowSES_{iht}$ is a dummy for lower-SES type, NM_{iht} is a dummy for non-top student type, and $Post_{iht}$ is a dummy for the post-reform period. The coefficient of interest is β_1 and it captures the effect of the reform on application behavior.

The outcome y_{iht} for both the event study and triple differences is the most selective *public* program choice on the platform. This is chosen as the main measure because the upward gamble induced as more options are chosen (Chade and Smith 2006), indicates that some of the higher listed and more selective choices will be the last to be chosen, and the first to be removed when the list size restriction becomes more binding.

I measure selectivity of each program as its historical cutoff score, set at year 2013. This decision is made to avoid changes in cutoff scores year-to-year from affecting the estimates when they may be a result of more or less difficult national exams, or years with more or fewer applicants. Importantly, this decision sidesteps changes in equilibrium cutoffs that may be as a result of changes in the capacity of private programs at the time of incorporation into the public match. The historical cutoff score serves as a measure of reputation of a program, which, unlike cutoff scores, takes longer to update.

4.3 Results and Discussion

Application sizes and list-filling

The first step to evaluating the effects of the reform on the strategic incentives of students is to establish the extent to which list-filling behavior changed after the reform and the extent of crowd-out of public programs by private programs in application lists. Figure 1 provides a visual description of the change and Table 4 formally tests the change. Before the reform, the average number of programs to which both high and lower-SES students applied was 8.7 (table 4) with approximately 65% of students filling their lists. After the reform, over 80% of students from both high and lower-SES backgrounds filled the application lists, not statistically differently from each other (first column of Table 4), with the average number of applications submitted at 9.4 per student. The crowd-out effect of the platform expansion, however, induced more students from high SES background to apply to more private programs than did students from lower-SES backgrounds. The second and third columns of Table 4 indicate that after the reform, lower-SES students had, on average, one 1.2 more public programs listed than did higher SES students (or about 13% more of their average application list). This is true even when restricting to applicants who filled their lists (fourth and fifth columns of Table 4), indicating that a larger share of the portfolio for those of higher SES that filled their lists was made up of private programs than for lower-SES students. The post-reform changes in portfolio composition is not merely due to students adding private programs to their application lists, but rather replacing some of the public programs they would otherwise include with private programs. This can be seen in the fact that the average count of public programs declines from 8.7 before the reform to 7.6 after for high-SES students.

Table 4
Effect of reform on application counts

	All applicants			Applicants who filled lists	
	Count of programs listed	Count public	Share public	Count public	Share public
Lower SES	0.010	0.040	0.004	0.048	0.005
	(0.134)	(0.126)	(0.003)	(0.030)	(0.003)
Lower SES x Post-Reform	-0.014	1.165***	0.133***	1.191***	0.119***
	(0.087)	(0.204)	(0.019)	(0.161)	(0.016)
High SES Mean Pre	8.680	8.680	1.000	10.000	1.000
High SES Mean Post	9.360	7.610	0.802	8.214	0.821
Adjusted R^2	0.087	0.064	0.194	0.195	0.195
Observations	109,092	109,092	109,092	81,252	81,252
Condition on Score	✓	✓	✓	✓	✓
District FEs	✓	✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓

Notes: Count public is the number of public programs listed in the application. Share public is the share of programs listed that are in public universities. The first three columns show the difference in list sizes, count public and share public for all applicants. The last two columns restrict the sample to those applicants who filled their lists. Sample includes all applicants for years 2013-2019, excluding 2018. Data from year 2018 are excluded from this table as I only have applications to public programs for that year. Standard errors are clustered at the district level.

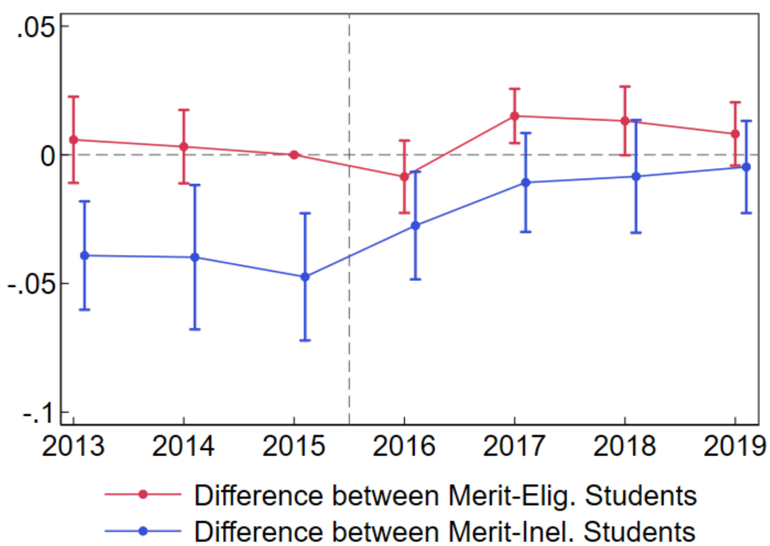
One potential reason for this could be that the composition of students changed. For example, students could have been induced to apply that prefer private programs to any public ones and would only or mostly apply to private programs through the platform. As shown in Appendix Figure A-2, the rate of application through the platform increased for both types of SES backgrounds by only 5pp, and even if all the marginal applicants had applied to only private programs, int could not explain the decline in the average number of public programs in the lists.

Even if changes in composition do not explain the changes in application counts, it is possible that these changes in applications are merely a result of the ability to express preferences over a larger number of programs on the platform. That is, replacing a public program with a private program may be consistent with truth-telling if the least preferred public programs are the ones being replaced with the private programs. In what follows, I will show evidence that it is in fact the most selective programs, and those highest ranked that are most likely to be replaced by private programs in the portfolio.

Application selectivity

The first exercise of this section is to measure the effect of the reform on the selectivity of highly ranked options. I use the most selective option included in the list as a proxy for the most preferred option. This is because in the pre-reform period, where lists are ranked, the top ranked option is the most selective in 94% of the applications. I estimate the event study specification and show the double differences for each period in Figure 3.

Figure 3
Event Study of Double Differences in Selectivity of Most Selective Public Programs
Top and Non-Top Students



Note: This chart plots the differences in selectivity of most selective public programs on the centralized application between private high school and public high school students. Negative differences reflect less selective choices at the top of portfolio for public high school students. Regressions control for average exam score and include district FE. Standard errors clustered at district level.

After the reform, the gap between the selectivities of the most selective choices of higher and lower-SES students closes. Appendix Figure A-5 shows the results of the event study triple differences, or the difference between the selectivity gap for top-performing high and lower-SES students and non-top high and lower-SES students. The application gap for top students between SES groups is small before the reform due to targeted scholarships for high performing lower-SES students, which leads both these groups to have similar access to outside options. This gap remains small after the reform. On the other hand, for lower performing students, the gap between high and lower-SES students is high before the reform, but declines after. This exercise suggests that after the reform students' top choices are not as selective as they were before the reform, and the gap in most selective choices between SES group declines, suggesting that the closure in the gap is not due to lower-SES students increasing the selectivity of their top option after the reform, but rather it is due to higher SES students decreasing the selectivity of their top option by more than the lower-SES students do. While the gap and its closure are most pronounced for most selective public choices in students' portfolios, I present event study results for average portfolio selectivity for public programs in Appendix Figure A-6. Overall, average selectivity of portfolios for higher SES students declines more than for lower-SES students, consistent with theoretical predictions.

In additional regressions, I check the robustness of the above results to two alternative measures of selectivity. First, I measure the selectivity of a program as the cutoff score of the program and center and scale the distribution of selectivities to have mean 0 and standard deviation 1 in each year[17]. This choice circumvents issues with year-to-year changes in cutoffs that reflect changes in difficulty of end-of-high school exams or changes in the average performance of the pool of candidates. The results of these robustness checks are presented in Appendix Figure A-8 and are qualitatively the same as those with the main selectivity measurement. In a second set of alternative results I measure the selectivity of a program as its rank in the given year and estimate the event study specification. Appendix Figure A-9 displays results similar to all specifications above.

Finally, in the triple differences analysis, I formally quantify the effect of the reform on program selectivity. The double differences pass tests of parallel trends in several specifications as shown in Appendix Table A-5. Results from the triple difference specification for most selective public program are shown in Table 5. The removal of outside options decreases the

[17]For programs with empty seats at the end of the admissions process, I set the cutoff score to the lowest average score a student can achieve in the national exams.

selectivity of public programs at the top of applications by 0.1 standard deviations. On average, portfolio selectivity declines by 0.04 standard deviations. These estimates are robust to alternative specifications and alternative measurements of program selectivity.

Table 5
Triple Difference Estimate of Exposure to Contraction of Outside Options on Selectivity of "Reach" Programs Chosen

	Including All Years			Excluding 2016		
	(1)	(2)	(3)	(4)	(5)	(6)
Public HS × Non-top × Post-reform	0.067	0.094**	0.078*	0.093**	0.115***	0.096*
	(0.042)	(0.041)	(0.046)	(0.047)	(0.044)	(0.050)
Public HS × Non-top	-0.132***	-0.152***	-0.156***	-0.130***	-0.149***	-0.153***
	(0.050)	(0.047)	(0.044)	(0.050)	(0.048)	(0.044)
Public HS × Post-reform	0.016	-0.016	0.013	0.006	-0.024	0.011
	(0.032)	(0.036)	(0.031)	(0.033)	(0.033)	(0.028)
Non-top × Post-reform	-0.039	-0.100**	-0.092*	-0.068	-0.130**	-0.119**
	(0.049)	(0.046)	(0.051)	(0.053)	(0.051)	(0.058)
Public HS	0.022	0.048**	-0.007	0.022	0.047*	-0.013
	(0.029)	(0.024)	(0.027)	(0.029)	(0.024)	(0.030)
Non-top	0.038	0.210***	0.219***	0.056	0.215***	0.224***
	(0.061)	(0.048)	(0.046)	(0.057)	(0.045)	(0.045)
Post-reform	-0.309***	-0.191***	-0.211***	-0.338***	-0.210***	-0.238***
	(0.042)	(0.045)	(0.039)	(0.041)	(0.042)	(0.032)
Adjusted R^2	0.335	0.412	0.427	0.338	0.415	0.434
Observations	132,079	132,079	132,079	111,909	111,909	111,909
Score Controls	Yes	Yes	Yes	Yes	Yes	Yes
Score and HS Path Controls	No	Yes	Yes	No	Yes	Yes
District FEs	No	No	Yes	No	No	Yes

Notes: Scores represent the weighted average of end-of-high-school exam scores. HS Path is a binary variable that represents whether the path chosen in the second year of high school is "scientific" or "social", which affects the weights programs give to the elective exams. 2018 applications are included in these regressions as the outcome variable only requires a selectivity measure for public programs in the application and those data are available for 2018. Standard errors are clustered at the district level and are shown in parentheses.

Next, I investigate the mechanisms that drive the closure in the application gap. Appendix Figure A-4 shows the event study estimates for the four groups of students. Estimates show that overall the reform lead to a reduction in the selectivity of the most selective choice listed in the applications of all types of students, but the reduction is largest for high-SES students that are lower performing. This is precisely what is expected to be the effect of the reform: higher SES students in the non-top performing group see their outside options restricted more than lower-SES students and reduce the selectivity of their applications to public schools more in response to this change.

5 Model of Application and Decisions on Waitlists

5.1 Model Primitives

5.1.1 Timing and Sequence of Decisions

1. *National Exams and Decision to Apply*—Students finishing their senior year of high school take end-of-high school exams and learn their scores from the exams. After learning their scores, those who pass the national exams decide whether to apply to colleges through the common application.

2. *Portfolio Selection*—If student i decides to apply, she must select an application portfolio R of ten programs (college-major pairs) from the available set of over 500 such that $R \subset \mathcal{J} = \{1, ..., J\}$, $|R| = 10$. All programs accept applications only through the common application and there are no institutions, public or private that conduct their admissions outside of the common app. Students are not required or encouraged to rank programs in any particular order and the ordering within application will not matter for admissions.

3. *Student Priority Ranking*—Applications are received by each program, and applicants are ranked according to their scores and a pre-determined and pre-announced formula. The first round of acceptances is made in each program for as many students at the top of the program's list as there are seats available. The platform forms waitlists for every program and everyone observes the state of all waitlists in $\mathcal{J}$. Formally, in round 1 of admissions, program j makes offers to the top $q_{1,j}$ students on its waitlist, where $q_{1,j}$ is the number of seats at program j.

4. *Offers and Enrollment Decisions*—At the beginning of each round $t \in \{1, ..., T\}$ of admissions (in the data, $T = 7$), program j makes offers to the top $q_{t,j}$ students on its waitlist, where $q_{t,j}$ is the number of available seats at program j at the beginning of wave t. The state of all waitlists is common knowledge at the beginning and end of each round. Each student with offer set $A_{t,i} \subseteq R_i$ makes a decision to accept a single offer from the set $A_{t,i}$ or reject all round t offers and remain on the waitlists of all programs in R_i to which she have has been offered admission yet, $R_i \backslash \bigcup_{s \leq t} A_{s,i}$.

5. *Waitlist Evolution*—If i receives and accepts an offer from program j in round t, then a j seat is allocated to i and i forgoes all potential future offers to programs she has not yet been admitted to as of round t. Program j has one fewer seats available in round $t+1$ and student i is removed from all waitlists. If i rejects all round t offers, she is removed from all waitlists in $A_{t,i}$, but remains in all waitlists of $j' \in R_i$ from which she have not yet received an offer. Once the final offers are made in round $T = 7$, any remaining seats are allocated on a first-come first-served basis.

5.1.2 Student Preferences for Programs

I model the indirect utility that is realized from attending a program as a function of observed and unobserved student characteristics, and observed program characteristics. The utility of student i from attending a program j is given by:

$$v_{ij} = u(z_i, x_j, \omega_{ij}, \varepsilon_{ij}; \theta) \tag{3}$$

where z_i is a vector of characteristics for student i, x_j is a vector of characteristics for program j and ω_{ij} is a vector of pair-specific characteristics, and ε_{ij} is an idiosyncratic taste shock for program j unobserved to the econometrician, but observed by the student at the time of deciding whether to apply to college. I assume that the distribution of the idiosyncratic taste shocks is known to the econometrician and each is drawn i.i.d. from a type-1 extreme value distribution. The distributional assumption of the idiosyncratic shock normalizes the scale and location of the utility. In addition, a key restriction imposed by the independence assumption is that the ε_{ij} shocks are independent from student characteristics, in particular distance to programs. This rules out location choices that are correlated with preferences for program. I further parameterize the utility function as follows:

$$\begin{aligned} v_{ij} &= \beta_{c(z_i)} x_j + \gamma^d_{c(z_i)} d_{ij} + \gamma^p_{c(z_i)} \text{price}_{ij} + \sum_k \lambda_{c(z_i)} x_{j,k} z_i + \varepsilon_{ij} \\ v_{i0} &= \varepsilon_{i0} \end{aligned} \tag{4}$$

where d_{ij} denotes distance to program j and price_{ij} is the student-specific out-of-pocket price, calculated by subtracting the scholarship a student is eligible for from the list price of program j. Scholarship eligibility and amount is primarily determined by the weighted average score for each student.[18]

Preference parameters are specific to each of four mutually exclusive groups of students in cells $c(z_i) \in \{\text{High-SES, Lower-SES}\}$ and allow for heterogeneity in preferences for observed program characteristics for students from different socio-economic backgrounds and high school subject path. Heterogeneity along the SES dimension for all program characteristics will be crucial in capturing the distributional consequences of alternative market designs, as strategic portfolio choices and enrollments will depend in part on preferences for programs inside and outside the centralized system. In addition, preferences for all program characteristics are allowed to be heterogeneous along high school academic path. In particular, students that chose the social science path at the beginning of tenth grade might care differentially about characteristics of programs such as selectivity and field of study compared to students who chose the science path.

Student characteristics z_i include average score on the end-of-high-school exams and urban/rural location. Program characteristics x_j include selectivity, private/public status, field of study in one of four categories (science and applied science, health, social science and humanities, business and economics), an indicator variable for whether j is located in the capital.

Finally, the value from the outside option is given by v_{i0} and represents the value of not enrolling in any college in the current year and is known to students at the time of application. This may include the value from entering the labor force without a college degree or waiting to apply the following year. This is without loss of generality because the choice of portfolio and decision to enroll will depend on differences with the non-college option and not on the value of the non-college option itself. Alternative admissions designs are assumed not to affect the value of the non-college option. This may not hold if changes in the way this market operates affect the expected utility from applying the following year, which is included in the non-college option.

This formulation does not allow for systematically different value from the outside option for high-SES and lower-SES students. Because preferences for programs are defined separately for each SES group and for application and enrollment decisions, only the differences between

[18] There are very few scholarships offered in each private program for non-merit categories (no more than one or two per program) and I cannot identify the students eligible for these scholarships in the data. The number of students whose choices would be affected by scholarship eligibility through the above categories is very small and its impact on parameter estimates negligible.

the value of programs and the value of outside options matter and if high-SES students have better non-college outside options, this will be captured by less strong preferences for any college alternative. Since the value of the outside option does not change in alternative market designs, this specification choice will not affect the distributional consequences of counterfactual policies.

5.1.3 Information and Beliefs Over Program Cutoffs

Before applying, students learn their score and form beliefs about their probability of admission to each program. For each student entering the application stage, each possible application portfolio is a lottery over entrance to one of the programs in the portfolio and not enrolling in college. In their portfolio selection, students take into account not only their preferences for each available program, but also the probabilities over possible outcomes induced by their choices. Because the admissions process involves multiple waves, the initial portfolio choices contain information not only about student preferences, but also about beliefs over the possible outcome paths generated by the portfolio and choices to enroll or wait in each of the admissions waves. A student may form beliefs about the state of waitlists in the each of the waves and distribution of outcomes in the next wave for each possible choice made in each possible induced state of the waitlists. The interdependence of waitlists down the admissions process generates an extremely large number of possible states and a different distribution of perceived possible outcomes for following rounds for each possible state, which makes estimation of all of the belief objects infeasible.

To circumvent this issue, I rely on the specification of the perceived payoff function, which only takes in the probabilities of clearing the final cutoff of each program. That is, the only relevant object in determining the perceived uncertainty of each outcome is $P(c_{Tj} < \text{score}_i)$ for all programs j, where c_{Tj} is the cutoff of program j at final wave T. To support this assumption, it helps to recall the institutional details of the information that students receive from the central admissions authority. At the time of application, students have access only to information about the previous year's *final* score cutoffs for each program (c_{Tj}). This feature is important as students are lacking any public information about the program cutoffs in intermediate waves.[19]

In addition to providing support for the objects over which students form beliefs, the in-

[19] In fact, according to the admissions agency, this decision is intentional so as to simplify the information given to applicants and to focus their attention on final cutoffs rather than intermediate ones.

formation provided by the admissions authority supports a model in which students form beliefs over the distribution of cutoffs for programs that arise form the previous year's cutoffs.[20] Therefore, I model below the perceived distribution of possible realizations of the cutoff c_j, $\tilde{c}_{c(i),j}$ as normally distributed and parameterized as follows:

$$\tilde{c}_j \sim N(\bar{c}_{j,c(i)}, \sigma_j^2)$$

where

$$\bar{c}_{j,c(i)} = \text{cutoff}_j^{y-1} + \mu_{0,c(i)} + \mu_{1,c(i)}\text{cutoff}_j^{y-1}$$

and

$$\sigma_j = \log(1 + \exp(\sigma_0 + \sigma_1 \text{cutoff}_j^{y-1})).$$

The specification above assumes that the belief object is centered around the shifted previous year's cutoff with cutoff_j^{y-1} denoting the previous year's $(y-1)$ cutoff, $\mu_{0,c(i)}$ denoting the shift intercept and $\mu_{1,c(i)}$ denoting the shift slope on program cutoff. This reflects the fact that the mean of the perceived possible distribution of cutoffs may depart systematically as students may believe that admissions in the year they are applying are more or less competitive in general than admissions in the year before and perhaps differentially so for more competitive programs. In addition, I model the standard deviation of the distribution of beliefs around the cutoff as a function of cutoff. This allows for the dispersion perceived distribution of cutoffs to be heterogeneous among students with different academic performance facing a set of choices that are more or less selective. Finally, all belief parameters to be estimated that are indexed by $c(i)$ are allowed to vary by SES. In practice, this parameterization with heterogeneous beliefs by SES, in addition to the preference specification, allow for separate estimation of both taste and belief parameters for the two groups of students.

With this model of perceived distributions of possible program cutoffs, we can derive the perceived probability of student i to clear the final cutoff of program j as:

$$\begin{aligned} P(c_{Tj} < s_i) &= \Phi\Big(\frac{1}{\sigma_j}(\text{score}_i - \tilde{c}_j)\Big) \\ &= \Phi\Big(\frac{1}{\log(1 + \exp(\sigma_0 + \sigma_1 \text{cutoff}_j^{y-1}))}(\text{score}_i - \text{cutoff}_j^{y-1} - \mu_{0,c(i)} - \mu_{1,c(i)}\text{cutoff}_j^{y-1})\Big) \end{aligned}$$

[20]This approach to modeling perceived admission probabilities as arising from beliefs over the distribution of cutoffs is widely used in the school choice literature. See for example Agarwal and Somaini (2020).

The probability of admission at program j will be 50% for students with scores at exactly the value $(\bar{c}_{j,c(i)})$. For anyone with scores to the right of that point, the perceived probability of being admitted to program j will be greater than 50%, but will be declining with increasing variance in perceived possible cutoff distribution. The opposite is true of students on the left side of $(\bar{c}_{j,c(i)})$. Their perceived probability of admission to program j is less than 50% but is increasing with increasing variance σ_j^2. To see the implications of allowing heterogeneity in variance for students with different performance, consider a higher scoring student (A) and a lower scoring student (B) who are looking at schools with past cutoffs equally far from each of them ($\text{score}_A - \text{cutoff}_j = \text{score}_B - \text{cutoff}_{j'}$). If they both fall on the $< 50\%$ probability of admission, A will think admission to j is less likely than B thinks is his admission probability to j' if A has less uncertainty than B. On the other side, if they are quite likely to be admitted, student A will be more certain of that than student B.

5.2 Choice Problem

5.2.1 Portfolio Choice

Applicants in the centralized system take in the vector of utilities v_i and cutoff clearance probabilities p_i and submit application portfolios of size $|R| = 10$ to maximize the following objective function:

$$\begin{aligned} V(R') &= v_{R'(1)} \cdot p_{R'(1)} + (1 - p_{R'(1)}) \cdot v_{R'(2)} \cdot p_{R'(2)} + \ldots + \prod_{l=1}^{|R|-1} (1 - p_{R'(l)}) \cdot v_{R'(|R'|)} \cdot p_{R'(|R'|)} \\ &= \sum_{l=1}^{|R'|} v_{R'(l)} \cdot p_{R'(l)} \prod_{k=1}^{|R'|-1} (1 - p_{R'(k)}). \end{aligned}$$

where $v_{R'(l)}$ denotes the utility from the lth most preferred program in the portfolio and $p_{R'(l)} = P(c_{R'(l)j} < \text{score}_i)$ denotes the probability of clearing the final cutoff of the lth most preferred program in the portfolio. The restriction on students filling their lists completely is a reasonable approximation in my empirical context as it is founded on the fact that the vast majority of students fill their lists ($> 80\%$) in the estimation period and less than 10% of students submit fewer than 7 programs in their application. In addition, the restriction of list-filling implies that some students may apply to programs they will not choose over the outside option which will help explain waitlist decisions to exit the system altogether than enroll in some available option. The probability of individual i choosing portfolio R_i is then:

$$\ell_i(R_i|\theta, z_i, \omega_i) = \mathbb{P}(V_i(R) > V_i(R')\forall R' \neq R).$$

5.2.2 Probability of Observed Choice in the First Round

At the first wave of admissions, students learn the set of programs A_{1i} they have been initially admitted to and their waitlist position in the programs to which they have not been admitted yet. They decide whether to enroll in one of the options to which they are admitted (a_j, s.t. $j \in A_{1i} \subseteq R_i$), forgo all such options and wait for a more preferred program in their list (w) or exit the admissions system altogether without enrolling in a program (e). In the offer acceptance and rejection stage of the first wave W_1, applicants take one of the possible actions $d_{1i} = \{a_j, w, e\}$ such that

$$d_{1i} = \underset{a_j,w,e}{\text{argmax}}\ V(W_1) = \begin{cases} v_{ij} & \text{if } d_{1i} = a_j \text{ s.t. } j \in A_{1i} \\ v_{i0} & \text{if } d_{1i} = e \\ V(R_i \backslash A_{1i}) & \text{if } d_{1i} = w \end{cases}$$

where

$$V(R_i \backslash A_{1i}) = \sum_{l=1}^{|R_i \backslash A_{1i}|} v_{R_i \backslash A_{1i}(l)} \cdot p_{R_i \backslash A_{1i}(l)} \prod_{k=1}^{|R_i \backslash A_{1i}|-1} (1 - p_{R_i \backslash A_{1i}(k)})$$

is the value of forgoing a current offer and waiting for future rounds. The expression for the value of future rounds comes directly from the perceived probabilities of clearing the admissions cutoffs for the programs that students had at the time of application. This expression makes the assumption that applicants' beliefs about the final cutoffs of programs they are waitlisted for do not change once they observe the cutoff for the first round. Two facts support this assumption: first, there are six additional waves of admission after the first and cutoffs in practice move dramatically between the first and the seventh wave for most programs.

Then the probability of observing d_{1i} conditional on the portfolio choice R_i and round 1 admission set A_{1i} is:

$$\ell_i(d_{1i}|A_{1i}, R_i, \theta, z_i, \omega_i) = P(d_{1i} = \text{argmax}\ V(W_1))$$

5.3 Identification argument

The identification challenge this paper faces is that of separating preferences for programs from beliefs about probabilities of admission to various programs. My strategy utilizes the three stages of decision-making modeled above: (1) decision to apply, (2) portfolio choice, and (3) waitlist decision to enroll or wait. On its own, none of the three stages allow separate identification of beliefs and preferences. In particular, portfolio choices arise from a combination of both beliefs and preferences and can be rationalized with many such combinations.

Separating preferences and beliefs: Identification of preferences and beliefs is done jointly using both portfolio choices and waitlist decisions. Waitlist decisions, in particular, help identification in two key ways. First, enrollment choices among students with multiple offers provide crucial information on preferences over programs by a simple revealed preference argument. It is important to clarify, however, that this is not sufficient for understanding whether and by how much the program that the student enrolled in is more preferred among all other programs given that each matriculation choice is made among a small and selected set of programs.[21] Incorporating the portfolio selection stage helps discipline the nature of selection into the waitlist stage choice sets. Second, it is helpful to observe waitlist decisions when students face both programs they've been admitted to and programs for which they remain on the waitlist. To illustrate why such decisions aid the separate identification of beliefs and preferences, consider a case in which (1) some students face a choice between program j, from which they have an offer among others, and waiting for program j', for which they are on the waitlist, and (2) other students face a choice between enrolling in either j or j', having been admitted to both. Differences in the extent to which students facing uncertainty for program j' are more likely to accept j (conditional on the application portfolio and observables) is determined by the probability of admission to j', which helps pin down $p_{j'}$.

In addition to waitlist choices, portfolio choices are used to identify substitution patterns across programs and heterogeneity in preferences. With relevance to heterogeneity by SES,

[21] The information obtained from matriculation choices is similar in nature to that obtained from observing the specific rankings of programs in a rank-order list under common DA implementations. Under the common assumption that students list the included programs in order of preference, an observed list provides information about the relative preference ordering of programs on the list, but not those programs among all in the choice set. This is the crux of the demand estimation challenge in school choice (see Agarwal and Somaini (2020) for a review of the evolution of approaches the school choice literature has taken to estimate demand).

observing more private university choices in the portfolios of higher-SES students than is predicted by lower-SES students' tastes implies a stronger taste for private education. More generally, the covariance between student and program characteristics identifies heterogeneous tastes for observable program characteristics. While the specification of my model does not include unobserved heterogeneity in preferences for program characteristics, it is quite possible to include such heterogeneity, which would be identified from seeing portfolios from observably similar students that have a concentration of programs with certain characteristics, or a starker absence of programs of other characteristics to an extent greater than can be predicted by a model without unobserved preference heterogeneity. Finally, it is worth noting the identification of the price coefficient. As discussed in Section 4, the policies on scholarships to private programs available to students with scores above a certain threshold provide exogenous variation in price for students on either side of the threshold,[22] which helps identify the price coefficient in equation (4).

Separating slope terms from intercepts in belief parameters: The main source of identification of slope parameters will come from choices across students of different performance. The variation in the distribution of selectivity of programs that are selected will be informative as the only way perceived admission probabilities will change across applicants facing choices in different ranges of school selectivities is through the slopes.[23] The extent to which higher performing students choose portfolios that are more selective relative to their own performance than lower performing students will be telling of the steepness of slopes for both mean shift and uncertainty parameters. To this end, choices on the waitlist across students of different performance will similarly aid identification. For example, the choice of higher performing students to wait for more selective schools relative to own performance than lower performing students would, pins down the differences in perceptions for higher cutoff schools.[24]

[22]In order to take advantage of this variation, I compute the price that would need to be paid by each student that clears the scholarship cutoffs for various scholarship amounts for each school. Using this constructed price as the price the students face for each private program assumes that they know they would be able to obtain these scholarships. This may not be an innocuous assumption if applicants are distrusting of the offer viability or exact amount.

[23]My estimation strategy does not involve computing the optimal portfolio. Instead this insight would be reflected as a higher likelihood of observing a deviation from the chosen portfolio toward a less selective reach program. See more details on estimation in Section 5.4.

[24]Variation in portfolio entries within student is also useful here. In particular, the extent of expansion on either side of the cutoff distribution relative to own score helps pin down how much admissions probabilities change along the cutoff line relative to own score. The addition of more selective reach choices than would be predicted optimal under a given set of beliefs indicates that students have either a more optimistic mean shift or a greater belief variance (assuming their scores fall below the shifted mean of the reach program

Separating shift parameters from variance parameters: The final challenge in identification is separating the mean shift in perceived distribution of potential cutoffs from the variance of this distribution. To help illustrate the identification point, Figure 4 shows how features of the selectivity of the chosen portfolio would vary with varying μ_0 and σ_0 holding slope parameters of the belief distribution constant. The key insight of the figure is that the most selective program, average selectivity, and least selective programs respond at different rates to changing σ_0 for the same mean shift and at different rates to changing μ_0 for the same variance. Even though there are two variance parameters for each mean shift that would yield the same *reach* program selectivity (panel (a)), only one variance guess of the two would be closer to yielding the data-observed selectivity of the *safety* school (panel (b)). In addition, across mean shifts some can be ruled out as fitting the data and among those that cannot, the likelihood-maximizing shift-variance pair will be one best-fitting all choices in the portfolio according to their likelihood contributions.

5.4 Estimation

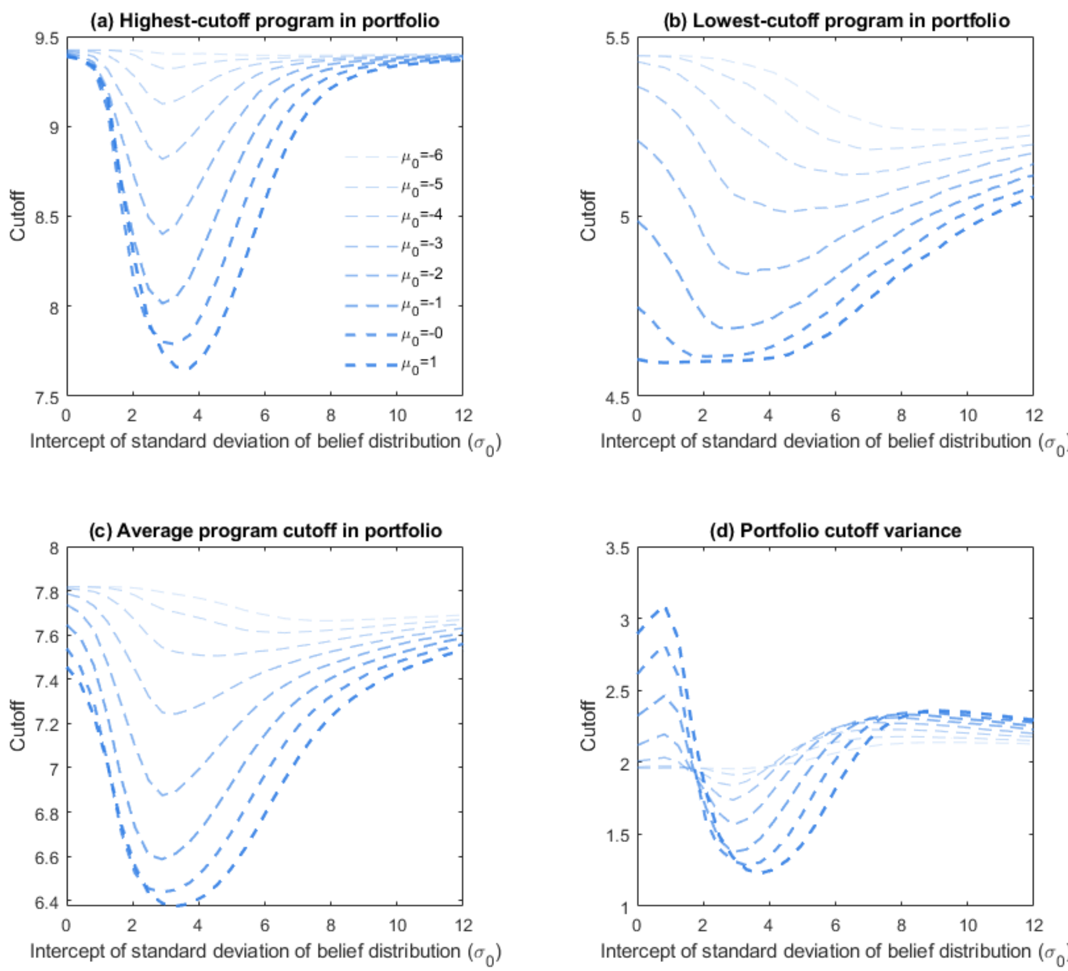
Bringing together the likelihood of observing student i making the decision of whether to apply, the likelihood of observing student i choosing the observed portfolio and the likelihood of student i making the observed wave 1 decision yields the following expression of the likelihood of observing the sequence of choices in the data for each student i:

$$\ell_i(\theta) = \int \ell_i(\text{apply}_i|\theta, z_i, \omega_i)\ell_i(R_i|\theta, z_i, \omega_i, \text{apply}_i)\ell_i(d_1|A_1, R_i, \theta, z_i, \omega_i, \text{apply}_i)dG_i(\epsilon)$$

Preferences and beliefs are jointly estimated via Simulated Maximum Likelihood (SMLE). The expressions for the likelihood function is not closed-form and an estimation method of this kind would call for drawing utilities many more times than there are portfolio alternatives given a guess of the parameter vector, computing the optimal portfolio and the probability that this portfolio is chosen among all possible candidates, and then computing the conditional probability of choice in the second stage. An obvious issue arises here. Given

included) about such choices. On the other end, if students are adding less selective safeties than would be predicted by a set of beliefs, then beliefs around less selective programs have a less optimistic mean or a higher variance (assuming student score falls above shifted mean of such programs). While it's not clear in principle whether the likelihood contributions of choices at the extremes of the portfolio are significant, the general argument holds for less extreme portfolio choices with cutoffs closer around the student's own performance: the distribution of selectivities of portfolio choices relative to own performance is a source of variation that helps identify the extent of decline in probability of admission along the cutoff line.

Figure 4
Portfolio selectivity features under varying belief parameters



Note: The figure shows portfolio selectivity features when varying belief parameters μ_0 and σ_0. The remainder of both preference and belief parameters are held constant. The x-axis varies the intercept of the standard deviation of the belief distribution (σ_0) while each line corresponds to a different value for the mean shift intercept (μ_0). Legend for all panels is in panel (a).

the number of possible programs to choose from, there are 10^{20} possible portfolio choices making computation of probabilities infeasible. I address it using a simplification introduced in Larroucau and Rios (2020). For rank order lists, the insight implies that it is sufficient for optimality that a rank order list is preferred to all portfolios created by a single-shot replacement of each ranked program in the list with a program not ranked. I adapt this insight to unordered portfolios and show that it holds for unordered portfolios too. Formally, using notation from Larroucau and Rios (2020):

Proposition 1. *Let $C = \{j_1, ..., j_k\}$ be an unordered application list of length at most K, i.e. $k \leq K$. Without loss of generality, let $u_{j_1} \geq u_{j_2} \geq ... \geq u_{j_k}$ so that the utility from submitting application portfolio C is*

$$V(C) = p_{j_1}u_{j_1} + (1 - p_{j_1})p_{j_2}u_{j_2} + ... + \left(\prod_{l=1}^{l=k-1}(1 - p_{j_l})\right)p_{j_k}u_{j_1}$$

If $\mathcal{S}(C)$ is the set of one-shot swaps of portfolio C and

$$V(C) \geq V(C') \ \forall \ C' \in \mathcal{S}(C),$$

then

$$V(C) \geq V(C') \ \forall \ \ C' \ \ s.t. \ \ |C'| = K.$$

Proof. See Appendix D.1 □

Estimation sample: Computation of the likelihood function remains challenging even after drastically reducing the choice set. Given that there are over 500 programs in the market, the likelihood of a portfolio being optimal even among its one-shot deviations must be calculated among more than 5000 alternatives.[25] Some simplifications are in order. I restrict the set of programs in each person's choice set to those in the region that the student went to high school and the capital. For example, I restrict the choice set of students from the north to the programs offered by the regional university in the north and the capital. This consists of 5 public universities (1 regional and 4 in the capital) and all of the private ones. For students in the south of the country, this implies 6 public universities (2 regional and 4 in the capital) and for students in the bigger central districts, this implies 6 public universities.[26]

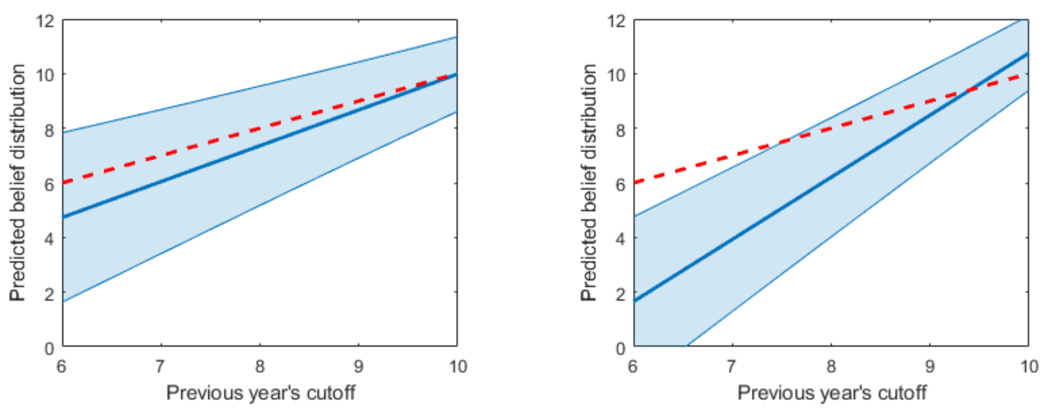
[25] The count of alternatives in the "choice set" is $|R'| \times (|\mathcal{J}| - |R'|)$, which in context would imply $10 \times (517 - 10)$ alternatives.

[26] This reflects the fact that in the Albanian context, students from the north will almost never apply

6 Results from the model

Predicted beliefs using model estimates are shown in Figure 5 and the remainder of the parameter estimates can be found in Appendix Table A-7. I find a substantially dispersed distribution of possible cutoffs. In fact, the distribution of cutoffs falls outside the range of possible grades, $[4, 10]$ for programs with cutoffs at the extremes of the range. Naturally, no student would believe that the cutoff of a program could be above the range of possible scores with any non-zero probability. I interpret the existence of probability mass outside of the extremes of possible scores as simply and indication that the highest scoring students do not believe they have a probability 1 of being admitted to all programs, and similarly, that the lowest scoring students do not believe they have 0 probability of getting into any programs. In addition, I find that lower-SES students are more optimistic about programs with low cutoffs and more pessimistic about programs with high cutoffs. The variance of the distributions is held constant across groups as estimates were performing poorly for the high-SES group when estimated separately.

Figure 5
Prediction of perceived program cutoff distribution using model estimates



Note: Plot shows the model-predicted distribution of perceived program cutoffs as a function of the previous year's cutoff. High-SES students' beliefs are plotted on the left panel and lower-SES students' beliefs on the right. The blue line in the middle of the shaded area is the mean of the distribution calculated as $\bar{c} = \text{cutoff}^{y-1} + \hat{\mu_0} + \hat{\mu_1} \times \text{cutoff}^{y-1}$. The shaded area reflects the scores that are within a standard deviation of the mean for the estimated standard deviation of the beliefs. I compute the bounds of the shaded area as $\bar{c} \pm \log(1 + \exp(\hat{\sigma}_0 + \hat{\sigma}_1 \times \text{cutoff}^{y-1}))$. I add a y=x line as a reference.

to regional universities in the south of the country and vice-versa. In addition, the choice sets within the relevant universities are restricted to reflect the differential choice sets relevant for students of different high school tracks. Those of the science track will never apply to certain humanities degrees in any of the years in the data, both before and after the reform. In addition, students from a social science high school track will never be observed applying to certain science degrees.

7 Counterfactual results

7.1 Evaluating trade-offs of partially and fully centralized admissions

Using the estimates for the structural model, I simulate choice and allocations in counterfactual designs and evaluate the role of private outside options on the choices, assignment outcomes and welfare of students graduating high school in Albania. First, I describe the setup for the counterfactual simulations and the measure of welfare I use and then present results from the simulations.

7.1.1 Counterfactual setup

I conduct counterfactual simulations using the sample of students who graduate high school in 2019. The preference parameters, including the distribution of the random taste parameters are held fixed in the simulations. Similarly, the distribution of beliefs is also held invariant to policy changes. The beliefs in estimation are captured in reduced form and aim to isolate the level of uncertainty and bias carried year-to-year in this market. I compare the performance of two market structures. The baseline structure is one in which students are allocated to college seats in two procedures simultaneously: the centralized assignment to public programs with restricted lists and a centralized assignment to private programs where there is no constraint on the number of applications one can submit. The alternative market structure is one in which students are allocated to college seats in a single procedure where their list sizes are constrained and they apply to both public programs and private programs in the same platform. I describe the two market structures in more detail below.

Centralized Public Match with Private Outside Options (baseline configuration): The baseline market is one that assigns students to seats according to a Deferred Acceptance mechanism in two parallel assignments. The mechanism proceeds as follows:

1. Students apply to ten programs in the public match and to all available private programs in a simultaneous private match.

2. Applicants to the public match are ranked according to the pre-determined formulas of each program and an initial placement to public programs is made through a DA algorithm.

3. Private programs simultaneously rank all students according to their formulas and determine student priorities

4. Once students have their initial placement in a public program, the private options begin proposing to students in order of score priority and the outside option proposes to everyone. Once private programs and outside options begin proposing, the allocation evolves as follows

   - If the initial public placement is a student's first choice among the set of all private programs, the outside option, and the public programs to which the student has applied then this initial placement is the student's final assignment.

   - If one of the private programs proposing is a student's first choice in the set above, then the student enrolls in that program, foregoing her placement in the public match. Similarly, if the outside option is the best option in the set then the student exits the mechanism foregoing all inside options. If the private proposal is better than the public proposal, the student temporarily holds the private offer and forgoes the public offer.

   - The public match fills vacancies created in the first stage of private program proposal by reassigning all students except those with final assignments in the first stage.

   - Private programs similarly reassign all temporarily assigned or unassigned students

   - The second stage of outside option proposal proceeds the same way as the first. Private programs and outside options propose anew and students accept only if the program proposing is their favorite program. Otherwise they hold their best offer temporarily.

   - The rounds of assignments proceed until either each student is enrolled in their best option that they applied to or private option conditional on clearing the program's cutoff, or the student has exited the mechanism, or has remained unassigned and no program that they applied to or any private program they prefer to the outside option would admit them.

A few notes are worth making about the above mechanism. First, the assignment procedure above is frictionless. That is, I assume a centralized market for outside options that is coordinated with the public match. This excludes the possibility that a student may receive an offer that they accept from a private program and there is a vacancy they leave behind in the public match that never gets filled. It also excludes the possibility of congestion in the private market and abstracts away from a situation in which offers may be made in the private market to students who apply first and may not be as qualified as students who may apply later do not get accepted due to the timing of their application to a private program. These types of aftermarket frictions are outside of the scope of this paper. The goal of the counterfactual exercises presented is to understand the effect of *strategic applications* that result from the quality of outside options and list size constraints.[27] In fact, the mechanism described above is equivalent to a one-stage deferred acceptance mechanism in which students submit a single application list to the platform in which they are allowed to include up to 10 public programs and all private programs.

Proposition 2. *A parallel mechanism for public programs with a maximum list size of 10 and unrestricted list size for private options is equivalent to a DA algorithm with 10 slots reserved for public programs and an unrestricted number reserved for private programs.*

The only difference of this mechanism with an unrestricted DA is what arises at the application stage due to list size restrictions, incorrect beliefs, and preferences for private options.

The final note that needs to be made about the counterfactual simulations is the assumption that beliefs are invariant to policy changes. In this counterfactual, I assume that both beliefs about the cutoffs of public programs and private programs are the same as those estimated. This has implications for applications not only directly through how students weigh the probability of admission to programs in the restricted-list application, but also through the value from the private options. Results are qualitatively the same in two alternative counterfactuals: (1) one in which the students' value from the private outside options is the utility from their favorite private option assuming that they do not consider the probabilities of admission to private programs, and (2) a counterfactual in which students carry rational expectations beliefs about the outside options. Next, I describe the second main counterfactual, the "all-in" configuration:

[27] A different study, Kapor et al. (2022) studies the role of aftermarket frictions on the allocation of students in partially centralized admissions. They find that outside options generate frictions in the centralized match through creating vacancies that need to be filled by creating other vacancies through extracting students from their assigned seats. My paper abstracts away from these aftermarket frictions and assumes that the chains of reassignment happen through a whole-market reassignment with no frictions.

The All-In Match (no-private outside options configuration): This no-private outside options configuration allows students to apply to college only through the single restricted-list match. The match assigns students to seats via a Deferred Acceptance mechanism. The only non-platform option is the no-college outside options. This procedure is more straightforward than the one with outside options as all options are on-platform and no assumptions have to be made about the allocations to programs outside the platform. The assignment proceeds as follows:

1. Students apply to ten programs only among both public and private programs in a single application on the "all-in" platform.

2. Applicants are ranked according to the pre-determined formulas within each program and an initial placement is made through a DA algorithm.

3. Once students have their initial placement in a public program, the outside option proposes and students who prefer the outside option exit the mechanism altogether

With the details of the two counterfactual assignment mechanisms established, it is worth emphasizing that the only difference in assignment generated by differences in applications, rather than any differences in the assignment process. This is unlike assignment processes with centralized public and decentralized private markets, but I make this choice in order to isolate the changes in assignments only through changes in application. In the end, final assignments may change in the case with outside options if congestion effects cause unstable matches.

7.2 Effects of incorporating all programs into the same platform

In this section, I report the descriptives of application patterns under the two alternative market structures, the final assignments with the assignment algorithms described above and the final welfare results.

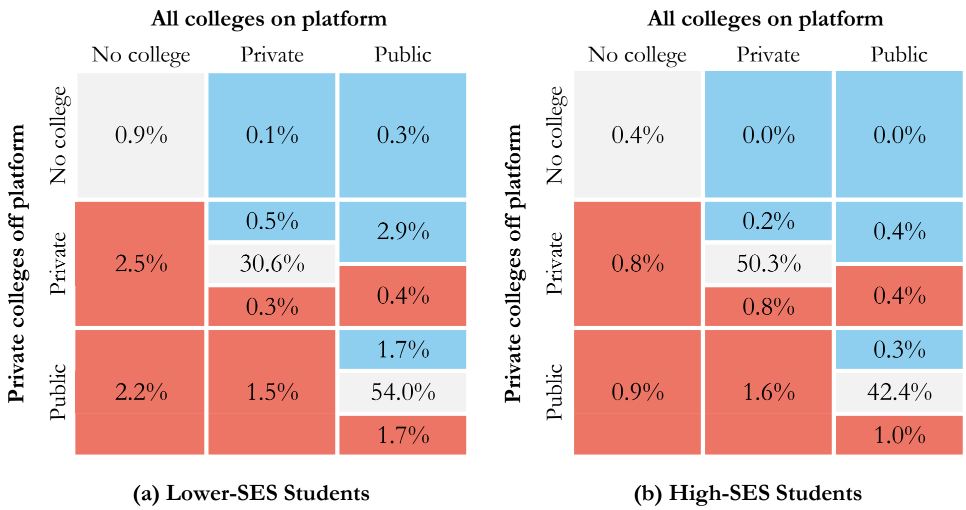
7.2.1 Changes in the allocation of students

The assignment results show that changing the market structure to an all-in system reduces matching efficiency relative to the counterfactual with a partitioned market, but more so for high-SES students, thus improving equity. Overall, the assignment results show that there is a decrease of 4.2pp in enrollment to college for lower-SES students and 1.7pp for high-SES students. In addition a net of 2.9% more high-SES students had worse assignments than had better ones whereas 1.1% more lower-SES students had better assignments than had worse assignments. The assignment results show that more people lose than gain in both groups, and this happens through two channels: (1) students whose applications became less selective and who excluded programs they would have been admitted to and that they preferred relative to the one they were assigned to; and (2) students who were displaced because more people are applying to less selective colleges. The reduction in enrollment is more prevalent for lower-SES students while the worsening of matches in an all-in system happens on net for the group of high-SES students alone, who have more inside options they are willing to accept inside the platform, but who are more constrained in their applications.

Figure 6 shows more detail on the above results and carefully documents differences in matches when going from a partitioned system to an all-in system. When the system changes, there are both winners and losers that result directly from changes in applications as well as indirectly from changes in the applications of others. Among lower-SES students, 4.6% are induced to not enroll in college and among high-SES students, 1.7% are induced to end up unenrolled. Even though applications changed for this group, the spillover effects from others dominate the conservativeness of their applications and they end up not being enrolled anywhere, whereas they would have attended college in a partitioned system where others would have applied to and gone to different, potentially more selective colleges allowing them to not get pushed out of going to college. The spillover effects serve not only to induce people to end up not enrolling in college, but also to induce some people who would have otherwise not gone to college to enroll. These benefits accrue more to lower-SES students who are more likely to keep applying the same way and are only exposed to the spillover effects from others: 0.4% of lower-SES students go to college in an all-in configuration that otherwise would have not. The final net effects of an all-in policy are to reduce college enrollment through constraints that induce students to change their applications as well as spillovers onto others of such changes.

By similar arguments, even the set of people who would enroll in college in both configurations observe both winners and losers. Of particular note in Figure 6 is the set of students who go from enrolling in a public program to enrolling in a private program. All the students in this group experience a worse assignment from the policy change. In the partitioned market they would have rejected their best possible private option for the offered public option,

Figure 6
Assignment results changing market structure from partitioned to all-in



Note: This figure shows the share of winners (blue) and losers (red) among graduating high school students when moving from a partitioned market structure to an all-in one. The share of students with an identical outcome under both regimes are shown in gray.

but in the all-in case they either never applied to this public option, or were pushed out of it. 1.6% of high-SES students and 1.5% of lower-SES students experience such a change, which is due to constrained applications–high-SES students must substitute away from public programs in their applications and toward private programs more frequently than lower-SES students and as such forgo a public assignment more frequently.

Finally, I discuss the outcomes of those that enroll in the public system in both structures. Spillover effects and application effects net out for lower-SES students, but a net of 0.7% high-SES students lose. The losses here come from reducing the size of the public-only subset of one's application and changing its composition. The constraints on this subset of the application are more binding for high-SES students who also prefer private programs more and will have to incorporate them to a larger extent in their applications.

In summary, strategic applications cause a net loss when incorporating outside option colleges in a central system. A higher share of high-SES students lose because their application behavior is the most affected. Lower-SES lose mostly through lower enrollment, but gain conditional on matching due to spillovers from others. Three forces are at play. First, applications to public programs become less selective due to both the worsening of outside options and a more binding list-size constraint. This results in some assignments that are worse because students have removed from their applications programs they would have preferred and been admitted to. Second, the new restriction on private applications induces misrepresentation of preferences among private options. Third, students are crowding their applications more toward lower selectivity programs causing spillover effects that push out of admission students that would otherwise have been admitted to certain programs. In all, these forces end up causing more losers than winners, on net 3.1% for the lower-SES group and 4.6% for the high-SES group.

7.2.2 Changes in welfare

I compute the average student welfare as the average ex-post utility from assignment under each regime. Formally, welfare is computed as:

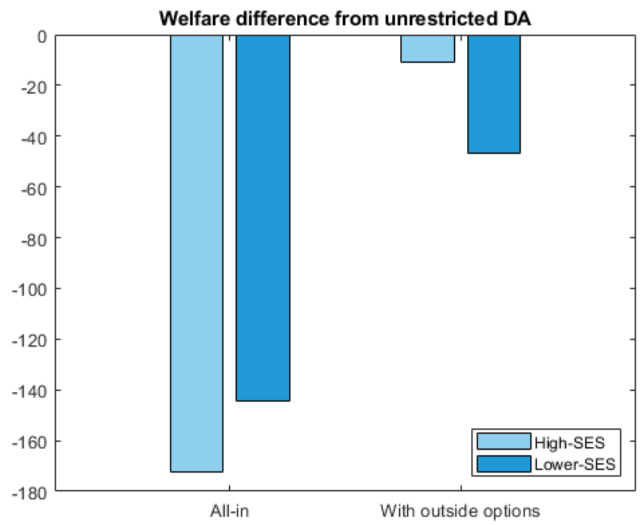
$$W(M) = \frac{1}{N} \sum_{i=1}^{N} E\left[v_{i,f(i)}\right]$$

where M is the market structure, $M \in \{\text{partitioned}, \text{all-in}\}$ and $f(i)$ is student i's final assignment $f(i) \in \varnothing \cup \mathcal{J}$.

Figure 7 presents welfare differences relative to welfare realized under an unrestricted DA algorithm. In the partitioned market, high-SES students achieve more of the gains possible than lower-SES students, but this difference reverses when moving to an all-in system. This reversal in relative gains, comes at an efficiency cost: the gap in realized gains relative to the unrestricted DA for both groups increases. For high-SES students, in particular, the welfare gap increases by €160, while for lower-SES students it increases by €98. The findings from the welfare calculations imply that a policy that aims to improve equity through disallowing a market outside of the centralized match to exist comes at an efficiency cost due to the application and information constraints.

A note of caution should be added here about interpreting these results. They do not reflect equilibrium behavior because in equilibrium students may update their beliefs about admission probabilities and reoptimize in a way that may change the direction of these qualitative results. Nevertheless, given the heterogeneity of beliefs by SES group and substantial uncertainty about cutoffs, it is hard to imagine that expectations are rational. Second, the counterfactual exercise abstracts away from improvements in efficiency from the reductions in

Figure 7
Welfare results under two market structures by SES



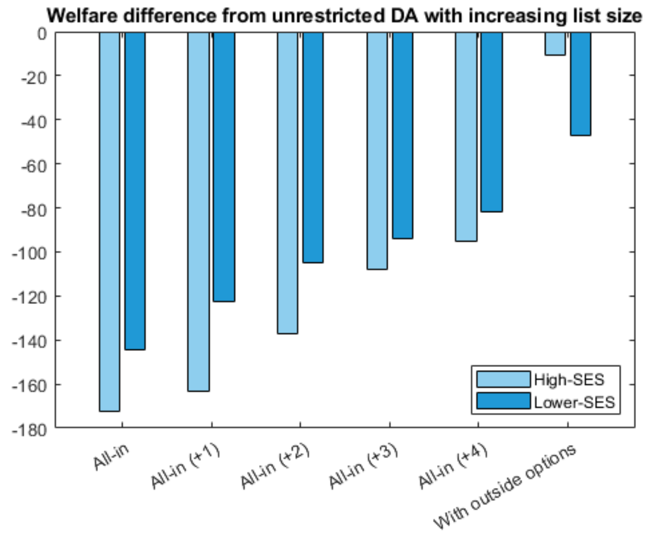
Note: The figure shows welfare differences in Euros between outcomes from an unrestricted DA mechanism and each of the two counterfactual market structures by SES group.

matching frictions that come from centralization. These improvements are well-documented in the literature (Abdulkadiroğlu et al. 2017; Kapor et al. 2022). The lesson is to highlight a channel of behavior that market designers must take into account when choosing their school choice or college admissions system. In the Albanian context, despite substantial uncertainty about cutoffs, outside options have only a small role in choice.

Welfare under alternative list sizes: I use the estimated model to evaluate alternative market structures that may improve student equity at a lower efficiency cost for the market. I evaluate alternative list sizes the size constraint with an all-in configuration. In principle, strategic effects coming from list size restrictions can be completely neutralized with an unrestricted list, but policymakers may not want to give unlimited choice to students.[28] Figure 8 shows the welfare effects of adding choices to the application list incrementally. Allowing just four more choices in the all-in system closes the welfare gap with the unrestricted DA system for lower-SES students by more than half relative to a partitioned system. The gains for high-SES students need more options to converge to a partitioned system because they value unlimited choice of private options more than lower-SES students do.

[28]There could be many reasons for this, which are not modeled in this paper. Allowing an unlimited number of choices may impose a cognitive burden on students searching for programs and may advantage those who have more resources and help and are able to obtain more information and better able to rank many choices. These forces are not modeled in the paper and I assume that allowing a few more marginal choices does not impose additional cost on students searching.

Figure 8
Welfare results by SES under varying list sizes and an all-in structure



Note: This chart shows welfare differences from welfare achieved under an unrestricted DA with an all-in structure for assignments under varying list sizes for high and lower-SES groups.

8 Conclusion

In this paper I evaluate the role of market structure in strategic choice and allocation in partially centralized college admissions with list size restrictions and private schools as outside options. The insight is that for applications within the match, it may matter strategically what outside options a student has. Students with better outside options can apply more ambitiously within the centralized platform and ultimately be assigned to more preferred programs than students with worse outside options. This strategic response may have significant efficiency and equity consequences, given that private outside options are expensive and offer higher value to students with higher socioeconomic status (SES). To evaluate the effects of outside options in strategic applications, I use novel data from Albania and a policy that incorporated all private colleges into the public centralized assignment while maintaining the same list size restriction. I combine a reduced-form analysis and structural model to evaluate strategic behavior, strategic advantages of students with better outside options, and their effect in welfare and equity.

I first provide descriptive evidence of differences in the applications of high and lower-SES students and the quality of public institution that they attend in the centralized system with restricted application sizes when private programs are excluded from the platform. Higher SES students apply to and enroll in more selective free public institutions than their lower-

SES peers with the same high school performance. This striking fact motivates the rest of my analysis.

I analyze a policy change implemented in the centralized match in Albania that enforced participation by all colleges, public and private, in the centralized platform. I use an event study design to measure the effects of centralizing all available alternatives on application behavior and match outcomes. For this event study analysis, I compare applications of high-SES students to lower-SES students before and after the reform. In addition, to reduce concerns over confounding effects of other market-wide changes that may affect different SES groups differently, I take advantage of merit-based scholarships that ensure that top students of all backgrounds have equal access to private college seats to argue that top students from high- and lower-SES backgrounds apply to similarly competitive programs before and after the match expansion and they are not differentially affected by other aspects of the reform. For all lower-performing students, the desirability of private decentralized options depends on SES status, so the theoretical framework predicts that the removal of private colleges as outside alternatives to the match has differential effects on application behavior of high-SES and lower-SES students.

The event study captures the differential changes in application behavior of high- and lower-SES students that come from changes in the desirability of outside alternatives. This exercise offers several suggestive findings. First, outside alternatives affect application behavior and matches within the centralized portion of the market. Second, when outside alternatives are private and costly, their desirability will break along SES lines, yielding different strategic application behavior for high- and lower-SES students. Outside alternatives are more desirable for high-SES students, giving them not only higher direct value from choosing these options, but also a strategic advantage over lower-SES students within the match to public colleges and majors. Third, participation in the match by all colleges reduces differences in application behavior, primarily driven by a reduction in aggressiveness of applications of high-SES students.

In the second part of the paper, I quantify the welfare and distributional impacts of the existence of private outside options and list size restrictions. I build a structural model of student applications and matriculation choices and estimate it using data and institutional features from both before the policy change and after it. My model captures student choices that balance preferences for college-major pairs and beliefs about probability of admission to each option. I advance the literature by relaxing assumptions of truth-telling or rational expectations, and instead rely on institutional features to separately identify expectation formation and preferences. Specifically, I develop an estimation procedure to identify preferences for college-majors using post-reform data, when market-clearing procedures change, and students clear the market in rounds of observing multiple offers and choosing to enroll or wait for a better offer. The setting allows me to observe direct choice between options and estimate preferences using standard revealed preference methods.

The estimated structural model enables me to conduct counterfactual analyses. First, I quantify the heterogeneous welfare and distributional effects of a centralizing policy change when all outside alternatives are private. Even further, extrapolating outside of my project's empirical setting, I simulate counterfactual policies that bring lower-SES students closer to the first best without sacrificing the welfare of high-SES students.

References

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A. Pathak, "The welfare effects of coordinated assignment: Evidence from the New York city high school match," *American Economic Review*, 2017, *107* (12), 3635–3689.

Agarwal, Nikhil and Paulo Somaini, "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism," *Econometrica*, 2018, *86* (2), 391–444.

_ **and** _ , "Revealed preference analysis of school choice models," *Annual Review of Economics*, 2020, *12*, 471–501.

Ajayi, Kehinde and Modibo Sidibe, "School Choice Under Imperfect Information," *SSRN Electronic Journal*, 2020.

Akbarpour, Mohammad, Adam Kapor, Christopher Neilson, Winnie Van Dijk, and Seth Zimmerman, "Centralized School Choice with Unequal Outside," 2021, (December).

Andersson, Tommy, Umut Dur, Sinan Ertemel, and Onur Kesten, "Sequential School Choice with Public and Private Schools," *Working Papers*, 2019, (2018:39), 1–40.

Calsamiglia, Caterina, Chao Fu, and Maia Güell, "Structural estimation of a model of school choices: The boston mechanism versus its alternatives," *Journal of Political Economy*, 2020, *128* (2), 642–680.

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

Chen, Yan and Onur Kesten, "Chinese college admissions and school choice reforms: A theoretical analysis," *Journal of Political Economy*, 2017, *125* (1), 99–139.

Fack, Gabrielle, Julien Grenet, and Yinghua He, "Beyond truth-telling: Preference estimation with centralized school choice and college admissions," *American Economic Review*, 2019, *109* (4), 1486–1529.

Idoux, Clemence, "Integrating New York City Schools: The Role of Admissions Criteria and Family Preferences," 2022.

Kapor, Adam J., Christopher A. Neilson, and Seth D. Zimmerman, "Heterogeneous beliefs and school choice mechanisms†," *American Economic Review*, 2020, *110* (5), 1274–1315.

Kapor, Adam, Mohit Karnani, and Christopher Neilson, "Aftermarket Frictions and the Cost of Off-Platform Options in Centralized Assignment Mechanisms," Technical Report 2022.

Larroucau, Tomás and Ignacio Rios, "Do "Short-List" Students Report Truthfully? Strategic Behavior in the Chilean College Admissions Problem," 2020, (2018), 1–65.

Luflade, Margaux, "The Value of Information in Centralized School Choice Systems," *Working paper*, 2017, (August), 1–72.

Otero, Sebastian, Nano Barahona, and Caue Dobbin, "Affirmative Action in Centralized College Admission Systems: Evidence from Brazil," 2021.

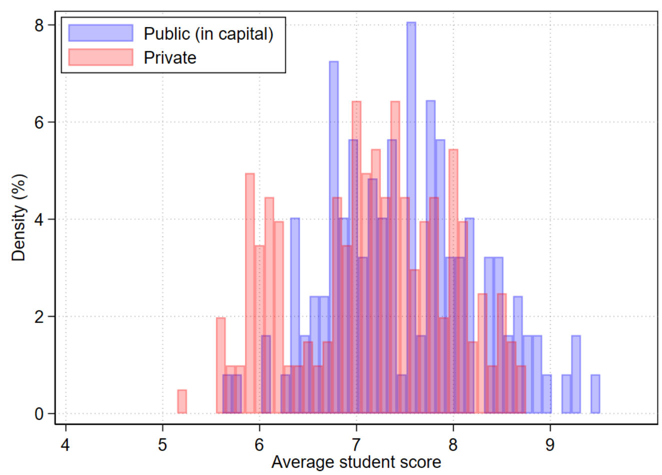
Waldinger, Daniel, "Targeting in-kind transfers through market design: A revealed preference analysis of public housing allocation," *American Economic Review*, 2021, *111* (8), 2660–2696.

APPENDIX

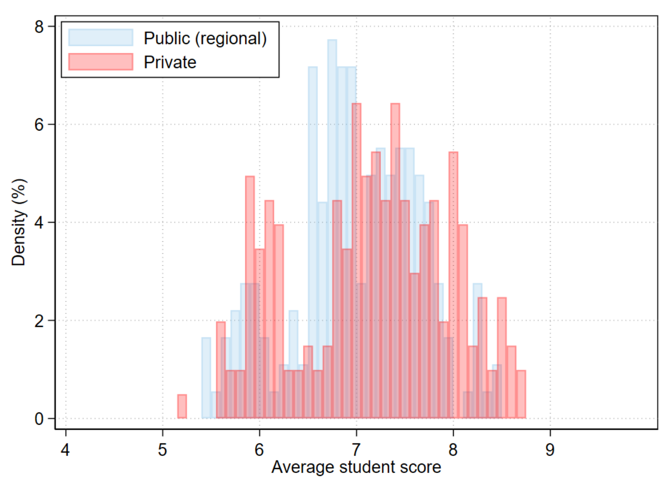
A Additional Tables and Figures

Figure A-1
Quality of students enrolled in programs by public status and geography

(a) Public programs in capital and all private programs

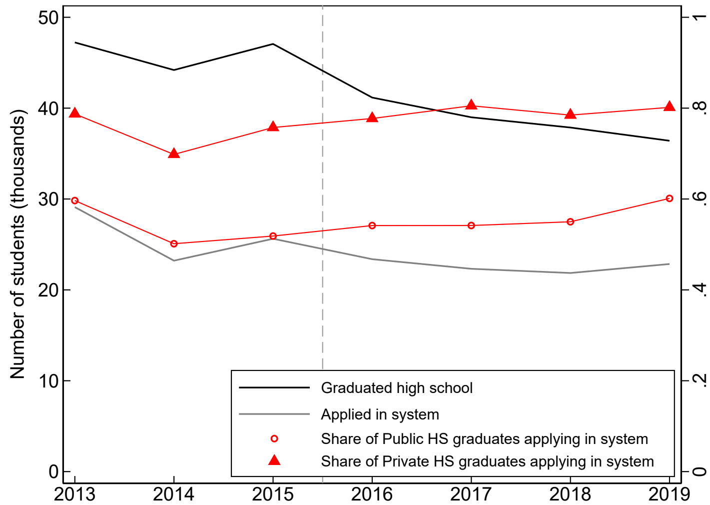


(b) Regional public programs and all private programs



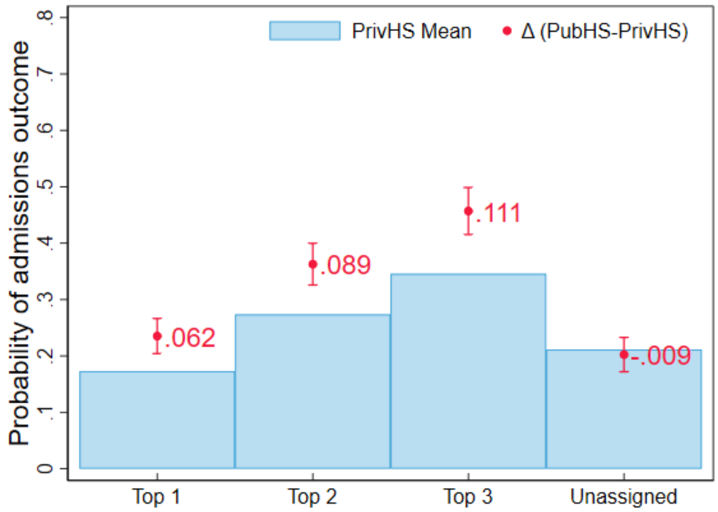
Note: This chart displays the distribution of the simple average of the weighted average score for enrollees in public university programs (blue) and private university programs (pink) in 2019. The top panel compares public programs in the capital to all private programs. 25 of 26 private universities are located in the capital and 7 out of 12 public universities are located in the capital. The bottom panel compares the average score of enrollees in private programs with enrollees in regional public universities.

Figure A-2
High School Graduates and College Applicants over Time



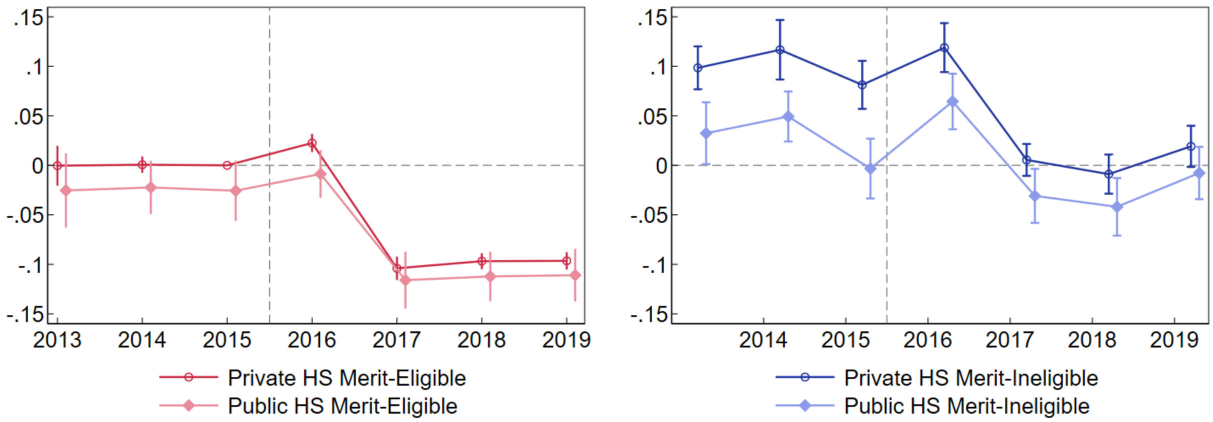
Note: Chart shows trends in number of students graduating high school and those applying to college through the centralized system. In shades of gray are the total number of students who graduated high school and the number that applied through the platform; in red, the share of graduating public HS students and private HS students separately that applied through the centralized platform.

Figure A-3
Final assignment outcomes in capital



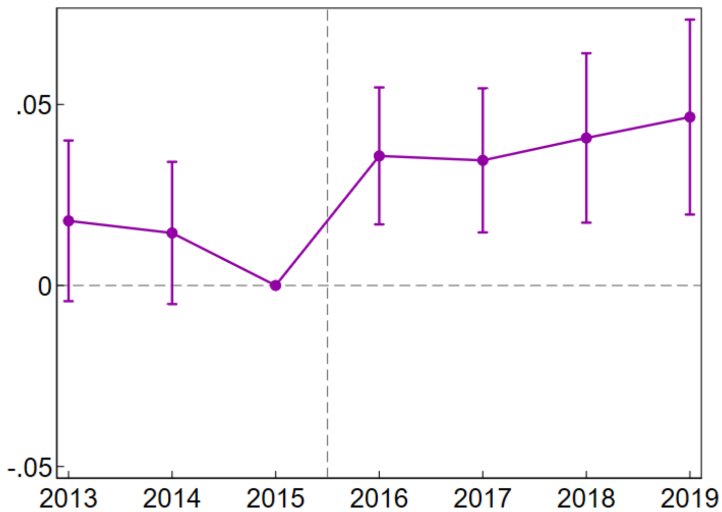
Note: Differences are conditional on score and district FE. Sample includes years 2013-2015. 95% confidence intervals are shown with standard errors clustered at the high school level. Score is the weighted average of test scores in Math, Language, choice subjects, and HS GPA.

Figure A-4
Selectivity of Most Selective Public Program



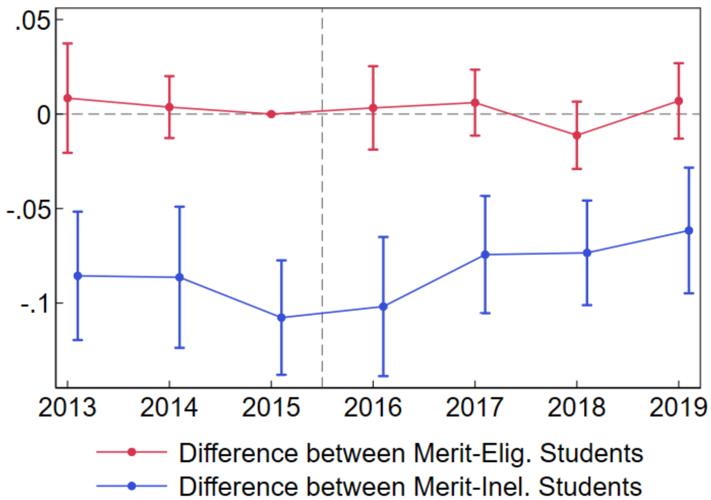
Note: Charts show selectivity of most selective public program for each of private high school top students, public high school top students, private high school non-top students, public high school non-top students. Top students are the set of students that would qualify for merit scholarships at private institutions based on their exam scores.

Figure A-5
Event Study of Triple Differences in Selectivity of Most Selective Public Programs
Top and Non-Top Students



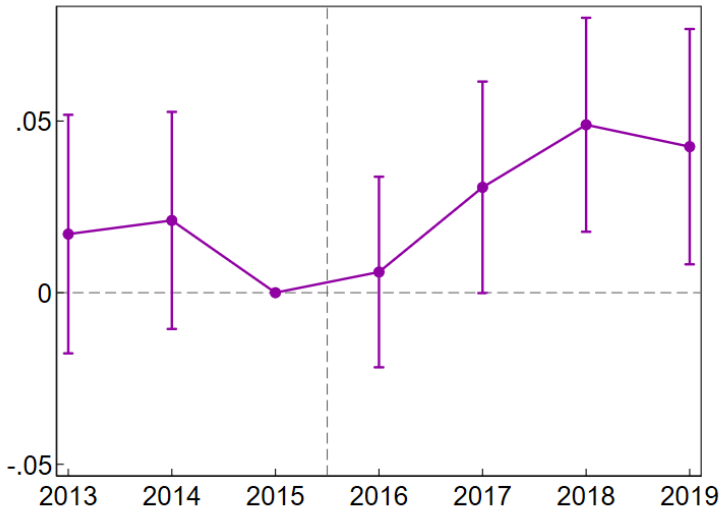
Note: This chart plots the differences in selectivity of most selective public programs on the centralized application between private high school and public high school students relative to the difference in the year before the reform (2015). Higher values reflect a reduction of the gap between the top choices of high- vs. lower-SES students. Regressions control for average exam score and include district FE. Standard errors clustered at district level.

Figure A-6
Outcome: average historical selectivity of programs in application
Outcome measure: program's standardized cutoff score in 2013



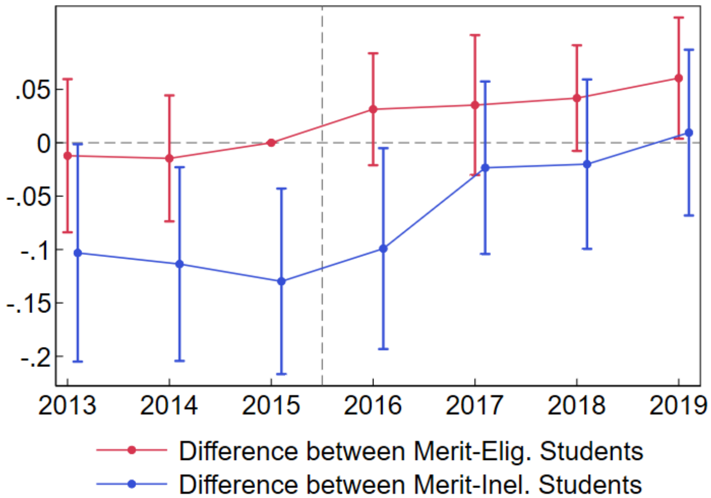
Note: This chart plots the differences in average selectivity of public programs included in the portfolio on the centralized application between private high school and public high school students. Negative differences reflect less selective portfolio choices for public high school students. Regressions control for average exam score and include district FE.

Figure A-7
Double differences: portfolio selectivity



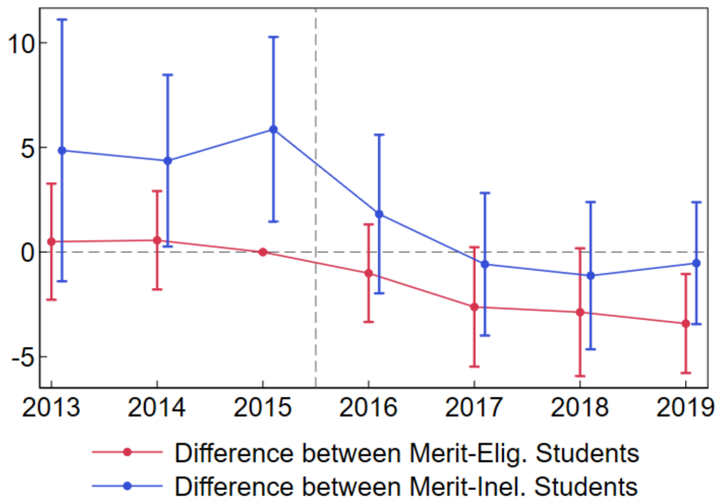
Note: This chart plots the differences in average selectivity of the public portion of the programs included in the portfolio on the centralized application between private high school and public high school students relative to the difference in the year before the reform (2015). Higher values reflect a reduction of the gap between the choices of high- vs. lower-SES students. Regressions control for average exam score and include district FE. Standard errors clustered at district level.

Figure A-8
First differences: most selective program
Outcome measure: standardized previous year's cutoff



Note: This chart plots the differences in most selected public programs included in the portfolio on the centralized application between private high school and public high school students. Negative differences reflect less selective portfolio choices for public high school students. The alternative measure of selectivity used in this graph is the standardized previous year's cutoff for each program. Regressions control for average exam score and include district FE.

Figure A-9
First differences: most selective program
Outcome measure: previous year's rank



Note: This chart plots the differences in most selected public programs included in the portfolio on the centralized application between private high school and public high school students. Negative differences reflect less selective portfolio choices for public high school students. The alternative measure of selectivity used in this graph is the previous year's rank for each program. Regressions control for average exam score and include district FE.

Table A-1
Relationship Between Attending a Public HS and Assignment Outcomes
Robustness Check for Initial Assignment Outcome

	National Outcomes			Outcomes in Capital		
	Assigned to First Choice	Assigned to One of Top Three Choices	Unassigned	Assigned to First Choice	Assigned to One of Top Three Choices	Unassigned
Public HS	0.032***	0.038***	-0.017	0.022	0.029*	-0.006
	(0.011)	(0.012)	(0.010)	(0.017)	(0.016)	(0.020)
Private HS Mean	0.298	0.540	0.188	0.192	0.404	0.288
Adjusted R^2	0.125	0.185	0.263	0.128	0.248	0.314
Observations	84,931	84,931	84,931	17,336	17,336	17,336
Condition on Score	✓	✓	✓	✓	✓	✓
District FEs	✓	✓	✓			
Year FEs	✓	✓	✓	✓	✓	✓

Notes: Sample includes years 2013-2015. Private HS Mean is the unconditional mean of the outcome variable for students attending private high schools. Standard errors are clustered at the district level for the national sample and are robust for the capital-only sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A-2
Relationship Between Attending a Public HS and Assignment Outcomes
Robustness Check for Final Assignment Outcome of Those Who Did not Reject Centralized Offer

	National Outcomes			Outcomes in Capital		
	Assigned to First Choice	Assigned to One of Top Three Choices	Unassigned	Assigned to First Choice	Assigned to One of Top Three Choices	Unassigned
Public HS	0.028**	0.042***	-0.030***	0.026	0.043**	-0.029
	(0.012)	(0.012)	(0.010)	(0.021)	(0.021)	(0.019)
Private HS Mean	0.340	0.588	0.159	0.218	0.436	0.268
Adjusted R^2	0.124	0.181	0.249	0.142	0.255	0.307
Observations	76,556	76,556	76,556	15,498	15,498	15,498
Condition on Score	✓	✓	✓	✓	✓	✓
District FEs	✓	✓	✓			
Year FEs	✓	✓	✓	✓	✓	✓

Notes: Sample includes years 2013-2015. Private HS Mean is the unconditional mean of the outcome variable for students attending private high schools. Standard errors are clustered at the district level for the national sample and are robust for the capital-only sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A-3
Relationship Between Attending a Public HS Application Selectivity Measures

	National Sample					Applicants from Capital				
	Listed Med.	Top Ranked	Second Ranked	Third Ranked	App. Sel.	Listed Med.	Top Ranked	Second Ranked	Third Ranked	App. Sel.
Public HS	-0.013***	-0.118***	-0.121***	-0.106***	-0.092***	-0.026***	-0.128***	-0.117***	-0.065***	-0.061***
	(0.005)	(0.023)	(0.021)	(0.024)	(0.018)	(0.006)	(0.019)	(0.020)	(0.019)	(0.009)
Score	0.094***	0.555***	0.571***	0.525***	0.500***	0.099***	0.460***	0.499***	0.454***	0.430***
	(0.002)	(0.024)	(0.020)	(0.020)	(0.019)	(0.002)	(0.006)	(0.006)	(0.006)	(0.003)
Priv. HS Mean	0.141	6.782	6.742	6.661	6.587	0.139	6.985	6.967	6.875	6.792
Adjusted R^2	0.170	0.371	0.377	0.362	0.631	0.170	0.301	0.314	0.299	0.614
Obs.	84,931	84,563	84,073	83,244	84,909	17,336	17,307	17,145	16,985	17,336
District FEs	✓	✓	✓	✓	✓					
Year FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Sample includes years 2013-2015. Private HS Mean is the unconditional mean of the outcome variable for students attending private high schools. Standard errors are clustered at the district level for the national sample and are heteroskedasticity robust for the capital-only sample. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A-4
Relationship Between High School GPA and Results in the Matura Exams by Type of High School

	Exam mean	Math score	Literature score	First elective score	Second elective score
hs_gpa	0.707***	0.569***	0.911***	0.652***	0.694***
	(0.021)	(0.018)	(0.030)	(0.024)	(0.022)
Public HS	0.932***	0.325*	1.416***	0.815***	1.175***
	(0.178)	(0.185)	(0.248)	(0.212)	(0.194)
gpa_pub	-0.095***	-0.009	-0.161***	-0.086***	-0.123***
	(0.023)	(0.019)	(0.034)	(0.026)	(0.025)
Adjusted R^2	0.621	0.524	0.518	0.359	0.363
Obs.	84929.000	84901.000	84896.000	84875.000	84842.000
Year FEs	✓	✓	✓	✓	✓

Notes: Sample includes years 2013-2015. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A-5
Tests of the Parallel Trends Assumption on Selectivity of "Reach" Schools

	(1)	(2)	(3)
Public HS × Non-Top × Year	-0.030	-0.009	0.012
	(0.021)	(0.022)	(0.029)
Public HS × Non-top	-0.060	-0.106	-0.145*
	(0.084)	(0.079)	(0.079)
Public HS × Year	0.014	-0.004	-0.023
	(0.022)	(0.017)	(0.015)
Non-Top × Year	-0.075***	-0.071***	-0.086***
	(0.022)	(0.021)	(0.026)
Public HS	-0.030	0.015	-0.052
	(0.047)	(0.034)	(0.071)
Non-top	0.324***	0.346***	0.382***
	(0.092)	(0.079)	(0.069)
Year	0.009	0.006	0.015
	(0.024)	(0.020)	(0.012)
Adjusted R^2	0.352	0.405	0.474
Observations	50,029	50,029	50,029
Score Controls	Yes	Yes	Yes
Score and HS Path Controls	No	Yes	Yes
District FEs	No	No	Yes

Notes: Regression are run on data for the application cycles in years 2013-2015, immediately before the 2016 reform. The three years of data are coded Year 1 through 4. The three specifications test for parallel trends in the double difference across performance groups and types of high school. Standard errors are clustered at the locality level and are shown in parentheses.

Table A-6
Triple Difference Estimate of Exposure to Contraction of Outside Options
on Selectivity of "Reach" Programs Chosen

	Including All Years			Excluding 2016		
	(1)	(2)	(3)	(4)	(5)	(6)
Public HS × Non-top × Post-reform	0.115***	0.132***	0.109**	0.102**	0.119***	0.095**
	(0.039)	(0.036)	(0.043)	(0.041)	(0.038)	(0.045)
Public HS × Non-top	-0.143***	-0.153***	-0.155***	-0.137***	-0.150***	-0.151***
	(0.049)	(0.047)	(0.042)	(0.050)	(0.048)	(0.043)
Public HS × Post-reform	0.007	-0.014	0.025	0.008	-0.015	0.032
	(0.030)	(0.029)	(0.027)	(0.030)	(0.030)	(0.026)
Non-top × Post-reform	0.230***	0.196***	0.212***	0.190***	0.144***	0.164***
	(0.037)	(0.034)	(0.039)	(0.037)	(0.036)	(0.043)
Public HS	0.019	0.035	-0.055	0.020	0.040	-0.054
	(0.029)	(0.026)	(0.042)	(0.029)	(0.025)	(0.043)
Non-top	-0.093*	-0.006	0.004	-0.028	0.084**	0.092**
	(0.049)	(0.040)	(0.039)	(0.046)	(0.038)	(0.038)
Post-reform	-0.458***	-0.380***	-0.410***	-0.367***	-0.264***	-0.303***
	(0.029)	(0.029)	(0.019)	(0.032)	(0.032)	(0.019)
Adjusted R^2	0.279	0.332	0.379	0.290	0.354	0.400
Observations	131,926	131,926	131,926	111,900	111,900	111,900
Score Controls	Yes	Yes	Yes	Yes	Yes	Yes
Score and HS Path Controls	No	Yes	Yes	No	Yes	Yes
District FEs	No	No	Yes	No	No	Yes

Notes: Scores represent the weighted average of end-of-high-school exam scores. HS Path is a binary variable that represents whether the path chosen in the second year of high school is "scientific" or "social", which affects the weights programs give to the elective exams. Standard errors are clustered at the locality level and are shown in parentheses.

Table A-7
Parameter estimates

		Estimate	SE
a. Preference parameters			
γ	cutoff	0.320	0.021
γ	capital	0.706	0.067
γ	public	-2.270	0.060
γ	private	-1.790	0.003
γ	price	-0.001	0.000
γ	distance	-0.011	0.000
λ	hises x public	-0.168	0.064
λ	hises x private	-3.421	0.063
λ	hises x capital	0.157	0.111
λ	hises x price	0.001	0.000
λ	hises x cutoff	-0.082	0.007
λ	hises x dist	0.005	0.001
γ	applied science	-0.217	0.047
γ	health	1.009	0.048
γ	social science and humanities	-0.891	0.063
b. Belief parameters			
σ_0	Standard deviation intercept	6.040	0.074
σ_1	Standard deviation slope on cutoff	-0.497	0.012
$\mu_{0,lowses}$	Mean shift intercept (lower-SES)	-12.004	0.104
$\mu_{0,highses}$	Mean shift intercept (high-SES)	-3.368	0.096
$\mu_{1,lowses}$	Mean shift slope on cutoff (lower-SES)	1.276	0.033
$\mu_{1,highses}$	Mean shift slope on cutoff (high-SES)	0.314	0.082

Notes: The table shows estimated parameters from the model.

B Description of pre and post-reform admissions procedures

B.1 Pre-reform mechanism: deferred acceptance

In the pre-reform period, the Center for Educational Services conducted admissions to all public programs through a DA algorithm. The stages of the application process are as follows:

1. In June of each year, students take national exams, two obligatory exams in math and literature, and two elective exams in subjects chosen by each student.
2. In July, grades from each of the exams become public through a set of lists in which students are ranked in each of the exams from best to worst performing.
3. In August, the admissions process for the public programs begins, with each student applying to up to ten programs through the centralized platform, and ranking programs in the order of most to least preferred.
4. Round 1, phase 1: A DA algorithm runs and assigns each student to a seat.
    (a) Step 1: Each student proposes to their first choice. Then each program tentatively assigns its seats to its proposers in descending order of program-specific weighted scores. All of the other students are tentatively rejected.
    (b) Step 2: All students that were rejected from their first choice propose to their second choice. Any of the programs that have seats left and are proposed to in step 2 assign their remaining seats to proposers in descending order of priority.
5. After learning the initial assignment, each student chooses one of three options: (1) to enroll in the given assignment, foregoing a reassignment round where they can be assigned to a program ranked at least as high in their list as the one they were initially assigned to, (2) exit the centralized assignment process and enroll in a private university, or (3) participate in a reassignment round where they are guaranteed to be assigned in a program they ranked at least as high as the program they were initially assigned to
6. Round 1, phase 2 (the reassignment round): students who decide to participate in the reassignment round are assigned a seat among the seats remaining in the programs that were not filled in the initial round of assignment again through a DA procedure. The allocation in round 2 is the final assignment for each student who participated

in the centralized assignment. At this stage, students can choose to enroll in their assigned program, enroll in a private program, or reject their assignment and wait for the second round.

7. Round 2: this round mainly serves for students who failed to qualify for university admission in the main round, those who were unassigned in the main round, and students who rejected their assignment in phase 2 of the main round. I do not describe this round because it is not relevant for the paper as the set of students who participate in this round would not be eligible to enroll in university at all in the post-reform period.

B.2 Post-reform mechanism: dynamic multi-offer

1. In June of each year starting in 2016, students take national exams, three obligatory exams in math, Albanian language and literature, and a foreign language, and one elective exam in a subject chosen by each student.
2. In July, grades from each of the exams become public through a set of lists.
3. In August, the admissions process for all programs begins, with each student applying to up to ten programs through the centralized platform, and submitting *unordered* portfolios.
4. The admissions procedure with 7 phases lasting 48 hours each unfolds:
   (a) Phase 1: Ranked lists of applicants in decreasing order of weighted average score are published by the mechanism for each program and students observe their position on each list and whether they have cleared the cutoff for each program in this phase. The student also observes the last person to be admitted by each of the programs. At this stage a few possible scenarios may happen:
      - Students who have cleared the cutoff of at least one of the programs but not all, face three choices. The first is to accept any of the offers received in phase 1, and forgo all other options in their original portfolio. The second option is to forgo all options received in phase 1 and wait for results of the next phase for the remainder of the programs in their portfolio. Third, they may choose to exit the mechanism unmatched.
      - Students who have cleared the cutoff of all their programs may either choose to enroll in one of the programs, or exit the mechanism and forgo all offers received.

At the end of phase 1, all seats but those taken by students who decided to accept an offer free up for students in the next phase. All rejected offers are removed from students' lists.

(b) Phases 2-6: At the beginning of each phase n between 2 and 6, students are ranked by each of the remaining programs in their portfolio that have empty seats left. They observe their new ranking relative to the remaining applicants in each of the programs, observe their phase-n offers in that program and observe the phase n cutoffs. They make their enrollment or waiting decisions as in phase 1.

(c) Phase 7: The final offers of the main round realize and students make their last-chance enrollment decisions for the round.

5. Round 2: this round mainly serves for students who failed were unassigned in the main round. I do not describe this round.

C Data appendix

C.1 Merit scholarship policies in private universities

Figure A-10
An example of a university posting its scholarship policy on the website

Bachelor/Integrated study program Scholarship

Epoka University grants merit-based scholarships for the best ones for the entire normal duration of their studies as follows:

Scholarships based on the results of high school studies and State Matura Exam of the Republic of Albania:

Point Range	Scholarship
6600 or more	100%
6400 – 6599.99	75%
6200 – 6399.99	50%

Scholarships based on SAT Exam results:

Point Range	Scholarship
1350 or more	100%
1290 - 1349	75%
1220 - 1289	50%

The number of scholarships is limited (10 - 15 scholarships for each study program). Applications after the aforementioned position will benefit from a scholarship with minus 25% of the respective scholarship amount. Candidates interested to receive a scholarship should apply at the earliest in order to benefit from the above-mentioned scholarships since their allocation will be carried out according to the date of application.

Note: Are eligible to apply for scholarship the candidates that have high school average at least 9.0

Click here to calculate your Matura points.

Social Scholarship

- Two scholarships for candidates who belong to the " children of Albanian policemen killed in line of duty"
- Two scholarships for candidates who are sportsmen with very high results in the respective Olympic sports category in Albania
- A scholarship for candidates who belong to the Roma/Egyptian community from Albania.
- Two scholarships for candidates who belong to the "orphan coming from a low-income family.

Note: This figure shows scholarship policies posted by Epoka University. https://admissions.epoka.edu.al/home-bachelor-integrated-study-program-scholarship-2745-497.html

D Simplifying the likelihood function

D.1 One-shot swaps

Proposition 1 of Larroucau and Rios (2020) shows that for a portfolio selection problem with ordered lists where probabilities of admission to each program are independent, it suffices to show that the chosen portfolio is preferred to all its one-shot swaps for the portfolio to be optimal. The following reformulation of the proposition is applicable to portfolios in settings where the ordering of the list does not matter for payoffs.

Theorem D.1 (adapted from Larroucou and Rios)**.** *Let $C = \{j_1, ..., j_k\}$ be an unordered application list of length at most K, i.e. $k \leq K$. Without loss of generality, let $u_{j_1} \geq u_{j_2} \geq ... \geq u_{j_k}$ so that the utility from submitting application portfolio C is*

$$V(C) = p_{j_1}u_{j_1} + (1 - p_{j_1})p_{j_2}u_{j_2} + ... + \left(\prod_{l=1}^{l=k-1}(1 - p_{j_l})\right)p_{j_k}u_{j_1}$$

If $\mathcal{S}(C)$ is the set of one-shot swaps of portfolio C and

$$V(C) \geq V(C') \; \forall \; C' \in \mathcal{S}(C), \tag{A-1}$$

then

$$V(C) \geq V(C') \; \forall \;\; C' \;\; s.t. \;\; |C'| = K. \tag{A-2}$$

Discussion: In the context of rank-ordered lists, unprofitability of one-shot swaps implies more restrictions than in the case of unordered lists. Take for example a case with lists of size 2. The set of one-shot swaps of a portfolio $C_o = (j_1, j_2)$ are $\mathcal{S}(C_o) = \{(j_1, x), (x, j_1), (x, j_2), (j_2, x)\}$ whereas the set of one-shot swaps for an unordered portfolio $C_u = \{j_1, j_2\}$ are $\mathcal{S}(C_u) = \{\{j_1, x\}, \{j_2, x\}\}$. I show below that even starting with fewer inequalities on OSS as in the case of unordered portfolios, optimality of a portfolio given the unprofitability of its OSS is satisfied.

Proof. Proof for the DA rank ordered lists can be found in Larroucau and Rios (2020). See below for a fast sketch of the proof for the case of unordered portfolios that follows the original proof. The proof is done by induction. The case for $K = 1$ is obvious. For $K = 2$, suppose that $C = \{j_1, j_2\}$ and let $\mathcal{S}(C) = \Big\{\{x, j_1\}, \{x, j_2\}\Big\} \forall x \in \mathcal{J} \backslash C$. Let $C \succeq \{x, j_1\}$ and

$C \succeq \{x, j_2\}\ \forall x \in \mathcal{J}$. Then for all $\{x, y\}$, suppose WLOG that $u_x \geq u_y$ and $u_{j_1} \geq u_{j_2}$. We have a few cases:

Case 1: $u_{j_1} \geq u_y \geq u_{j_2}$: Using the fact that $\{j_1, j_2\} \succeq \{j_1, y\}$ implies that $p_{j_2}u_{j_2} \geq p_y u_y$,

$$V(\{x, y\}) = p_x u_x + (1 - p_x)p_y u_y \leq p_x u_x + (1 - p_x)p_{j_2}u_{j_2} = V(\{x, j_2\}) \leq V(\{j_1, j_2\}) = V(C).$$

Case 2a: $u_{j_1} \geq u_{j_2} \geq u_y$ and $u_x \geq u_{j_2}$:

$$V(\{x, y\}) = p_x u_x + (1 - p_x)p_y u_y \leq p_x u_x + (1 - p_x)p_{j_2}u_{j_2} = V(\{j_2, x\}) \leq V(\{j_1, j_2\}) = V(C)$$

Case 2b: $u_{j_1} \geq u_{j_2} \geq u_y$ and $u_x \leq u_{j_2}$:

$$\begin{aligned} V(\{x, y\}) = p_x u_x + (1 - p_x)p_y u_y &\leq p_x u_x + (1 - p_x)p_{j_2}u_{j_2} \\ &\leq p_{j_2}u_{j_2} + (1 - p_{j_2})p_x u_x = V(\{j_2, x\}) \leq V(\{j_1, j_2\}) = V(C) \end{aligned}$$

The first inequality holds because $\{j_1, j_2\} \succeq \{j_1, y\} \implies p_{j_2}u_{j_2} \geq p_y u_y$. The second inequality holds because $j_2 \succeq x$ so in the case of admission to both, the applicant will choose j_2.

Case 3: $u_y \geq u_{j_1} \geq u_{j_2}$:

$$V(\{x, y\}) = p_x u_x + (1 - p_x)p_y u_y \leq p_x u_x + (1 - p_x)p_{j_2}u_{j_2} = V(\{x, j_2\}) \leq V(\{j_1, j_2\}) = V(C)$$

For the inductive step, assume that the theorem holds for portfolios of length k. It remains to show that the theorem holds for portfolios of length $k + 1$. The rest of the proof goes as follows: suppose portfolio C_{k+1} satisfies $C_{k+1} \succeq C'_{k+1}\ \forall\ C'_{k+1} \in \mathcal{S}(C_{k+1})$. First show that, C_k, the portfolio that has highest utility among the k-sized subsets of C_{k+1} satisfies $C_k \succeq C'_k\ \forall\ C'_k \in \mathcal{S}(C_k)$. This implies that C_k is the optimal portfolio among all k-sized portfolios by the inductive assumption. Then show that the remaining element added to form C_{k+1} is added by Marginal Improvement Algorithm (which **?** show to be optimal), which implies that the final C_{k+1} is the optimal portfolio of size $k + 1$. □