Dynamic College Admissions and the Determinants of Students' College Retention*

Tomás Larroucau[†] Ignacio Rios[‡]

May 14, 2021

Most recent version

Abstract

We analyze the effects of centralized assignment mechanisms on downstream outcomes in higher education. To do so, we study the relevance of incorporating dynamic incentives and eliciting private information about students' preferences to improve their welfare and outcomes beyond their initial assignment, including their decisions to switch or drop out. We show that the most common assignment mechanism, the Deferred Acceptance (DA) algorithm, can result in significant inefficiencies as it fails to elicit cardinal information on students' preferences. We collect novel data about students' preferences, their beliefs on admission chances, and their college outcomes for the Chilean college system. We analyze two main behavioral channels that explain students' dynamic decisions. First, by exploiting discontinuities on admission cutoffs, we show that not being assigned to ones' top-reported preference has a positive causal effect on the probability of re-applying to the centralized system and switching one's major/college, suggesting that students switch to more preferred programs due to initial *mismatches*. Second, we find that a significant fraction of students change their preferences during their college progression, and that these changes are correlated with their grades, suggesting that students may *learn* about their match-quality. Based on these facts, we build and estimate a structural model of students' college progression in the presence of a centralized admission system, allowing students to learn about their match-quality over time and re-apply to the system. We use the estimated model to disentangle how much of students' switching behavior is due to initial mismatches and learning, and we analyze the impact of changing the assignment mechanism and the re-application rules on the efficiency of the system. Our counterfactual results show that policies that provide score bonuses that elicit information on students' cardinal preferences and leverage dynamic incentives can significantly decrease switchings, dropouts, and increase students' overall welfare.

*We are very grateful for the guidance and helpful comments from Nikhil Agarwal, Francesco Agostinelli, Juan Pablo Atal, Eduardo Azevedo, Gorkem Bostanci, Juan Camilo Castillo, Stuart Craig, Hanming Fang, Joao Granja, Margaux Luflade, Petra Moser, Christopher Neilson, Aviv Nevo, Sergio Ocampo, Holger Sieg, Andrew Shephard, Michela Tincani, Gabrielle Vasey, Sergio Villalvazo, Rakesh Vohra, and seminar participants at the Empirical Micro Seminar at UPenn for their valuable feedback. This paper would not have been possible without the support from MINEDUC, SUA, DEMRE, and the Vicerrectoría de Tecnologías de la Information (VTI) from the University of Chile. We especially thank Leonor Varas, María Elena Gonzalez, and José Piquer for their collaboration and support. We also thank Manuel Martínez and Consilium Bots for their help designing and implementing our nationwide surveys; Giorgio Parra for his help and support at many points of this project; and Alejandra Venegas for running the regressions involving college grades. Any errors or omissions are our own.

[†]University of Pennsylvania, Department of Economics. Email: tomasl@sas.upenn.edu

[‡]University of Texas at Dallas, Naveen Jindal School of Management. Email: ignacio.riosuribe@utdallas.edu

1 Introduction

Higher education is a scarce resource that exhibits high returns and plays a crucial role in countries' development.[1] However, many higher education systems experience low retention and on-time graduation rates. According to the OECD (2019), only 40% of full-time bachelor students graduate by the theoretical duration of their programs, and a significant fraction of students drop out without graduating. This low yield can be particularly severe for developing countries. For instance, the on-time graduation rate in Chile is 16%—the lowest among all countries in the OECD—, 30% of students switch their programs, and close 30% drop out from the college system.

According to Kapor et al. (2020a), at least 46 countries around the world use a centralized system to organize their admissions to college, including Turkey, Taiwan, Tunisia, Hungary, and Chile.[2] Although extensive literature analyzes the pros and cons of different mechanisms to perform the allocation, their effect on downstream outcomes that are policy-relevant (beyond the initial assignment) is unclear. For instance, in organ transplant systems, one of the primary goals is to maximize patient survival (Agarwal et al., 2019). In school choice, policymakers often care about achieving social mobility, meritocracy, and equal access of opportunities (Tanaka et al., 2020). In higher-education, colleges primarily focus on the academic progression (e.g., grades and retention) and the labor market outcomes of their students. As these examples illustrate, there are important downstream outcomes that should be taken into account when allocating these resources, and centralized assignment mechanisms may help to improve them.

To understand the effects of centralized mechanisms on outcomes, it is essential to account for some features that characterize real-life applications and that are mostly overlooked in the literature. In college admissions, one such feature is that these systems are typically dynamic.[3] For instance, students can learn over time about their match-quality with programs, re-apply to the system each year, switch from their initial assignment if they are assigned to a more preferred program, and they can also drop out at any point in their college progression. Another feature is that, to decide the assignment, centralized systems typically use information about students' academic performance, such as admission tests, and information about their ordinal preferences. However, students may have private information about their preferences that could affect their future outcomes and the higher education system's efficiency. For instance, students' intrinsic motivation or vocation, which would be captured by their cardinal preferences,[4] could affect their per-

[1] Among countries in the OECD, a graduate with a bachelor's degree earns 44% more than those with an upper secondary education.

[2] Centralized assignment systems are widely used in other settings, including school choice, higher education, labor markets, the allocation of organs from deceased donors, kidney exchange, among many others.

[3] This feature is also present in other contexts. In the allocation of organs from deceased donors, a patient can reject a donor and remain in the wait-list for an organ of higher quality (Agarwal et al., 2019). In school choice, many systems—including that in NYC (Abdulkadiroğlu et al., 2005a), and (Narita, 2018), Boston (Abdulkadiroğlu et al., 2005b) and Chile (Correa et al., 2019)—have multiple rounds, and after each round, students/families can either accept their assignment or reject it and re-apply to the system in the next one.

[4] The literature also refers to this as the *intensity* of students' preferences. In the rest of the paper, we

sistence in their programs, impacting the overall college retention rates and the system's efficiency. Therefore, designing admission systems that consider the dynamic nature of incentives and elicit information about students' cardinal preferences can be critical to improve students' outcomes and the efficiency of the higher education system.

In this paper, we study how centralized assignment mechanisms can impact students' welfare and downstream outcomes, including their college grades (achievement), on-time graduation rates, and retention. We accomplish this by incorporating dynamic incentives and eliciting information about students' cardinal preferences. To motivate the relevance of these features, we start with a stylized model that illustrates how the most common assignment mechanism, the Deferred Acceptance (DA) algorithm, can lead to inefficiencies in a dynamic context. In particular, we show that if students can re-apply and switch to more preferred programs, a system that elicits cardinal preferences—e.g., imposing limits on the length of applicants' lists or penalizing re-applicants—can improve colleges' yield and students' welfare. The intuition is that, by including these strategic considerations (either through changes in the mechanism or in the re-application rules), students face trade-offs that incentivize them to choose programs for which the intensity of their preferences is higher. However, these policy changes may increase inefficiencies in the allocation if other reasons explain students' switchings. For instance, if switchings are due to students' learning about their match-quality, it may be welfare-enhancing to encourage exploration and reduce switching costs to avoid ex-post mismatches. Hence, it is essential to understand the main drivers of switchings and dropouts to design a mechanism that can improve students' welfare and outcomes.

We conjecture the existence of two behavioral channels that can explain students' dynamic decisions. The first channel, called the *learning* channel, states that students may receive new information during their college experience and learn about their match-quality with programs. This new information could modify their consumption values while in college and their labor market returns upon graduation, motivating them to switch or drop out to avoid ex-post mismatches. The second channel, called the *initial mismatch* channel, states that students who have dynamic considerations and face uncertainty in their admissions may switch in the future if they were not initially assigned to their most desired option, as they may try to improve their preference of assignment by participating again in the assignment process. Notice that students can benefit from enrolling in less preferred programs, even if they are likely to switch in the future, because they can improve their outside option. However, if the system is in excess demand and colleges care about retaining their students, this behavior generates a congestion externality. As a result, it might be ex-ante inefficient to assign some students to less preferred programs if they face low retention probabilities.[5]

Our empirical application uses data from the centralized admissions system in Chile, which uses a variant of the Deferred Acceptance (DA) algorithm (Ríos et al., 2020) to assign more than 250,000 students each year. This setting is suitable for our study for multiple reasons. First, students directly apply to programs, which are combinations

use these interchangeably.

[5]A similar congestion externality can occur when students entail in repeated test-taking behavior (see Krishna et al. (2018)).

of major/campus/institution, and they must re-apply to the centralized system if they want to switch (transfers between programs are relatively rare). As a result, students who switch tend to spend more years in college, lowering on-time graduation rates. Second, many over-demanded programs exhibit low retention rates. As previously discussed, due to congestion externalities, this is costly for other students who were displaced and assigned to less preferred options. Moreover, these switches are costly for universities, as they lose seats that cannot be re-allocated efficiently, resulting in forgone tuition. Finally, the Chilean context is exciting because these inefficiencies may also have fairness and equity effects. There are significant differences in students' switching and dropout behavior depending on their socioeconomic background and other observable characteristics,[6] and these differences may exacerbate other prevalent inequities in the system. For these reasons, it is essential to understand the determinants of students' switching and dropout behavior to account for them in the assignment process and address the inefficiencies mentioned above.

Combining administrative data and two nationwide surveys that we designed and conducted, we show that the two behavioral channels play a significant role in the Chilean system. More specifically, by exploiting the discontinuities generated by admission cutoffs, we show that there is a positive causal effect of being assigned to lower reported preferences on the probabilities of re-applying to the centralized system and switching majors/colleges, supporting the existence of the *mismatch* channel. As previously shown in Larroucau and Ríos (2018), we confirm that students do not report their preferences truthfully, and we also show that students' top-true preferences for programs—elicited through our surveys—change over time. Finally, we show that switching probabilities are negatively correlated with students' grades in college, which suggests that students *learn* about their match-quality through their grades.

To account for these findings, we introduce a structural model that captures the application behavior of students, as well as their decisions to enroll, re-take the admission tests, re-apply, switch, and drop out, allowing students to learn about their unobserved abilities—match-quality—during their academic progression. In particular, we assume that students make their application and enrollment decisions considering both the value of studying each program, the continuation value of re-taking the admission tests and re-applying to the system, and their labor market prospects. As they progress in college, students observe noisy signals of their unobserved ability from their grades, and they use this information to update their continuation values for each program. Based on this, students decide whether to continue in their current program, re-apply to the system, or drop out and choose their outside option. Finally, students face exogenous probabilities of graduating, enter the labor force, and receive pecuniary and non-pecuniary values from the labor market.

The main challenge to estimate our model is to separately identify the *learning* and the *mismatch* channels. To identify the learning channel parameters, we use the correlation between students' college grades and their decisions and outcomes, including their re-applications and changes in the composition of their preference lists, switchings, and

[6]Low-income students are less likely to switch due to credit constraints. However, they are significantly more likely to drop out.

dropout decisions. On the other hand, to identify the mismatch channel, we combine two sources of variation: (1) students' beliefs on their current and future admission probabilities and (2) the persistence of students' preferences and the relation between students' preference of assignment and their outcomes, which we obtain from our two nationwide surveys and rich administrative data. Our results suggest that initial mismatches explain close to a third of switching decisions, while learning about abilities explain the remaining switches and part of the dropout decisions.[7]

After estimating the structural model, we assess whether changes in the assignment process—either through changes in the re-application rules or changes in the assignment mechanism—can affect students' outcomes. Both approaches can elicit students' cardinal preferences, as they introduce opportunity costs that students must take into account when making their applications. We find that giving a score bonus for all first-year applicants—as it is the case in Finland—or allowing students to signal one of their preferences to get a bonus in that specific program—in the spirit of the signaling mechanism in the Economics job market—can significantly affect students' outcomes, namely, reduce switching and dropout rates, while at the same time increase students' welfare. For instance, the Finish policy increases colleges' retention and students' welfare by 6.2% and 9.2%, respectively.[8] We also find that these effects are robust to changes in the fraction of participants that behave strategically, as opposed to other approaches such as constraining the length of application lists. Moreover, we observe considerable heterogeneity in how these policies affect students depending on their income level and gender, with low income and male students benefiting the most. Our results show that these policies must be carefully designed, as they have a non-linear effect on students' outcomes depending on the magnitude of the bonuses or penalties applied, and because they can target and benefit different groups of students depending on how they are implemented. Our counterfactual experiments stress the importance of correctly balancing the effects of the two behavioral channels: allowing students to learn through experimentation and reducing the congestion externality caused by initial mismatches. Overall, our results show that incorporating dynamic incentives and eliciting students' cardinal preferences through changes in the re-applications rules and the assignment mechanisms can significantly affect students' outcomes and their overall welfare.

The paper is organized as follows. In Section 2, we discuss the most closely related literature. In Section 3, we provide a stylized-example that illustrates how eliciting cardinal preferences can affect students' outcomes. In Section 4, we describe the Chilean college admissions system and provide empirical evidence for the aforementioned behavioral channels. In Section 5, we present our model. In Section 6, we describe our identification strategy. In Section 7, we describe the estimation approach and its results. In Section 8, we report our counterfactual results. Finally, in Section 9 we conclude.

[7]The residual is explained by students' learning through their random preference shocks.

[8]In Section 8 we explain in detail how to compute this measure of welfare.

2 LITERATURE

Our paper brings together two strands of the literature: (i) the empirical analysis on centralized assignment mechanisms, and (ii) the empirical analysis of college/major choices under uncertainty.

The first strand of the literature focuses on (1) understanding the incentives that centralized assignment mechanisms introduce, (2) how to use the data generated from these mechanisms to identify and estimate students' preferences/beliefs, and (3) measuring the welfare effects of changing assignment mechanisms in different settings. Depending on the available data and the incentives students face, researchers have developed various methodologies to identify and estimate students' beliefs and preferences.[9] For instance, Fack et al. (2019) and Abdulkadiroğlu et al. (2017) analyze the case when the mechanism used is *strategy-proof*, while He (2012), Agarwal and Somaini (2018), Calsamiglia et al. (2020), Kapor et al. (2020b), among others, analyze the opposite case (for instance, when the mechanism used is IA instead of DA). Other studies analyze the case when the rules of the system introduce strategic considerations, including Ajayi and Sidibe (2017), Artemov et al. (2017), Fack et al. (2019), Luflade (2017)). Furthermore, some studies allow students to play weakly-dominant strategies (see Fack et al. (2019), He (2012) and Larroucau and Ríos (2018)), and make strategic mistakes in their applications (see Kapor et al. (2020b) and Artemov et al. (2017)).

After estimating the primitives that govern students' application behavior (students' preferences and beliefs), researchers typically analyze the ex-ante welfare effects of changing the assignment mechanism (Agarwal and Somaini (2018), Calsamiglia et al. (2020), He (2012), Hwang (2016), Kapor et al. (2020b), among others), the application rules (Ajayi and Sidibe (2017), Luflade (2017), Hernández-Chanto (2017), Larroucau and Ríos (2018), Carvalho et al. (2019), among others), or students' priorities (Fack et al. (2019), Larroucau and Ríos (2018), among others). The current evidence of the effects of changing the assignment mechanism and application rules on students' welfare has mixed results. Researchers have found that mechanisms that elicit the intensity of students' preferences can achieve higher ex-ante welfare (Agarwal and Somaini (2018), Calsamiglia et al. (2020), He (2012), Hwang (2016), among others), but this heavily depends on the assumptions on students' sophistication (see Kapor et al. (2020b)), which suggests that the appropriate mechanism depends on the specific setting.

Despite the progress on understanding the role of assignment mechanisms and their impact on students' welfare, the aforementioned studies consider static settings or assume that preferences do not vary over time. Taking a dynamic approach to analyze centralized assignment mechanisms can give new insights to the classical trade-off between *strategy-proof* mechanisms (such as DA) and mechanisms that elicit the intensity on students' preferences (such as IA). For instance, when students have repeated interactions with the assignment mechanism (i.e., through re-applications), ignoring the system's dynamics can lead to biased estimates of the welfare effects of changing the assignment mechanism. The reason is that, in static settings, researchers estimate the indirect utilities that

[9]See Agarwal and Somaini (2019) for a recent and more exhaustive survey.

students receive for being—initially—assigned to schools/colleges as primitives to the model. In this sense, the indirect utilities are invariant to policy changes, such as changes in the assignment mechanism or application rules. However, in a dynamic setting where students can re-apply to the centralized system over time, the indirect utilities of being—initially—assigned can change in the counterfactual because the continuation values are affected by the future interactions with the assignment mechanism. Moreover, static approaches do not allow researchers to evaluate alternative policies that could enhance welfare, such as modifying re-application rules, as is the case in Finland and Turkey. Finally, it is crucial to understand the implications of changing assignment mechanisms on students' outcomes beyond their initial assignment, such as students' achievement, persistence, and their labor-market outcomes.

To our knowledge, the only exception to this is Narita (2018), who analyzes theoretically and empirically the welfare performance of dynamic centralized school-choice mechanisms when demand evolves over time. The author uses data from the NYC school-choice system and shows that families' choices change after their initial match. He then develops an empirical model of evolving demand for schools under learning, allowing for endowment effects in response to prior assignments, and switching costs. He uses some particular features of the school choice setting to separately identify these components, and he estimates that the initial match's welfare performance is heavily affected by demand-side frictions, primarily by switching costs. Finally, the author investigates improvements to NYC's discretionary re-application process by using dynamic centralized mechanisms.

Although the dynamics and learning processes are related, our paper is substantially different, as there are essential differences between school-choice and college admissions systems that affect both the research questions and the identification strategies. In our setting, "switching" costs naturally arise since students bear an opportunity cost when they switch programs and delay their graduation. These switching costs are not present in school-choice systems and produce a negative externality that affects the system's efficiency. Given these differences, we focus on changing the (dynamic) centralized assignment mechanism by eliciting preference intensity and modifying re-application rules on students' welfare and their college outcomes such as achievement, persistence, and the system's efficiency. To the extent of our knowledge, none of the previous studies have evaluated the welfare consequences of changing the assignment mechanism and re-application rules on students' outcomes when students can re-apply to the centralized system, and their preferences and admission probabilities can change over time.

The second strand of the literature studies individual education and occupation choices, stressing the role of the human capital specificity, uncertainty about preferences and abilities, and how students' choices impact their educational outcomes and labor market returns. We refer the reader to the comprehensive reviews by Altonji et al. (2012) and Altonji et al. (2016). Almost all papers in this literature focus on decentralized college markets or ignore any rationing mechanism that could play a role in college admissions (an exception is Bordon and Fu (2015)). We use insights from the seminal work by Arcidiacono (2005) and the recent work by Arcidiacono et al. (2016) to model students' learning process and their labor-market outcomes, and augment their methodology by micro-founding the college/major choice process in the presence of a centralized admission system, taking into account students' strategic behavior.

Within this strand of the literature, the closest paper to ours is Bordon and Fu (2015). It analyzes the effects of changing the Chilean university system from students choosing college and major at the same time (J system) to choosing college first and then major (S system). The authors model students' enrollment and dropout decisions and consider potential peer effects. They estimate students' preferences and compare the J and S systems. The authors find that match-quality and students' welfare would increase under the S system compared to the J one.[10] Our paper's main difference is that we model the entire application and switching behavior of students and use the information in their reported Ranked Ordered List (ROL) over time, their grade records, and survey responses to separately identify the persistence on students' preferences from learning. These differences allow us to rely less on the model's particular structure to identify the model primitives. However, we do not consider peer effects in the analysis, and we do not have access to a panel of students' future wages. Our counterfactual experiments also differ in nature. Instead of changing the university system's structure and affecting the *learning* channel, we focus on changes to the assignment mechanism and re-application rules—affecting the *mismatching* channel—and we evaluate these changes on different outcomes such as achievement, switchings, and on-time graduations.

Our work is complementary to these two strands of the literature, as we provide new insights on the effects of centralized assignment mechanisms from a dynamic perspective. To the extent of our knowledge, ours is the first paper that structurally measures the effects of centralized assignment mechanisms and re-application rules on students' college outcomes beyond their initial assignment, including achievement, college retention, and on-time graduation rates. Finally, we also contribute to the literature by revisiting the trade-off between eliciting intensity on students' preferences and guaranteeing *strategy-proofness*, but we do so in a dynamic context.

3 Motivating Example

We first analyze whether it is—theoretically—possible to increase aggregate students' welfare and increase the system's yield by changing the assignment mechanism and re-application rules. Furthermore, we provide intuition on how switching behavior can be affected by the assignment mechanism in a dynamic setting.

If students face uncertainty over their admission chances, either because of uncertainty about admission cutoffs or their future application scores, switchings can endogenously occur over time. As students do not know their ex-post choice sets, they could choose to enroll in a program in the first year and switch in the following year to a more preferred

[10] Malamud (2011) also analyzes the trade-offs students face when they have to specialize early in their college education. The author argues that if the rate of field switching in systems with an early specialization is high, this can be seen as evidence that education provides valuable information on match-quality and that match-quality has a large impact on education returns.

program if their choice set allows them to. Moreover, if students are uncertain about their match-quality with programs, and after enrollment, they learn about their preferences/abilities, they could choose to switch programs or drop out to avoid ex-post mismatches. Regardless of which mechanism dominates, individual switchings and dropouts impose an externality on universities and on other students. Given the sequential nature of colleges' academic progression, when a student switches at the end of the academic year, the resulting vacancy is lost for the next year, and, in the absence of a proper transfer system that allows students to switch at different stages of their college progression, this vacancy can not be reallocated to another student.

To illustrate how switches may arise endogenously, even in the absence of learning, consider a centralized college admissions problem with re-applications and two periods. Let $S = \{A, B\}$ and $C = \{I, II\}$ be the sets of students and colleges, respectively. We assume that students are expected utility maximizers, i.e., they submit a preference list that maximizes their expected utility conditional on their preferences and beliefs on admission probabilities. Let R_i^t be the preference list submitted by student i at time t. After students submit their applications, colleges post their first-year vacancies. Let q_j^t be the first-year vacancies posted by college j at time t. To add uncertainty on students' admission chances, we assume that the number of vacancies is uncertain, unknown ex-ante by students, and distributed according to

$$\mathbb{P}(q_j^t = 1) = \mathbb{P}(q_j^t = 0) = \frac{1}{2}, \quad \forall t, j \in \{I, II\}.$$

Students' preferences are given over their expected assignment. We assume that students' utilities for being assigned to each college are given by[11]:

$$A: \quad u_I^A \gg u_{II}^A > 0, \qquad\qquad B: \quad u_{II}^B \gg u_I^B > 0$$

Each student $i \in S$ has an application score s_i^j for every college $j \in C$, which determines their position in colleges' preference lists. In particular, we assume that

$$I: \quad s_I^B > s_I^A, \qquad\qquad II: \quad s_{II}^A > s_{II}^B.$$

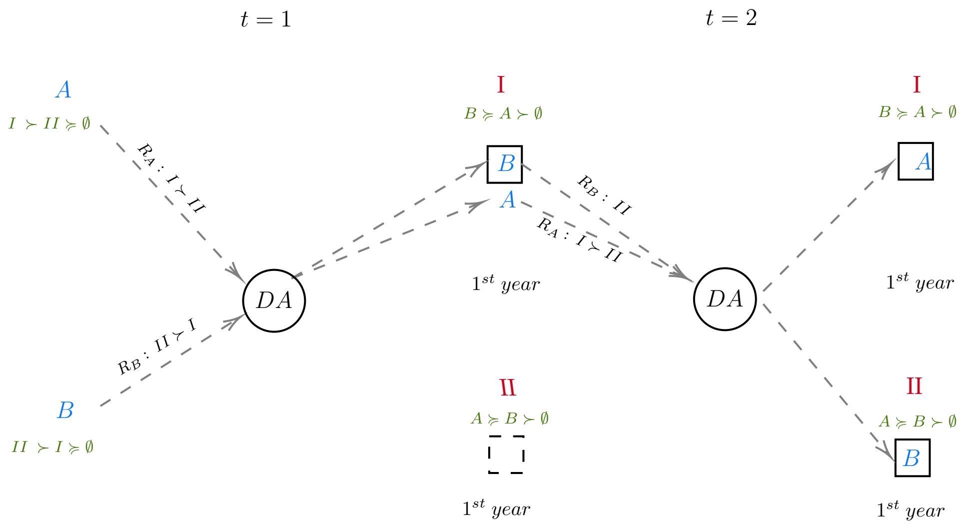
Finally, we assume that colleges care about students' persistence and bear a cost τ per student that does not remain enrolled. This cost captures the idea that colleges make investments in their students and that the vacancy (and the corresponding future tuition payments) is lost when students switch.

To illustrate the impact of the assignment mechanism used, we compare the outcomes of two alternative mechanisms: (i) Deferred Acceptance (DA), where students can apply to as many programs as they want; and (ii) Constrained Deferred Acceptance (CDA), where students can apply to at most one college. If DA is the mechanism in use, both students apply according to their true preferences. If there is only one seat in the system, as illustrated in Figure 3.1 where $q_I^1 = 1$ and $q_{II}^1 = 0$, we observe that both students compete for that seat, and the student with the highest score is assigned, while the other

[11] We assume that both students considerably prefer more their top preference than their second preference.

student remains unassigned. Given the setup of this example, the student that is assigned gets her second choice, and thus will have incentives to re-apply in the second period and try to switch to her top preference.[12]

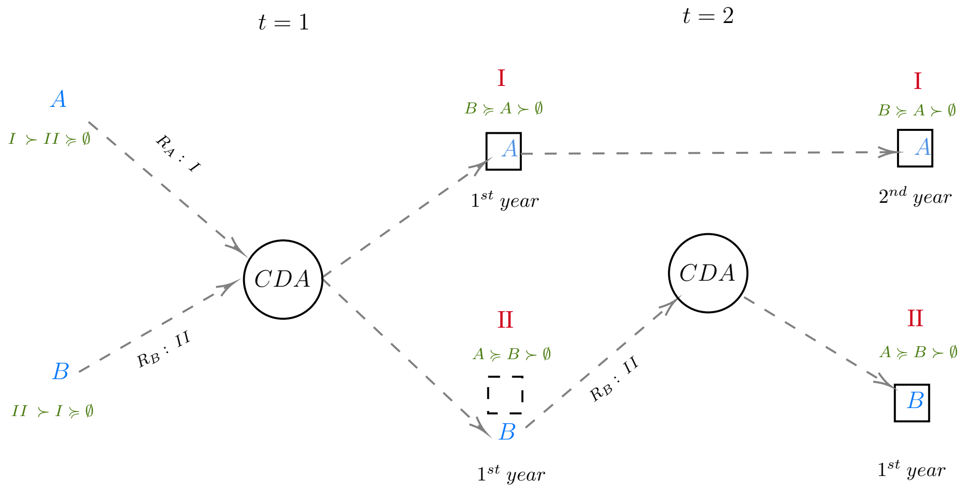
Figure 3.1: Dynamic inefficiencies under DA



On the other hand, Figure 3.2 describes the students' progression when CDA is in place in the same case as Figure 3.1 (i.e., when $q_I^1 = 1$ and $q_{II}^1 = 0$). Since students can submit at most one preference, student A applies to I and B applies to II. Hence, student A is assigned to I, and B results unassigned. Since the former student is assigned to her top choice—in the absence of learning—only the latter will re-apply in the second period, and thus there are no switchings in the system. Thus, using CDA reduces the probability of switchings, improving the efficiency and yield of the system.

[12]In Appendix A, we show that a significant fraction of students who are assigned close to the admission cutoffs, do not rank those programs as their top preference, similar to the current example.

Figure 3.2: Reducing dynamic inefficiencies under CDA



In Appendix A we show that the aggregate ex-ante welfare is given by

$$W^{CDA} \approx \frac{\left(u_I^A + u_{II}^B\right)(8+12\beta)}{16} > \frac{\left(u_I^A + u_{II}^B\right)(4+9\beta)}{16} - \frac{1}{2}\tau \approx W^{DA}.$$

Moreover, as the game is symmetric, CDA leads to a *Pareto*-improvement in *ex-ante* expected utility, and switching behavior is eliminated, lowering the costs for universities compared to the outcome of DA. Interestingly, the same outcome obtained by CDA can be achieved by changing re-application rules without changing the assignment mechanism itself. For instance, if we penalize enough re-applications from initially assigned students (Turkish mechanism), students would also only apply to their top preference in every period. In summary, this example shows how mechanisms and re-application rules that elicit intensity on students' preferences can affect students' applications, their assignments, and their switching decisions, increasing colleges' yield and the overall welfare of students.

Discussion: The stylized model above does not allow students to learn about their match-quality with colleges, and their preferences are fixed over time. How much of students' switching behavior is due to initial mismatches versus learning about their match-quality over time is an empirical question and our main identification challenge. Notice that the consequences of both channels can be quite different. In the first case, if students' preferences are persistent over time, it may be desirable to restrict re-applications and force students to internalize the negative externality they impose on other students and colleges. On the other hand, if most of the switches are due to students' learning about their match-quality, it may be welfare-improving to facilitate switching behavior to avoid ex-post mismatches. Hence, the welfare implications are unclear.

4 College Admissions in Chile

Tertiary education in Chile is offered by 156 institutions that can be classified into four types: (i) Universities (60), which have the exclusive right to award academic degrees—Bachelor, Master, and Doctorate—and offer academic programs that require a previous degree, such as Medicine and Law; (ii) Professional Institutes (IP) (43), which offer professional/technical programs that lead to a professional/technician qualification; (iii) Technical Schooling Centers (CFT) (46), which exclusively offer vocational programs leading to a technician qualification; and military and police academies (FFAA) (7).

The admissions process to these institutions is semi-centralized, with the most selective universities having a centralized system and the remaining institutions carrying their admission processes independently. This paper's empirical application focuses on the centralized part of the system, known as *Sistema Único de Admisión* (SUA). This part of the system is organized by the *Consejo de Rectores de las Universidades Chilenas* (CRUCH), and its admission process is operated by the *Departamento de Evaluación, Medición y Registro Educacional* (DEMRE).

4.1 Centralized System

To apply to any of the close to 1,500 academic programs held by the 41 universities that are part of the centralized system, students must undergo a series of standardized tests (*Prueba de Selección Universitaria* or PSU). These tests include Math, Language, and a choice between Science or History, providing a score for each of them. The performance of students during high-school gives two additional scores, one obtained from the average grade during high-school (*Notas de Enseñanza Media* or NEM) and a second that depends on the relative position of the student among his/her cohort (*Ranking de Notas* or Rank). A distinctive feature of the system is that the admission to programs is solely based on these *admission factors*.[13]

After scores are published, students can submit a list with no more than ten academic programs, ranked in strict order of preference. We refer to these lists as Rank Order Lists (ROLs). Notice that students directly apply to an academic program, i.e., they must list pairs of university-major in their ROL. In the remainder of the paper, we refer to these pairs simply as programs. Besides, it is important to highlight that there is no monetary cost for submitting an application.

On the other side of the market, each program announces its vacancies, the weights on each admission factor to compute application scores, and the set of additional requirements they will consider for applications to be valid. For instance, universities may require a minimum application score or a minimum score in some of the PSU tests, among other requirements. Each program's preference list is defined by first filtering all applicants that do not meet the specific requirements. Students are then ordered based on their

[13]Some programs such as music, arts, and acting, may require additional aptitude tests.

application scores, which are computed as the weighted sum of the applicants' scores and the weights pre-defined by each program.

Considering the vacancies and the preference lists of the applicants and programs, DEMRE runs an assignment algorithm to match students to programs. The mechanism used is a variant of the student-proposing Deferred Acceptance algorithm, where all tied students for the last seat of a program must be admitted. A thorough description of the assignment mechanism can be found in Ríos et al. (2020), and in Appendix B.1, we provide an overview of it. As a result of the assignment process, each program is associated with a cutoff such that all students whose weighted score is above it are granted admission, whereas all students with scores below the cutoff are wait-listed and thus may have to enroll in a lower-ranked preference. This property is known as the *cutoff structure*.

The enrollment process starts right after the assignment results are published. This process considers two rounds. In the first round, only assigned students can enroll in their preference of assignment, while in the second, programs with seats left after the first stage can call students in their wait-lists and offer them the chance to enroll. Also, at any point, applicants can apply and potentially enroll in a program outside the centralized admission system, and they also have the chance to join the labor force directly. Moreover, students can participate in the admission process as many times as they want, and they can use the scores obtained in the previous year as part of their application.[14]

4.2 Data

We combine administrative data of the Chilean college admissions process with massive records on students' college grades for every student enrolled in a program in the centralized system and a unique data set obtained from surveys about students' preferences and beliefs on admission probabilities. Our dataset spans from 2012 to 2016 and includes information provided by DEMRE, the Ministry of Education (MINEDUC), two surveys designed and conducted in collaboration with CRUCH and DEMRE, and grade records facilitated by CRUCH.

Admission process. We have information on the admission processes from 2012 to 2016, including students' scores, admission weights and requirements for each program, and the final assignment. In addition, we have data on students' socioeconomic characteristics, including self-reported family income, parents' education, the municipality where the student lives, among others.

Enrollment. MINEDUC provided data on students' enrollment decisions for the entire universe of programs in the university system. This information is matched to schol-

[14]To compute the application score, each program uses the weighted average score considering the pool of scores of the current year and the pool of scores of the previous year (if any). Then, the maximum between these two application scores is considered as part of the application.

arship records and all the information on applications, scores, and socioeconomic characteristics provided by DEMRE. We also have data from MIFUTURO, including programs' and universities' characteristics inside and outside the centralized system, such as their tuition, duration, major, and the program's location.

LABOR MARKET. SIES and MIFUTURO provide aggregate information about the labor market prospects of each program. More specifically, we have estimates for average wages at the program level, for the fourth year after graduation, and the overall employment probability one year after graduation. Moreover, MIFUTURO and SIES also provide information collapsed at the major level, including average from the first to the fifth year after graduation; five points in the distribution of average wages at the first year and fifth year after graduation (percentiles 10th, 25th, 50th, 75th, and 90th), employment probabilities at the first and second year after graduation; and the evolution of average wages from the first to the tenth year after graduation.

GRADES. Since 2013, CRUCH has been gathering grades' records for all students who enroll in a program that is part of the centralized system. This information has been used by CRUCH to test the predictive validity of the admission factors. To our knowledge, this is the first structural project that uses this data.

SURVEYS - 2019 AND 2020. In 2019 and 2020, we designed and conducted, in collaboration with CRUCH and DEMRE, surveys to gather information on students' preferences for programs, and their beliefs on admission probabilities.[15] These surveys were sent to all students that participated in the PSU tests (more than 150,000 each year), and it was sent at the end of the application process.[16] We ask students about their top-true preference, their beliefs on admission probabilities on each program in their ROL, and also regarding their top-true preference (if not in ROL), the probability of enrolling, and the probability of remaining enrolled, conditional on the preference of assignment, among other questions. The structure of the survey is a repeated cross-section. However, as many students re-apply to the centralized system after a year, we have information about students' preferences and their beliefs for a small panel of re-applicants in the survey. To our knowledge, this is the first time that data on beliefs about admission probabilities and college persistence is collected for a centralized college admissions system.

4.3 EMPIRICAL FACTS

The motivating example in Section 3 suggests that there are two possible behavioral channels to explain switching behavior: (i) *mismatching*, whereby students assigned to less preferred programs re-apply to improve their allocation; and (ii) *learning*, whereby

[15] See the Electronic Companion for details on the survey.

[16] Students can update and submit their application as many times as they want during the application time window.

students learn about their abilities and preferences over time and potentially decide to move to other programs. In this section, we provide empirical evidence supporting the existence of these two channels. More specifically, in Section 4.3.1 we show that students assigned to lower reported preferences are less likely to enroll in their assigned program, and they have a higher probability of switching and delaying their graduation than students assigned in higher reported preferences. This fact suggests that students assigned to less preferred programs—mismatched students—are less likely to continue and graduate from their assigned program. On the other hand, in Section 4.3.2, we show that students change their top-true preference over time and re-apply and enroll in programs that are either less preferred or even not present in their initial ROL, suggesting that students' preferences change over time. We conclude this section reporting evidence that this learning process can be through the grades students obtain in the first year of college.

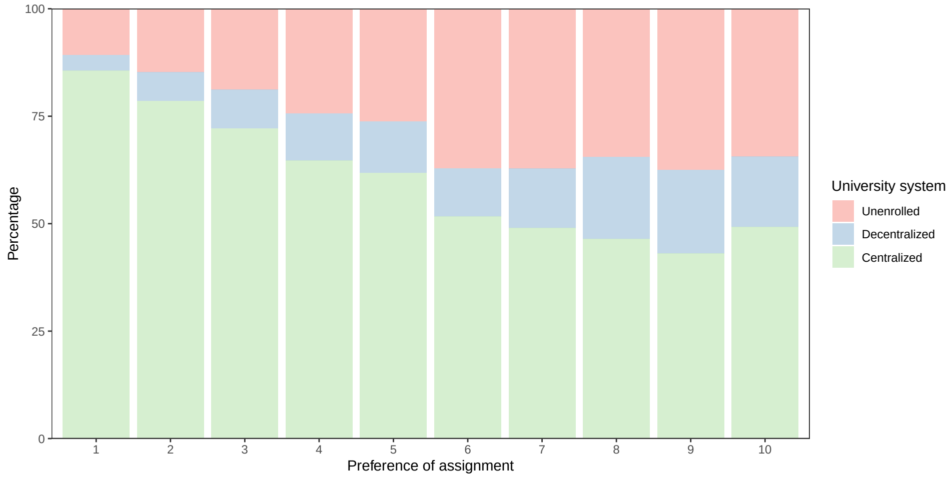
One of the main challenges to disentangle these two behavioral channels is that we do not have cardinal information regarding students' preferences, as we only observe their characteristics and their submitted ROLs. Moreover, students' reports may not be truthful, as some students tend to skip programs for which their admission chances are relatively low (Larroucau and Ríos, 2018). Despite this, we claim that reported ROLs still shed some light on the intensity of students' preferences. For instance, we know that listing a program in a higher position of the ROL implies a higher preference intensity than programs listed in lower preferences (Haeringer and Klijn (2009)). Moreover, not listing a program for which the probability of admission is high enough implies that the ROL programs are preferred (see Larroucau and Ríos (2018) for a detailed discussion). Finally, apart from the information that we can extract from students' ROLs, adding dynamics can help identify preferences' intensity. For example, students who decide to re-apply must have higher intensity in their preferences than students who remain in their program (conditional on observable characteristics and in the absence of learning). Similar information can be inferred from switchings and dropout decisions.

4.3.1 Mismatching

Enrollment. Figure 4.1 shows the fraction of students that enrolls within and outside the centralized system by preference of assignment.[17] We see a steep decreasing pattern, as students assigned to lower reported preferences are less likely to enroll compared to students assigned in higher reported preferences. For instance, close to 90% of students assigned to their top reported preference enroll within the centralized system, whereas close than 50% do so among those assigned to their tenth reported preference.

[17]In Figure B.4 in Appendix B.4 we report the distribution of preference of assignment. We observe that more than 50% of students are assigned to their top reported preference, and this fraction is decreasing in the preference. Also, Figure B.5 shows the distribution of the number of different majors and universities included in each student's ROL. We observe that there is heterogeneity in the ROLs, as an important fraction of students includes two or more different majors and universities. See Larroucau and Ríos (2018) for more details on the application process.

Figure 4.1: Enrollment probability by preference of assignment



Note: Percentage of enrolled students in the centralized and decentralized systems, by preference of assignment in the centralized system.

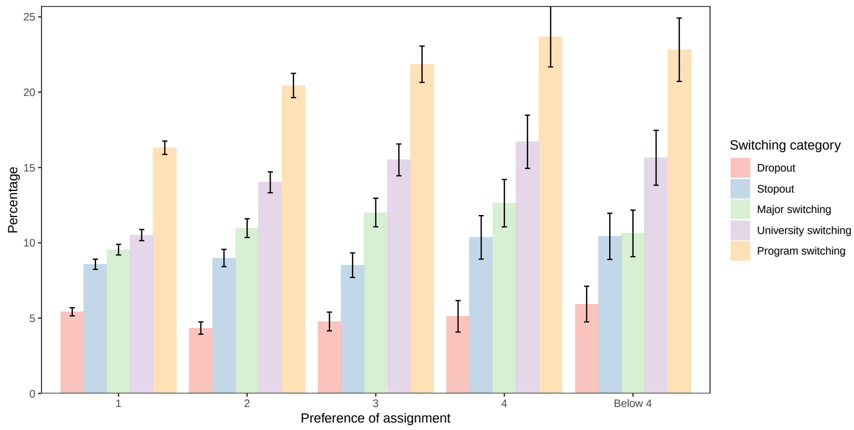
Switchings, Re-Applications, and Dropout. In Table 4.1, we report the average switching and dropout rates separating by income level—high or low—and gender. First, comparing switching and dropout rates by gender (within an income level), we observe that women are more persistent in their academic progression, as their switching and dropout rates are lower than those for men. On the other hand, comparing these rates by income level (within gender), we observe that low-income students are significantly less likely to switch programs during their academic progression. However, we also observe that low-income students are significantly more likely to drop out. One potential explanation is that low-income students have less flexibility to switch programs and delay their graduation due to budget constraints, and at the same time, face a more challenging time in college due to their disadvantageous background, which increases their chances of dropping out. These results suggest that there are significant differences in switching and dropout rates by gender and income.

Table 4.1: Switchings and Dropout by Gender and Income

		Switches				
	Income	Program	University	Major	Math type	Dropout
Men	Low	0.135	0.0656	0.0289	0.0582	0.102
	High	0.180	0.0756	0.0381	0.0721	0.0513
Women	Low	0.0915	0.0513	0.0316	0.0363	0.105
	High	0.150	0.0737	0.0499	0.0532	0.0403

To assess whether the preference of assignment impacts student outcomes, Figure 4.2 shows switching and dropout rates conditional on students' preference of assignment in their first year. We observe that students assigned to lower reported preferences switch majors[18] and universities in higher proportions compared to students assigned to their top reported preference. Indeed, among students assigned to their top reported preference, close to 15% switch programs during the time-span of the sample, compared to almost 25% who are assigned to their fourth choice. The evidence regarding dropouts and stop-outs is less conclusive than switchings. These results suggest a strong corre-

Figure 4.2: Switchings and dropout



Note: the "switching" categories do not include stop out.

lation between the preference of assignment and the probability of switching majors or universities during students' college path. One potential explanation is that there are observable differences between students assigned to lower and higher preferences. For instance, students with low scores are systematically assigned to lower preferences, generating a positive correlation between assignment preference and switching rates. Similarly, programs listed in lower preferences are more likely to be of lower quality, which may incentivize students to try to switch[19] To make a causal claim, we use a regression discontinuity design that exploits the algorithm's cutoff structure to perform the allocation.

Causal effects. If we assume that students around the cutoff are similar and only differ in their right to enroll in a higher preference, we can estimate the causal effect of interest.[20] In Figure 4.3, we display binned means of different outcomes as a function of the distance between the cutoffs and the students' scores considering the applications of students to their most preferred listed program.[21] Figure 4.3a shows that students right below the cutoff are close to 11% more likely to re-apply in the following year, which corresponds

[18]We refer to *majors* as the fields of education provided by the International Standard Classification of Education (ISCED) (UNESCO (2012)) which is adapted to Chile. The modified version of the ISCED fields used in Chile classifies programs into Farming, Art and Architecture, Science, Social Sciences, Law, Humanities, Education, Technology, Health, Management and Commerce.

[19]In Appendix B.2, we estimate *logit* models of the probability of switching majors, universities, or dropping out from the university system, on the preference of assignment, controlling by a rich set of observable characteristics. We find that the previous correlation patterns remain.

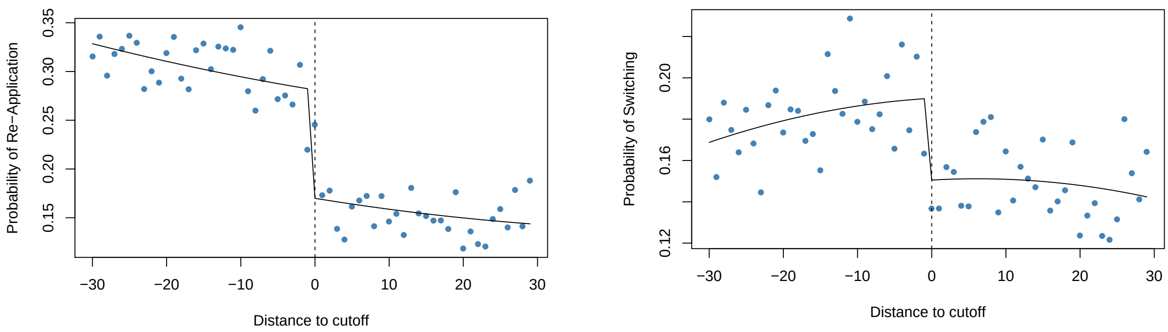
[20]A detailed discussion of this analysis and its potential selection issues are provided in Appendix B.3.

[21]In Appendix B.3.1 we report the results of a similar analysis considering students' top true preferences. The results are relatively the same.

to a relative change of close to 100%.[22] Figure 4.3b shows that students below the cutoff are close to 3.96% more likely to switch programs within the centralized system, which corresponds to a relative change of more than 22%.[23] These results confirm our previous findings, i.e., that students assigned in lower preferences are more likely to re-apply and switch programs in the following year.

Figure 4.3: Effect of Cutoff Crossing

(a) Re-Application (b) Switching



Perceived persistence and preference of assignment. The previous empirical facts show a causal effect of the preference of assignment on students' persistence with respect to their initial assignments. To show that this is partially explained by the mismatch channel, we use the survey on students' preferences and beliefs of 2020. We show that a significant fraction of students know—before enrolling in their assigned programs—that they will be less likely to remain enrolled in the same program if they are assigned to lower reported preferences.[24]

Figure 4.4 shows the average "perceived" probability of remaining enrolled in the same program after one year, by the preference of enrollment. We observe that there is a significantly lower "perceived" probability of enrollment for lower-ranked preferences. On average, students believe that there is an 85% probability of remaining in the same program after a year for their first reported preference, whereas it is close to 65% for programs ranked below the fourth choice. Figure 4.4 also provides evidence of forward-looking behavior (similar to the data patterns observed for students' switching probabilities),[25]

[22]On average, the fraction of students that re-apply is 15.52% and 30.38% for students above and below the cutoff, respectively.

[23]On average, the probability of switching is 14.87% and 18.16% for students above and below the cutoff, respectively.

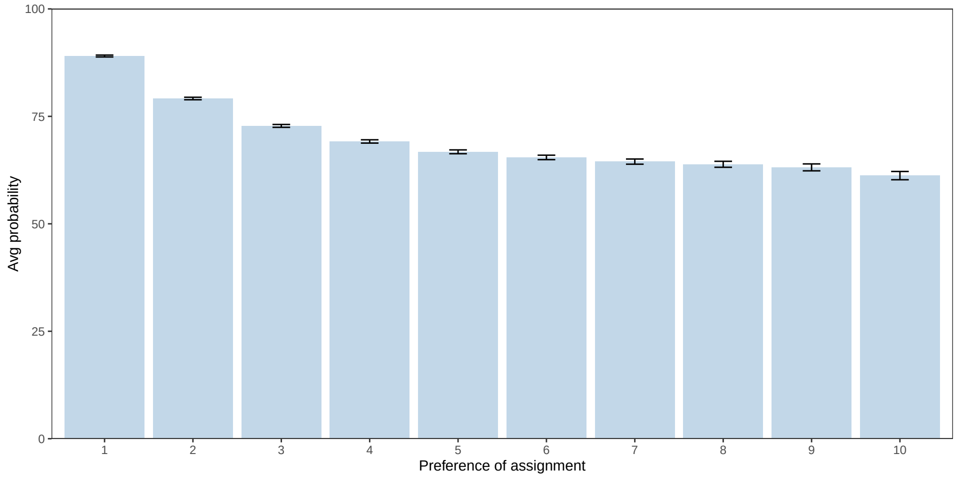
[24]The survey of 2020 included the following question, tailored to elicit students' beliefs about their persistence probabilities, conditional on their preference of enrollment:

The objective of the next question is to know about your preferences and expectations. We remind you that this question is completely hypothetical and will not affect your application or your chances of admission.
With respect to "program i", and in the case that you enroll in this program in the process of 2020, what is the probability that you will enroll in the same program the next year (2021 process)? On a scale from 0 to 100, where 0 is "entirely sure that I will NOT enroll in the same program", and 100 is "entirely sure that I WILL enroll in the same program".

[25]Notice that stop-out and dropout probabilities do not exhibit a positive correlation with the preference

which suggests that–on the aggregate–students' subjective beliefs are close to rational expectations beliefs.

Figure 4.4: Average "perceived" probability of remaining enrolled in the same program, by preference of enrollment



The previous evidence does not guarantee that there are match-effects between students and programs that are correlated with college persistence. For instance, a similar pattern could be observed if all students agree on their preference rankings over programs, and most of the correlation between reported preferences and college persistence was due to programs' characteristics. To rule this out and give evidence of match-effects, we exploit the panel structure of students ROLs, as we observe the perceived persistence probability for every program listed in the ROL. We consider the following specification:

$$P_{ij} = \alpha_i + \alpha_j + X_{ij}\beta + \beta_R R_i(j) + \varepsilon_{ij}, \tag{4.1}$$

where P_{ij} is the perceived persistence probability of student i in program j, α_i is students i's fixed effect, α_j is program j's fixed effect, X_{ij} are student-program characteristics, that include a third-degree polynomial of the application score of student i in program j, $R_i(j)$ is the position of program j in ROL R_i, and ε_{ij} is an i.i.d shock. Table 4.2 shows the estimation results. The preference of enrollment has a significant and strong effect on the perceived probability of persistence. We conclude that there are match-effects in the setting, which exhibit a strong correlation with students' college persistence.

The results reported so far show that (1) there is a clear effect of the preference of assignment on the switching behavior of students, that (2) a significant fraction of students forecast this, and that (3) these results cannot be explained by students or programs' characteristics solely.

of assignment. Thus, they can not drive most of the correlations shown in Figure 4.4.

Table 4.2: Two-way Fixed Effects Regression Results

	Dependent variable: Prob. of Persistence
Preference 2	-9.891^{***}
Preference 3	-16.844^{***}
Preference 4	-21.355^{***}
Preference 5	-24.831^{***}
Preference 6	-27.148^{***}
Preference 7	-29.164^{***}
Preference 8	-30.329^{***}
Preference 9	-31.995^{***}
Preference 10	-34.757^{***}
Constant	89.181^{***}
Observations	159,894
R^2	0.095
Adjusted R^2	0.095

Note: Significance reported: *p<0.1; **p<0.05; ***p<0.01.

4.3.2 Learning

Re-Application and switchings. Students' preferences may change during their first year in college, which could affect their re-applications. We analyze students' re-applications and initial applications and classify switchings in three categories: (i) moving *Up*, (ii) moving *Down*, and (iii) moving *Out*. A student moves *Up* (*Down*) if she switches in 2015 to a program listed above (below) her initial enrollment in 2014 with respect to her application list of 2014. A student moves *Out* if she switches in 2015 to a program not listed in her application list of 2014. We focus on students not assigned to their top choice. We find that among students who switch in first year, 15.8% move *Down*, 21% move *Up*, and 63% move *Out*. Moreover, in Appendix B.9 we show that half of students who move *Out* do so to more selective programs, i.e., programs with higher admission cutoffs compared to their initial enrollment. These results suggest that both channels explain students' switchings significantly. Students who move *Down* or *Out* to less selective programs are likely to have changed their preferences (learning channel), while students' who move *Up* or *Out* to more selective programs may be trying to find a better match (mismatch channel).

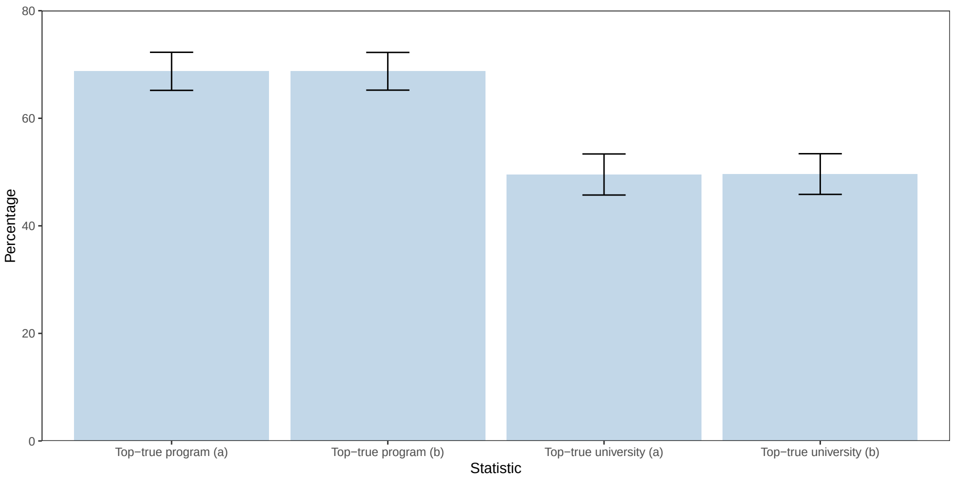
Changes in Top True Preference. To show that students' top true preferences may also change over time, we construct an index of preference variation over time using the two surveys conducted in 2019 and 2020. Given that a significant fraction of students re-apply to the centralized system immediately after their first year of college, we create a panel with close to 2,600 students that participated in both surveys. Students were asked about their top-true preferences in one of two ways:

(a) *If the Admissions Process did not depend on your PSU score, nor your NEM scores, nor your Ranking scores. What would have been your top program choice?*

(b) *If the Admissions Process did not depend on your PSU score, nor your NEM scores, nor your Ranking scores, nor the tuition or enrollment costs. What would have been your top program choice?*

Notice that in version (b), we ask students to report their top-true preference for programs, but hypothetically assuming that they do not have to pay for college. The two versions were randomized at the student level in each survey. Given this randomization, we end up with close to 1,300 students who replied to the same version of both surveys. Figure 4.5 shows the percentage of re-applicants that change their top-true program or the university of their top-true program between 2019 and 2020.

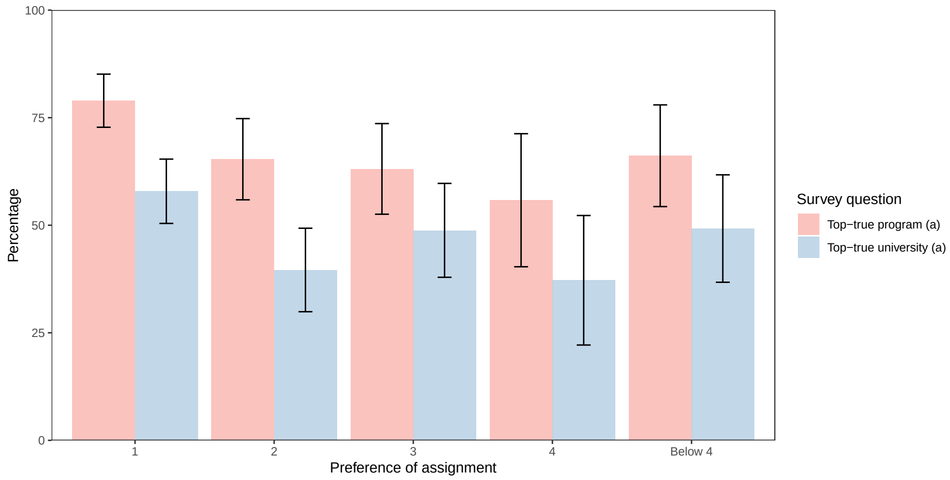
Figure 4.5: Percentage of re-applicants that change their top-true preference



We observe that close to 65% of students who re-applied immediately after their first year of enrollment reports a different top-true preference. Moreover, close to 77% of these changes involve a change in the university of their top-true preference, which suggests that the *learning* mechanism is also present, and it explains students' re-application and switching behavior.

Finally, to disentangle whether the correlation patterns between switching probabilities—realized and forecasted by students—and the preference of assignment are not driven by an increasing prevalence of learning in lower reported preferences compared to the top preference, we compute the percentage of re-applicants that change their top-true preference, by the preference of assignment. Figure 4.6 shows the percentage of re-applicants who change their top-true program or the university of their top-true program between 2019 and 2020. We do not find evidence that lower preferences of assignment are correlated with a higher incidence of learning. Indeed, when looking at students assigned to their top-reported preferences in 2019 and that re-applied in 2020, we observe that a higher share report having changed their top-true preference in the survey compared to students assigned to a lower reported preference. This result is consistent with our claim that the *mismatch* and the *learning* channels affect students' college persistence, as students initially assigned to their top-reported preferences have a lower probability of being *mismatched*, and thus their re-application suggests that they learned about their preferences during their first year in college.

Figure 4.6: Percentage of re-applicants that change their top-true preference, by preference of assignment in 2019



Learning through Achievement. In Table 4.3 we report the results of logit regressions aiming to measure the effect of grades on different outcomes, including the decision to re-take the PSU tests, re-apply in the centralized system, and switch to a different program (either within or outside the centralized system). In all these models, we control for demographics (gender, income, school type), scores (NEM and average between Language and Math), university and major fixed effects, and the preference of assignment in the initial year. Columns (1), (3), and (5) include the entire sample, while columns (2), (4), and (6) focus on students with a GPA greater than or equal to 4.0. Since 4.0 is the pass/-fail threshold (the scale is from 1.0 to 7.0), by focusing on students with GPA above 4.0 we rule out the explanation that all students who switch were forced to leave their initial programs.

We observe the coefficient of GPA is negative and significant for all specifications, meaning that a higher GPA reduces the probability of re-taking the PSU tests, re-applying, and finally changing programs. Besides, we observe that, even after controlling for grades, the preference of assignment has a significant effect on these outcomes, and these results are consistent with those described above.

The last column (7) in Table 4.3 reports the results of regressing the average GPA in the first year of college on the same variables used for the logit models. We observe that all coefficients are negative and most of them are significant, implying that being assigned to a lower preference is slightly correlated with a lower average GPA.

Table 4.3: Effect of Grades on Outcomes

	Re-Take PSU		Re-Apply SUA		Change Program		GPA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GPA	-0.904^{***}	-0.437^{***}	-0.903^{***}	-0.437^{***}	-1.221^{***}	-0.357^{***}	-
	(0.018)	(0.042)	(0.018)	(0.042)	(0.019)	(0.039)	-
Preference 2	0.653^{***}	0.870^{***}	0.651^{***}	0.868^{***}	0.163^{***}	0.275^{***}	-0.057^{***}
	(0.040)	(0.048)	(0.040)	(0.048)	(0.038)	(0.046)	(0.011)
Preference 3	0.922^{***}	1.141^{***}	0.923^{***}	1.142^{***}	0.352^{***}	0.437^{***}	-0.061^{***}
	(0.050)	(0.060)	(0.050)	(0.060)	(0.050)	(0.060)	(0.015)
Preference 4	1.201^{***}	1.387^{***}	1.202^{***}	1.388^{***}	0.562^{***}	0.630^{***}	-0.070^{***}
	(0.070)	(0.083)	(0.070)	(0.083)	(0.071)	(0.083)	(0.022)
Preference 5	1.116^{***}	1.366^{***}	1.116^{***}	1.366^{***}	0.523^{***}	0.621^{***}	-0.013
	(0.103)	(0.117)	(0.103)	(0.117)	(0.102)	(0.117)	(0.032)
Preference Below 5	1.098^{***}	1.334^{***}	1.099^{***}	1.334^{***}	0.454^{***}	0.633^{***}	-0.113^{***}
	(0.112)	(0.132)	(0.112)	(0.132)	(0.115)	(0.134)	(0.035)
GPA ≥ 4.0	No	Yes	No	Yes	No	Yes	No
Observations	39,275	31,976	39,275	31,976	39,275	31,976	39,275

Note: We use data on grades from the cohort that graduated from high-school in 2014 and enrolled in 2015 the program they were assigned in the centralized system. GPA is measured on a scale of 1 to 7, and failing grades are below 4.0. Significance reported: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

5 MODEL

This section describes our model of students' applications, enrollment, and dropout decisions, including learning about their match-quality with programs over time. The goal is to have a model that encompasses the empirical evidence described in the previous sections, allowing us to measure how much of students' switching behavior is explained by learning over time vs. initial mismatches, and assess whether students' outcomes can be affected by changing the mechanism and re-application rules.

Throughout the model, we assume that abilities are multidimensional and partially known by students. In particular, students receive signals of their unknown abilities through their college GPA and—based on this information—they update their beliefs. Given their updated beliefs, students choose to (i) continue in their enrolled programs, (ii) re-apply to the centralized system expecting to switch programs, or (iii) dropout from the centralized system. Finally, we model labor market returns as a function of the major, students' abilities and observable characteristics, and their path through college.

5.1 MODEL OVERVIEW

For estimation purposes, we consider a three-period model. Periods 1 and 2 correspond to the first and second years of college after graduating from high-school. Period 3 starts at the beginning of the third year of college and collapses the later years until graduation, with the discounted payoffs received in the labor market. Every period has several stages that we explain in detail below.

Period 1. In period 1, students who have just graduated from high-school make their application decisions, receive their enrollment, choose to whether re-take the PSU again or not, obtain their college grades at the end of the first year, and update their beliefs about their unknown ability.

(i) *Applications:* given students' preferences, beliefs over their admission and enrollment probabilities, prior beliefs about their unknown abilities, and the labor market return of studying each option, students make application decisions to the centralized system. After knowing their preference shocks' realizations, students choose a ROL that maximizes their expected utility.

(ii) *Assignment:* once applications are made, a matching algorithm computes students' assignment to each program. In particular, this process is approximated by drawing a set of admission cutoffs from the joint distribution of cutoffs and assigning students according to the matching algorithm's cutoff structure.

(iii) *Enrollment:* once the assignment is realized, students face exogenous probabilities of enrollment in their assigned program or choosing the outside option.

(iv) *PSU preparation:* at the beginning of students' first year of college, or in the outside option, students choose whether to prepare and re-take the PSU tests. This decision affects their flow utility while in college and can improve the set of potential programs they can enroll in the second period if they choose to re-apply to the system.

(v) *Grades:* at the end of the year, students receive their college grades—which are noisy signals of their unknown abilities—and update their beliefs.

Period 2. In period 2, students decide whether to re-apply to the centralized system, and depending on their assignment and enrollment status, choose between remaining in their current enrollment, switching to their new assigned program, or dropping out from college.

(i) *Re-applications:* at the beginning of period 2, students observe the realization of new preference shocks and PSU scores and, given their updated beliefs on their unknown abilities, decide whether to re-apply[26] to the centralized system.

(ii) *Assignment:* once applications are made, a matching algorithm computes students' assignment to each program. In particular, this process is approximated by drawing a set of admission cutoffs from the joint distribution of cutoffs and assigning students according to the matching algorithm's cutoff structure.

[26]Students can also apply for the first time in period 2.

(iii) *Enrollment:* once the assignment is realized, students face exogenous probabilities of enrollment in their assigned program. If students do not enroll in their assigned program, they can choose between staying in their current enrollment or dropping out of college.

(v) *Grades:* at the end of the year, students receive their college grades and update their beliefs regarding their unkonwn abilities.

Period 3. Students face exogenous graduation probabilities and enter the labor market.

(i) *Expected graduation:* students face an exogenous graduation probability for every year they are enrolled after completed the formal duration of their programs.

(ii) *Labor market:* students who graduate receive the lifetime earnings and non-pecuniary payoffs given their college decisions. Students who do not graduate receive the value function of students who dropped out.

5.2 LABOR MARKET

For the labor market stage of the model we follow Arcidiacono (2004) and Arcidiacono (2005). The labor market is an absorbing state, and utility while in the workforce is given by the present value of lifetime earnings and non-pecuniary utility. We further assume that utility is separable over time. In particular, we assume the following specification

$$V_{ijt}^{w} = \underbrace{\alpha_1^w \alpha_{im_j} + \alpha_2^w A_{ij} + \alpha_3^w \bar{A}_{k_j} + \alpha_4^w A_{ij}^u}_{\text{non-pecuniary}} + \underbrace{\alpha_5^w \log\left(E_w\left[\sum_{\tau=0}^{T-t} \beta^{\tau} P_{m_j\tau}^{w} w_{ij\tau}\right]\right)}_{\text{pecuniary}}, \tag{5.1}$$

where the first four terms capture the non-pecuniary payoff that individuals perceive from working in a job associated with program j. We allow these payoffs to vary with the student's observed ability in program j, A_{ij}, the average observed ability of students in college k_j, $\bar{A}_{k_j}$, and the student's unknown ability A_{ij}^u. We also include student i's random coefficient for *major* m_j, α_{im_j}.[27] By incorporating random coefficients we introduce persistence over time on students' unobserved preferences, which can affect both their flow utility and their utility in the work force.[28] The fifth term captures the pecuniary payoff that students receive in the work force, with $w_{ij\tau}$ representing the earnings for student i with tenure τ, graduating from program j. In addtion, T is the retirement

[27]We refer to *majors* as the fields of education provided by the International Standard Classification of Education (ISCED) (UNESCO (2012)) that is adapted to Chile. The modified version of the ISCED fields used in Chile classifies programs into: Farming, Art and Architecture, Science, Social Sciences, Law, Humanities, Education, Technology, Health, and Management and Commerce. We further group these categories into four broad majors: Science (Science, Farming, and Technology), Social Sciences (Social Sciences, Art and Architecture, and Law), Education and Humanities (Education and Humanities), and Health (Health).

[28]In Section 5.3.1 we describe how we model the random coefficients.

date (which varies by gender), t corresponds to the year—period—in which the student graduates from college and enters the work force, β is a common discount factor, and $P^w_{m_j\tau}$ is the employment probability in *major* m_j for an individual with tenure τ. Notice that student i receives this continuation value only if she graduates from her program. If, instead, student i drops out in period t, we assume that she receives a continuation value given by V_{i0t} that depends only on her observable characteristics X_i.[29] This is formalized in Assumption 1.

Assumption 1. If student i graduates from program j in period t, she obtains a continuation value equal to V^w_{ijt}. In contrast, student i receives a continuation value equal to V_{i0t} if she drops out from her program in period t.

We specify the wage that students receive conditional on graduation as a function of their tenure, their *major* m_j, their observable characteristics Z^w_i, their average grades in college $\bar{G}_{ij}$, and the average ability of their classmates $\bar{A}_{k_j}$.[30] More specifically, we assume that the log earnings for student i with tenure τ, graduating from program j in period t, can be written as

$$\log(w_{ij\tau}) = \lambda_{1m_j} + \lambda_2\bar{A}_{k_j} + \lambda_3\bar{G}_{ij} + \lambda_4 Z^w_i + \Lambda_{m_j\tau} + \epsilon_{ij\tau}, \tag{5.2}$$

where $\Lambda_{m_j\tau}$ is a function that specifies how wages in *major* m_j depend on tenure τ. In particular we consider a parsimonious specification:

$$\Lambda_{m_j\tau} = \lambda_{5m_j}\tau + \lambda_{6m_j}\tau^2. \tag{5.3}$$

5.3 Academic Progression

During their academic progression, students receive their flow utility from attending college and observe their grades, which provide them a signal of their unknown abilities. As we discussed in the previous section, students take into account their ability when computing their labor market returns, and thus the information obtained from their grades is highly valuable. Students may use this information to decide whether to re-apply in the next period, continue enrolled in the same program, or drop out of college. In this section, we model the flow utility obtained in each period in college, the learning process, and the graduation process. We defer the model of re-applications to Section 5.4.

[29] In Section 5.3.3 we describe the dropout and graduation process. In addition, we estimate V_{i0t} as

$$V_{i0t} = V_0\left(X_{i0}, t\right),$$

where X_{i0} is a vector of observable characteristics of student i at the beginning of the horizon. The vector X_{i0} includes student's gender.

[30] More details about the ability are reported in Section 5.3.2.

5.3.1 FLOW UTILITY

Let u_{ijt} be the flow utility that student i receives for attending program j at time t,[31]

$$u_{ijt} = \alpha_{im_j} + \alpha_{ik_j} + Z_{ij}^u \alpha - C_{ijt} + \varepsilon_{ijt}, \tag{5.4}$$

where α_{im_j} and α_{ik_j} are student i's random coefficients for *major* and *university type*, respectively;[32] $Z_{ij}^u \alpha$ captures the effect of student and program characteristics that are time invariant,

$$Z_{ij}^u \alpha = \alpha_1 A_{ij} + \alpha_2 \bar{A}_j + \alpha_3 D_{ij} + \alpha_4 \frac{(A_{ij} - \bar{A}_j)}{\bar{\sigma}_j} \tag{5.5}$$

where D_{ij} is the distance between student i's and program j's municipalities; A_{ij} is student i's observed ability in program j, $\bar{A}_j$ is the average observed ability for students assigned in program j in the previous calendar year (program's selectivity), and $\bar{\sigma}_j$ is its standard deviation.[33] Finally, C_{ijt} captures the monetary cost for student i to enroll in program j at time t and it is given by

$$C_{ijt} = \alpha_{c0} c_{jt} + \alpha_{c1} c_{jt} \mathbb{1}_{(\text{low income})} + \alpha_{c2} c_{jt} \mathbb{1}_{(\bar{s}_i \geq 500)}, \tag{5.6}$$

where c_{jt} is program j's yearly tuition. We allow for different price sensitivities depending on the level of income and students' scores. In this way, we capture potential credit constraints that may affect low-income families, as well as potential scholarships or financial aids that high-achieving students may have access to and that are not included in the data.

We follow Larroucau and Ríos (2018) and model the random coefficients as a multivariate regression on a set of students' observable characteristics. In particular,

$$\alpha_{im} = \Delta^m Z_i^m + \chi_i^m, \quad \alpha_{ik} = \Delta^k Z_i^k + \chi_i^k, \tag{5.7}$$

where Δ^m and Δ^k are matrices of coefficients to be estimated, $\chi_i^m \sim N(0, V_\alpha^m)$ and $\chi_i^m \sim N(0, V_\alpha^m)$ are vectors of idiosyncratic shocks with mean zero and variance-covariance matrices $V_\alpha^m = \sigma_\alpha^{2m}\mathbb{I}$ and $V_\alpha^k = \sigma_\alpha^{2k}\mathbb{I}$, respectively; and Z_i^m and Z_i^k are matrices of observable characteristics, where the former includes a constant and students' gender, while the latter includes a constant and students' family-income type.[34] Finally, ε_{ijt} is an idiosyncratic preference shock that is distributed i.i.d type I extreme value with a scale parameter of one. We specify a location normalization, and we set the systematic value of the outside option (not enrolling in a program within the centralized system) to be $\bar{u}_{i0t} = 0$.

---

[31] See Larroucau and Ríos (2018) for a similar specification.

[32] We classify universities in three categories: CRUCH-Public, CRUCH-Private and Non-CRUCH.

[33] The coefficient α_1 captures the possibility that students with high scores (observed abilities) could perceive a higher flow utility of enrolling in the centralized system in particular programs, compared to students with low scores (observed abilities). The coefficient α_2 captures how much students care about the level of selectivity of their enrolled program (which can be seen as a proxy for programs' qualities). The coefficient α_4 captures how much students like a program depending on their ability relative to students assigned in the previous year.

[34] We classify students' as *low-income* if their self-reported family income is below the median of the family income distribution, and as *high-income* otherwise.

5.3.2 LEARNING

As described in Equation 5.2, students' (pecuniary) labor market returns depend on their grades, which in turn depend on their abilities. We assume that these abilities have two components, one that is directly observable and known by students (and the econometrician), and another that is unknown and learned from the grades obtained during college. More specifically, we assume that students have beliefs on their abilities, and they update them as they observe their grades according to Bayes rule. To formalize these ideas, we start modeling students' abilities. Then, we model the grade equation, and we finish this section by modeling beliefs and the updating process.

Ability. Each student i has an observed subject-specific ability vector $A_i = (A_{is_m}, A_{is_v})$; an unobserved (to the student and to the econometrician) subject-specific ability vector $A_i^u = (A_{is_m}^u, A_{is_v}^u)$; and a *major*-specific ability $A_{im_j}^u$ for each *major* m_j. Each component of these ability vectors captures the student's known and unknown abilities in math and verbal, indexed by s_m and s_v, respectively. We assume that student i's (un)observed ability in program j is given by the weighted sum of her (un)observed abilities, i.e.,

$$A_{ij} = \sum_{k \in \{s_m, s_v\}} \omega_{jk} A_{ik}, \quad \text{and} \quad A_{ij}^u = A_{im_j}^u + \sum_{k \in \{s_m, s_v\}} \omega_{jk} A_{ik}^u, \tag{5.8}$$

where ω_{jk} is the admission weight of factor k in program j. Even though the subject-specific components do not vary across programs, there is still variation on students weighted abilities—even within a specific *major*—due to the heterogeneity on programs' specific weights, ω.

Grades. As described above, we assume that students observe their grades at the end of each of the first two periods and, based on these signals, they update their beliefs on their unknown abilities.

We assume that grades depend on the *major* (m_j) of the program where the student is enrolled, on the known (A_{ij}) and unknown abilities (A_{ij}^u), and on a set of observable characteristics (Z_i^g).[35] Also, to capture that students' initial preferences may affect their performance, we include student i's random coefficients for *major*, α_{im_j}, and *university type*, α_{ik_j}.

Following Arcidiacono (2004), we assume that the grade equation for the first period is given by

$$G_{ij1} = \gamma_{1m_j} + \gamma_2 A_{ij} + \gamma_3 Z_i^g + \gamma_4 \alpha_{im_j} + \gamma_5 \alpha_{ik_j} + A_{ij}^u + \varepsilon_{ij1}^g, \tag{5.9}$$

where ε_{ij1}^g is a white noise distributed $N(0, \sigma_g^2)$. In addition, for the second-year grade equation we allow a different intercept and slope for those students who are in their second academic year,[36] but the relative importance of each component remains the same,

[35] The estimation results described in Section 7 only include gender.

[36] Students who either switch programs after first year, or who enrolled for the first time in the second year, are in their first academic year in the second period.

that is,

$$G_{ij2} = (1 + Y_2\gamma_6)\,\gamma_{1m_j} + (1 + Y_2\gamma_7)\left(\gamma_2 A_{ij} + \gamma_3 Z_i^g + \gamma_4\alpha_{im_j} + \gamma_5\alpha_{ik_j} + A_{ij}^u\right) + \varepsilon_{ij2}^g, \quad (5.10)$$

where Y_2 indicates whether the student is in her second academic year in program j, and ε_{ij2}^g is a white noise normally distributed with mean zero and variance σ_g^2.

Beliefs and Updating. We assume that students are rational and update their beliefs using the signals about their unknown abilities that come with their grades according to Bayes rule. In particular, we assume that students' initial prior about their unobserved major-specific ability is normally distributed with mean zero and variance σ_m^2 for all students and majors. Similarly, we assume that students' prior about their unobserved subject-specific abilities is also normally distributed with mean zero and variance σ_s^2 for all students and subjects. We formalize this in Assumption 2.

Assumption 2. Students initial priors on their unobserved major and subject-specific abilities are normally distributed with means zero and variances σ_m^2 and σ_s^2, respectively. These priors are common to all students.

A direct consequence of this assumption is that the posterior distribution of the overall unknown ability in Equation 5.8 will also follow a normal distribution. Let $\mu_t(A_{ij}^u)$ and $\sigma_t(A_{ij}^u)$ be the prior mean and standard deviation of A_{ij}^u at the beginning of period t. When clear from the context, we will remove the argument and simply write them as μ_{ijt} and σ_{ijt}, respectively. Hence, Assumption 2 implies that

$$\mu_{ij1} = 0, \quad \text{and} \quad \sigma_{ij1}^2 = \sigma_m^2 + \sum_{k\in\{s_m,s_v\}} \omega_{jk}^2\sigma_s^2.$$

In Proposition 1, we show how to compute the posterior mean and variance of the overall unobserved ability after observing a signal a_{ijt}. We defer the proof to Appendix C.1.

Proposition 1. Suppose that student i is enrolled in program j in period t, and that she observes a signal a_{ijt}. Then, she will update her mean unobserved ability in each program j' according to:

$$\mu_{ij't+1} = E_t\left(A_{ij'}^u \middle| a_{ijt}\right) = \begin{cases} \left(\sigma_{ijt}^2 + \sigma_g^2\right)^{-1} \cdot \left[\sum_{l\in\{s_m,s_v\}} \omega_{j'l}\omega_{jl}\sigma_s^2 a_{ijt}\right] & \text{if } m_{j'} \neq m_j \\ \left(\sigma_{ijt}^2 + \sigma_g^2\right)^{-1} \cdot \left[\sum_{l\in\{s_m,s_v\}} \omega_{j'l}\omega_{jl}\sigma_s^2 a_{ijt} + \sigma_m^2 a_{ijt}\right] & \text{if } m_{j'} = m_j \end{cases}$$
$$\sigma_{ij't+1}^2 = V_t\left[A_{ij}^u \mid a_{ijt}\right] = \left((\sigma_{ijt}^2)^{-1} + (\sigma_g^2)^{-1}\right)^{-1} \quad (5.11)$$

Intuitively, students will learn more about similar programs to the ones they are currently enrolled in, especially for programs that belong to the same major and that place similar weights in the admissions' scores. It is crucial to notice that, according to our model, only those students who are enrolled in a program observe a signal of their abilities. Hence, we assume that students who are not enrolled do not update their prior.[37]

[37]We make this assumption because we do not have data on students' grades outside the centralized system.

5.3.3 Dropout and Graduation

We assume that the academic progression concludes with students either (i) graduating from their program (after period 2) or (ii) dropping out. We assume that these outcomes are exogenously given so that the probabilities of graduating and dropping out depend only on the students' observable characteristics and on their first and second-year choices. This is formalized in Assumption 3.

Assumption 3. Students have rational expectations over their graduation and dropout probabilities. Moreover, we assume that this decision follows a multinomial logit model, i.e.,

$$P^g_{ij\tau} = \frac{\exp\left(X_{ij\tau}\psi^g\right)}{1+\sum_{a\in\{g,d\}}\exp\left(X_{ij\tau}\psi^a\right)}, \quad \text{and} \quad P^d_{ij\tau} = \frac{\exp\left(X_{ij\tau}\psi^d\right)}{1+\sum_{a\in\{g,d\}}\exp\left(X_{ij\tau}\psi^a\right)} \tag{5.12}$$

where $P^g_{ij\tau}$ and $P^d_{ij\tau}$ represent the probabilities that student i decides to graduate and dropout from program j after τ periods enrolled in the program, respectively; $X_{ij\tau}$ is a vector of observable characteristics,[38] and ψ^g, ψ^d are vectors of parameters that need to be estimated.

5.4 Admission Process

Every time a student decides to (re-)apply, we assume that they go through the following steps: (i) PSU tests, (ii) application, and (iii) enrollment.

5.4.1 PSU Tests

As described in Section 4.1, the assignment is based on a series of admission factors, which include the PSU tests and two additional scores related to the students' performance during high-school. Let $\mathcal{L} = \{1,\ldots,L\}$ the set of admission factors, and let $\vec{s}_{it} = \{s_{itl}\}_{l\in\mathcal{L}}$ be the vector of scores of student i in period t. In addition, let ω_{jtl} be the weight that program j assigns to factor $l\in\mathcal{L}$ in period t. Then, the application score of student i in program j and period t is given by

$$s_{ijt} = \sum_{l\in\mathcal{L}} \omega_{jtl}\cdot s_{itl}.$$

Since students can re-take the PSU tests and re-apply, we need to model (i) the evolution of their scores and (ii) the evolution of their beliefs on the admission weights that programs will use in the future. To model the former, we assume that the scores of student i in period $t+1$, $\vec{s}_{it+1}$, are exogenously given conditional on the scores of the student in

[38] The vector includes a constant, student's gender, a dummy variable that identifies if the student's family income is below the median of the income distribution, and student's High-school GPA.

period t, $\vec{s}_{it}$, and the observable characteristics, X_i. To address the latter, we assume that students correctly forecast future weights. This assumption is likely to hold in practice as admission weights are relatively stable over time. These considerations are formalized in Assumption 4.

Assumption 4. Conditional of re-taking the exam, the scores of student i in period $t+1$ are exogenously given and distributed according to

$$\vec{s}_{it+1} \sim F_{\vec{s}_{it},X_i}(s), \quad \forall i \tag{5.13}$$

where $F_{\vec{s}_{it},X_i}(s)$ is the distribution of scores conditional on the initial vector of scores $\vec{s}_{it}$ and the observable characteristics X_i. In addition, we assume that students correctly forecast the admission weights $\{\omega_{jtl}\}_{l\in\mathcal{L}}$ used by each program j in each period t.

As a simplifying assumption, we further assume that the evolution of scores in each admission factor is proportional to the student's current scores, as described in Assumption 5.[39]

Assumption 5. The scores of student i evolve according to the following process:

$$\vec{s}_{it+1}|\vec{s}_{it}, X_i \sim \max\{s_{it}, \tilde{s}_{it+1}\} \tag{5.14}$$

with

$$\tilde{s}_{ilt+1} = \begin{cases} \alpha_l(1+\nu_{it+1})s_{ilt} & \text{if } s_{ilt} > 0 \\ \alpha_{0l}(1+\nu_{it+1})\bar{s}_{it} & \text{if } s_{ilt} = 0 \end{cases} \quad \text{and} \quad \nu_{it+1} \sim N(0, \sigma_{psu}),$$

where s_{itl} is the score of student i in exam l in period t, $\bar{s}_{it}$ is the average Math-Verbal score of student i in period t, and $\{\alpha_l, \alpha_{0l}\}_{l\in\mathcal{L}}$, and σ_{psu} are parameters to be estimated.

Finally, we assume that students must pay a cost for re-taking the PSU tests. This cost accounts for the direct cost of taking the exam and the time spent to prepare for it. Since we do not have information regarding preparation time, we assume that this cost is a constant C^{psu}.

5.4.2 Application

Once students get their scores—either the first time they take the exams or after re-taking them—they must decide which programs to include in their ROL. We assume that students' application behavior can be classified as one of two types: (i) weak truth-tellers, and (ii) strategic.[40] These types are exogenously given, with students being weak-truth-tellers with probability ρ and strategic with probability $1-\rho$. We assume that weak

[39] This specification captures the fact that students use the maximum application score from both pools of test scores, for each program they apply to.

[40] Figures B.6 and B.7 in Appendix B.4 show evidence that this mixture provides a good approximation of the observed behavior of students. More specifically, Figure B.6 shows that the cutoff of students' top true preference is in most of the cases higher than that of the top reported preference, which suggests that students take into account their admission chances when deciding where to apply. On the other hand, Figure B.7 shows that a large fraction of students do not report their true preferences, even when the constraint on the length of the ROL is not binding.

truth-tellers report their true preferences as long as they exceed the outside option, while strategic students submit a ROL that maximizes their expected value. Following Chade and Smith (2006), we assume that this process can be modeled as an optimal portfolio problem. Each student i that applies in period t considers a vector of indirect utilities $\{v_{ijt}\}_{j\in M}$ and a vector of beliefs on admission probabilities $\{p_{ijt}\}_{j\in M}$, and the submitted ROL R_{it} satisfies

$$R_{it} \in \underset{R'\in\mathcal{R},|R'|\leq K}{\operatorname{argmax}} U(R') - c(R'). \tag{5.15}$$

where

$$U(R) = z_{R(1)} + (1 - p_{R(1)}) \cdot z_{R(2)} + \ldots + \prod_{l=1}^{k-1}(1 - p_{R(l)}) \cdot z_{R(K)}, \tag{5.16}$$

$z_{R(k)} = p_{R(k)} \cdot v_{R(k)t}$ represents the expected utility (over the assignment) obtained from the k-th preference in the ROL, and $c(R)$ is the cost of submitting the ROL R.[41]

This model relies on the assumption that students perceive their admission chances as independent across programs. Also, to simplify the analysis, we further assume that students do not include programs in their ROL unless it is strictly profitable. This assumption implies that strategic students will not add programs for which their admission probability is zero. Finally, we assume that students have rational expectations regarding their admission probabilities.[42] These assumptions are formalized in Assumption 6.

Assumption 6. Students take the distributions over cutoffs to be independent across programs. In addition, students have rational expectations regarding their admission probabilities, and they include programs in their portfolio only if it is strictly profitable to do so.[43]

Discussion: we include the mixture of application behavior because it better fits the application patterns in the data. However, the parameter ρ should not be interpreted as a primitive of the model, as we expect to vary with the counterfactuals. The reason is that, in the baseline, it could be payoff equivalent to report a ROL as a weak truth-teller or strategically. However, if we change the assignment mechanism or the re-application rules, acting as a weak truth-teller may lead to a payoff relevant strategic mistake. As we do not model the latter, in our counterfactual analysis we consider two scenarios: (i) all students behave strategically,[44] and (ii) a fraction $(1-\rho)$ behaves strategically.

5.4.3 Assignment and Enrollment

Once students submit their optimal ROLs, they observe a draw from the joint distribution of cutoffs. Let $\vec{s}^{\,t} = \left\{\bar{s}^t_j\right\}_{j\in M}$ the vector of realized cutoffs in period t. Based on the

[41]In our setting students face no monetary costs for submitting a ROL, thus, we assume that $c(\cdot) = 0$.

[42]This is a common assumption in the literature (Agarwal and Somaini, 2018; Larroucau and Ríos, 2018). Moreover, as Figure B.8 in Appendix B.4 shows, students tend to forecast the cutoff of the current year correctly.

[43]We discuss the implications of Assumption 6 and how to solve the optimal portfolio problem in Appendix C.3.

[44]This would hold if information policies that give precise information about admission probabilities to students were implemented.

mechanism's cutoff structure, the allocation can be easily obtained by assigning each student to the highest preference for which their application score is greater than or equal to the realized cutoff.

After the assignment results are released, students decide whether to enroll in their assigned preference, remain enrolled in their current program if they are re-applying, or take the outside option. For simplicity, we do not model this decision and simply assume that students enroll in their preference of assignment with some exogenous probability P^e_{it} that depends on their observable characteristics.[45] This is formalized in Assumption 7.

Assumption 7. Student i enrolls in her assigned program in period t with probability P^e_{it}, which is given by

$$P^e_{it} = \frac{\exp\left(X^e_i \psi^e\right)}{1 + \exp\left(X^e_i \psi^e\right)}, \tag{5.17}$$

where X^e_i is a vector of observable characteristics.[46]

If students do not enroll in their new assignment, we allow them to choose the best alternative between remaining in their current program for one more period or choosing the outside option. In Appendix C.2 we show that the solution to the student's problem can be obtained via Backward Induction.

6 IDENTIFICATION

In this section we describe our identification strategy and how we use the data described in Section 4.2 to this end.

Labor Market. As discussed in Section 4.2, we only have information about wages aggregated at the program and major levels. We identify the wage equation parameters (λ) by exploiting variation across programs on students' average wages and their correlation with students' and programs' characteristics.[47] The non-pecuniary labor market parameters (α^w) are identified by the correlation between student observable characteristics and graduation probabilities. To identify the effect of students' random coefficient on the non-pecuniary labor market utility (α^w_1), we use the correlation between students' reported preferences and graduation probabilities.[48]

As we do not have information on wages for students who dropped out of college, we model the value functions of dropping out as a function of students' observable charac-

[45]Students pay an enrollment cost C^e for the first time they enroll in a program, which captures both administrative and potentially psychological costs of first-time enrollment process.

[46]The vector includes a constant, student's gender, a dummy variable that identifies if the student's family income is below the median of the income distribution, and student's High-school GPA.

[47]In Appendix D.1, we provide an example of how we can identify these parameters using aggregate data.

[48]In Section E.1 we show how we use this correlation in estimation.

teristics. Intuitively, these value functions' parameters are identified by the share of students who dropped out conditional on their observable characteristics, including gender, income level, among others.

Flow utility. The identification challenge of separately identifying the parameters that govern the unobserved preferences' from those related to the learning process is that both channels affect students' choices over time and are unobserved by the econometrician. However, due to the rational expectations assumption and the assumption on common prior beliefs about students' unknown abilities, students' initial application decisions are informative of students' unobserved preferences because students have not received any signal about their unknown abilities when they submit their initial applications. Hence, we can identify the flow utility parameters using students' initial choices and the correlation between students' characteristics and the characteristics of the programs they list and enroll. In particular, to identify the major and university-type specific parameters ($\alpha_{im_j}, \alpha_{ik_j}$ in 5.4), we leverage the heterogeneity in terms of major and college types within students' ROLs.[49] Then, we use these values as moments to be matched in the estimation procedure. On the other hand, to identify the parameters in Equation 5.5, we use the variation on students' observed ability (to identify α_1), the variation on the level of selectivity (as a proxy for programs' quality to identify α_2), and the variation on students' observed ability compositions across programs—particularly variation in the standard deviation of students' observed ability across programs— (to identify α_4). The variation on the distance between students and programs identifies the coefficient α_3, and the variation on programs' tuition levels and students' socioeconomic characteristics identifies the coefficients of the cost function C. Finally, as standard practice, we normalize the logit shocks' scale to one, the mean utility of the outside option to zero, and we further consider a discount factor β equal to 0.9.

Grades and Learning. According to Equations 5.9 and 5.10, grades are functions of observed characteristics, students' unobserved preferences for majors/colleges, students' unknown abilities, and the signal's noise. To identify the effect of unobserved preferences for majors/colleges, we use the correlation between grades and students' preferences and their assignment, and we also use the correlation between students' application composition—share of different majors and share of different college types—and grades. Intuitively, if students' unobserved preferences for majors positively affect their college grades, there should be a positive correlation between students' reported preferences and their first-year college grades. Similarly, we expect that students whose ROLs imply a high preference for a particular major—i.e., having a high share of programs that belong to the same major—should also have higher first-year grades than other students. As students' reports always preserve their indirect utilities' relative order, we expect a higher preference intensity on students' assigned to top reported preferences compared to students assigned to lower reported preferences. On the other hand, to separate the

[49] For each student, we compute the fraction of preferences belonging to each major and university type, and then we compute the average across students for each major and college type. For example, if a given student applies to Medicine in PUC, Medicine in UCH, and Engineering in UCH, then the share of Health in the student's ROL is 2/3, the share of Technology is 1/3, and the share is 0 for all the other majors.

impact of students' learning about their unknown abilities from the grade noise, we compare the law of motion between students' first-year grades and second-year grades for switchers and non-switchers (Arcidiacono et al. (2016)), and the correlation between students' first-year grades and the change in students' ROL composition for majors and college-types. To get the intuition of our identification argument, consider the following equation that defines student i's posterior unknown ability for program j:

$$\mu_{ij't+1} = \frac{\left(\omega_{s_{j'}}\omega_{s_j} + (1-\omega_{s_{j'}})(1-\omega_{s_j})\right)\sigma_s^2 a_{ijt} + 1_{\{m_{j'}=m_j\}}\sigma_m^2 a_{ijt}}{\sigma_g^2 + \sigma_m^2 + \left(\omega_{s_j}^2 + (1-\omega_{s_j})^2\right)\sigma_s^2} \tag{6.1}$$

where ω_{s_j} and $\omega_{s_{j'}}$ are the weights that programs j and j' use for math, a_{ijt} is the signal that student i receives from her grades in program j at time t, and σ_m^2, σ_s^2, and σ_g^2 are the variances of the major-unknown ability, subject unknown ability, and the grade noise, which are the parameters of interest that we want to identify. At the left hand side of the equation, $\mu_{ij't+1}$ is the unknown ability of student i in program j' at time $t+1$. The posterior unknown ability affects students' switching and dropout decisions and their re-applications. Intuitively, if students' grades have a very low correlation with their switching, dropout, or reapplication choices, the signal is not very informative about the students' unknown abilities, and most of the signal is noise (high σ_g^2). On the other hand, if there is a high (negative) correlation between students' first-year grades (signals) and their switching and reapplication choices, particularly changing majors or math-types, that tells us that the signal is highly informative about the unknown abilities for major (high σ_m^2) and subjects (high σ_s^2) respectively[50] Using these insights, we include as moments the correlation between students' first-year grades and their switching choices, and the change in the composition of their re-applications with respect to their initial applications, in terms of majors and math-types[51].

Application. Two key components affect students' application behavior: students' indirect utilities over the expected assignment, and students' beliefs on admission probabilities. Using a large-market assumption and assuming rational expectations, beliefs on admission probabilities can be estimated from the data. Given beliefs, indirect utilities can be non-parametrically identified by using the variation on admission probabilities over time or by incorporating a special regressor in the flow utility function (Agarwal and Somaini, 2018). Since the cutoff distributions do not vary much for the years considered in our sample, we assume that distance from students' home municipality to programs' municipalities is exogenous, giving us exogenous variation that shifts the distribution of indirect utilities. To estimate the probability that students are either truth-tellers or strategic (see Section 5.4.2), we use the results of the survey on students' true preferences and the ROLs submitted to construct moments that allow us to identify this parameter. In particular, we use the share of students' applications for which their top-reported choice

[50]The value of the signal is also affected by the effect of grades on wages (λ_3) and by the effect of the unknown ability on the non-pecuniary work utility (α_4^w). These parameters directly affect switching and dropout probabilities but do not affect the signal's scale in the grade equation of the first period.

[51]The underlying identification assumption is that students' past signals (which are a function of their grades) are a sufficient statistic about how their unknown abilities affect their choices.

has zero admission probability. Finally, we add additional identifying information from students that re-apply to college.[52] We use the panel of repeated respondents in the 2019 and 2020 surveys and compute the share of re-applicants that report a different top-true preference for programs, majors, and college-types. As we have direct information on top-true preferences, the variation in students' responses gives us an additional information source that helps us identify students' learning.[53]

7 Estimation

In this section we describe our estimation strategy. We start in Section 7.1 describing the sample considered and how we aggregate the data to reduce the computational complexity of performing the estimation. Then, we describe in Section7.2 our main estimation procedure.

7.1 Sample Selection and Aggregation

To perform the estimation, we focus on students living in the Metropolitan region and restrict attention to programs located in the same region. This sample restriction reduces the number of programs to less than half (471) while keeping the richness of students' choice-sets[54]. We draw a random sample of 5,000 students who graduated in the year 2013 and have valid scores in the admission process of 2014.

Broad Majors and College Types. As discussed in Section 5.2, we group majors in four *broad majors*—Science (Science, Farming, and Technology), Social Sciences (Social Sciences, Art and Architecture, and Law), Education and Humanities (Education and Humanities), and Health (Health)—to reduce the number of parameters to be estimated. In addition, we consider three types of college: CRUCH-Public, CRUCH-Private and Non-CRUCH.

Subjects. To further facilitate the estimation, we classify programs in two types depending on their admission weights $\{\omega_{js_m}, \omega_{js_v}\}$: (i) math intensive, which includes programs for which the weight on math is higher than that on verbal, $\omega_{js_m} > \omega_{js_v}$; and (ii) verbal intensive in the converse case. In a slight abuse of notation, we denote by s_j the type of program j, and we say that $s_j = s_m(s_v)$ if program j is math (verbal) intensive. Then, instead of considering the weights of each program, we use the average

[52]In Appendix D.2 we formalize these ideas and present a proposition that provides testable implications to measure the extent of preference variation that can be inferred from re-applications.

[53]We do not have grade information for these cohorts. Thus we can not construct correlations between students' true preferences for programs and their college grades.

[54]Close to 80% of applications from students living in the Metropolitan region, include only programs located in the Metropolitan region.

math weight among all programs that belong to the same type. As a result, the unknown ability of student i in program j becomes:

$$A^u_{ij} = A^u_{im_j} + \bar{\omega}_{s_j} A^u_{is_m} + (1 - \bar{\omega}_{s_j}) A^u_{is_v} \tag{7.1}$$

where $\bar{\omega}_{s_j}$ is the average weight on math for programs of type $s_j \in \{s_m, s_v\}$.

7.2 Estimation Procedure

We estimate the model parameters, θ, via Indirect Inference (II). The idea behind II is to choose a statistical model that gives a rich description of the data patterns (Bruins et al. (2018)), allowing us to identify the model parameters. This statistical model—also known as *auxiliary* model—is estimated both on the data and on simulated data from the structural model. The II estimator minimizes an objective function that compares the distance between the estimated data parameters and the parameters estimated from the simulated data.[55] In this sense, the Simulated Method of Moments is a particular case of II, where the *auxiliary* model is just a vector of moments.

We choose to follow this estimation strategy for the following reasons:

(i) We only have remote access to the data on students' grades, and CRUCH only allowed us to obtain regression results and summary statistics at the aggregate level, making it difficult to estimate a likelihood-based estimator. However, II allows us to estimate a rich statistical representation of the data on students' grades and use the estimated parameters to construct moment conditions to estimate the model's structural parameters.

(ii) The parameters involving the grade equation and wage equation have a clear reduced-form representation in the data.

(iii) Estimating students' preferences in a portfolio setting is computationally challenging for likelihood-based estimation methods (see Larroucau and Ríos (2018)). However, given a model parameters' guess, simulating data from our structural model is relatively fast because under Assumption 6, we can simulate strategic ROLs efficiently using the Marginal Improvement Algorithm (Chade and Smith (2006)).

We now introduce the estimator, following closely Bruins et al. (2018). Let $y_i := (y_{it}, ..., y_{iQ})$, be a collection of Q outcomes for student i', and let $\mathbf{y} := \{y_i\}_{i=1}^N$ to denote the aggregate outcomes of all students $i \in \{1, ..., N\}$. Similarly, let x_i and $\mathbf{x}$ be the individual and aggregate students' and programs' characteristics, and η_i and η be the individual and aggregate random shocks. Let $\hat{\beta}_n$ be the vector of parameter estimates of the auxiliary model, that is,

$$\hat{\beta} := \underset{\beta}{\operatorname{argmax}}\, \mathcal{L}(\mathbf{y}, \mathbf{x}; \beta) = \underset{\beta}{\operatorname{argmax}} \frac{1}{N} \sum_{i=1}^{N} l(y_i, x_i; \beta), \tag{7.2}$$

[55] This is known as the Wald approach to indirect inference. Other criterion functions can also be used for estimation.

where $l(\cdot;\beta)$ is the log-likelihood function given the vector β. Let $\eta^b := \{\eta_i^b\}_{i=1}^N$ denote a set of simulated draws for the random shocks of the structural model for simulations $s = 1, ..., S$, where each set of draws is simulated independently from each other. Let $\theta \in \Theta$ be a vector of parameters from the structural model, with $d_\theta \leq d_\beta$. Given the observable characteristics $\mathbf{x}$ and a parameter vector $\theta \in \Theta$, we can use the structural model to simulate data $\mathbf{y}^b(\theta) := \{y_i^b(\theta)\}_{i=1}^N$, and estimate the auxiliary model on each simulated sample:

$$\hat{\beta}^b := \underset{\beta}{\operatorname{argmax}}\, \mathcal{L}(\mathbf{y}^b(\theta), \mathbf{x}; \beta). \tag{7.3}$$

Let $\bar{\beta}(\theta)$ be the average of these estimates, i.e., $\bar{\beta}(\theta) := \frac{1}{B}\sum_{s=1}^S \hat{\beta}^b(\theta)$. Then, the II estimator minimizes the following function:

$$Q(\theta) := \left(\bar{\beta}(\theta) - \hat{\beta}\right)^T W \left(\bar{\beta}(\theta) - \hat{\beta}\right) \tag{7.4}$$

where W is a positive-definite weighting matrix.

For a given value of the parameters θ, and given the first stage estimates—i.e., students' beliefs and enrollment, dropout, graduation, and employment probabilities—, computing the objective function $Q(\theta)$ involves solving the model via backward-induction and then forward-simulating outcomes.[56] Solving the model is computationally expensive, especially computing the continuation value terms, as they depend on the realization of the random coefficients $\{\alpha_i\}_{i=1}^N$ (which are known to the students), which restricts the number of draws of the random coefficients we can use to evaluate the objective function. To reduce the noise due to a small number of draws, we consider a larger number of draws for those shocks that do not affect the backward-induction. In Algorithm 1, Appendix E, we describe in detail how we perform the estimation, and we discuss some related technical considerations.

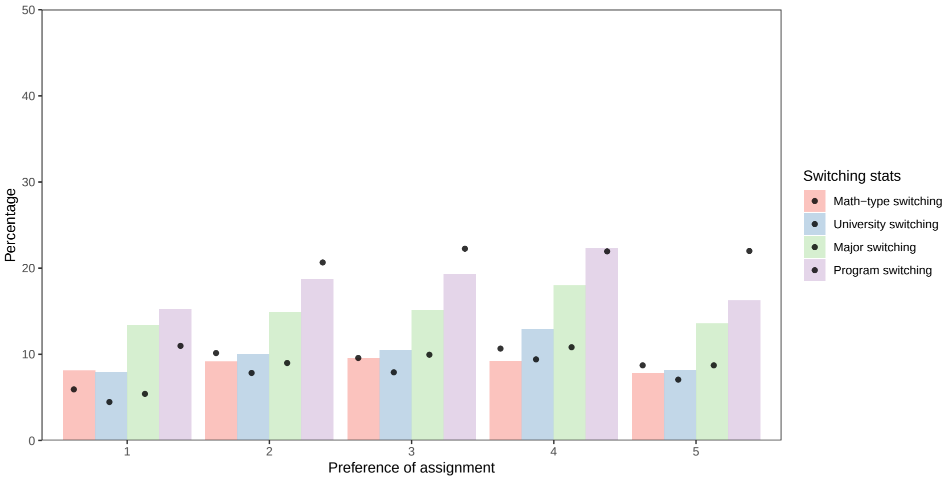
AUXILIARY MODEL. We use as an *auxiliary* model a combination of regression models—including data analogs of the grade equations, wage equations, linear probabilities models of graduation, and linear probability models of switching and dropout—and a vector of moment conditions. The parameters of the model are identified jointly by the moment conditions generated with the auxiliary model. However, some sets of parameters are directly linked to particular moment conditions. We describe this auxiliary model and the matrix of weights in detail in Appendix E.1.

7.3 RESULTS

Table E.2 shows the estimated parameters. We observe that the estimated share of students who apply strategically is 0.89. Thus, a significant fraction of students behaves as weak-truth-tellers. We also observe that the correlation and persistence of students' preferences by major are relatively high ($\sigma_\alpha^{2m} = 14.70$), considering that the variance of

[56] Where we have suppressed the dependency on the first-stage estimators for readability.

Figure 7.1: Summary Statistics by Preference of Assignment



Notes: All switching statistics are regarding first year students, and are computed conditional on enrolling in the centralized system in 2014.

students' idiosyncratic preference shocks is normalized to $\pi^2/6$. Additionally, we observe that the prior variance for the subject-specific abilities (σ_s^2) is bigger than the prior variance for the major-specific ability (σ_m^2), which implies that students' signals are more informative about their subject specific-abilities than their major-specific ability. However, the magnitudes of the prior variances should not be interpreted in isolation because the signal's value is affected by the importance of the unknown ability in the non-pecuniary work utility plus the effect of students' grades on their future wages. Thus, we analyze the importance of students' learning regarding their effects on outcomes in the counterfactual experiments.

In Tables E.4 and E.5 (see Appendix E) we compute the moments and coefficients predicted from our model with their data counterparts. We observe that, in most of the cases, the fit found in the model matches closely the values observed in the data, suggesting that our model captures the richness of the data relatively well. The main discrepancies between the values predicted by the model and the data are related to re-applications—our model over-estimates their incidence—and to reported true preferences—our model underestimates the fraction of students that include their top-true preference in their ROLs.

To highlight some of the most relevant outcomes, in Figure 7.1, we plot some relevant statistics by the preference of assignment. The bars represent values predicted by the model, while the dots represent the data's corresponding values. We observe that the model predictions regarding university switchings and math type switchings are very close to the data's values. Moreover, we capture the increasing pattern of switching probabilities by the preference of assignment (for the top four preferences). However, we also observe that major switching rates are over-estimated.

8 COUNTERFACTUALS

We now present our counterfactual analysis. As discussed in Section 1, our counterfactuals aim to serve two purposes:

1. Assess to which extent students' switching and dropout decisions are explained by the behavioral channels previously described, namely, initial mismatches and learning.
2. To evaluate if different policies oriented to elicit cardinal preferences may help to improve students' outcomes and the system's efficiency.

As discussed in Section 5.4.2, we first evaluate each counterfactual assuming that students behave strategically. This assumption is reasonable if precise information about admission probabilities is provided to students, highlighting the trade-offs involved in choosing a ROL. The main challenge to perform this analysis is that these interventions can modify the aggregate distributions of cutoffs, affecting students' beliefs over their admission probabilities. To take this into account, we need to solve for the equilibrium distribution of cutoffs, but this is not straightforward because there may be multiple equilibria due to students' strategic behavior. To address this, we select an equilibrium following a similar approach than Kapor et al. (2020b). However, our case differs from theirs in that (i) we have to solve for a stationary distribution in the dynamic application problem, creating a mixture of applicants and re-applicants that participate in the same admission process (whereas their setting does not consider re-applications), and that (ii) students need to form believes over a large set of cutoff distributions. To reduce the numerical complexity of solving for the equilibrium in this large-scale problem, we include the following assumption.

Assumption 8. For each counterfactual, and for each program $j \in M$, let π_j be the distribution of cutoffs, and let π_j^B be the distribution of cutoffs for the baseline model. Then,

$$\pi_j = \xi \cdot \pi_j^B \tag{8.1}$$

pointwise, where ξ is a constant to be determine that is counterfactual specific.

Under Assumption 8, students believe that all expected cutoffs change in the same direction and–proportional–magnitude relative to the baseline. Algorithms 2 and 3 in Appendix F describe the algorithms to estimate students' equilibrium beliefs over the cutoff distributions, with and without imposing Assumption 8, respectively.

8.1 UNDERSTANDING BEHAVIORAL CHANNELS

To accomplish the first goal, we consider three different counterfactuals:

1. No Systematic Learning: sets the value of the standard deviation of each unknown ability to zero. Hence, there are no unknown abilities.

2. No Mismatch: assigns each student to their top preference, independent of programs' capacities. As a result, programs' capacities may be exceeded. This counterfactual eliminates completely initial mismatches, allowing us to isolate the learning channel.

3. No Mismatch nor Systematic Learning: combines the two previous counterfactuals, allowing us to isolate the learning channel from the effects of the idiosyncratic shocks (random learning).

The first column in Table 8.1 reports the results of the baseline model, which includes the two main behavioral channels. The next three columns match the three counterfactuals mentioned above. We group the first two columns as *With Mismatches* and the last two columns as *No Mismatches* to highlight that in the latter, the mismatch channel is not present. Notice that in the case with no mismatches the number of seats offered by each program may be exceeded. Finally, each row represents an outcome of interest, including statistics regarding re-applications, switchings, dropout, enrollment, on-time graduation, among others.

Table 8.1: Results Counterfactuals - Behavioral Channels

	With Mismatches		No Mismatches	
	Baseline	No Systematic Learning	Baseline	No Systematic Learning
Re-applicants [%]	20.78	16.19	17.19	12.80
Program switchings [%]	4.13	1.21	2.70	0.01
Major switchings [%]	3.68	0.74	2.65	0.00
Math-type switchings [%]	2.20	0.40	1.69	0.00
Dropouts [%]	9.09	3.31	11.98	3.41
Dropouts - first year [%]	5.77	0.01	8.48	0.04
First enrollment in second period [%]	9.35	11.24	8.40	8.55
First year in second period [%]	13.53	12.48	11.12	8.59
Second year in second period [%]	25.12	26.39	37.20	41.20
Applicants in first period [%]	52.88	41.90	69.35	58.49
Applicants in second period [%]	22.12	21.01	17.19	13.36
Enroll in same program [%]	25.16	26.42	37.22	41.23
Graduate - first enrollment [%]	19.89	21.18	28.91	31.99
Program switchings or Dropouts [%]	13.23	4.52	14.67	3.42
Unassigned in first period [%]	51.44	62.34	30.65	41.51
Graduate on-time [%]	4.87	5.67	6.45	6.82

Notes: switching and dropout rates are computed with respect to the total sample of participants.

WITH MISMATCHES. We start focusing on the first two columns. First, we observe that having no learning decreases the number of re-applications, program switches, and dropout rates but increases the number of unassigned students in the first period. By shutting down the learning process, we increase the persistence of students' preferences over time, which translates into higher college retention rates. Additionally, without the gains from learning, the value from enrolling in the centralized system drops. Therefore, a higher fraction of students choose the outside option. Finally, we observe that the systematic learning channel explains close to two-thirds of students' switching and dropout behavior.

NO MISMATCHES. We now focus on the case with no mismatches. Recall that, in this counterfactual, all students are assigned to their most desired preference, possibly exceeding the vacancies of programs. For this reason, the fraction of students that are unassigned decreases considerably, and thus these results are not directly comparable to those previously described. However, comparing the two columns labeled as "Baseline" provides an idea of the benefits of not having initial mismatches. In particular, we observe that the fraction of students that re-applies is considerably smaller (78% of the baseline), and so are the switching rates (close to 65% of the baseline). The reason is that this counterfactual assigns students to their most desired program, eliminating initial mismatches and thus reducing the incentives for students to re-apply or switch. On the other hand, we observe an increase in the dropout rates at the end of the first year and within the first two years. However, notice that this rate is computed relative to the entire population, so, naturally, this increases as more students are assigned under this counterfactual.[57] We also observe that eliminating mismatches improves on-time graduation rates, which is mainly affected by the reduction in switching rates. These results suggest that eliminating initial mismatches is a sensitive approach to reduce switchings and increase on-time graduation rates, improving the system's yield.

Finally, comparing the third and fourth columns—i.e., within the group of *No Mismatches*—we find that eliminating the systematic learning decreases dropouts to close to 28% of the baseline without mismatches, and virtually eliminates switchings. This result suggests that preference shocks explain a significant fraction of dropout decisions but do not explain switchings without the presence of systematic learning.

8.2 ASSIGNMENT MECHANISMS AND RE-APPLICATIONS RULES

To assess if changes in the mechanism and re-application rules can affect students' outcomes and their welfare, we implement two families of counterfactuals: (i) modifying the assignment mechanism, and (ii) modifying re-application rules. As before, we evaluate these policies considering different measures of switchings, dropout rates, re-applications, on-time graduation, among others. Moreover, for these counterfactuals, we add a measure of students' welfare, given by the relative change in future wages that would be necessary to achieve the same level of average ex-post utility in the baseline scenario.[58]

8.2.1 ASSIGNMENT MECHANISMS.

We evaluate the effects of eliciting intensity on students' preferences by changing the assignment mechanism. In particular, we evaluate two mechanisms:

[57] Indeed, if we compute the dropout rate among those students who are assigned, we observe that the rates are relatively similar for both columns labeled as Baseline ($11.98/(100-30.65) = 17\%$ vs. $9.09/(100-51.44) = 18.7\%$, respectively).

[58] Ex-post utilities are computed at the end of period two, adding the discounted value function of period three, i.e., after students have made all their choices in the model.

1. Constrained Deferred Acceptance (CDA): change the constraint in the length of the ROLs, K. We evaluate $K \in \{1, 2, 3\}$, since most students submit a ROL with length less than or equal to 3.
2. Choice-Augmented Deferred Acceptance with score bonus (CADA): students can signal one program in their submitted ROLs, receiving a bonus φ in their application score for that specific program. We implement this mechanism only for first period applicants, and therefore students who apply in the second period do not receive the bonus.[59]

Both mechanisms elicit the intensity of students' preferences as they introduce opportunity costs that students must take into account when submitting their applications. In the case of CDA, constraining the length of applicants list limits students from including other programs in their ROLs, and thus they must account for the opportunity cost of including each program. In the case of CADA, students can signal only one program, and thus they must carefully decide which program to target to get the bonus. However, notice that eliciting the intensity of students' preferences may not necessarily lead to an overall reduction of switchings and dropouts. On the one hand, if eliciting this information decreases initial mismatches, we would expect to reduce inefficient switchings. On the other hand, if the assignment mechanism also elicits the intensity of preferences among students who re-apply to the system and these preferences change considerably due to learning, we would see an increase in efficient switchings due to an increase in the value of re-applications. In this sense, we expect that in the case of CADA— which provides a score bonus only in the first period—switchings would decrease more than in the case where the score bonus is applied in both periods.

In Table 8.2 we report the results of these counterfactuals. As in the previous section, the first column includes the results of the baseline model. The next three columns report the results of constrained DA considering values $K \in \{1, 2, 3\}$ in decreasing order, while the last three columns report the results of CADA with score bonus $\varphi \in \{10\%, 20\%, 30\%\}$.

First, we observe that CDA increases the fraction of re-applicants if K is sufficiently low. This result is relatively intuitive, as reducing the maximum size of the ROLs increases the risk of being unassigned, increasing the incentives to re-apply in the next year. On the other hand, we observe that limiting the size of the ROLs is not very effective at decreasing the overall number of switches and dropouts. Finally, we observe that the effect on welfare is non-monotonic, as it is higher with $K = 3$ while it is smaller with $K = 1$ relative to the baseline.

On the other hand, we observe that CADA effectively decreases students' switchings and dropouts. Also, we observe that CADA increases the fraction of students who apply in the first period and can increase the fraction of students remaining in their programs. As a result, this mechanism leads to higher persistence in programs and slightly

[59] See Abdulkadiroğlu et al. (2015) for details. We choose to implement CADA only in the first period to avoid solving for the continuation values under this mechanism, which would add a high computational burden to the model. To implement this mechanism, we need to specify how to find the optimal ROL for each student, given their preferences and beliefs. Algorithm 4 in Appendix F describes a procedure to accomplish this.

Table 8.2: Results Counterfactuals -Mechanisms

	Baseline	Constrained DA $K=3$	Constrained DA $K=2$	Constrained DA $K=1$	CADA $\varphi=10\%$	CADA $\varphi=20\%$	CADA $\varphi=30\%$
Re-applicants [%]	19.65	19.56	19.94	21.71	19.59	19.99	20.66
Program switchings [%]	4.87	4.88	4.92	4.53	4.45	4.12	3.92
Major switchings [%]	4.13	4.15	4.16	3.80	3.84	3.66	3.50
Math-type switchings [%]	2.47	2.49	2.51	2.27	2.32	2.17	2.03
Dropouts [%]	9.53	9.44	9.57	9.14	9.71	9.56	9.44
Dropouts - first year [%]	6.11	6.04	6.05	5.59	6.28	6.13	5.92
First enrollment in second period [%]	9.55	9.55	9.62	10.87	9.69	10.17	10.93
First year in second period [%]	14.50	14.52	14.64	15.47	14.23	14.38	14.93
Second year in second period [%]	24.13	24.15	23.97	22.34	24.67	24.56	24.05
Applicants in first period [%]	49.90	49.87	49.88	49.94	51.32	52.46	53.31
Applicants in second period [%]	21.44	21.35	21.72	23.43	20.84	20.97	21.54
Enroll in same program [%]	24.21	24.24	24.08	22.41	24.76	24.65	24.14
Graduate - first enrollment [%]	19.20	19.25	19.15	17.94	19.50	19.45	18.99
Program switchings or Dropouts [%]	14.39	14.32	14.49	13.67	14.17	13.68	13.36
Unassigned in first period [%]	51.39	51.39	51.64	55.40	50.72	51.45	52.63
Graduate on-time [%]	4.50	4.60	4.58	4.38	4.97	4.70	4.96
	Difference in Ex-Post Welfare Relative to Baseline						
Overall	-	2.70	-1.40	-33.50	7.70	1.90	-8.60
Women	-	1.00	-3.80	-35.00	2.60	-1.90	-10.80
Men	-	4.90	2.40	-31.40	16.10	7.80	-2.50
Low-income	-	6.20	2.60	-29.90	14.90	16.90	10.80
High-income	-	0.00	-5.10	-35.50	2.00	-11.40	-21.70

Notes: switching and dropout rates are computed with respect to the total sample of participants.

lower shares of delayed graduation. Furthermore, we observe that CADA considerably increases students' welfare compared to both the baseline and constrained DA. In particular, we observe that the group of students that benefit the most from CADA are male students and students from low-income families. One possible reason for the former is that the bonus is applied to the scores related to the high-school grades, which are on average lower for men.[60] On the other hand, one possible reason for the higher effect on low-income students is that high-income students are almost twice as likely to switch programs than low-income students. Thus, eliciting intensity increases the probability of admission among low-income students relatively more than among high-income students. Finally, we observe that the overall impact of the bonus is non-monotonic, increasing welfare compared to the baseline in the case where $\varphi = 10\%$ and $\varphi = 20\%$, but decreasing welfare compared to the baseline when $\varphi = 30\%$. These results suggest that CADA with an appropriate score bonus could be a sensible policy to reduce switches and increase students' welfare.

8.2.2 Reapplication Rules.

Another policy to reduce the incentives to switch is to provide bonuses to students applying for the first time to the system, or to penalize students who re-apply and try to switch programs. These policies have been implemented in Finland and Turkey, respectively. To our knowledge, none of these policies has been analyzed in terms of their impact on students' outcomes. To analyze this, we consider the following two families of policies:

[60] We conjecture that, if instead of applying the bonus to the scores related to high-school grades we apply it to PSU scores, the benefit for women would be higher than for men. We are planning to add these results to future versions of the paper.

(i) Turkish re-application rule: applicants receive a penalty ψ in the scores related to their high-school GPA (thus, affecting the application score of their re-applications) if they are currently enrolled in the centralized system when they submit their ROLs.

(ii) Finnish re-application rule: students receive a bonus φ in the scores related to their high-school GPA (thus, affecting the application score of their re-applications) the first time they submit a ROL to the centralized system.

Even though both policies aim to reduce the incentives for switching, they affect students' applications and re-applications in different ways. On the one hand, the Finnish policy directly reduces the incentives to re-apply to the system, regardless of the programs that students include in their ROLs. As a result, the Finnish policy increases the continuation value of choosing the outside option, and thus increases the fraction of students that wait an extra year to submit their first application. On the other hand, the Turkish policy reduces the incentives to re-apply if students previously enrolled in a program in the system, i.e., it reduces the incentives to apply to programs if they are very likely to switch from them in the future (e.g., programs for which students have low preference intensity). Hence, the Turkish policy may decrease the fraction of students enrolling in the first period in less preferred programs. Despite these differences, we expect that both policies would decrease the frequency of re-applications and switches. In contrast, the welfare effects of these policies is unclear. Students may benefit from these policies as both the penalty and the bonus help to address the negative externality that switchers generate in the system. However, since under these policies students face more barriers for switching, the benefits of learning become lower, and thus, students' welfare may decrease.

In Table 8.3 we report the results of these counterfactuals. As expected, we observe that the Turkish policy reduces the re-application and switching rates, and the magnitude of the effect is increasing in the magnitude of the penalty. Moreover, we observe that dropout rates for the first year slightly increase as we increase the penalty. A potential explanation for this is that the Turkish policy increases switching costs. Thus, students who receive low signals about their match-qualities with their enrolled programs face lower probabilities for switching than the baseline, increasing their incentives to drop out instead. Finally, we observe that the ex-post welfare decreases compared to the baseline as we increase the penalty value. These results suggest that the gains from learning and having the option of switching can exceed the negative externality imposed by students switching and displacing other students who may have stronger preferences for those programs. Moreover, these findings confirm that largely reducing switchings can also be inefficient for the system.

On the other hand, we observe that the Finnish policy has a similar effect on students' outcomes, but the magnitude of the effect varies relative to the Turkish policy. For instance, we observe that the latter is better at reducing re-applications and students' switches but increases dropouts slightly, while the former is better at improving on-time graduation and increasing students' welfare. Additionally, the Finnish policy increases the fraction of students unassigned in the first period and the fraction of students who decide to delay their tertiary education entry. An explanation for this is that students

Table 8.3: Results Counterfactuals - Re-Application Rules

		Turkish Rules			Finnish Rules		
	Baseline	$\psi = 10\%$	$\psi = 20\%$	$\psi = 30\%$	$\varphi = 10\%$	$\varphi = 20\%$	$\varphi = 30\%$
Re-applicants [%]	19.65	18.14	17.10	16.55	18.71	19.22	19.50
Program switchings [%]	4.87	4.04	3.46	3.10	4.32	3.99	3.70
Major switchings [%]	4.13	3.62	3.25	2.93	3.78	3.53	3.32
Math-type switchings [%]	2.47	2.09	1.87	1.67	2.25	2.14	2.01
Dropouts [%]	9.53	9.58	9.63	9.59	9.56	9.42	9.23
Dropouts - first year [%]	6.11	6.23	6.31	6.38	5.98	5.79	5.38
First enrollment in second period [%]	9.55	9.62	9.65	9.73	10.42	11.57	12.70
First year in second period [%]	14.50	13.74	13.17	12.88	14.84	15.65	16.47
Second year in second period [%]	24.13	24.82	25.43	25.65	24.29	23.50	22.98
Applicants in first period [%]	49.90	49.59	49.56	49.47	49.61	49.94	49.98
Applicants in second period [%]	21.44	20.21	19.22	18.76	21.67	22.68	23.71
Enroll in same program [%]	24.21	24.90	25.49	25.70	24.40	23.59	23.05
Graduate - first enrollment [%]	19.20	19.73	20.18	20.42	19.19	18.67	18.12
Program switchings or Dropouts [%]	14.39	13.62	13.09	12.69	13.88	13.41	12.93
Unassigned in first period [%]	51.39	51.35	51.25	51.39	51.99	53.77	55.83
Graduate on-time [%]	4.50	4.71	4.74	4.80	4.67	4.94	4.84
		Difference in Ex-Post Welfare Relative to Baseline					
Overall	-	-1.40	-1.40	-4.10	9.20	-6.80	-13.20
Women	-	-6.30	-3.80	-6.70	5.50	-13.00	-21.30
Men	-	7.30	2.40	1.50	15.10	5.40	7.30
Low-income	-	2.10	1.50	2.10	9.20	-1.30	4.60
High-income	-	-4.10	-4.10	-8.90	9.60	-10.80	-22.60

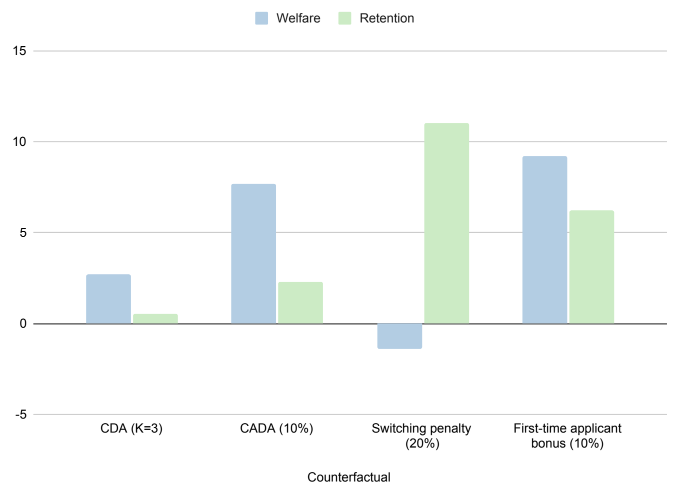
Notes: switching and dropout rates are computed with respect to the total sample of participants.

that know their preferences but do not have good enough scores in the first period are better off waiting a year to retake the exams and improve their application score instead of enrolling in the first year and try to switch later. These results suggest that the most desired policy depends on the objective to be addressed. If the goal is to decrease switchings and improve the systems' yield, then the Turkish policy seems to be the best option. In contrast, if the goal is to increase students' welfare, then the Finnish policy leads to better outcomes.

Figure 8.1 shows a summary of the counterfactual results.[61] Overall, the results of our counterfactual analyses show that some changes in the mechanism—e.g., Choice-Augmented DA or the Finnish re-application rule—can be effective at improving college retention rates while at the same time increasing the welfare of students. However, these policies must be carefully designed, as these can also negatively affect students' welfare if their underlying parameters are not correctly set.

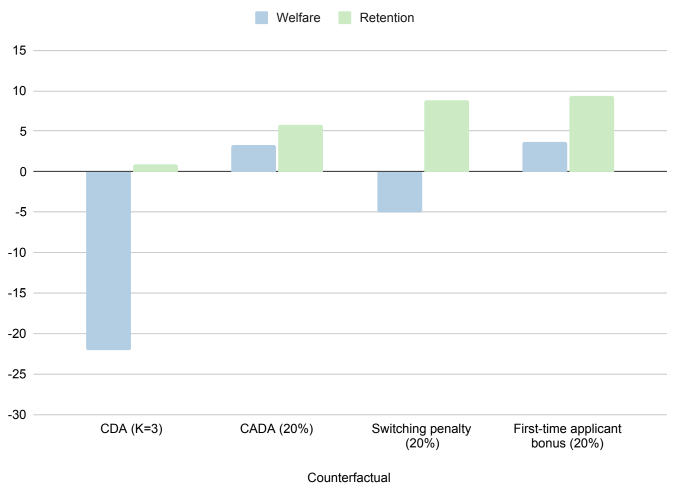
[61] For Figures 8.1 and 8.2 we compute retention considering switchings and first-year dropouts.

Figure 8.1: Summary of Counterfactuals: Strategic Behavior



SENSITIVITY TO NON-STRATEGIC STUDENTS. It is important to highlight that all the aforementioned counterfactual analyses were conducted assuming that all students are strategic. However, many students in practice are not strategic, i.e., they report their true preferences. As a robustness check, we conducted the same analysis assuming that 10% of students are non-strategic—similar to the estimation results in the baseline model—and we find that the results are directionally the same. However, the magnitude of the effects changes significantly. Figure 8.2 shows a summary of the results. In particular, we observe that the overall welfare under CDA decreases as we make the constraint on the length of ROLs more binding. On the other hand, we observe that ex-post welfare also increases for the Finnish re-application rule and CADA relative to the baseline, although the magnitude of the improvement is smaller than when all students are strategic. These results suggest that the latter two policies—Finnish re-application rule and CADA—are more robust to deal with students that may not report their preferences strategically.

Figure 8.2: Summary of Counterfactuals: Mixture of Strategic and Truth-Telling Behavior



9 CONCLUSIONS

In this paper, we analyze the effects of centralized assignment mechanisms on downstream outcomes such as students' decisions to switch or dropout from college. To accomplish this, we study the relevance of eliciting information on participants' cardinal preferences and incorporating their dynamic incentives in the design of the assignment process, features that have been mostly overlooked by the literature.

Using data from the Chilean college admissions system and two nationwide surveys that we designed and conducted, we provide empirical evidence suggesting that two central behavioral channels explain students' dynamic decisions. The first channel, called the *initial mismatch* channel, predicts that students may have incentives to switch programs if they were initially assigned to less desired preferences. The second channel, called the *learning* channel, suggests that students may learn about their match-qualities during their college progression, and thus may decide to switch to programs where their match-qualities—and their expected outcomes in the labor market—are higher. Moreover, we find significant differences in switch and dropout decisions depending on gender, income level, and preference of assignments.

Considering these findings, we introduce a structural model that captures students' decisions during their academic progression, allowing them to learn about their match-quality from their grades. We use the estimated structural model in two ways. First, we use it to disentangle the extent to which each of the two behavioral channels explains students' switching and dropout decisions. We find that both initial mismatches and learning play a significant role, with the former explaining close to a third of students' switchings. Second, we use the model to analyze the effect of a set of counterfactual policies aiming to elicit the intensity of students' preferences and account for their dynamic incentives. We evaluate changes in the re-applications rules—implementing those used in Turkey and Finland—and the assignment mechanism—adding further constraints on the length of lists and adding the option for students to signal one of the programs in their preferences to obtain a score bonus. Our results show that these policies are effective to increase college retention rates while at the same time increasing students' welfare, particularly for low-income students. Moreover, these effects are robust to changes in the fraction of participants that behave strategically, as opposed to other approaches such as constraining the length of the lists. However, lack of sophistication in students' ranking strategies undermines the effectiveness of these policies, which stress the importance of giving students correct information about their admission probabilities and helping them in choosing optimal application lists.

Overall, our results show that incorporating dynamic incentives and eliciting information on participants' cardinal preferences can significantly increase students' welfare and downstream outcomes. These insights can be informative to improve the design of many matching markets that exhibit similar features. For instance, in entry-level labor markets, agents have private information about their preferences, learn about their match-qualities through experience, and face dynamic considerations, such as deciding when to enter the market (apply), re-enter (re-apply), exit (dropout), and re-match (switch). Our key insight is that market designers should correctly balance the gains from learning through experimentation and the congestion externality produced by initial mismatches to improve the efficiency of these markets.

References

Abdulkadiroğlu, A., Agarwal, N., and Pathak, P. A. (2017). The Welfare Effects of Coordinated Assignment: The New York City High School Match. *American Economic Review*, 95(2):364–367.

Abdulkadiroğlu, A., Che, Y. K., and Yasuda, Y. (2015). Expanding "choice" in school choice. *American Economic Journal: Microeconomics*, 7(1):1–42.

Abdulkadiroğlu, A., Pathak, P. A., and Roth, A. E. (2005a). The New York City high school match. *American Economic Review*, 95:364–367.

Abdulkadiroğlu, A., Pathak, P. A., Roth, A. E., and Sönmez, T. (2005b). The Boston public school match. *American Economic Review*, 95:368–371.

Agarwal, N., Ashlagi, I., Rees, M., Somaini, P., and Waldinger, D. (2019). Equilibrium allocations under alternative waitlist designs: Evidence from deceased donor kidneys. *SSRN Electronic Journal*.

Agarwal, N. and Somaini, P. (2018). Demand analysis using strategic reports: An application to a school choice mechanism. *Econometrica*, 86(2):391–444.

Agarwal, N. and Somaini, P. (2019). Revealed preference analysis of school choice models. *Annual Review of Economics*, 12.

Ajayi, K. and Sidibe, M. (2017). An Empirical Analysis of School Choice under Uncertainty.

Altonji, J. G., Arcidiacono, P., and Maurel, A. (2016). The Analysis of Field Choice in College and Graduate School. Determinants and Wage Effects. In *Handbook of the Economics of Education*, volume 5, pages 305–396. Elsevier.

Altonji, J. G., Blom, E., and Meghir, C. (2012). Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers. *Annu. Rev. Econ*, 4:185–223.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1-2):343–375.

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

Arcidiacono, P., Aucejo, E., Maurel, A., and Ransom, T. (2016). College attrition and the dynamics of information revelation. Technical report, National Bureau of Economic Research.

Artemov, G., He, Y., and Che, Y.-K. (2017). Strategic " Mistakes ": Implications for Market Design.

Bordon, P. and Fu, C. (2015). College-major choice to college-then-major choice. *Review of Economic Studies*, 82(4):1247–1288.

Bruins, M., Duffy, J. A., Keane, M. P., and Smith Jr, A. A. (2018). Generalized indirect inference for discrete choice models. *Journal of econometrics*, 205(1):177–203.

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

Carvalho, J. R., Magnac, T., and Xiong, Q. (2019). College choice, selection, and allocation mechanisms: A structural empirical analysis. *Quantitative Economics*, 10(3):1233–1277.

Chade, H. and Smith, L. (2006). Simultaneous Search. *Econometrica*, 74(5):1293–1307.

Correa, J., Epstein, R., Escobar, J., Rios, I., Bahamondes, B., Bonet, C., Epstein, N., Aramayo, N., Castillo, M., Cristi, A., and et al. (2019). School choice in chile. *Proceedings of the 2019 ACM Conference on Economics and Computation*.

DeGroot, M. H. (2005). *Optimal statistical decisions*, volume 82. John Wiley & Sons.

Dong, Y. (2017). Regression discontinuity designs with sample selection. *Journal of Business & Economic Statistics*, 37(1):171–186.

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

Haeringer, G. and Klijn, F. (2009). Constrained school choice. *Journal of Economic Theory*, 144(144):1921–1947.

He, Y. (2012). Gaming the Boston School Choice Mechanism in Beijing. *Manuscript, Toulouse School of Economics*, 2012(May).

Hernández-Chanto, A. (2017). Centralized assignment of students to majors: Evidence from the university of costa rica.

Hwang, S. (2016). A robust redesign of high school match. Technical report, Working Paper.

Kapor, A., Karnani, M., and Neilson, C. (2020a). Aftermarket frictions and the cost of off-platform options in centralized assignment mechanisms.

Kapor, A. J., Neilson, C. A., and Zimmerman, S. D. (2020b). Heterogeneous beliefs and school choice mechanisms. *American Economic Review*, 110(5):1274–1315.

Krishna, K., Lychagin, S., and Frisancho, V. (2018). Retaking in high stakes exams: Is less more? *International Economic Review*, 59(2):449–477.

Larroucau, T. and Ríos, I. (2018). Do "Short-List" Students Report Truthfully? Strategic Behavior in the Chilean College Admissions Problem.

Luflade, M. (2017). The value of information in centralized school choice systems.

Malamud, O. (2011). Discovering one's talent: Learning from academic specialization. *Industrial and Labor Relations Review*, 64(2):375–405.

Narita, Y. (2018). Match or Mismatch? Learning and Inertia in School Choice. *SSRN Electronic Journal*.

Olszewski, W. and Vohra, R. (2016). Simultaneous selection. *Discrete Applied Mathematics*, 200:161–169.

Ríos, I., Larroucau, T., Parra, G., and Cominetti, R. (2020). Improving the Chilean College Admissions System. *Operations Research (forthcoming)*.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, pages 999–1033.

Schlueter, M., Egea, J., and Banga, J. (2009). Extended ant colony optimization for non-convex mixed integer nonlinear programming. *Computers and Operations Research*, 36(7):2217–2229.

Schlueter, M., Erb, S., Gerdts, M., Kemble, S., and Ruckmann, J. (2013). Midaco on minlp space applications. *Advances in Space Research*, 51(7):1116–1131.

Tanaka, M., Narita, Y., and Moriguchi, C. (2020). Meritocracy and its discontent: Long-run effects of repeated school admission reforms.

Appendix

A Appendix for Section 3

We first analyze whether it is–theoretically possible–to increase aggregate students' welfare and reduce inefficiencies in the system by changing the assignment mechanism. Furthermore, we provide some intuition on how switching behavior can be affected by the assignment mechanism in place in a dynamic setting.

Consider a college admissions problem where students can reapply once a year. If students face uncertainty over their admission chances, either because of uncertainty on cutoffs' realizations or uncertainty over their future application scores, switchings can endogenously occur over time. As students do not know whether their choice sets could change or not in the following years, they could choose to enroll in a program in the first year and switch in the following year to a more preferred program if their choice sets change. Moreover, if students are uncertain about their match-quality with programs, and after enrollment, they learn and update their majors' or universities' preferences, allowing for switchings could reduce mismatch problems. Notice that, regardless of which mechanism dominates, individual switchings and dropouts impose an externality on universities and students. Given the sequential nature of colleges' academic progression, when a student switches at the end of the academic year, the vacancy she was using is lost for the next year, and, in the absence of a proper transfer system that allows students to switch at different stages of their college progression, this vacancy can not be reassigned to other students.

The following example shows how students' switching behavior can emerge in a dynamic matching setting, even in the absence of learning about their match-quality with their enrolled programs. Furthermore, the example shows how matching mechanisms–that elicit intensity on students' preferences–can affect students' applications, their assignments, and their following switching decisions, increasing overall welfare and reducing inefficiencies in the centralized system.

Example 1 (Constrained DA vs Unconstrained DA)**.**

Consider a centralized college admissions problem with re-applications and two periods. Let the set of students $S = \{A, B\}$, and the set of colleges $C = \{I, II\}$. Students are expected utility maximizers. Given their preferences and beliefs about their admission chances, students make their application decisions. Let R_i^t be the ROL submitted by student i at time t. After students submit their applications, colleges post their first-year vacancies. Let q_j^t be the first-year vacancies posted by college j at time t. In order to incorporate uncertainty on students' admission chances, we model colleges' vacancies decisions as following a random process, with

$$\mathbb{P}(q_j^t = 1) = \mathbb{P}(q_j^t = 0) = \frac{1}{2}, \quad \forall t, j \in \{I, II\}.$$

Notice that, in this example, there are four equally-likely states of the world in every period regarding vacacies posted.

Students' preferences are given over their expected assignments. Assume students' preferences are given by

$$A: \quad u_I^A \gg u_{II}^A > 0, \qquad\qquad B: \quad u_{II}^B \gg u_I^B > 0$$

Notice that we have assumed that student A prefers to be assigned to college I considerably more than what she prefers to be assigned to college II, and the opposite is true for student B. Moreover, to simplify computations, we assume that each student's payoff from being assigned to their second preference is strictly positive but close to zero. Each student $i \in S$ has an application score s_i^j for every college $j \in C$. Colleges' preferences are defined over the application scores of students assigned to them. Assume that colleges' preferences are given by

$$I: \quad s_I^B > s_I^A, \qquad\qquad II: \quad s_{II}^A > s_{II}^B$$

We further assume that colleges care about students' persistence and perceive a cost of τ per student who drops out. This cost captures the idea that colleges make investments in their students and that when they switch after the first period, that vacancy is lost due to the sequential nature of college progression.

After students make their application decisions and colleges post their vacancies, a clearinghouse runs a matching algorithm. We analyze students' college outcomes induced by two different mechanisms: (i) unconstrained student-proposing DA, and (ii) constrained student-proposing DA. In both cases, we allow students to apply and switch colleges in every period at no cost.

Unconstrained student-proposing DA

When the assignment mechanism is unconstrained student-proposing DA, it is a weakly-dominant strategy for students to report their true preferences, regardless of the uncertainty they face about their current or future admission chances. In every period t, students can apply and, if they do so, submit

$$\begin{array}{llll} R_A^{t=1}: & I \succ II & R_A^{t=2}: & I \succ II \\ R_B^{t=1}: & II \succ I & R_B^{t=2}: & II \succ I \end{array}$$

Notice, though, that given the payoffs we have assumed for students' preferences, students will only re-apply in the second period if they were unassigned in the first period or assigned to their second preference in the first period. Let $\mu^t(A, B)$ be the expected assignment in period t for students A and B. Given students' applications, college preferences, uncertainty on admission chances, and the assignment mechanism, the expected assignment at period $t = 1$ is given by

$$\mu^{t=1}(A, B) = \frac{1}{4} \circ (\emptyset, \emptyset) + \frac{1}{4} \circ (I, II) + \frac{1}{4} \circ (II, \emptyset) + \frac{1}{4} \circ (\emptyset, I)$$

Given their first period assignment, students receive their payoffs and make re-application decisions. Let $\mu^{t=2}\left(A,B|\mu^{t=1}\left(A,B\right)\right)$ be the second period expected assignment, conditional on the realized first period assignment. The second period conditional expected assignment is then given by

$$\begin{aligned}
\mu^{t=2}\left(A,B|(\emptyset,\emptyset)\right) =& \frac{1}{4}\circ(\emptyset,\emptyset)+\frac{1}{4}\circ(I,II)+\frac{1}{4}\circ(II,\emptyset)+\frac{1}{4}\circ(\emptyset,I)\\
\mu^{t=2}\left(A,B|(I,II)\right) =& 1\circ(I,II)\\
\mu^{t=2}\left(A,B|(II,\emptyset)\right) =& \frac{1}{4}\circ(I,II)+\frac{1}{4}\circ(II,\emptyset)+\frac{1}{4}\circ(II,II)+\frac{1}{4}\circ(II,I)\\
\mu^{t=2}\left(A,B|(\emptyset,I)\right) =& \frac{1}{4}\circ(I,II)+\frac{1}{4}\circ(\emptyset,I)+\frac{1}{4}\circ(I,I)+\frac{1}{4}\circ(II,I)
\end{aligned}$$

Notice that when students were assigned in the first period to their first preference or when both were unassigned, there are no switchings in the second period. However, when there was only one vacancy in one of the colleges, both students compete for that spot in the first period, and the student with the highest application score is assigned. We have chosen scores and preferences such that when these states of the world realize, the assigned student is assigned to her second preference. This creates incentives for that student to re-apply in the second period and try to switch to her first preference. Figure 3.1 depicts this scenario, and its implications to students college persistence and welfare.

Let W^{DA} be the aggregate expected welfare for unconstrained student-proposing DA. We then get that

$$W^{DA}\approx\frac{\left(u_I^A+u_{II}^B\right)(4+9\beta)}{16}-\frac{1}{2}\tau \tag{A.1}$$

Were the term $\frac{1}{2}\tau$ comes from the expected cost that colleges face when students switch.

Constrained student-proposing DA (K=1)

Suppose now that students face a constraint in the length of their reported lists. Under this mechanism, students can include only one program in their ROLs, so they need to take into account their admission probabilities. Their optimal ROL includes the college that gives them the highest expected payoff. Given the payoffs we have specified, the optimal ROLs are given by:

$$\begin{aligned}
R_A^{t=1}:\quad I \qquad R_A^{t=2}:\quad I\\
R_B^{t=1}:\quad II \qquad R_B^{t=2}:\quad II
\end{aligned}$$

As each student values their first preference significantly more than their second preference, they choose to include in their ROLs only their first preference, even though their admission is uncertain. Under this mechanism, the first period expected assignment is given by

$$\mu^{t=1}(A,B)=\frac{1}{4}\circ(\emptyset,\emptyset)+\frac{1}{4}\circ(I,II)+\frac{1}{4}\circ(I,\emptyset)+\frac{1}{4}\circ(\emptyset,II)$$

Notice that in this case, no student is assigned to their second preference, therefore, no student wants to switch in the second period. The conditional second period expected assignment is then given by

$$\begin{aligned}
\mu^{t=2}\left(A, B|(\emptyset,\emptyset)\right) =& \frac{1}{4}\circ(\emptyset,\emptyset) + \frac{1}{4}\circ(I,II) + \frac{1}{4}\circ(I,\emptyset) + \frac{1}{4}\circ(\emptyset,II) \\
\mu^{t=2}\left(A, B|(I,II)\right) =& 1\circ(I,II) \\
\mu^{t=2}\left(A, B|(\emptyset,II)\right) =& \frac{1}{4}\circ(I,II) + \frac{1}{4}\circ(I,II) + \frac{1}{4}\circ(\emptyset,II) + \frac{1}{4}\circ(\emptyset,II)
\end{aligned}$$

Figure 3.2 shows how Constrained DA or restricting switchings can reduce the externality aforementioned, improving college persistence and aggregate welfare.

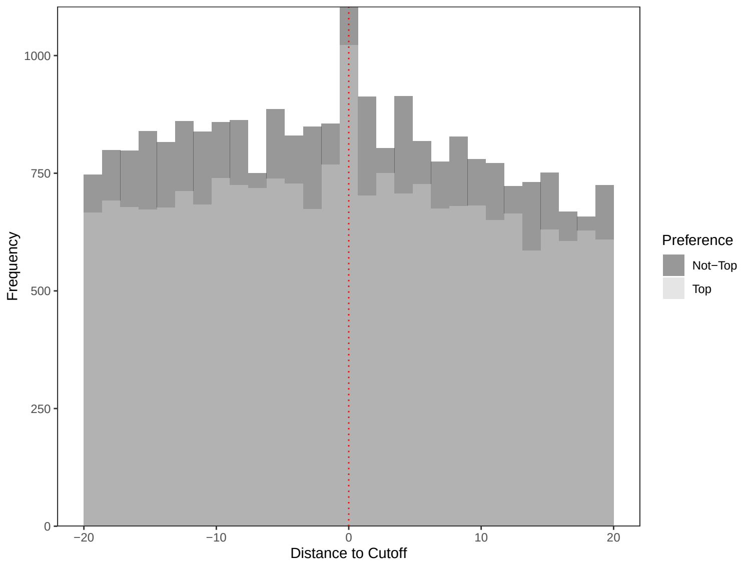
Let W^{CDA} be the aggregate expected welfare for constrained student-proposing DA. We then get that

$$W^{CDA} \approx \frac{\left(u_I^A + u_{II}^B\right)(8+12\beta)}{16} > W^{DA} \tag{A.2}$$

In the previous example, the expected welfare achieved under constrained DA is higher than the expected welfare achieved under unconstrained DA. Moreover, as the game is symmetric, constrained DA gives a *pareto*-improvement in *ex-ante* expected utility. Finally, switching behavior is eliminated, lowering the costs for universities compared to the outcome of student-proposing DA.

This theoretical example assumes that we can find students like A and B in the data, that is, students with similar application scores but different assignment preferences. Figure A.1 shows the distribution of preference of assignment around admission cutoffs. We observe that a significant fraction of students assigned just above admission cutoffs do not rank those programs as their top choices.

Figure A.1: Distribution of preference of assignment around admission cutoffs



B Appendix to Section 4

B.1 The Chilean Mechanism

The Chilean mechanism is a variant of the student-proposing deferred acceptance algorithm[62] in which tied students in the last seat of a program are not rejected if vacancies are exceeded. More formally, the allocation rule can be described as follows:

Step 1. Each student proposes to his first choice according to their submitted ROL. Each program rejects any unacceptable student, and if the number of proposals exceeds its vacancies (q), rejects all students whose scores are strictly less than the q-th most preferred student.

Step $k \geq 2$. Any student rejected in step $k-1$ proposes to the next program in their submitted ROL. Each program rejects any unacceptable student, and if the number of proposals exceeds its vacancies (q), rejects all students whose score is strictly less than the q-th most preferred student.

The algorithm terminates either when there are no new proposals or when all rejected students have exhausted their preference lists. The final allocation is obtained by assigning each student to the most preferred program in his ROL that did not reject him. As a side outcome of this, the algorithm generates a set of cutoffs $\{\bar{s}_j\}_{j\in M}$, where $\bar{s}_j$ is the minimum application score among students matched to program $j \in M$. Hence, for any student i with ROL R_i and set of scores $\{s_{ij}\}_{j\in M}$, the allocation rule can be described as

$$i \text{ is assigned to } j \Leftrightarrow j \in R_i,\ s_{ij} \geq \bar{s}_j \text{ and } s_{ij'} < \bar{s}_{j'}\, \forall j' \in R_i\ st.\ j' \succ_{R_i} j,$$

where $\succ_{R_i}$ is a total order induced by R_i over the set $\{j : j \in R_i\}$, such that $j' \succ_{R_i} j$ if and only if program j' is ranked above program j in R_i.

B.2 Odds ratios

We estimate *logit* models of the probability of switching majors, universities, or dropping out from the university system, on the preference of assignment. We control for a comprehensive set of students' characteristics, including application scores, family background, gender, access to scholarships, relative position in the program, distance from student's municipality to program's municipality, number of different majors, and number of different universities listed in the ROL, among others; and programs' fixed effects.

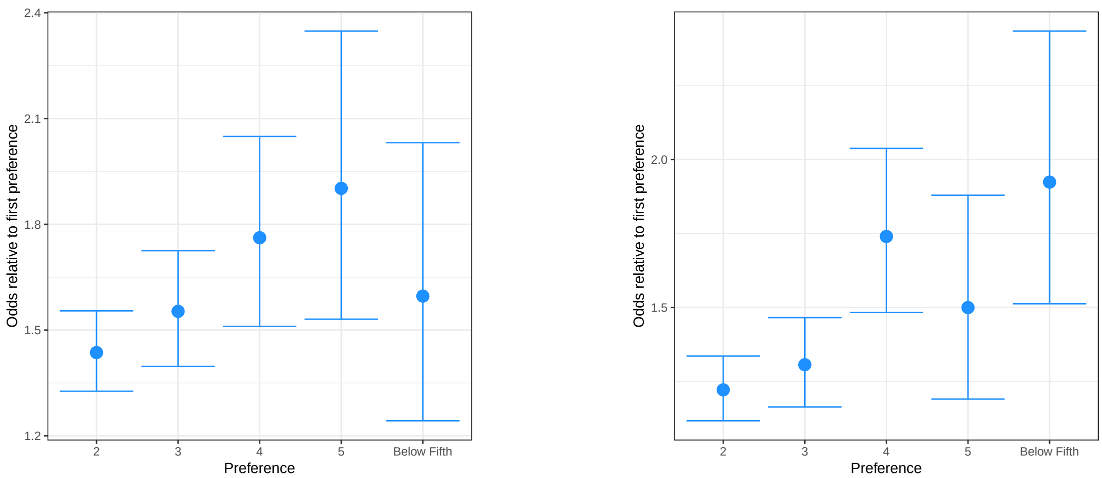
In Figure B.1, we show the odds ratios for the probability of switching majors and universities, conditional on the preference of assignment (relative to the top reported preference). We observe that being assigned to a lower preference significantly increases

[62] Before 2014; the algorithm used was the university-proposing version. The assignment differences between both implementations of the algorithm are negligible Ríos et al. (2020).

students' probability of switching majors and universities. This finding suggests that, conditional on observable characteristics, initial preferences play a significant role in students' switching choices.

Figure B.1: Odds of major and university switching

(a) Majors (b) Universities



Note: Odds ratios of the probability of switching majors (left) and universities (right) by the preference of assignment for the 2012 cohort. The reference category is being enrolled in their first reported preference.

B.3 Regression Discontinuities

This section provides causal evidence that the preference of assignment affects different outcomes of interest. We use a regression discontinuity design that exploits the algorithm's cutoff structure to perform the allocation. As a result of the assignment process, each program gets a cutoff such that all students whose weighted score is above it are granted admission, whereas all students with scores below the cutoff are wait-listed and thus may have to enroll in a lower-ranked preference. Hence, if we assume that students around the cutoff are similar and only differ in their right to enroll in a higher preference, we can estimate the causal effect of interest.

Formally, we estimate the effect of being assigned in a higher preference using the following specification:

$$y_{bp} = f_p(d_{bp}) + \delta_p \cdot Z_{bp} + \epsilon_{bp}, \tag{B.1}$$

where y_{bp} is the average outcome of interest for students in bin of distance b applying to preference p, f_p is a smooth function of the distance d_{bp} between the bin's score and the cutoff of their preference p, Z_{bp} is an indicator function equal to 1 if bin b's score is greater than or equal to the cutoff of their p-th preference, and 0 otherwise; and ϵ_{bp} is an error term.[63]

[63] Similar results are obtained running these models at the student-preference level. We report the results at the bin-preference level to match the plots included.

Notice that many of the outcomes we consider—e.g., switches, dropouts, stopouts, among others—rely on students enrolling in the centralized system. If there are significant differences in the enrollment rates among students right above and below the cutoff, then the two samples would not be directly comparable. In that case, there would be a selection problem, and thus we would not be able to—point—estimate the causal effect of the preference of assignment on the outcomes of interest (Dong, 2017). To show that this is not the case, in Figure B.2b we show the binned means of the probability of enrolling in the centralized system as a function of the distance to the cutoff. In addition, the line represents the predicted values obtained from estimating the regression discontinuity model described in (B.1) considering as dependent variable the probability of enrolling in the centralized system. As Figure B.2b and column (1) in Table B.1 show, there are no significant differences in the enrollment probabilities among students above and below the cutoff, so we conclude that the potential selection problem is not a concern in our case.

To assess the causal effect of the preference of assignment on other outcomes, we focus on students that applied to at least two programs in the centralized system, and we restrict the analysis to the top preference of each student for simplicity.[64] In Figure 4.3 we display binned means of different outcomes as a function of the distance between the cutoffs in their top preference and the students' scores, while in Table B.1 we report the corresponding estimation results.

Figure B.2b shows the probability of enrolling in the top preference. As reported in column (2) in Table B.1, exceeding the cutoff increases the probability of enrollment in the top preference by 51.3%. Notice that this is not 100% for two reasons: (i) students whose score exceeds the cutoff may decide not to enroll, and (ii) students whose score was below the cutoff may end-up enrolling after being pulled from the wait-list. Figures B.2c and B.2d are discussed in Section 4.3, and show that being above the cutoff significantly reduces the probability of re-applying and switching programs within the system. These results are confirmed in columns (3) and (4) in Table B.1. Figure B.2e and column (5) in Table B.1 show a similar pattern, as it shows that the probability of major switching also decreases among students above the cutoff. In contrast, we observe no significant difference in university switchings. Finally, in Figure B.2g, we show that there is no effect of exceeding the cutoff on dropout rates. In contrast, we observe that students that exceed the cutoff are 2.2% more likely to stop out compared to students below the cutoff, as shown in Figure B.2h and column (7) in Table B.1. One potential reason for this effect is that, if students agree in their preferences for programs, there will be a discontinuous effect in the composition of peers of students above and below the cutoff. For instance, if a student gets admitted just above the cutoff, she will be at the bottom of the application scores' distribution within her assigned program. However, if she falls below the cutoff of her first preference, she could be at the top of her assigned program's distribution of application scores (second preference). If the student's relative position matters for her academic achievement, students who are just above the cutoffs could be more likely to

[64] Notice that we could perform the RD analysis for every cutoff, i.e., we could compute for every program the causal effect of being assigned to that program when it is listed as a top reported preference. In this sense, the causal effect that we estimate under the current specification is the average of causal effects across all programs that are listed as a top preference.

receive lower grades and be expelled from their programs, which could be observed in the data as a stop out.

Table B.1: Effect of Crossing Cutoff

	Enroll	Enroll Top Pref.	Re-App	Switch	Switch Major	Switch University	Dropout	Stopout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Z_{ip}	0.031	0.522***	−0.111***	−0.040***	−0.032**	0.001	−0.006	0.023***
	(0.040)	(0.035)	(0.013)	(0.009)	(0.010)	(0.009)	(0.005)	(0.006)
Observations	60	60	60	60	60	60	60	60
R^2	0.070	0.968	0.907	0.500	0.389	0.181	0.352	0.246
Adjusted R^2	0.065	0.966	0.902	0.472	0.356	0.138	0.317	0.206
Note:						*p<0.1; **p<0.05; ***p<0.01		

B.3.1 Regression Discontinuities with True Preferences.

Our previous analysis focuses on the effect of being above or below the cutoff of the top reported preferences on different outcomes. Using the cohort of 2019 and our nationwide survey, we can perform a similar analysis to estimate the causal effect of being above or below the cutoff of students' top-true preferences on their outcomes. In Figures B.3a and B.3b being above the cutoff significantly reduces the probability of re-applying to the system and being assigned to a different program in the next year. These results are consistent with those reported in Figure B.2, with the effect on re-applications being slightly smaller and that in switching being somewhat larger in magnitude compared to the analysis above.

Figure B.3: Effect of Cutoff Crossing for Top True Preference

(a) Re-Application (b) Re-Assignment

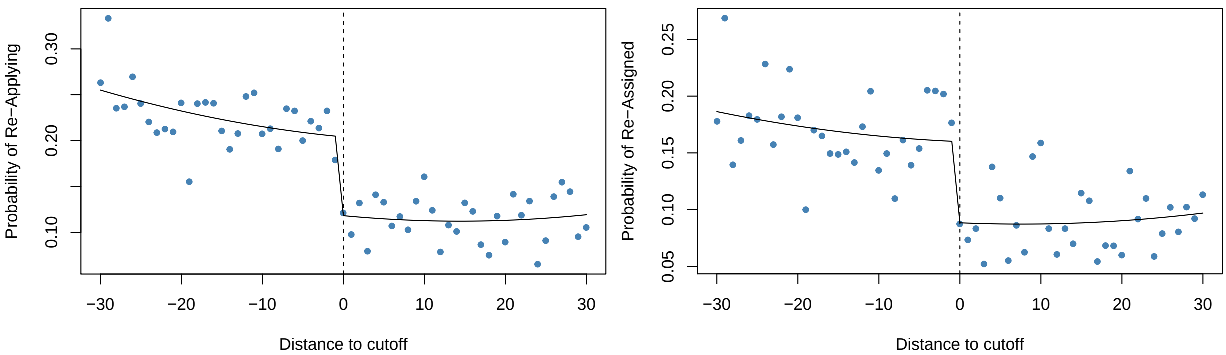
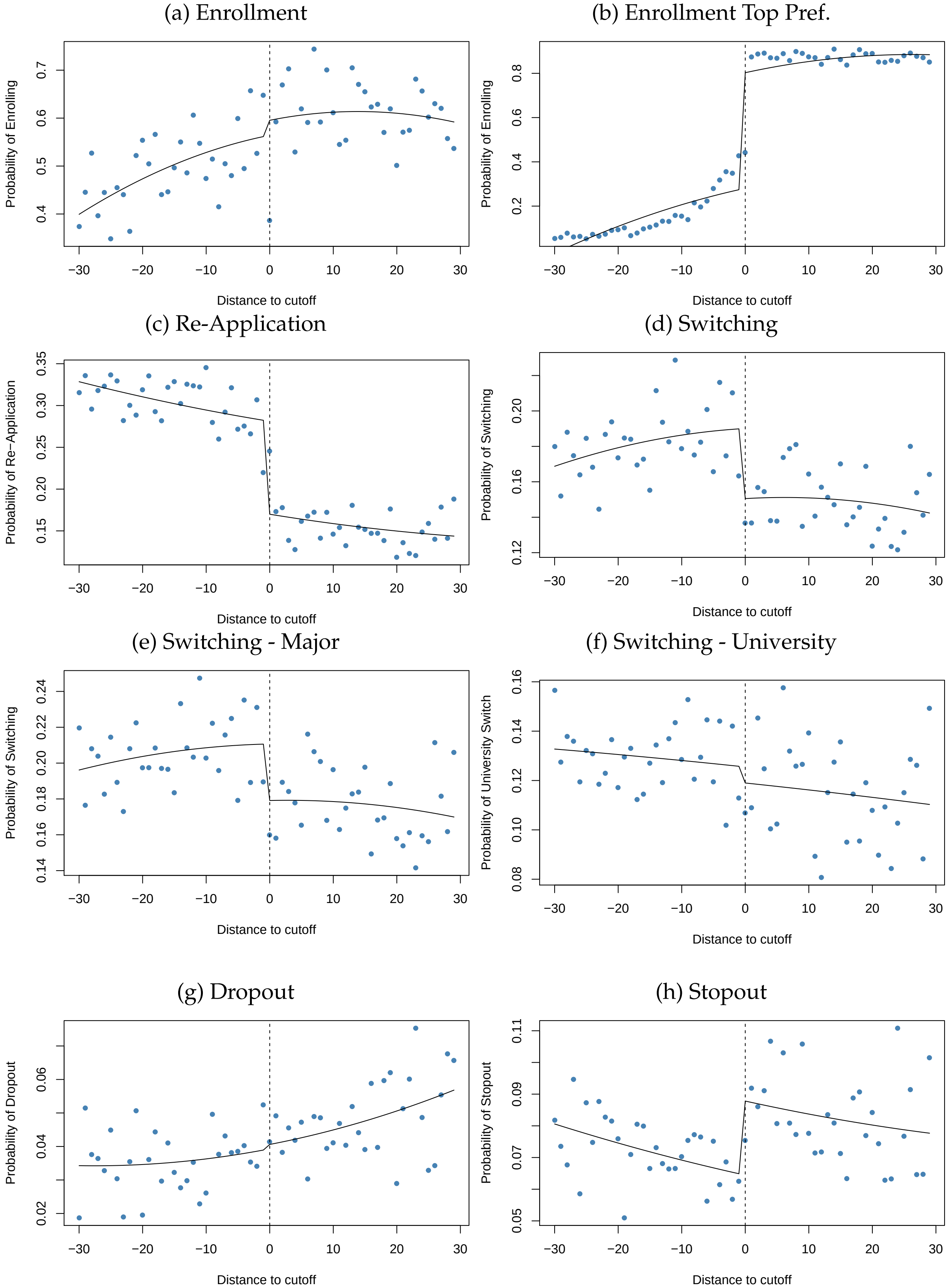


Figure B.2: Effect of Cutoff Crossing



(a) Enrollment

(b) Enrollment Top Pref.

(c) Re-Application

(d) Switching

(e) Switching - Major

(f) Switching - University

(g) Dropout

(h) Stopout

B.4 ADDITIONAL EVIDENCE

Figure B.4: Distribution Preference of Assignment

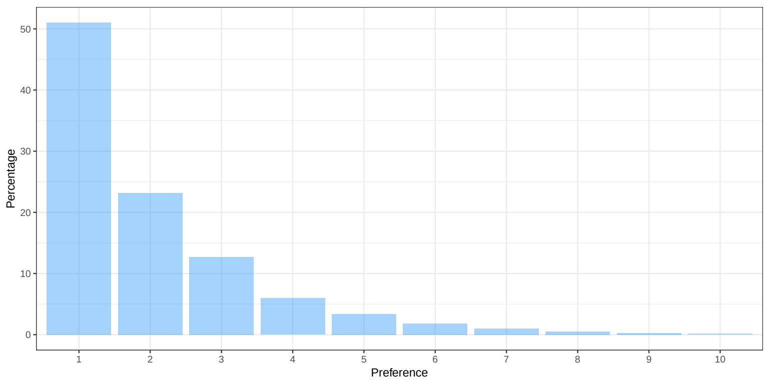


Figure B.5: Distribution of the number of different majors and universities in a ROL

(a) Majors

(b) Universities

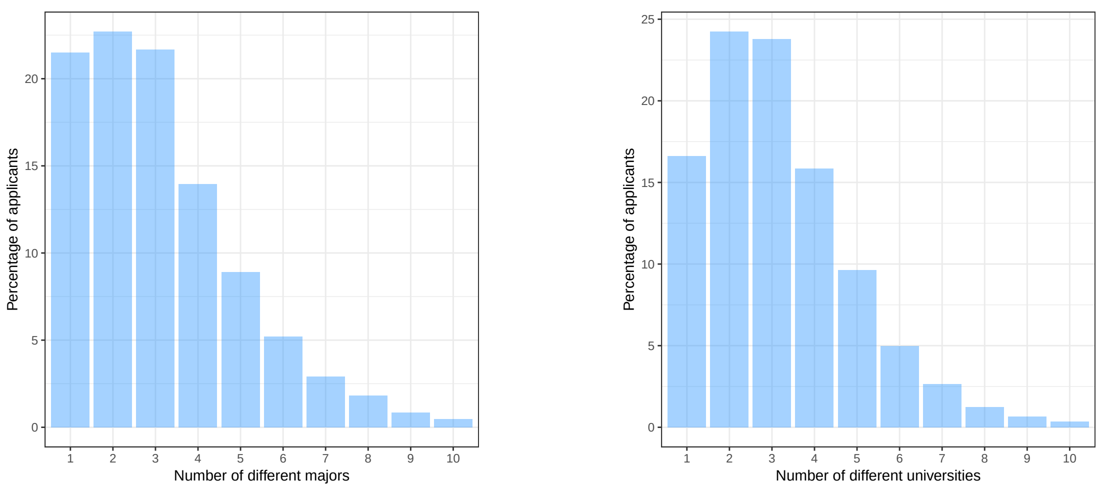


Figure B.6: Difference between expected cutoff for first true preference and expected cutoff for first listed preference

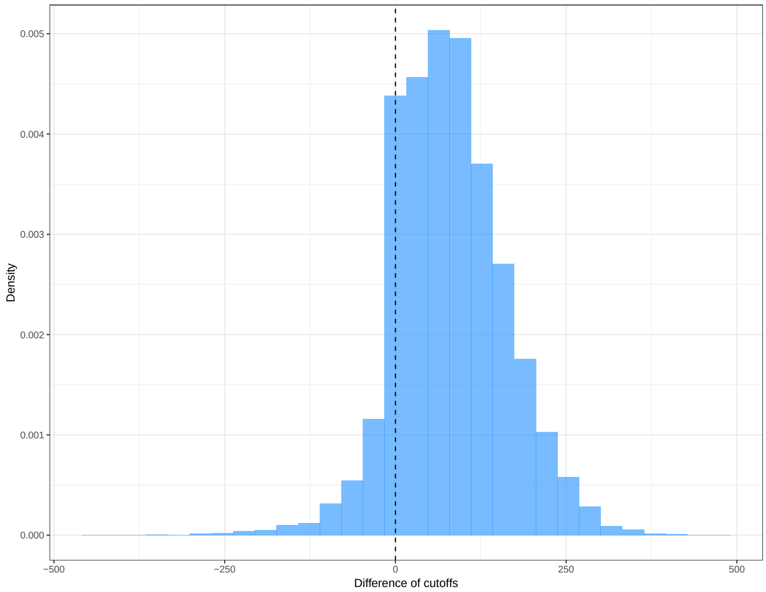


Figure B.7: Share of Truth-Tellers among Constrained vs. Short-List Students

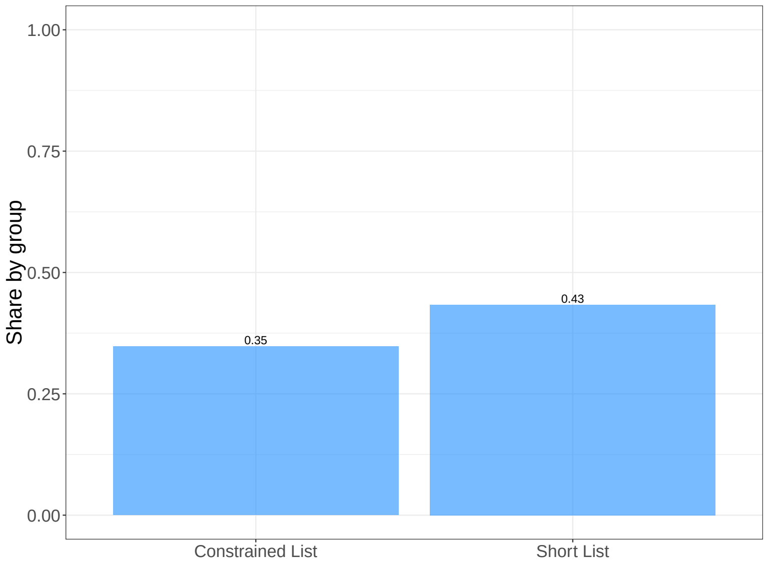


Figure B.8: Distribution of Optimism

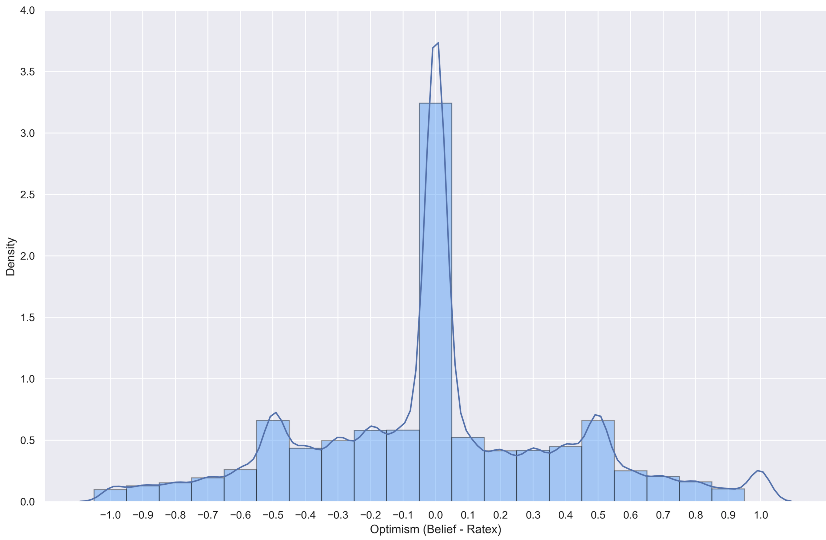
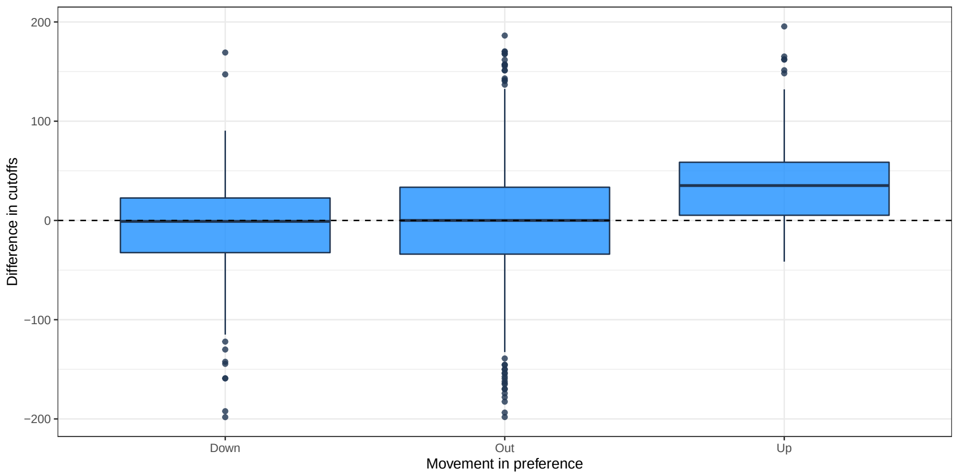


Figure B.9: Differences in cutoffs between switched and initial program by switching category



C Appendix for Section 5

C.1 Learning

Proposition 1 allows us to obtain the posterior mean and variance for student i's unknown ability in any program j'. We show how the student's statistical problem can be re-written to make inference about each component in A^u_{ij}.

To make inference about $A^u_{im_j}$ we can write

$$\begin{aligned} A^u_{ijt} &= A^u_{ij} + \varepsilon^g_{ijt} \\ \Leftrightarrow A^u_{ijt} &= A^u_{im_j} + \sum_k \omega_{jk} A^u_{ik} + \varepsilon^g_{ijt} \\ \Leftrightarrow A^u_{ijt} &= \mathbb{E}_{t-1}\left[\sum_k \omega_{jk} A^u_{ik}\right] + \nu_{gti} \\ \Leftrightarrow A^u_{ijt} - \mathbb{E}_{t-1}\left[\sum_k \omega_{jk} A^u_{ik}\right] &= \nu_{gti} \end{aligned}$$

where

$$\nu_{gti} \sim N\left(A^u_{im_j}, \sigma^2_g + \sum_k \omega^2_{jk}\sigma^2_s\right) \tag{C.1}$$

where now we treat $A^u_{ijt} - \mathbb{E}_{t-1}\left[\sum_k \omega_{jk} A^u_{ik}\right]$ as the new signal, and we make inference about $A^u_{im_j}$. This statistical problem now fits into DeGroot (2005)'s framework. We can similarly make inference about each A^u_{il} as follows

$$\begin{aligned} A^u_{ijt} &= A^u_{ij} + \varepsilon^g_{ijt} \\ \Leftrightarrow A^u_{ijt} &= A^u_{im_j} + \sum_k \omega_{jk} A^u_{ik} + \varepsilon^g_{ijt} \\ \Leftrightarrow A^u_{ijt} &= \mathbb{E}_{t-1}\left[A^u_{im_j}\right] + \mathbb{E}_{t-1}\left[\sum_{k\neq l} \omega_{jk} A^u_{ik}\right] + \xi_{gtil} \\ \Leftrightarrow A^u_{ijt} - \left(\mathbb{E}_{t-1}\left[A^u_{im_j}\right] + \mathbb{E}_{t-1}\left[\sum_{k\neq l} \omega_{jk} A^u_{ik}\right]\right) &= \xi_{gtil} \\ \Leftrightarrow \frac{A^u_{ijt} - \left(\mathbb{E}_{t-1}\left[A^u_{im_j}\right] + \mathbb{E}_{t-1}\left[\sum_{k\neq l} \omega_{jk} A^u_{ik}\right]\right)}{\omega_{jl}} &= \frac{\xi_{gtil}}{\omega_{jl}} \end{aligned}$$

where

$$\frac{\xi_{gtil}}{\omega_{jl}} \sim N\left(A^u_{il}, \frac{\sigma^2_g + \sigma^2_m + \sum_{k\neq l}\omega^2_{jk}\sigma^2_s}{\omega^2_{jl}}\right) \tag{C.2}$$

We can now write the posterior mean unknown ability if the student i enrolls in program j' in the second period, after receiving the first period signal a_{ij1} in program j:

$$\begin{aligned} E_1\left(A^u_{ij'} \middle| a_{ij1}\right) &= E_1\left(A^u_{im_{j'}} + \sum_k \omega_{j'k} A^u_{ik} \middle| a_{ij1}\right) \\ &= E_1\left(A^u_{im_{j'}} \middle| a_{ij1}\right) + \sum_k \omega_{j'k} E_1\left(A^u_{ik} \middle| a_{ij1}\right). \end{aligned}$$

Notice that if the student switches majors, i.e $m_{j'} \neq m_j$, she learns nothing about her major-specific unknown ability in her new program. This implies that the posterior mean equals the prior, that is,

$$E_1\left(A^u_{im_{j'}}\middle|a_{ij1}\right) = 0.$$

So the posterior mean is given by

$$E_1\left(A^u_{im_{j'}}\middle|a_{ij1}\right) = \begin{cases} 0 & \text{if } m_{j'} \neq m_j \\ \frac{\sigma^2_m a_{ij1}}{\sigma^2_g + \sum_k \omega^2_{jk}\sigma^2_s + \sigma^2_m} & \text{o.w} \end{cases} \tag{C.3}$$

We now turn to compute the posterior mean for the subject-specific unknown ability, i.e, $E_1\left(A^u_{ik}\middle|a_{ij1}\right)\ \forall k$. Although the student's subject-specific unknown ability does not vary across programs, given that grades depend on the average ability, and average ability varies depending on the program-specific admission weights ω_j, the amount of subject-specific unknown ability learned by the student will be program-specific.

The subject-specific posterior unknown ability is given by

$$E_1\left(A^u_{il}\middle|a_{ij1}\right) = \frac{\omega_{jl}\sigma^2_s a_{ij1}}{\sigma^2_g + \sigma^2_m + \sum_k \omega^2_{jk}\sigma^2_s} \tag{C.4}$$

Finally, we can write the posterior mean for the unknown ability in any program j' by

$$E_1\left(A^u_{ij'}\middle|a_{ij1}\right) = \begin{cases} \frac{\sum_l \omega_{j'l}\omega_{jl}\sigma^2_s a_{ij1}}{\sigma^2_g + \sigma^2_m + \sum_k \omega^2_{jk}\sigma^2_s} & \text{if } m_{j'} \neq m_j \\ \frac{\sigma^2_m a_{ij1}}{\sigma^2_g + \sum_k \omega^2_{jk}\sigma^2_s + \sigma^2_m} + \frac{\sum_l \omega_{j'l}\omega_{jl}\sigma^2_s a_{ij1}}{\sigma^2_g + \sigma^2_m + \sum_k \omega^2_{jk}\sigma^2_s} & \text{o.w} \end{cases} \tag{C.5}$$

Intuitively, the posterior mean places more weight on the signal for the subjects with a higher admission weight in ω_j. In this sense, student i learns more about her math ability if she enrolls in Engineering, which has a high admission weight on math.

C.2 MODEL SOLUTION

In this subsection, we describe the solution of the model via Backward Induction.

In period three, the terminal value function is given by

$$\begin{aligned} V_{ijt}(\mu_{ij2}, \tau_{ijt}) = {} & E_t\left[\sum_{t'=\tau_{ijt}+1}^{T_f} P^g_{ijt'}\left(\mathbb{E}_\varepsilon\left[\sum_{t''=0}^{t'-(\tau_{ijt}+1)} \beta^{t''} u_{ij(t+t'')}\right] + \beta^{t'-\tau_{ijt}} \underbrace{V^w_{ij(t+t'-\tau_{ijt})}(\mu_{ij2})}_{\text{Value fcn Labor market}}\right)\right] \\ & + E_t\left[\sum_{t'=\tau_{ijt}+1}^{T_f} P^d_{ijt'}\left(\mathbb{E}_\varepsilon\left[\sum_{t''=0}^{t'-(\tau_{ijt}+1)} \beta^{t''} u_{ij(t+t'')}\right] + \beta^{t'-\tau_{ijt}} \underbrace{V_{i0(t+t'-\tau_{ijt})}}_{\text{Value fcn Dropout}}\right)\right], \end{aligned}$$

where μ_{ij2} is the posterior unknown ability of student i in program j after observing the period one signal, and τ_{ijt} is the number of academic years the student has completed in program j, at the beginning of period three. In period three, there are no decisions to be made, and the value functions can be collapsed to the period two value functions of enrolling in program j in the following way:

$$V_{ijt}(\mu_{ij2}, \tau_{ijt}) = u_{ijt} - \mathbb{1}_{\{(j\neq 0)\cap(\tau_{ijt}=0)\}}C^e + \beta\mathbb{E}_\varepsilon\left[V_{ijt+1}(\mu_{ij2}, \tau_{ijt+1})\right]$$

where C^e is a first-time enrollment cost.

The value function in period one is then given by

$$V_{ijt}(\mu_{ij1}, \tau_{ijt}, \vec{s}_{it}) = \max_{d^s_{it}} E_0\Big[u_{ijt} - d^s_{it}C^{psu} - \mathbb{1}_{\{j\neq 0\}}C^e +$$
$$\beta\int_{a_{ij1}}\int_{\vec{s}_{it+1}} \underbrace{EmaxROL(\tau_{ijt}+1, \vec{s}_{it+1}, \mu_{i2}(a_{ij1}))}_{\text{continuation value of reapplications}}\underbrace{d\pi(a_{ij1})}_{\text{signal}}\underbrace{dF(\vec{s}_{it+1}|\vec{s}_{it}, d^s_{it})}_{\text{future scores}}\Big].$$

Notice that in period one, the value function of enrolling in program j considers that the student will update her beliefs about her unknown abilities for every program (μ_{i2}), that in the next period, her scores could change if she retakes the PSU ($d^s_{it} = 1$), and that she will have the option of submitting an optimal application in the second period ($EmaxROL(\cdot)$). In Appendix C.4 we derive analytical expressions for the continuation value of re-applications.

Actions. In periods one and two, students can choose to submit an application list. The indirect utility over assignment for student i to program k in period t, given her current enrollment in program j, $v_{ikt}|j$, can be written as:

$$v_{ikt}|j = P^e_{it}\cdot V_{ikt} + (1-P^e_{it})\cdot \max\{V_{i0t}, V_{ijt}\}$$

Given students' indirect utilities over the assignment and their beliefs about admission probabilities, students choose an application list—depending on their application type—as detailed in Section 5.4.2.

C.3 MIA

Chade and Smith (2006) shows that the optimal portfolio problem is NP-Hard. However, when admission probabilities are independent[65] and the cost of applying to a subset of

[65]Notice that in our case we have assumed in Assumption 6 independence of beliefs on admission probabilities.

programs S only depends on its cardinality, i.e., $c_i(S) = c(|S|)$ for some function c, the unconstrained problem is Downward Recursive, and the optimal solution is given by a greedy algorithm called Marginal Improvement Algorithm (MIA).

MIA: Marginal Improvement Algorithm (Chade and Smith (2006))

- Initialize $S_0 = \emptyset$
- Select $j_n = \arg\max_{j \in M \setminus S_{n-1}} \{U(S_{n-1} \cup j)\}$
- If $U(S_{n-1} \cup j_n) - U(S_{n-1}) < c\,(S_{n-1} \cup j_n) - c\,(S_{n-1})$, then STOP.
- Set $S_n = S_{n-1} \cup j_n$

MIA recursively adds programs that give the highest marginal improvement to the portfolio, as long as they exceed the marginal cost of adding them. Olszewski and Vohra (2016) show that MIA also returns the optimal ROL when the number of applications is constrained and when $c(S)$ is supermodular. If Assumption 6 does not hold, the strict inequality in MIA's stopping criteria becomes a weak inequality. In this case, if students face degenerate admission probabilities, there could be multiplicity of best response (He (2012)). We discuss this potential identification threat in Larroucau and Ríos (2018). Assumption 6 is a sufficient condition to rule out the multiplicity of best response.

C.4 EmaxROL

In this subsection, we analyze the problem of computing the expected value of reporting an optimal ROL in the centralized system, where the expectation is taken over next period preference shocks. Formally, let the utility of being assigned to program j by $u_j = \bar{u}_j + \varepsilon_j$, then define the $EmaxROL$ by

$$EmaxROL := \mathbb{E}_{\varepsilon}\left[U(R_{max}) := \max_{R' \in \mathcal{R}, |R'| \leq K} U(R') - c(R')\right] \tag{C.6}$$

where, given Assumption 6 and a ROL $R = \{r_1, \ldots, r_k\}$,

$$U(R) = z_{r_1} + (1 - p_{r_1}) \cdot z_{r_2} + \ldots + \prod_{l=1}^{k-1} (1 - p_{r_l}) \cdot z_{r_k}, \tag{C.7}$$

where $z_j = p_j \cdot u_j = p_j \cdot (\bar{u}_j + \varepsilon_j)$ for each $j \in M$.

Finding a—potentially—closed-form solution to this problem is relevant because it allows us to characterize the continuation value in any dynamic model where students can make application decisions over time. However, to the extent of our knowledge, this problem has not been analyzed in the literature yet. The following example shows why this problem is different from computing the continuation value in a dynamic discrete choice model, usually referred to as $Emax$ operator, or inclusive value.

Example 2 (EmaxROL).

Consider a portfolio problem where students can submit ROLs of length $K = 1$ and there is no cost of application, i.e, $c(R) = 0, \forall R \in \mathcal{R}$. In this case, the $EmaxROL$ problem simplifies to the expectation—over the preference shocks—of choosing the program that gives the highest expected utility over assignment, that is

$$\begin{aligned} EmaxROL &:= \mathbb{E}_\varepsilon \left[U(R_{max}) := \max_{R' \in \mathcal{R}, |R'| \leq K} U(R') - c(R') \right] \\ &= \mathbb{E}_\varepsilon \left[\max_{j' \in \mathcal{J}} p_j(\bar{u}_j + \varepsilon_j) \right] \\ &= \mathbb{E}_\varepsilon \left[\max_{j' \in \mathcal{J}} p_j \bar{u}_j + p_j \varepsilon_j \right] \end{aligned}$$

Event though in this case the $EmaxROL$ reduces to the expectation of choosing the best alternative in a discrete choice problem, now the preference shocks are weighted by—potentially different—admission probabilities $\{p_j\}$. This key difference—compared to a traditional discrete choice problem—makes that, even if we assume that preference shocks are distributed i.i.d type-I extreme value, the resulting random shocks, $p_j\varepsilon_j$, won't have equal variance. This implies that the inclusive value formulas derived in Rust (1987) do not hold.

The previous example shows that, in general, the $EmaxROL$ will not have a closed-form solution, even when preference shocks are distributed i.i.d type-I extreme value. We show now sufficient conditions under which the $EmaxROL$ can be efficiently approximated.

C.4.1 Pairwise-Stability

Under mild assumptions, Fack et al. (2019) shows that the allocation outcome from constrained DA satisfies pair-wise stability with respect to students' true preferences. We can exploit this fact for efficiently computing the $EmaxROL$.

When the allocation satisfies pair-wise stability, the problem of the student reduces to choosing the most preferred program among the programs for which she is *ex-post* admissible. That is, given a realization of programs' cutoffs, $\{P_j\}_{\mathcal{J}}$, student i's allocation , $\mu(i|\{P_j\}_{j\in\mathcal{J}})$, satisfies pair-wise stability if and only if

$$\mu(i|\varepsilon_i, \{P_j\}_{j\in\mathcal{J}}) = \underset{j \in J_i(\{P_j\}_{j\in\mathcal{J}})}{\text{argmax}} \; \bar{u}_{ij} + \varepsilon_{ij} \tag{C.8}$$

where

$$J_i(\{P_j\}_{j\in\mathcal{J}}) := \{j \in \mathcal{J} : s_{ij} \geq P_j\} \bigcup \{j = 0\} \tag{C.9}$$

This implies that we can write the $EmaxROL$ as follows

$$\begin{aligned} EmaxROL &:= \mathbb{E}_{\varepsilon}\left[U(R_{max}) := \max_{R' \in \mathcal{R}, |R'| \leq K} U(R') - c(R')\right] \\ &= \mathbb{E}_{\{P_j\}_{j\in\mathcal{J}}}\left[\mathbb{E}_{\varepsilon_i}\left[\max_{j\in J_i(\{P_j\}_{j\in\mathcal{J}})} \bar{u}_{ij} + \varepsilon_{ij} \Big| \{P_j\}_{j\in\mathcal{J}}\right]\right] \end{aligned}$$

and when ε_{ij} are distributed i.i.d type-I extreme value, the previous expression reduces to

$$\begin{aligned} EmaxROL &= \mathbb{E}_{\{P_j\}_{j\in\mathcal{J}}}\left[\mathbb{E}_{\varepsilon_i}\left[\max_{j\in J_i(\{P_j\}_{j\in\mathcal{J}})} \bar{u}_{ij} + \varepsilon_{ij} \Big| \{P_j\}_{j\in\mathcal{J}}\right]\right] \\ &= \mathbb{E}_{\{P_j\}_{j\in\mathcal{J}}}\left[\log\left(\sum_{j\in J_i(\{P_j\}_{j\in\mathcal{J}})} \exp\left(\bar{u}_{ij}\right)\right) + \gamma\right], \end{aligned}$$

where γ is the Euler's constant.

If we take the distribution of cutoffs to be invariant to students' individual reports, following a similar argument than Agarwal and Somaini (2018); we can estimate in a first stage the distribution of cutoffs $\{\hat{P}_j\}_{j\in\mathcal{J}}$ and then estimate the structural parameters of the model. This implies that we can compute the frequency of the random sets by using the bootstrap realizations of the cutoffs $J_i(\{P_j^{\tilde{b}}\}_{j\in\mathcal{J}})$ just once, where $\tilde{b} = 1, ..., \tilde{B}$ is a random sample of the bootstrap realizations of the cutoffs. We can then approximate the $EmaxROL$ by doing

$$\begin{aligned} EmaxROL &= \mathbb{E}_{\{P_j\}_{j\in\mathcal{J}}}\left[\log\left(\sum_{j\in J_i(\{P_j\}_{j\in\mathcal{J}})} \exp\left(\bar{u}_{ij}\right)\right) + \gamma\right] \\ &\approx \frac{\sum_{\tilde{b}\in\tilde{B}} \log\left(\sum_{j\in J_i(\{P_j^{\tilde{b}}\}_{j\in\mathcal{J}})} \exp\left(\bar{u}_{ij}\right)\right) + \gamma}{\tilde{B}} \end{aligned}$$

Pairwise stability in the dynamic problem. We can a follow similar calculations than before and give an expression for the expected value of reporting a ROL in our dynamic setting. The expected value over assignment to program k, given that student i is currently enrolled in program j, is given by

$$\begin{aligned} v_{ikt} &= P_{it}^e V_{ikt} + (1 - P_{it}^e) \max\{V_{i0t}, V_{ijt}\} \\ &= P_{it}^e \left(\bar{V}_{ikt} + \varepsilon_{ikt}\right) + (1 - P_{it}^e) \max\{\bar{V}_{i0t} + \varepsilon_{i0t}, \bar{V}_{ijt} + \varepsilon_{ijt}\} \end{aligned}$$

then we can write

$$\begin{aligned} \mathbb{E}_{\varepsilon}\left[\max_k v_{ikt}\right] &= \mathbb{E}_{\varepsilon}\left[\max_k P_{it}^e \left(\bar{V}_{ikt} + \varepsilon_{ikt}\right) + (1 - P_{it}^e) \max\{\bar{V}_{i0t} + \varepsilon_{i0t}, \bar{V}_{ijt} + \varepsilon_{ijt}\}\right] \\ &= P_{it}^e \mathbb{E}_{\varepsilon}\left[\max_k \left(\bar{V}_{ikt} + \varepsilon_{ikt}\right)\right] + (1 - P_{it}^e)\, \mathbb{E}_{\varepsilon}\left[\max\{\bar{V}_{i0t} + \varepsilon_{i0t}, \bar{V}_{ijt} + \varepsilon_{ijt}\}\right] \end{aligned}$$

and we get that

$$EmaxROL(\tau_{ijt}, \vec{s}_{it}, a_{ij1}) \approx$$
$$\frac{\sum_{\tilde{b}\in\tilde{B}} P_i^e \log\left(\sum_{k\in J_i(\{P_k^{\tilde{b}}\}_{k\in\mathcal{J}}, \vec{s}_{it})} \exp\left(\bar{V}_{ikt}\right)\right)}{\tilde{B}} + (1-P_i^e)\log\left(\exp\left(\bar{V}_{ijt}\right) + \exp\left(\bar{V}_{i0t}\right)\right) + \gamma$$

Finally, when students re-take the PSU in the first period, $d^s_{it-1} = 1$, we also need to integrate over students' future scores. Under Assumption 5 and using Gauss-Hermite polynomials, we can approximate the integral with stochastic scores over $EmaxROL$ by

$$\int_{\vec{s}_{it}} EmaxROL(\tau_{ijt}, \vec{s}_{it}, a_{ij1}) \underbrace{dF(\vec{s}_{it}|\vec{s}_{it-1}, d^s_{it-1})}_{\text{future scores}} \approx$$
$$\frac{1}{\sqrt{\pi}}\sum_{k=1}^{n_w} w_k \left(\frac{\sum_{\tilde{b}\in\tilde{B}} P_i^e \log\left(\sum_{l\in J_i\left(\{P_l^{\tilde{b}}\}_{l\in\mathcal{J}}, \vec{s^k}_{it}\right)} \exp\left(\bar{V}_{til}\right)\right)}{\tilde{B}}\right) + (1-P_i^e)\log\left(\exp\left(\bar{V}_{ijt}\right) + \exp\left(\bar{V}_{i0t}\right)\right) + \gamma$$

where

$$\vec{s^k}_{it} = \max\{\vec{s^k}_{it-1}, \tilde{s^k}_{it}\} \tag{C.10}$$

with

$$\tilde{s^k}_{ilt} = \begin{cases} \alpha_l(1+\sqrt{2}\sigma_{psu}x_k)s_{ilt} & \text{if } s_{ilt} > 0 \\ \alpha_{0l}(1+\sqrt{2}\sigma_{psu}x_k)\bar{s}_{it} & \text{if } s_{ilt} = 0 \end{cases}$$

where n_w is the number of nodes at which we evaluate the integrand and w_k is the k-th integration weight for the k-th integration node x_k, given by the Gauss-Hermite formula. The accuracy of the previous approximation will depend on the number of nodes used to approximate the integral, n_w, and the number of joint draws of the cutoff scores, $\tilde{B}$.

D Appendix for Section 6

D.1 Identification of Wage Parameters with Aggregate Data

The following example gives intuition in how the wage equation's parameters can be identified with aggregate-level data.

Example 3 (Identification with aggregate labor market information).

Consider the following simpler log-wage equation:

$$\log(w_{ij\tau}) = \lambda_{1m_j} + \lambda_{k_j2} + \lambda_{4m_j}\bar{G}_{ij} + \epsilon_{ij\tau} \tag{D.1}$$

We can compute the average log-wage for each program j as

$$\log(w_{ij\tau}) = \lambda_{1m_j} + \lambda_{k_j2} + \lambda_{4m_j}\bar{G}_{ij} + \epsilon_{ij\tau}$$
$$\frac{\sum_{i=1}^{N_j}\log(w_{ij\tau})}{N_j} = \lambda_{1m_j} + \lambda_{k_j2} + \lambda_{4m_j}\bar{G}_j + \frac{\sum_{i=1}^{N_j}\epsilon_{ij\tau}}{N_j}$$
$$\log(\bar{w}_{j\tau}) = \lambda_{1m_j} + \lambda_{k_j2} + \lambda_{4m_j}\bar{G}_j + \epsilon^{\bar{w}}_{j\tau}$$

Variation in the average of log wages across programs, identifies the parameters. In particular, the terms λ_{1m_j} and λ_{k_j2} capture the mean wage for programs within a *broad major* and within a *university type*. Variation on programs' average grades $\bar{G}_j$ identifies the coefficient $\lambda_{4m_j}\bar{G}_j$.

D.2 Identification from Re-Applications

This subsection shows that we can use students' reapplications and a revealed preferences approach to obtain information about how much students' preferences change over time.

Analyzing ROLs when students are not truth-tellers is challenging because ROLs are the product of both beliefs on admission chances and preferences over programs. Observing two different ROLs in two periods does not immediately imply that students' preferences are changing because their beliefs on admission probabilities could also be changing. Moreover, if there is degeneracy on students' admission probabilities, there could be multiplicity of best response, which further complicates the analysis. Fortunately, Agarwal and Somaini (2018) show how we can identify the set of students' indirect utilities that are rationalized by a given ROL R if we can estimate beliefs in a first stage, and Larroucau and Ríos (2018) extend their methodology to construct this set in large scale portfolio problems like the Chilean College Admissions problem. We use their insights to construct a learning measure from observed reapplications, which does not involve parametric assumptions on the learning process nor the utility function. However, it does rely on estimating beliefs on admission probabilities in the first stage and assuming that students maximize their expected utility over the assignment given their subjective beliefs and preferences.

Let v_t be a vector of indirect utilities over programs at time t such that $v_t = \{v_{1t}, ..., v_{Jt}\}$, $v_{0t} \equiv 0$ to be the value of being unassigned to the centralized system at time t, $\{p_t\}$ to be the set of admission probabilities at time t, and R_t, to be a ROL submitted at time t. Larroucau and Ríos (2018) show that, under Assumption 6, the set of indirect utilities that rationalize R_t to be an optimal ROL, $C_v(R_t)$, is given by the solution to the following system of linear inequalities:

$$C_v(R_t) \equiv \{v_t : \Gamma_{R_t}(v_t - v_{0t}) \geq 0\} \quad \text{(D.2)}$$

where Γ_{R_t} is a matrix that encodes, by row, the implied admission probabilities of reporting R_t minus the implied admission probabilities of reporting a ROL $\tilde{R}_t$ that could dominate R_t in expected utility terms.

We can construct the corresponding set of indirect utilities that is implied by a reapplication R_{t+1}, when the outside option is given by $v_{0t+1} \equiv \max\{0, v_{kt+1}\}$, where v_{kt+1} is the indirect utility of being enrolled in program k at time $t+1$:

$$C_v(R_{t+1}) \equiv \{v_{t+1} : \Gamma_{R_{t+1}} (v_{t+1} - v_{0t+1}) \geq 0\} \tag{D.3}$$

We now give a sufficient condition to test for the variation of students' preferences over time:

Proposition 2 (Identification from reapplications**).**
Let $C_v(R_t)|v_{0t}, \{p_t\}$, and $C_v(R_{t+1})|v_{0t+1}, \{p_{t+1}\}$ be defined by Equations D.2 and D.3 respectively, then,

$$C_v(R_t) \cap C_v(R_{t+1}) = \emptyset \Rightarrow \mathbb{P}(v_t \neq v_{t+1}) = 1 \tag{D.4}$$

Proposition 2 has testable implications: we can construct a metric for measuring the change of preferences' over time, that is implied by students' reapplications, by characterizing the set $C_v(R_t) \cap C_v(R_{t+1})$. The set $C_v(R_t) \cap C_v(R_{t+1})$ is given by the solution to the following system of linear equations:

$$\begin{pmatrix} \Gamma_{R_t} \\ \Gamma_{R_{t+1}} \end{pmatrix} \begin{pmatrix} v \\ v - \max\{0, v_k\} \end{pmatrix} \geq 0 \tag{D.5}$$

thus, $C_v(R_t) \cap C_v(R_{t+1})$ is empty if, and only if, the previous system of linear equations has no solution.

E APPENDIX FOR SECTION 7

E.1 AUXILIARY MODEL AND WEIGHTING MATRIX

We describe below the regressions and moment conditions we use in the estimation and the sets of parameters that explain most of each moment's variation.

GRADE EQUATIONS. The auxiliary model that targets the grade equations' structural parameters (γ) are given by the regression analogs of Equations 5.9 and 5.10:

$$G_{ij1} = \beta^{\gamma}_{1m_j} + \beta^{\gamma}_2 A_{ij} + \beta^{\gamma}_3 Z^g_i + \beta^{\gamma}_4 \mathbb{1}\{j = R_{1i}(1)\} + \beta^{\gamma}_5 s_{1im_j} + \beta^{\gamma}_6 s_{1ik_j} + \varepsilon^g_{ij1}, \tag{E.1}$$

and

$$G_{ij2} = (\beta^{\gamma}_7 + \beta^{\gamma}_8 S)\, G_{ij1} + \beta^{\gamma}_9 + \gamma_{10} S + \varepsilon^g_{ij2}. \tag{E.2}$$

where s_{1im_j} and s_{1ik_j} are the shares of major m_j and college-type k_j in the ROL of student i in period 1 respectively, $\mathbb{1}\{j = R_{1i}(1)\}$ is an indicator function that equals to 1 if the student is assigned to her top-reported preference in period 1, and $S = 1$ if the student is in her second academic year, and $S = 0$ otherwise.

WAGE EQUATION. The auxiliary models that target the parameters in the wage equation (λ) are given by:

$$\log(\bar{w}_{j(\tau=4)}) = \beta^{\lambda}_{1m_j} + \beta^{\lambda}_2 \bar{A}_{k_j} + \beta^{\lambda}_3 \bar{G}_j + \beta^{\lambda}_4 \bar{Z}^w + \epsilon_{j(\tau=4)},$$

and

$$\log(\bar{w}_{m_j\tau}) = \beta^{\lambda}_{5m_j} + \beta^{\lambda}_{6m_j}\tau + \beta^{\lambda}_{7m_j}\tau^2 + \epsilon_{m_j\tau},$$

where τ is tenure after graduating[66].

NON-PECUNIARY LABOR MARKET PARAMETERS. The auxiliary model that targets the parameters that specify the non-pecuniary payoffs in the work force (α^w) is given by the following linear probability model:

$$y_{ij} = \beta^w_1 s_{1im_j} + \beta^w_2 \mathbb{1}\{j = R_{1i}(1)\} + \beta^w_3 A_{ij} + \beta^w_4 \bar{A}_{k_j} + \beta^w_5 Z^g_i + \varepsilon^w_{ij} \tag{E.3}$$

where

LEARNING PARAMETERS. The auxiliary models that target the parameters associated with students' learning process (σ^2_s, σ^2_m, σ^2_g, and α^w_4) are given by the following linear probability models of switchings and dropout:

For each outcome $O \in \{\text{switching major}, \text{switching math-type}, \text{switching program}, \text{dropping out}\}$

$$O_{ij} = \beta^o_{1m_j} + \beta^o_2 A_{ij} + \beta^o_3 Z^g_i + \beta^o_4 \mathbb{1}\{j = R_{1i}(1)\} + \beta^o_5 s_{1im_j} + \beta^o_6 s_{1ik_j} + \beta^o_7 G_{ij1} + \varepsilon^o_{ij}, \tag{E.4}$$

where O_{ij} equals one if student i enrolled in program j, switches major, switches math-type, switches program, or drops out respectively, and zero otherwise.

MOMENT CONDITIONS. We incorporate additional moment conditions to compute the objective function, capturing the identifying variations detailed in Section 6. Table E.1 summarizes the moment conditions and the targeted parameters:

WEIGHTING MATRIX AND STANDARD ERRORS. We use as a weighting matrix a diagonal matrix. Each element in the diagonal is the inverse of each data moment's variance, which we obtain via a bootstrap procedure. We weight up three moments in the weighting matrix that are key to identify the parameters involved in the learning process: the correlation between students' first-year college grades and the norm of the difference between the vectors of majors and broad-majors shares for students who reapply, and the norm of the difference between the vectors of ω shares for students who reapply. We do not use the optimal weighting matrix because of the numerical complexities involved in computing the derivatives of the objective function $Q(\theta)$. Therefore, our estimator will be unbiased but not efficient. Due to the first-stage estimation of students' beliefs, we need to estimate standard errors following a bootstrap procedure.

[66] See Section 4.2 for a description of the aggregate data on wages.

Table E.1: Estimation moments

Moment description	Targeted parameters
Share of students who retake the PSU	C^{psu}
Share of students who dropout by gender and income level	$\{\alpha_d\}_d$, α^w, C^e
Grade auxiliary model 1 coefficients	γ, σ_g^2
Grade auxiliary model 2 coefficients	γ, , σ_g^2
Wage auxiliary model coefficients	λ
Switchings and dropout auxiliary model coefficients	σ_g^2, σ_m^2, σ_s^2, α_4^w
Share of students who reapply	
Share of students who switch programs	σ_m^2, σ_s^2, V_{α^m}, V_{α^k}, C^e
Share of students who switch majors	σ_m^2, V_{α^m}
Share of students who switch college-types	V_{α^k}
Share of students who dropout at the end of the first year of college	α^w
Share of students who choose the outside option every year	α^w
Share of students who start college in the second year	
Share of students who remain in the same program after two years	
Share of top-true preferences by program, grouped by students' scores and income groups	α_1, α_2
Share of students whose top-reported preference is their top-true preference in R_1	ρ
Share of students whose top-reported preference is their top-true preference in R_2	ρ
Share of students whose top-reported preference has zero admission probability	ρ
Share of ROLs R_1 with length 10	ρ
Share of ROLs R_2 with length 10	ρ
Share of students who apply in the first year	
Share of students who apply in the second year	
Share of reapplications that change in their top-true preference	σ_m^2, σ_s^2, V_{α^m}, V_{α^k}
Shares of majors within R_1	V_{α^m}
Shares of college-types within R_1	V_{α^k}
Shares of majors within R_2	V_{α^m}
Shares of college-types within R_2	V_{α^k}
[1] Norm of the difference between the vectors of college-type shares for students who reapply	V_{α^k}
Norm of the difference between the vectors of major shares for students who reapply	σ_m^2, V_{α^m}
[1] Norm of the difference between the vectors of ω shares for students who reapply	σ_s^2, V_{α^m}, V_{α^k}
Correlation between first-year grades and the norm of the difference between the vectors of major shares for students who reapply	σ_m^2, σ_g^2
[1] Correlation between first-year grades and the norm of the difference between the vectors of ω shares for students who reapply	σ_s^2, σ_g^2
Share of applications by major and college-type, grouped by gender in R_1	Δ^m, Δ^k
Share of applications by major and college-type, grouped by gender in R_2	Δ^m, Δ^k
Share reapplications from top-reported preferences	
Share reapplications from top-true preferences	
Mean and variance of tuition for top-reported preferences, grouped by students' scores and income groups	$\{\alpha_c\}_c$
Mean and variance of tuition for top-true preferences, grouped by students' scores and income groups	$\{\alpha_c\}_c$
Mean and variance of distance for top-reported preferences	α_3
Mean and variance of distance for top-true preferences	α_3
Mean and variance of relative observed ability position for topreported preferences	α_4
Mean and variance of relative observed ability position for toptrue preferences	α_4
Mean and variance of $\log\left(\frac{s_{lt+1}}{s_{lt}}\right)$ for positive PSU scores	$\{\alpha_l\}_l$, σ_{psu}
Mean and variance of $\log\left(\frac{s_{lt+1}}{\bar{s}_t}\right)$ for PSU scores wit zero value in the first year	$\{\alpha_{0l}\}_l$,σ_{psu}

Note: [1] These statistics are computed conditional on the major/college-type of initial assignment.

E.2 TECHNICAL CONSIDERATIONS

The structural model has a mixture of continuous and discrete outcomes. This feature complicates the estimation procedure for a simulation-based method like II, because the objective function, $Q(\theta)$, becomes a multidimensional step function which inherits the discontinuities produced in the simulated data[67]. Bruins et al. (2018) propose a solution to overcome these computational difficulties by introducing noise and smoothing to the objective function. They refer to this estimation procedure as "Generalized Indirect Inference" (GII). With the smoothed objective function, the researcher can use a gradient-based optimization method to minimize the objective function, which tends to be faster than gradient-free optimization routines. We choose to avoid this smoothing procedure, and we estimate the objective function and find the global optimum using MIDACO solverSchlueter et al. (2013)[68]. We choose to do this because the model has close to 80 parameters to be estimated, and the gradient must be computed through numerical simulation. Thus, the evaluation of the gradient would take several minutes. The computational time of this approach could be significantly reduced by parallelizing the numerical approximation of the gradient. However, we have chosen to parallelize the objective function's computation instead and increase the number of draws in the forward-simulation stage to smooth the objective function. As solving the model and forward-simulating outcomes are completely independent across students, we parallelize the algorithm's outer loop to evaluate $Q(\theta)$.[69]

Algorithm 1 Computing $Q(\theta)$

Input. Value of the structural parameters θ, and first-stage estimates $\hat{p}$, $\hat{P}^e$, $\hat{P}^d$, $\hat{P}^g$, and $\hat{P}^w$.

Output. Value of the objective function $Q(\theta)$.

Step 1. For each student i, program j, and simulation b

Step 1.a. Draw a vector of random coefficients $\alpha_i^{m_{rc}}$,

Step 1.b. Solve the model by backward-induction,

Step 1.c. For each simulation in N_s and for each date, Draw a vector of preference shocks $\varepsilon_i^{m_s,m_{rc}}$, enrollment shocks $\varepsilon_i^{e,m_s,m_{rc}}$, wage shocks $\epsilon_i^{m_s,m_{rc}}$, vector of random cutoff scores $P^{m_s,m_{rc}}$ from the empirical distribution of cutoffs, vector of PSU score shocks $\nu_i^{m_s,m_{rc}}$, vector of unknown abilities $A_i^{u,m_s,m_{rc}}$, and grade shocks $\varepsilon_i^{g,m_s,m_{rc}}$

Step 1.d. Forward-simulate the model and obtain a set of outcomes $y_i^{m_s,m_{rc}}$,

Step 2. For each simulation, estimate the *auxiliary* model parameters, $\hat{\beta}^{m_s,m_{rc}}(\theta)$, on the simulated sample

Step 3. Compute $\bar{\beta}(\theta) = \frac{1}{N_{rc} \times N_s} \sum_{m_{rc}} \sum_{m_s} \hat{\beta}^{m_s,m_{rc}}(\theta)$

Step 4. Return $Q(\theta) := \left(\bar{\beta}(\theta) - \hat{\beta}\right)^T W \left(\bar{\beta}(\theta) - \hat{\beta}\right)$

---

[67]For a given realization of the random shocks, measures constructed from discrete outcomes of the model, change discontinuously when we change the value of the structural parameters.

[68]MIDACO uses an evolutionary hybrid algorithm based on the Ant Colony Optimization (ACO) metaheuristic (Schlueter et al. (2009))

[69]The model is coded in RcppArmadillo and parallelized with OpenMP.

E.3 RESULTS

F APPENDIX FOR SECTION 8

Index each counterfactual experiment and the baseline model by τ, then the Rational Expectations equilibrium cutoff distributions, $\hat{p}(\tau)$, can be computed with the following algorithm:

Algorithm 2 Computing $\hat{p}(\tau)$

Input. Structural parameter estimates $\hat{\theta}$, first-stage estimates $\hat{p}$, $\hat{P}^e$, $\hat{P}^d$, $\hat{P}^g$, and $\hat{P}^w$, and tolerance level ϵ_{tol}.
Output. Rational Expectations equilibrium cutoff distributions $\hat{p}(\tau)$
Step 1. For each program j
Step 1.a. Solve the model and simulate outcomes given the rules implied by counterfactual τ and the estimated objects $\left(\hat{\theta}, \hat{p}, \hat{P}^e, \hat{P}^d, \hat{P}^g, \hat{P}^w\right)$
Step 1.b. Obtain a set of simulated ROLs and scores $(R_1^0, R_2^0, s_1^0, s_2^0)$
Step 1.c. For each program j, estimate the mean and standard deviation of the cutoff distributions $\hat{\delta}_j^0 \equiv \left(\hat{\mu}_j^0, \hat{\sigma}_j^0\right)$
Step 2. $\delta_{diff} = 2\epsilon_{tol}$, $k = 1$, $\rho = 0.9$
Step 3. While $\delta_{diff} > \epsilon_{tol}$
Step 3.a. For each student i, solve the model via Backward Induction given τ, the estimated parameters $\left(\hat{\theta}, \hat{P}^e, \hat{P}^d, \hat{P}^g, \hat{P}^w\right)$, and cutoff distributions $\hat{p}^{k-1}$, and obtain the continuation values for each student and state
Step 3.b. Forward simulate first period ROL R_{i1}^k given τ, the estimated parameters $\left(\hat{\theta}, \hat{P}^e, \hat{P}^d, \hat{P}^g, \hat{P}^w\right)$, cutoff distributions $\hat{p}^{k-1}$, and continuation values
Step 3.c. For each program j, estimate the mean and standard deviation of the cutoff distributions $\hat{\delta}_j^0 \equiv \left(\hat{\mu}_j^0, \hat{\sigma}_j^0\right)$
Step 3.d. Given initial first period applications R_1^k, second period applications R_2^{k-1}, and students' scores s_1^k, and s_2^{k-1}, run the Chilean matching mechanism and obtain an allocation $\mu^k\left(R_1^k, R_2^{k-1}, s_1^k, s_2^{k-1}\right)$
Step 3.e. Given $\mu^k\left(R_1^k, R_2^{k-1}, s_1^k, s_2^{k-1}\right)$, τ, the estimated parameters $\left(\hat{\theta}, \hat{P}^e, \hat{P}^d, \hat{P}^g, \hat{P}^w\right)$, cutoff distributions $\hat{p}^{k-1}$, and continuation values, forward simulate second period ROLs R_{i2}^k
Step 3.f. Given $\left(R_1^k, R_2^{k-1}, s_1^k, s_2^{k-1}\right)$, run the boostrap procedure and estimate the Rational Expectations cutoffs distributions $\tilde{p}^k$ Take a convex combination of the realized cutoffs $\tilde{p}^k$ and $\hat{p}^{k-1}$ (point wise), i.e, $\hat{p}^k = \rho^k \hat{p}^{k-1} + \left(1 - \rho^k\right)\tilde{p}^k$
Step 3.g. Estimate the mean and standard deviation of the cutoff distributions $\hat{\delta}_j^k \equiv \left(\hat{\mu}_j^k, \hat{\sigma}_j^k\right)$
Step 3.h. Compute $\delta_{diff} = ||\hat{\delta}^k - \hat{\delta}^{k-1}||$ $\hat{p}(\tau) = \hat{p}^{k-1}$ $k++$

Under Assumption 8 the previous algorithm can be reduced to:

Table E.2: Estimation Results - Parameters

Parameters	Values
Application behavior and Dropout	
Share of strategic ROLs $(1-\rho)$	0.89
Cost of retaking PSU (C^{psu})	0.05
Dropout flow-utility for females ($\alpha^{dropout}_{female}$)	6.06
Dropout flow-utility for males ($\alpha^{dropout}_{male}$)	7.83
Dropout flow-utility for low-income ($\alpha^{dropout}_{low-income}$)	1.65
First-time enrollment cost (C^{e})	10.0
Flow-utility and Priors	
Tuition Common term (α_{c0})	-0.29
Tuition Low-income (α_{c1})	-1.59
Tuition Above-median score (α_{c2})	-0.92
Relative position (α_4)	0.354081
Distance (α_3)	-1.37695
Student observed ability (α_1)	3.37879
Program observed ability (α_2)	4.77871
Constant by major (Δ^m)	(3.58, 2.63, -0.24, 2.92)
Gender effect by major (Δ^m)	(0.37, -1.46, 2.96, 2.39)
Variance major random coefficient (σ^{2m}_{α})	14.70
Constant by college (Δ^k)	(-2.91, -0.65, 0.21)
Income effect by college (Δ^k)	(-2.72 4.39, 2.75)
Variance college random coefficient (σ^{2k}_{α})	13.17
Major prior variance (σ^2_m)	0.95
Subject prior variance (σ^2_s)	1.61
Grade equations	
Constant by major (γ_{1m_j})	(3.74, 3.07, 4.77, 3.05)
Student observed ability (γ_2)	0.27
Gender effect (γ_3)	1.39
Random coefficient effect on grades (major) (γ_4)	0.11
Random coefficient effect on grades (colleges) (γ_5)	0.03
Grade shock variance (σ^2_g)	0.62
Second year intercept (γ_6)	0.0
Second year slope (γ_7)	2.16
Evolution of scores	
Variance of ν (σ^2_{psu})	0.02
Mean proportional change (Verbal, Math, History, Science) ($\{\alpha_l\}_l$)	(1.07, 1.04, 1.06, 1.06)
Mean proportional change from zero score (History, Science) ($\{\alpha_{0l}\}_l$)	(1.04, 1.02)
Non-pecuniary work utility	
Major random coefficient (α^w_1)	31.13
Student observed ability (α^w_2)	7.32
College observed ability (α^w_3)	3.98
Non-pecuniary work value of unknown ability (α^w_4)	170.17
Pecuniary work utility parameter (α^w_5)	21.11
Wage parameters	
Constant by major (λ_{1m_j})	(-0.49, -0.24, -0.46, -0.35)
College observed ability (λ_2)	0.35
Grades (λ_3)	0.10
Gender effects by major (λ_4)	-0.27
Wage shock variance (σ^2_w)	0.80
Wage growth	
Linear term by major (λ_{5m_j})	(0.00, 0.15, 0.30, 0.32)
Quadratic term by major (λ_{6m_j})	(-0.06, -0.01, -0.06, -0.24)

Notes: the order of majors is Social Sciences, Science, Education and Humanities, and Health.
Notes: the order of colleges is CRUCH-Public, CRUCH-Private, and Non-CRUCH.

Table E.3: Estimation Results - Goodness of Fit (I)

Targets	Model	Data
Share retakers	0.273326	0.228102
Share dropouts	0.0811526	0.0497383
Share dropouts females	0.0797434	0.0427122
Share dropouts low-income	0.0724375	0.0474016
Share reapplicants	0.226216	0.116105
Share program switching	0.0568454	0.0554627
Share broad major switching	0.0137196	0.0161918
Share major switching	0.0478247	0.028331
Share university switching	0.0298804	0.0254052
Share college type switching	0.0146598	0.0139202
Share dropout end of first period	0.0563505	0.0233881
Share enrolls first in second period	0.0879897	0.359744
Share first year in second period	0.145701	0.150578
Share second year in second period	0.212934	0.26899
Share true pref. is top reported in ROL 1	0.419464	0.420469
Share true pref. is top reported in ROL 2	0.538836	0.47449
Share of ROLs of length 10 (first year)	0.118162	0.0632714
Share of ROLs of length 10 (second year)	0.113908	0.0652281
Share of students that apply (first year)	0.493095	0.513829
Share of students that apply (second year)	0.242447	0.184433
Share of students that change top true pref.	0.383797	0.653061
Share re-apps from top-reported prefs	0.600709	0.250274
Share re-apps from top-true prefs	0.22289	0.0668449
Mean tuition of top-reported prefs	3.60992	3.98598
Mean tuition of top-true prefs	3.73253	4.08897
Var tuition of top-reported prefs	0.803022	0.939992
Var tuition of top-true prefs	0.826379	1.00018
Mean distance of top-reported prefs	6.07096	13.206
Mean distance of top-true prefs	5.67945	12.4723
Var distance of top-reported prefs	26.6432	122.853
Var distance of top-true prefs	25.0246	136.235
Mean relative position of top-reported prefs	-0.541089	0.0308099
Mean relative position of top-true prefs	-2.45742	-1.6442e-18
Var relative position of top-reported prefs	6.86152	0.912193
Var relative position of top-true prefs	12.6331	0.934617
Mean average share math types ROL (year 1)	0.3612 0.6388	0.3935 0.6065
Mean average share math types ROL (year 2)	0.4091 0.5909	0.4667 0.5333
Mean norm diff broad major shares from outside option	0.0830831	0.336654
Mean norm diff broad major shares from outside option	0.180933	0.416953
Mean norm diff math types shares	0.348243	0.288687
Mean norm diff math types shares from outside option	0.161596	0.268223
Mean average dummy math types (year 1)	0.5040 0.7733	0.1982 0.8018
Mean average dummy math types (year 2)	0.5473 0.7155	0.2481 0.7519
Mean average dummy math types (year 1, females)	0.5655 0.7342	0.2411 0.7589

Notes: the order of majors is Social Sciences, Science, Education and Humanities, and Health.
Notes: the order of colleges is CRUCH-Public, CRUCH-Private, and Non-CRUCH.

Table E.4: Estimation Results - Goodness of Fit (II)

Targets	Model	Data
Mean corr norm broad majors grades (year 1)	-0.119489	-0.129867
Mean corr norm majors grades (year 1)	-0.216001	-0.161541
Mean corr norm math types grades (year 1)	-0.0888692	-0.0474172
Mean share top true broad majors changed	0.0555747	0.183673
Mean share top true majors changed	0.23003	0.265306
Mean share top true math types changed	0.14677	0.290816
Mean share top true prefs changed from outside option	0.0791454	0.301887
Mean share top true broad majors changed from outside option	0.0220033	0.188679
Mean share top true majors changed from outside option	0.0928286	0.226415
Mean share top true math types changed from outside option	0.0791454	0.301887
Mean share Math-type switchings [%]	0.0288454	0.0199898
Mean coeffs dropout 1	-0.1055 0.0166 -0.0304 0.8819 0.7950 1.0101 0.8303 0.0129 0.0125 -0.1228	-0.0337 0.0045 -0.0144 0.2980 0.3073 0.3136 0.2936 -0.0430 0.0107 0.0045
Mean coeffs switch program 1	0.0649 -0.0582 -0.0020 1.0564 0.9232 1.2043 0.8885 -0.1537 -0.0358 0.0189	0.0176 -0.0591 0.0101 0.6642 0.6696 0.6742 0.6821 -0.1098 -0.0810 0.0741
Mean coeffs switch broad major 1	0.0277 -0.0089 0.0162 0.4648 0.4421 0.5176 0.4515 -0.0420 -0.2870 0.0516	0.0089 -0.0120 0.0125 0.3792 0.4018 0.4028 0.3993 -0.0529 -0.1439 0.0237
Mean coeffs switch major 1	0.0654 -0.0327 0.0115 1.0131 0.8750 1.1491 0.7592 -0.1478 -0.0757 0.0321	0.0146 -0.0176 0.0127 0.4953 0.4933 0.5002 0.4875 -0.0739 -0.1331 0.0295
Mean coeffs switch math type 1	0.0362 -0.0144 0.0156 0.6614 0.4650 0.6956 0.5149 -0.0899 -0.0544 0.0064	-2.1948e-03 -5.0483e-05 5.5030e-03 2.2753e-01 2.1598e-01 2.1793e-01 2.5755e-01 -3.2688e-02 -4.2368e-02 -3.9812e-03
Mean tuition of top reported pref low income (year 1)	3.3024	3.68845
Mean tuition of top true pref low income (year 1)	3.45603	3.85171
Mean tuition of top reported pref above median (year 1)	3.67755	4.0299
Mean tuition of top true pref above median (year 1)	3.79301	4.14843
Mean observed ability scores of top reported pref (year 1)	0.978433	1.03799
Mean observed ability scores of top true pref (year 1)	0.975708	1.19079
Var. observed ability scores of top reported pref (year 1)	0.472234	0.554518
Var. observed ability scores of top true pref (year 1)	0.480734	0.785768
Mean observed ability scores program of top reported pref (year 1)	1.17649	1.02176
Mean observed ability scores program of top true pref (year 1)	1.63038	1.19079
Var. observed ability scores program of top reported pref (year 1)	0.315409	0.360733
Var. observed ability scores program of top true pref (year 1)	0.251461	0.262369
Mean share apply top reported with prob zero	0.103322	0.460174

Notes: the order of majors is Social Sciences, Science, Education and Humanities, and Health.
Notes: the order of colleges is CRUCH-Public, CRUCH-Private, and Non-CRUCH.

Table E.5: Estimation Results - Goodness of Fit (III)

	Moments: Evolution of Scores	
Targets	Model	Data
Mean scores evolution lang	0.039146	0.0406234
Mean scores evolution math	0.0244741	0.0409032
vars scores evolution lang	0.000396494	0.0107354
vars scores evolution math	0.0004052	0.0104524
Mean scores evolution hist nozero	0.0642348	0.0477539
vars scores evolution hist nozero	0.000378637	0.0104401
Mean scores evolution cien nozero	0.0629765	0.0658743
vars scores evolution cien nozero	0.000376386	0.0144259
Mean scores evolution hist zero	0.0728078	0.0863868
vars scores evolution hist zero	0.000371135	0.0159172
Mean scores evolution cien zero	0.0481128	0.0302498
vars scores evolution cien zero	0.000389301	0.017558
	Moments: Market Shares and Shares Within ROL	
Shares broad majors within ROL (year 1)	0.6730 0.1264 0.0713 0.0547 0.0746	0.6448 0.1364 0.1322 0.0311 0.0556
Shares broad majors within ROL (year 2)	0.2383 0.2991 0.0545 0.4082	0.3510 0.2545 0.1207 0.2671
Norm difference on broad major shares	0.16993	0.385261
Dummies broad majors within ROL (year 1)	0.4195 0.2448 0.1932 0.2779	0.4967 0.4430 0.1752 0.2784
Dummies broad majors within ROL (year 2)	0.3975 0.2183 0.1795 0.3028	0.4718 0.3798 0.2051 0.3361
Dummies broad majors within ROL (year 1, women)	0.4018 0.1517 0.2580 0.3157	0.5017 0.3094 0.2165 0.3795
Market shares by major (year 1)	0.6730 0.0678 0.0018 0.0236 0.0269 0.0320 0.0029 0.0335 0.0212 0.0746 0.0425	0.6406 0.0496 0.0105 0.0266 0.0245 0.0385 0.0217 0.0228 0.0083 0.0556 0.0972
Market shares by major (year 2)	0.6414 0.0644 0.0025 0.0252 0.0309 0.0411 0.0058 0.0303 0.0280 0.0891 0.0413	0.5807 0.0618 0.0121 0.0299 0.0230 0.0488 0.0243 0.0346 0.0089 0.0716 0.1013
Market shares by major (year 1, women)	0.6525 0.0745 0.0003 0.0188 0.0130 0.0347 0.0030 0.0476 0.0330 0.0913 0.0315	0.6574 0.0425 0.0111 0.0322 0.0226 0.0451 0.0208 0.0298 0.0102 0.0770 0.0470
Market shares by major (year 2, women)	0.6002 0.0720 0.0011 0.0224 0.0178 0.0436 0.0058 0.0456 0.0425 0.1178 0.0312	0.5814 0.0539 0.0133 0.0367 0.0218 0.0581 0.0233 0.0461 0.0106 0.1007 0.0509
	Auxiliary Model: Grade Equation 1	
Observed ability	0.187119	0.399753
Top-reported preference	0.0216734	0.0566592
Female	0.261718	0.180387
Broad Majors	4.2946 3.5410 5.1743 3.6789	4.0334 3.7081 4.2512 4.1799
Broad major share	0.421916	0.153988
College share	0.0191326	0.0717304
$\hat{\sigma}^2_{g1}$	2.60998	0.656285
	Auxiliary Model: Grade Equation 2	
Observed ability	0.118934	0.347962
Top-reported preference	-0.0321065	0.0520348
Female	0.236729	0.203915
Broad Majors	4.3463 2.9953 5.8516 3.1641	2.2540 1.8514 2.4193 2.4782
Second year student	4.74925	2.46812
Broad major share	0.500815	-0.293413
College share	0.0906893	-0.347303
$\hat{\sigma}^2_{g2}$	6.46581	1.4013
	Auxiliary Model: Time Series for Grades	
No switchers - constant	-0.337025	1.44215
No switchers - slope	1.73014	0.68989
Switchers - constant	2.80692	0.594362
Switchers - slope	0.479655	0.87654
	Auxiliary Model: Wage Equation	
Majors	-0.4977 0.5248 0.4445 -1.0565	0.0425 0.1793 -0.3449 0.2991
Grades	0.0430009	0.0141564
Observed ability college	0.0649156	0.0404434
Women	0.365717	0.149161
Standard error	0.203282	0.0594395
	Auxiliary Model: Wage Growth Equation	
Wage growth broad major dummies	0.9075 0.9261 0.8489 0.3161	-0.4839 -0.4124 -0.7613 -0.6654
Wage growth broad major-specific linear	0.1469 0.2276 0.4147 0.8361	0.0730 0.1543 0.1034 0.1269
Wage growth broad major-specific quadratic	-0.0624 -0.0227 -0.0613 -0.2478	0.0004 -0.0072 -0.0063 -0.0080
Wage growth standard error	0.000466272	0.109634
	Auxiliary Model: Non-Pecuniary Utility Equation	
Top-reported preference	0.0226872	0.0430703
Observed ability	0.0544531	0.164618
Observed ability college	-0.0753146	-0.0149746
Major dummies	0.6216 0.6419 0.6145 0.6597	0.4006 0.2304 0.3188 0.5111
Major-specific - women	0.0585 0.0331 0.0795 0.0422	0.0306 0.0395 0.1325 -0.0324
Broad major share	-0.0306212	0.0234728
Standard error	0.19072	0.220015

Notes: the order of majors is Social Sciences, Science, Education and Humanities, and Health.
Notes: the order of colleges is CRUCH-Public, CRUCH-Private, and Non-CRUCH.

Algorithm 3 Computing $\hat{p}(\tau)$

Input. Structural parameter estimates $\hat{\theta}$, first-stage estimates $\hat{p}$, $\hat{P}^e$, $\hat{P}^d$, $\hat{P}^g$, and $\hat{P}^w$, and tolerance level ϵ_{tol}.
Output. Rational Expectations equilibrium cutoff distributions $\hat{p}(\tau)$
Step 1. For each program j
Step 1.a. Solve the model and simulate outcomes given the rules implied by counterfactual τ and the estimated objects $\left(\hat{\theta}, \hat{p}, \hat{P}^e, \hat{P}^d, \hat{P}^g, \hat{P}^w\right)$
Step 1.b. Obtain a set of simulated ROLs and scores $(R_1^0, R_2^0, s_1^0, s_2^0)$
Step 1.c. For each program j, estimate the mean of its cutoff distribution and compute the average over programs, $\hat{\delta}^0 \equiv \sum_{j \in M} \frac{\hat{\mu}_j^0}{M}$
Step 2. $\delta_{diff} = 2\epsilon_{tol}$, $k = 1$, $\rho = 0.9$
Step 3. While $\delta_{diff} > \epsilon_{tol}$
Step 3.a. For each student i, solve the model via Backward Induction given τ, the estimated parameters $\left(\hat{\theta}, \hat{P}^e, \hat{P}^d, \hat{P}^g, \hat{P}^w\right)$, and cutoff distributions $\hat{p}^{k-1}$, and obtain the continuation values for each student and state
Step 3.b. Forward simulate first period ROL R_{i1}^k given τ, the estimated parameters $\left(\hat{\theta}, \hat{P}^e, \hat{P}^d, \hat{P}^g, \hat{P}^w\right)$, cutoff distributions $\hat{p}^{k-1}$, and continuation values
Step 3.c. For each program j, estimate the mean of its cutoff distribution and compute the average over programs, $\hat{\delta}^0 \equiv \sum_{j \in M} \frac{\hat{\mu}_j^0}{M}$
Step 3.d. Given initial first period applications R_1^k, second period applications R_2^{k-1}, and students' scores s_1^k, and s_2^{k-1}, run the Chilean matching mechanism and obtain an allocation $\mu^k\left(R_1^k, R_2^{k-1}, s_1^k, s_2^{k-1}\right)$
Step 3.e. Given $\mu^k\left(R_1^k, R_2^{k-1}, s_1^k, s_2^{k-1}\right)$, τ, the estimated parameters $\left(\hat{\theta}, \hat{P}^e, \hat{P}^d, \hat{P}^g, \hat{P}^w\right)$, cutoff distributions $\hat{p}^{k-1}$, and continuation values, forward simulate second period ROLs R_{i2}^k
Step 3.f. Given $\left(R_1^k, R_2^{k-1}, s_1^k, s_2^{k-1}\right)$, run the boostrap procedure and estimate the Rational Expectations cutoffs distributions $\tilde{p}^k$ Take a convex combination of the realized cutoffs $\tilde{p}^k$ and $\hat{p}^{k-1}$ (point wise), i.e, $\hat{p}^k = \rho^k \hat{p}^{k-1} + \left(1 - \rho^k\right) \tilde{p}^k$
Step 3.g. For each program j, estimate the mean of its cutoff distribution and compute the average over programs, $\hat{\delta}^k \equiv \sum_{j \in M} \frac{\hat{\mu}_j^k}{M}$
Step 3.h. Update cutoff distributions with proportional updating, i.e, $\hat{p}^k = \frac{\hat{\delta}^k}{\hat{\delta}^0} \hat{p}^0$
Step 3.i. Compute $\delta_{diff} = ||\hat{\delta}^k - \hat{\delta}^{k-1}||$ $\hat{p}(\tau) = \hat{p}^{k-1}$ $k++$

Algorithm 4 Constrained Deferred Acceptance with signal and bonus ψ

Input. Indirect utilities v, application scores s, cutoff distributions P, and application score bonus ψ

Output. Optimal ROL $R\left(v, s, P, \psi_{\tau}\right)$

Step 1. For each program j

Step 1.a. Compute admission probabilities given cutoff distributions P and application scores $\tilde{s}(j) = \{s_1, ...s_{j-1}, \psi_{\tau} s_j, s_{j+1}, ...s_J\}$

Step 1.b. Compute and store optimal ROL $R\left(v, \tilde{p}(j)\right)$ using MIA

Step 2. Compute optimal signal

$$s_j^* = \underset{j}{\operatorname{argmax}} \left\{R(v, \tilde{p}_j)\right\}$$

Step 3. Compute optimal ROL $R(v, \tilde{p}_j)$