Submitted to *Econometrica*

DYNAMIC COLLEGE ADMISSIONS

TOMAS LARROUCAU
Department of Economics, Arizona State University

IGNACIO RIOS
School of Management, The University of Texas at Dallas

We study the relevance of incorporating dynamic incentives and eliciting private information about students' preferences to improve their welfare and downstream outcomes in centralized assignment mechanisms. Using administrative data and two nationwide surveys, we identify two behavioral channels that largely explain students' dynamic decisions: (i) initial mismatches and (ii) learning. 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 preferences and abilities over time and reapply to the system. We use the estimated model to analyze the impact of changing the information environment and the assignment mechanism and reapplication rules on the efficiency of the system. Our counterfactual results show that policies that front-load learning can help to reduce switches, while providing score bonuses that elicit information on students' cardinal preferences and leverage dynamic incentives can significantly increase retention and students' overall welfare.

KEYWORDS: college admissions, dynamic matching, college retention.

Tomas Larroucau: Tomas.Larroucau@asu.edu
Ignacio Rios: iar200000@utdallas.com

We are very grateful for the guidance and helpful comments from our advisors and committee, Eduardo Azevedo, Hanming Fang, Margaux Luflade, Christopher Neilson, Rakesh Vohra, and Paulo Somaini; and seminar participants at UPenn, Western Ontario, TSE, NYU, ASU, UCH, PUC, Exeter, PSU, Cowles, Caltech, Duke, Wisconsin, and Columbia. This paper would not have been possible without 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, Alejandra Venegas, José Piquer, and Giorgio Parra for their collaboration and support. Any errors or omissions are our own.

1. INTRODUCTION

According to Kapor et al. (2024), at least 46 countries—including Turkey, Taiwan, Tunisia, Hungary, and Chile—use centralized systems to organize college admissions. Although extensive literature analyzes the pros and cons of different mechanisms used to perform the allocation, their effects on downstream policy-relevant outcomes is unclear. For instance, only 40% of full-time bachelor's students graduate on time in developed countries (OECD 2019), and in Chile, just 16% do so, with 30% switching programs and another 30% dropping out. These patterns illustrate potential inefficiencies in the assignment process, suggesting that better-designed assignment mechanisms and information environments could substantially improve downstream educational outcomes.

Improving outcomes requires accounting for features that characterize real-life applications and that are generally overlooked in the literature. One such feature is that matching markets are typically dynamic. For instance, students can learn over time about their preferences and abilities, and this may influence their decisions to reapply, switch to another program, or drop out. Another feature is that students may have private information that is not elicited in the admissions process and 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 (or the intensity of their) preferences, could affect their persistence in their programs and, thus, impact the system's yield. Therefore, designing admissions systems that consider the dynamic interplay between learning, decisions and elicit information about students' cardinal preferences can be critical for improving students' outcomes and the efficiency of the system.

In this paper, we study the design of matching markets where agents face dynamic considerations, learn about their abilities and preferences through experience, hold private information that affects outcomes, and may hold biased beliefs about their abilities and future earnings. We examine how changes to the information environment and the assignment mechanism affect students' welfare and downstream outcomes—including college performance, retention, and on-time graduation—by facilitating learning and incorporating dynamic incentives that leverage students' cardinal preferences.

By combining data from the Chilean college admissions system and multiple nationwide surveys that we designed and conducted, we show that two behavioral channels largely explain students' dynamic decisions. The first, called the *learning* channel, posits that students learn about their preferences and abilities during their college experience, affecting both their consumption value and expected returns, and potentially motivating them to switch programs or drop out to avoid ex post mismatches. The second, called the *initial mismatch* channel, suggests that students may initially enroll in less-preferred programs to improve their future outside options and later switch to more desirable ones.

These two channels may have different implications. On the one hand, if learning is limited and thus preferences are persistent over time, it may be desirable to discourage reapplications by forcing students to internalize the crowd-out externality they generate—that is, when a student is admitted to an over-demanded program, they may displace another student who would have otherwise secured that spot and is instead assigned to a less preferred option. If the first student later decides to switch programs, the seat they leave behind often goes unfilled, creating an externality that negatively affects others in the system. On the other hand, if learning explains most students' dynamic decisions, it may improve welfare to facilitate switching and avoid ex post mismatches. Hence, the welfare implications of limiting or encouraging switching are unclear. We illustrate this tension through a stylized example in Appendix A, which shows how both channels affect students' decisions and welfare, and how inefficiencies can be mitigated by eliciting cardinal preferences.

To evaluate the effects of these two channels, we introduce a structural model that captures the application behavior of students, as well as their decisions to enroll, retake the admission tests, reapply, switch, and drop out, allowing students to learn about their unknown abilities and their preferences during their academic progression. In particular, we assume that students base their application and enrollment decisions on both (i) the value of studying each program, (ii) the continuation value of retaking the admission tests and reapplying to the system, and (iii) their labor market prospects. Students have heterogeneous preferences—which evolve to capture persistent idiosyncratic shocks that occur each period—and potentially biased beliefs over their future college grades and labor market prospects—which they update as they observe noisy signals of their unknown ability from their grades. Based on their updated continuation values, students decide whether to continue in their current program, reapply to the system, or drop out and choose their outside option. Finally, students face graduation probabilities, and then enter the labor force and receive pecuniary and non-pecuniary values from the labor market.

The main challenge in estimating our model is separately identifying the *learning*—which includes learning about their abilities from grades and their preferences about their major of enrollment—and the *mismatch* channels. To identify the former, we leverage the longitudinal structure of the data, focusing on how students update their beliefs and change their preferences over time. We isolate learning about preferences by examining the composition of initial applications and reapplications following non-informative grade signals, where any preference change likely reflects evolving tastes rather than updated beliefs about ability. In contrast, we identify learning about unknown abilities by using elicited beliefs about academic performance (e.g., expected grades) and their correlation with (i) changes in the composition of preferences and (ii) decisions (e.g., switching or dropping out) after receiving first-year grades. The correlation between these belief updates and switching behavior and their heterogeneity across majors and subjects help us distinguish between belief-based learning and preference shifts. To identify the initial mismatch channel, we exploit quasi-experimental variation around admission cutoffs, which allows us to compare students who are initially assigned to different programs despite similar application behavior and observed characteristics. These initial differences in assignment—combined with subsequent outcomes such as switching, dropout, or reapplication—provide clean variation to identify the consequences of mismatch separately from endogenous learning dynamics. For instance, we show that there is a strong causal effect of not being assigned to the top-reported preference on the probabilities of reapplying and switching (63% and 51% increase, resp.), supporting the existence of the mismatch channel.

After estimating the structural model, we conduct a series of counterfactual analyses. First, we find that learning accounts for about one-third of all switching decisions, while initial mismatches and congestion explain the remainder and part of dropout decisions.

Next, we examine how changes in (i) the information environment and (ii) the assignment process—through changes in the reapplication rules or in the mechanism—affect students' outcomes. On the information side, we find that enabling students to front-load their learning process (through counseling during high-school or a mandatory one-year post-secondary program) can help reducing switching rates. On the mechanism side, we find that penalizing students who switch (as in Turkey); giving a score bonus for all first-year applicants (as in Finland); or allowing students to signal one of their preferences to get a score bonus (as in some job markets) can significantly reduce switching rates and increase students' welfare. Finally, we find that these effects are robust to changes in the share of strategic participants, unlike policies such as restricting application list lengths, which are widespread but far less effective and sometimes detrimental for students' outcomes.

Our counterfactual experiments highlight the importance of carefully balancing the two key behavioral forces at play: the benefits of learning through experimentation and the crowd-out externality created by initial mismatches. We show that *facilitating learning*, incorporating dynamic incentives, and eliciting students' cardinal preferences—through modifications to reapplication rules and the assignment mechanism—can meaningfully influence both educational outcomes and overall welfare. More broadly, these findings underscore the relevance of accounting for dynamic behavior and externalities in the design of matching markets, with potential implications for improving efficiency and equity in contexts such as school choice, organ allocation, and entry-level labor markets.

The remainder of the paper is organized as follows. Section 2 discusses the most closely related literature. Section 3 describes the Chilean college admissions system and provides empirical evidence for the two behavioral channels. Section 4 presents our model. Section 5 describes our identification strategy. Section 6 describes the estimation approach and its results. Section 7 reports our counterfactual results, and Section 8 concludes.

2. LITERATURE

Our paper combines two strands of the literature: (i) the empirical analysis of assignment mechanisms and (ii) the empirical analysis of college choices under uncertainty.

The first strand focuses on understanding the incentives that centralized assignment mechanisms introduce; how to use the data generated by these mechanisms to identify and estimate students' preferences/beliefs (Agarwal and Somaini, 2019); and measuring the welfare effects of changing assignment mechanisms in different settings (Agarwal and Somaini, 2018, Calsamiglia et al., 2020, He, 2012, Kapor et al., 2020, Luflade, 2017, Hernandez-Chanto, 2021).

Despite progress in this area, the aforementioned studies either consider static settings, assume fixed preferences over time, or simply ignore the effects of assignment mechanisms on downstream outcomes.[1] In response, a recent line of research has begun to examine dynamic mechanisms or settings in which agents interact repeatedly with the assignment process, including the allocation of public housing (Waldinger, 2021), hunting licenses (Reeling and Verdier, 2021), general practitioners (Huitfeldt et al., 2025), teachers' labor markets (Ederer, 2023), and kidneys from deceased donors (Agarwal et al., 2024a,b). In educational markets, dynamic models have been recently used to evaluate the impact of middle school placements on high school outcomes (Hahm and Park, 2022), daycare assignments (Kuno, 2023, De Groote and Rho, 2024), and sequential college admissions (De Groote et al., 2025). Among these, the closest to ours is (Narita, 2018), who analyzes theoretically and empirically the welfare effects of centralized school-choice mechanisms with rematching when demand evolves over time. We contribute to this literature by developing a dynamic framework for centralized college admissions that captures how different forms of learning, strategic behavior, and repeated interactions with the mechanism shape system yield, student outcomes, and welfare.

The second strand studies education and occupation choices, focusing on how students gather information and learn about their preferences, abilities and future earnings (Arcidiacono, 2005, Arcidiacono et al., 2025, Malamud, 2011) and how their choices impact outcomes and returns (Altonji et al., 2012, 2016). Within this strand, the closest paper to ours is Bordon and Fu (2015), who analyze the effects of switching from a system where stu-

[1] With some exceptions in school reforms (Tanaka et al., 2020), college admissions (Otero et al., 2021), kidney allocation (Agarwal et al., 2024a), and the medical match (Friedrich et al., 2024).

dents choose both college and major simultaneously to one where they choose college first and then major. Our paper complements their work by studying the effects of changing the assignment mechanism and reapplication rules in a dynamic context, allowing for heterogeneous and potentially biased subjective beliefs—identified with survey data—rather than imposing rational expectations, identifying their effect on students' learning about their abilities and their preferences, and measuring its effect on student outcomes.

Finally, within these two strands of the literature, we also contribute to recent papers exploiting survey data for identification of structural models in college admissions (Wiswall and Zafar, 2015a, 2021, Tincani et al., 2023) and school-choice (Kapor et al., 2020). We add to this literature by leveraging students' subjective and potentially biased beliefs over their college outcomes and exploiting this information to identify their learning process.

To the best of our knowledge, this is the first paper to structurally evaluate how the information environment and market design choices affect students' welfare and outcomes—such as achievement, retention, and on-time graduation—beyond the initial match. We also contribute by revisiting the trade-off between eliciting preference intensity and ensuring *strategy-proofness* in a dynamic framework that incorporates private information, learning about abilities and preferences, and heterogeneous, subjective, and potentially biased beliefs about ability and future outcomes. Accounting for these elements is essential to understand how changes in the assignment mechanism influence students' indirect utilities and welfare, and to evaluate intrinsically dynamic policies—such as reapplication rules—designed to improve both student outcomes and the system's yield.

3. COLLEGE ADMISSIONS IN CHILE

The college admissions process in Chile is semicentralized, with the most selective universities having a centralized system and the remaining institutions conducting their admissions processes independently. This paper's empirical application follows the cohort of 2014 and 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 admissions process is operated by the *Departamento de Evaluación, Medición y Registro Educacional* (DEMRE).

To apply to any of the close to 1,500 academic programs offered 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, and provide a score for each. Students' performance during high school yields two additional scores: one obtained from the average grade during high school (NEM) and a second that depends on the relative position of the student among their cohort (Rank). Admission to any program is solely based on these *admission factors*.

After scores are published, students can submit—at no cost—a list with no more than ten programs, ranked in strict order of preference. We refer to these lists as Rank Order Lists (ROLs). Notice that students directly apply to a program, i.e., they must list university-major pairs on their ROL. In the remainder of the paper, we refer to these pairs as programs.

On the other side of the market, each program announces its vacancies, the weights on each admission factor, 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 PSU tests, among other less common requirements. Each program's preference list is defined by first filtering all applicants who do not meet these requirements. Students are then ordered based on their application scores, which are computed as the weighted sum of the applicants' scores and the predefined weights.

Considering the vacancies and the preferences of the applicants and programs, DEMRE runs an algorithm to match students to programs. The mechanism used is a variant of the student-proposing Deferred Acceptance algorithm, in which all students tied for the last seat in a program are admitted (Rios et al., 2021). A description of the mechanism can be found in Appendix B.1. 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, and all students with scores below the cutoff are waitlisted and thus may have to enroll in a lower-ranked preference. This property is known as the *cutoff structure*.

Enrollment begins with the released of assignment results and unfolds in two rounds. In the first, only assigned students may enroll in their designated program. In the second, programs with remaining seats invite waitlisted students to enroll. Throughout the process, applicants may also choose to enroll in non-centralized programs or enter the workforce.

Students may retake the admissions test and reapply to the centralized system in subsequent years. However, retaking the test is not required for those who last took it in the previous year. For students who take the test in two consecutive years, the system allows them to choose the better of the two score pools—i.e., they cannot combine individual test scores across years—for each program to which they apply. Although some students switch programs without formally reapplying, the vast majority of switches—approximately 70%—occur through reapplication. Therefore, in the remainder of the paper, we focus on switches that follow a reapplication and abstract from internal transfers, capturing exactly how the centralized system works. Finally, seats that students vacate when they switch are not reallocated to other students nor included in the offer of the next year, so there may be programs that are under-utilized ex-post despite of being oversubscribed ex-ante.

3.1. *Data*

Our dataset includes (i) administrative data provided by DEMRE and the Ministry of Education (MINEDUC), (ii) several surveys designed and conducted in collaboration with CRUCH and DEMRE aiming to elicit students' preferences and beliefs about admissions probabilities, and (iii) grade records facilitated by CRUCH. Specifically:

- Admissions process: This includes students' socioeconomic characteristics—including self-reported family income, parents' education, and the municipality in which the student lives, among others—; scores; applications; final assignment; and enrollment decisions, spanning from 2007 to 2023. In addition, this includes data on programs characteristics, including their vacancies, weights, tuition, duration, major, location, etc.
- Labor market: This includes aggregate information about the labor market outcomes in each program—including average wages four years after graduation and overall employment probabilities one year after graduation—spanning from 2014 to 2018. We also have data at the major level, including average wage for the first five years after graduation; employment probabilities for the first and second year after graduation; and the evolution of average wages from the first ten years after graduation.

- Grades: This includes the cumulative GPA for their first three years of college for every student who enrolls in a program that is part of the centralized system in 2014 and 2015.
- Surveys post-application: This includes responses to surveys that we designed and conducted in 2019 and 2020 to gather information on students' preferences—specifically, about their top-true preference—and their beliefs about their admissions probabilities for each program on their ROL and the elicited top-true preference. As many students reapply to the centralized system after a year, these two surveys provide a panel for reapplicants' preferences and their beliefs (see Appendix B.2 for more details).
- Surveys pre and post-application: To complement the previous surveys, in 2023, we designed and conducted similar surveys before and after the application time window. In the latter, we ask students about their beliefs (about admission chances, grades, wages, among others) for programs both included in their ROL and also outside it, allowing us to capture beliefs on choices and also on counterfactual programs.

To our knowledge, this is the first paper to use the grades data described above and to incorporate nationwide-survey data on students' true preferences, beliefs about admission probabilities, college progression, and future labor-market outcomes.

Throughout the paper, we focus on a subset of the population to reduce computational complexity. Specifically, we focus on students who graduated from a high school within the Region Metropolitana (RM) in 2013, participated in the 2014 admissions process (i.e., took the PSU tests), had an average score between Math and Language above 475, and applied to programs located in RM. The latter allows us to reduce the number of programs to less than half (435) without major loss, as close to 80% of applications from students living in RM include only programs located in that region. Finally, we exclude students with average score below 475 (less than 13% of students who can apply) because they do not satisfy loan eligibility requirements. Contrary to previous evidence (Lochner and Monge-Naranjo, 2012, 2016, Solis, 2017), Card and Solis (2020) show that enrollment decisions and college completion rates among students whose average score exceeds 475 (thus, eligible for scholarships) are not highly sensitive to differences in prices. Hence, we can ignore potential effects of credit constraints on students' dynamic choices.

3.2. *Empirical Facts*

We posit that students' dynamic decisions are largely explained by two behavioral channels: (i) *initial-mismatching*, whereby students assigned to less preferred programs reapply to improve their allocation, and (ii) *learning*, whereby students learn about their match-quality and preferences over time and potentially decide to move to other programs. In this section, we provide empirical evidence that supports the existence of these two channels.

Mismatching. In Table I, we report the average switching and dropout rates in the first four years, separating by income level—high or low—and gender.[2] First, we observe that close to 21.2% of students switch from their first program of enrollment, and 15.3% drop out within the first four years. Second, comparing outcomes by gender (within an income level), we observe that women are more persistent in their academic progression, since 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 less likely to switch programs during their academic progression. However, we also observe that low-income students are significantly more likely to drop out. These patterns—suggesting large differences in switching and dropout rates by gender and income—are similar to those observed for first-year outcomes. Therefore, for simplicity, we restrict our analysis to switching and dropout within the first year. This choice is without major loss of generality, since close to 80% of switching takes place in the first two years, and roughly 2/3 of those happen in the first year (see Figure B.4 in Appendix B.4).

To evaluate the impact of assignment rank on student outcomes, Figure 1 presents first-year switching and dropout rates by preference of assignment. We observe that students assigned to lower reported preferences switch at higher rates compared with students assigned to their top reported preference. For instance, among students assigned to their top reported preference, 9.86% switch programs at the end of their first year, compared with

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

TABLE I

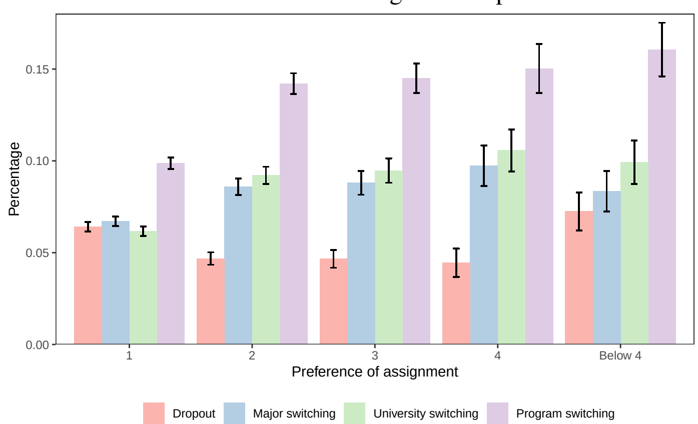
SWITCHING AND DROPOUT BY GENDER AND INCOME

		Switches				
	Income	Program	University	Major	Math type	Dropout
Men	Low	0.213	0.108	0.099	0.038	0.232
		(0.007)	(0.005)	(0.005)	(0.003)	(0.007)
	High	0.24	0.128	0.132	0.044	0.134
		(0.006)	(0.004)	(0.005)	(0.003)	(0.005)
Women	Low	0.173	0.089	0.095	0.042	0.186
		(0.006)	(0.005)	(0.005)	(0.003)	(0.007)
	High	0.209	0.102	0.125	0.063	0.096
		(0.006)	(0.004)	(0.005)	(0.003)	(0.004)
Overall		0.212	0.109	0.116	0.048	0.153
		(0.003)	(0.002)	(0.002)	(0.002)	(0.003)

Note: Standard errors reported in parenthesis. Program denotes any change in the university–major pair; University, Major, and Math-type switches are not mutually exclusive.

almost 15% of those assigned to their fourth choice. These results suggest a strong correlation between assignment rank and switching behavior.

FIGURE 1.—Switchings and dropout

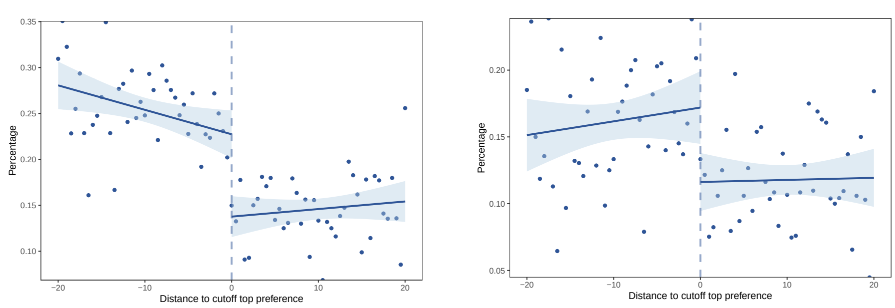


Note: Switching categories do not include stop out.

To make a causal claim, we use a regression discontinuity design that exploits the algorithm's cutoff structure to perform the allocation. 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. In Figure 2, we display binned means of different outcomes as a function of the distance between cutoffs and students' scores in their most preferred listed program. Figure 2a shows that students right below the cutoff are 9% more likely to reapply the following year, which corresponds to a relative change of close to 62.8% compared to those right above the cutoff. Figure 2b shows that students below the cutoff are 5.8% more likely to switch programs within the centralized system, which corresponds to a relative change of more than 50.9%. Similar results hold when we consider students' top true preferences instead of their top-reported preferences, as shown in Appendix B.3.1. These results confirm our previous findings, i.e., that students assigned to lower preferences are more likely to reapply and switch programs the following year. In Appendix B.3, we provide more details and additional evidence of the effect of crossing the cutoff of the top-reported preference on students' decisions. Importantly, we find no causal effect on enrollment, and, contrary to what Figure 1 suggests, we find no causal effect of cutoff crossing on dropout.

FIGURE 2.—Effect of Cutoff Crossing

(a) Reapplication (b) Switching



The previous empirical facts show a causal effect of the preference of assignment on students' persistence in their initial assignment. To show that the mismatch channel partially explains this, we use the 2020 survey on students' preferences and beliefs, in which we find that a significant fraction of students know, before enrolling in their assigned programs, that they will be less likely to remain in that program if they are assigned to a lower reported preference (see the details in Appendix B.4.1). These results cannot solely be explained by students' or programs' characteristics and, consequently, we consider them as evidence of match-effects (on students beliefs) about their future college progression.

Learning. Students' preferences and beliefs may change during their first year in college, which could affect their decision to switch. We analyze students' switching decisions and classify them into three categories: (i) *Up*, (ii) *Down*, and (iii) *Out*. Students move *Up* (*Down*) if they switch to a program listed above (below) their initial enrollment on their initial ROL. Students move *Out* if they switch to a program not listed on their initial ROL.

We find that, among students who switch in their first year, 17.3% move *Down*, 14.2% move *Up*, and 67.1% move *Out*. The remaining 1.4% of switchers were initially enrolled in programs outside their original ROL (through special admission channels) and subsequently switched. Moreover, more than half of the latter switches involve more selective programs, i.e., programs with higher cutoffs compared to their initial enrollment. These results suggest that both channels explain students' switching significantly. Students who move *Down* or *Out* to less selective programs are likely to have learned about their (poor) ability (learning channel), and students' who move *Up* or *Out* to more selective programs may be trying to find a better match given their initial preferences.

In Table II, we analyze the effect of the signal received at the end of the first year—i.e., the difference between the obtained GPA and the expected one—on reapplication and switching decisions. In all of these models, we control for demographics (gender, income); scores (NEM and the average between Language and Math); and the preference of assignment in the initial year. Odd columns include the entire sample, while even columns restrict the sample to 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), the latter rules out forced switches.[3]

We observe that the magnitude of the signal is negatively correlated with students' likelihood of reapplying, modifying the composition of their rank-ordered list, and switching programs. Moreover, the signal is negatively correlated with the probability of switching

[3]Our measure of GPA is aggregated, so it may not capture forced switches due to a single course (e.g., failing a mandatory class multiple times). However, students with high GPA who are forced to quit are uncommon.

TABLE II

EFFECT OF SIGNAL ON REAPPLICATION AND SWITCHING

	Reapplied		Changed comp.		Switch		Switch Down		Switch Up	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Signal	-0.980^{***}	-0.092	-0.850^{***}	0.077	-0.790^{***}	0.226^{***}	-0.883^{***}	-0.705^{***}	0.753^{***}	0.997^{***}
	(0.033)	(0.061)	(0.035)	(0.073)	(0.031)	(0.057)	(0.054)	(0.144)	(0.088)	(0.107)
Mean	0.119	0.085	0.08	0.057	0.131	0.099	0.022	0.015	0.024	0.026
Test	0.11^{***}	0.086^{***}	0.074^{***}	0.074^{***}	0.124^{***}	0.093^{***}	0.019^{***}	0.017^{***}	0.019^{***}	0.018^{***}
Demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Passed ($GPA \geq 4.0$)	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	14,395	12,591	14,395	12,591	14,395	12,591	14,395	12,591	14,395	12,591

Note: Sample includes all students from the cohort that graduated high school in 2014, enrolled in 2014 and 2015 in a program offered by a university within the system, and for whom first-year grade data is available. GPA is measured on a 1–7 scale, with failing grades defined as below 4.0. Test reports the results of t-test assuming no signal (all p-values are below 1×10^{-4}), i.e., testing whether the intercept is different than zero. Significance levels: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

to less preferred programs ("switching down") and positively correlated with switching to more preferred programs ("switching up"). These patterns suggest that students who receive stronger signals—indicating they performed worst than expected and thus revise their beliefs about their unobserved ability downward—are less inclined to pursue more desirable programs and more inclined to downgrade their choices. These findings indicate that students respond to signals in ways consistent with learning about ability.

While signals about ability play a central role in shaping students' decisions, our findings also suggest that students may learn about their preferences through their college experience. To illustrate this, Table II reports the results of t-tests on the predicted probabilities of reapplying, modifying the rank-ordered list (ROL), and switching behaviors, conditional on receiving a null signal—i.e., when actual grades are as expected. We find that even in the absence of a signal, students remain significantly likely to reapply, modify the composition of their ROL, and switch programs. This pattern is further supported by the summary statistics in Table III, which show that students who receive a neutral signal still engage in these behaviors at meaningful rates—generally falling between those observed for students receiving positive or negative signals. For instance, 8.6% of students with a neutral signal reapply, compared to 19.9% of those with a positive signal and 7.6% of those with a negative signal. Similarly, 6.0% of students with a neutral signal change the composition of their ROL, versus 13.5% for positive and 5.3% for negative signals.

TABLE III

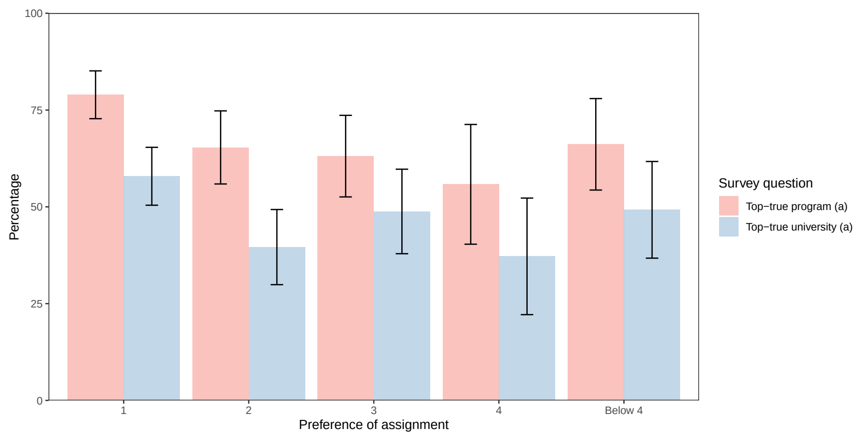
SWITCHING STATISTICS BY SIGNAL TYPE

		Reapply		Switch		Switch Down		Up
	N	All	Changed comp.	Program	Major	Program	Major	Program
Negative	5745	0.199	0.135	0.216	0.098	0.039	0.014	0.011
		(0.005)	(0.005)	(0.005)	(0.004)	(0.039)	(0.002)	(0.001)
Neutral	1665	0.086	0.060	0.100	0.043	0.019	0.007	0.016
		(0.007)	(0.006)	(0.007)	(0.005)	(0.019)	(0.002)	(0.003)
Positive	8325	0.076	0.053	0.105	0.041	0.012	0.003	0.033
		(0.003)	(0.002)	(0.003)	(0.002)	(0.012)	(0.001)	(0.002)

Note: Sample includes all students from the cohort that graduated high school in 2014, enrolled in 2014 in a program offered by a university within the system, and for whom first-year grade data is available. Positive / Negative signal considers students who received a sufficiently positive / negative signal, i.e., 0.1 points below or above expected, respectively. Neutral signal includes all students who did not receive a positive or negative signal.

Our previous results show that students' reported preferences may change during their first year in college. To rule out the possibility that these changes are driven by students' being more experienced in participating in the system (as opposed to them learning about their ability or preferences), we analyze changes in true preferences using the surveys conducted in 2019 and 2020. Specifically, we construct a panel that includes students who participated in both surveys and responded to the same questions (close to 1,300 students), and we compare the top true preference reported in each year. In Figure 3, we plot the fraction of students who changed their top true preference (for programs and also for universities) as a function of their initial preference of assignment (in 2019). First, we observe that on average, close to 65% of the students in the data changed their top true preference after their first year in college. Moreover, close to 50% of students even changed their most preferred university. Second, we observe that students initially assigned to lower preferences are less likely to change their top true preference for programs. This pattern is consistent with the existence of both the *mismatch* and *learning* channels: students assigned to lower-ranked programs are likely affected by initial mismatches and, thus, a higher proportion keeps their initial top preference. In contrast, students initially assigned to their top-reported preference have a lower probability of being *mismatched*; thus, their reapplication suggests that they learned about their ability during their first year in college.

FIGURE 3.—Percentage of reapplicants that change their top-true preference, by preference of assignment in 2019



4. MODEL

In this section, we present a dynamic model of students' application behavior and academic progression that incorporates how they learn about their abilities and preferences over time. The model adapts and extends the work from Arcidiacono (2004, 2005), Arcidiacono et al. (2025) to (i) capture the empirical patterns documented in the previous section; (ii) quantify the extent to which students' switching behavior is driven by initial mismatch versus subsequent learning; and (iii) assess how alternative learning environments and market design interventions affect students' outcomes in equilibrium.

Overview. We assume that students enter college with imperfect information about two key dimensions: (i) their match-specific abilities—as multidimensional skills relevant to labor market returns—and (ii) their preferences over majors of study. Abilities are partially revealed through college performance, while preferences evolve based on students' experience in their enrolled programs, independently of grades. This learning process shapes students' beliefs about future returns upon graduation and informs their decisions to continue in their current program, reapply to the centralized system, or drop out. Finally, upon graduation, students enter the labor market and receive utility that depends both on a pecuniary component corresponding to their lifetime labor market returns and a non-pecuniary one that captures the alignment between their program of graduation and their preferences.

Notation. We use $i \in \mathcal{I}$ to refer to students, $j \in \mathcal{J} \cup \{\emptyset\}$ to refer to programs ($\emptyset$ representing no enrollment), and $t \in \mathcal{T} = \{1, \ldots, T\}$ to refer to periods. In addition, let $\mathcal{M}$ be the set of majors and $m : \mathcal{J} \to \mathcal{M}$ be the function that maps programs to majors, i.e., $m(j)$ represents the major of program j. Finally, with a slight abuse of notation, we use functions to make explicit the underlying arguments and structure from which key elements of the model are derived (e.g., $G(\cdot)$ for grades); subscripts to denote their realized values (e.g., $G_{i,j,t}$); tilde notation to denote beliefs over these values (e.g., $\tilde{G}_{i,j,t}$); and hat notation to represent their estimated values (e.g., $\hat{G}_{i,j,t}$).

Organization. Section 4.1 describes the learning process during college. Section 4.2 specifies the flow utility from enrollment. Section 4.3 presents the workforce utility obtained upon graduation. Finally, Section 4.4 discusses implementation details, including timing, decisions, sources of heterogeneity, additional details considered in the estimation.

4.1. *Learning*

As previously discussed, we assume that students learn about (i) their preferences through their enrollment decisions, and (ii) their unknown abilities through their grades.

Learning about preferences. We allow students to learn about their preferences for majors. In particular, we assume that student i's idiosyncratic preferences for majors in period t are captured by a set of random coefficients $\{\alpha_{i,m,t}\}_{m \in \mathcal{M}}$, whose initial values are normally distributed with mean $\mu^{\alpha}_{i,m}$ and standard deviation $\sigma^{\alpha}_{i,m}$. Furthermore, we assume that students only learn about their preferences for their major of enrollment. Thus, student i's idiosyncratic preferences for major m, conditional on being enrolled in program j, evolves according to $\alpha_{i,m,t+1} = \alpha_{i,m,t} + \mathbb{1}_{\{m=m(j)\}} \cdot \varepsilon^{\alpha}_{i,m,t}$, where $\epsilon^{\alpha}_{i,m,t} \sim N\left(0, \left(\zeta^{\alpha}_{i,m}\right)^2\right)$. Finally, as we later discuss in Section 4.2, students observe preference shocks (related to their flow utility) that are not persistent over time and i.i.d. across periods, so we do not consider these as part of the (systematic) learning process.

Learning about abilities. Beyond learning about their preferences, we allow students to learn about their abilities. Specifically, we assume that the ability $A_{i,j}$ of student i in program j corresponds to the sum of three elements: (i) $A^o_{i,j}$, an observable component that is known by the student and the econometrician; (ii) $A^p_{i,j}$, a private component only known by the student; and (iii) A^u_{ij}, unknown by both the student and the econometrician but learned from the grades obtained during college. More specifically, let $\mathcal{S}$ be the set of subjects that are mandatory for all students in the standardized exam (i.e., $\mathcal{S} = \{s_m, s_v\}$, with s_m and s_v representing the math and verbal parts of the standardized test, respectively),[4] and let $\omega_{j,s}$ be the weight that program j places on subject $s \in \mathcal{S}$. Then, student i's ability in program j is given by

$$A_{ij} = \underbrace{\sum_{s\in\mathcal{S}} \omega_{j,s} A^o_{i,s}}_{:=A^o_{i,j}} + \underbrace{A^p_{i,m(j)} + \sum_{s\in\mathcal{S}} \omega_{j,s} A^p_{i,s}}_{:=A^p_{i,j}} + \underbrace{A^u_{i,m(j)} + \sum_{s\in\mathcal{S}} \omega_{j,s} A^u_{i,s}}_{:=A^u_{i,j}}. \tag{1}$$

Note that the private and unobserved components comprise two elements: (i) $A^l_{i,m(j)}$, which is major-specific, and (ii) $A^l_{i,s}$, which are subject-specific, with $l \in \{p, u\}$. For private abilities, we assume that these are randomly drawn at the beginning of the academic progression, with $A^p_{i,k} \sim N(\mu^p_{i,k}, \varsigma^2_{i,k})$ for $k \in \mathcal{M} \cup \mathcal{S}$, where $\mu^p_{i,k}$ is the mean—normalized to zero for majors—and $\varsigma_{i,k}$ is the standard deviation of student i's private ability in program j and component k. For unobserved abilities, we assume that the true distribution of $A^u_{i,k}$ is normal with mean zero and standard deviation $\sigma^2_{i,k}$ for $k \in \mathcal{M} \cup \mathcal{S}$, but we allow students to have subjective, heterogeneous and potentially biased prior beliefs about these elements; namely, we assume that student i's prior beliefs about $A^u_{i,k}$ in period t is normally dis-

[4]Elective test scores (Science/History) and high school performance measures (NEM/Rank) are excluded from the ability term because they are not mandatory and consistently available for all students. However, elective scores are updated for exam retakers, and all five admission factors are used when estimating admission probabilities.

tributed with mean $\tilde{\mu}_{i,k,t}$ and variance $\tilde{\sigma}^2_{i,k,t}$, with initial values $\tilde{\mu}_{i,k}$ and $\tilde{\sigma}^2_{i,k}$. Consequently, student i's prior mean and variance about their unobserved ability in program j in period t are given by $\tilde{\mu}_{i,j,t} = \tilde{\mu}_{i,m(j),t} + \sum_{s\in\mathcal{S}} \omega_{j,s} \cdot \tilde{\mu}_{i,s,t}$ and $\tilde{\sigma}^2_{i,j,t} = \tilde{\sigma}^2_{i,m(j),t} + \sum_{s\in\mathcal{S}} \omega_{j,s} \cdot \tilde{\sigma}^2_{i,s,t}$, respectively. Furthermore, we assume that students' grades are a function of a set of time invariant observable characteristics $X^G_{i,j}$— which include characteristics of student i (e.g., their observed ability) and program j (e.g., major)—, their private and unknown abilities $A^p_{i,j}, A^u_{i,j}$, and a noise term $\epsilon^G_{i,j,t} \sim N\left(0, \sigma^2_G\right)$, i.e.,

$$G_{i,j,t} = G\left(X^G_{i,j}, A^p_{i,j}, A^u_{i,j}, \epsilon^G_{i,j,t}\right). \tag{2}$$

Consequently, based on their prior beliefs about their unobserved ability, students construct beliefs about their grades, which we denote by $\tilde{G}_{i,j,t} := \mathbb{E}\left[G_{i,j,t}\right]$—the expectation is computed over $\tilde{A}^u_{i,j} \sim N\left(\tilde{\mu}_{i,j,t},\, \tilde{\sigma}_{i,j,t}\right)$ and $\epsilon^G_{i,j,t} \sim N(0, \sigma^2_G)$—and, upon observing their actual grades, they receive a noisy signal that they use to update their prior beliefs about their unknown abilities according to Bayes rule, as formalized in Proposition 1. We defer the proof to Appendix C.1.

PROPOSITION 1: *Let $\left\{\tilde{\mu}_{i,j',t},\, \tilde{\sigma}_{i,j',t}\right\}_{j'\in\mathcal{J}}$ be the sets of prior means and variances of student i's unobserved ability in programs $j' \in \mathcal{J}$ at time t, respectively. In addition, let $j \in \mathcal{J} \cup \{\emptyset\}$ be the program of enrollment of student i in period t, and $a_{i,j,t}$ be the observed signal (if $j \in \mathcal{J}$). Then, the posterior mean of student i in program $j' \in \mathcal{J}$ is given by*

$$\tilde{\mu}_{i,j',t+1} = \begin{cases} \tilde{\mu}_{i,j',t} + \frac{\left(a_{ijt} - \tilde{\mu}_{i,j,t}\right)}{\tilde{\sigma}^2_{ijt} + \sigma^2_G} \cdot \left[\tilde{\sigma}^2_{i,m(j),t} + \sum_{s\in\mathcal{S}} \omega_{j's}\omega_{js}\tilde{\sigma}^2_{i,s,t}\right] & \text{if } m\left(j'\right) = m\left(j\right),\ j \in \mathcal{J} \\ \tilde{\mu}_{i,j',t} + \frac{\left(a_{ijt} - \tilde{\mu}_{i,j,t}\right)}{\tilde{\sigma}^2_{ijt} + \sigma^2_G} \cdot \left[\sum_{s\in\mathcal{S}} \omega_{j's}\omega_{js}\tilde{\sigma}^2_{i,s,t}\right] & \text{if } m\left(j'\right) \neq m\left(j\right),\ j \in \mathcal{J} \\ \tilde{\mu}_{i,j',t} & \text{if } j \notin \mathcal{J} \end{cases}$$

Intuitively, students learn more about programs that resemble the one they are currently enrolled in—particularly those within the same major or with comparable weighting of admissions scores. Finally, we assume that students' beliefs are heterogeneous across students; in Table C.II in Appendix C.6, we show how these vary among several observable characteristics, including students' gender, income level, and observed ability.

4.2. *Flow Utility*

Let $U_{i,j,t}$ be the flow utility that student i receives for attending program j in period t. We assume that this utility can be modeled as:

$$U_{i,j,t} = U\left(X^U_{i,j}, \alpha_{i,m(j),t}, \epsilon^U_{i,j,t}\right), \tag{3}$$

where $X^U_{i,j}$ is a set of time-invariant observable characteristics that include characteristics of student i (e.g., gender, income, or observed ability), program j (e.g., fixed effects, quality, or selectivity), and interaction terms (e.g., the distance between student i's and program j's municipalities or the net cost that student i pays for program j); $\alpha_{i,m(j),t}$ is student i's idiosyncratic preference for major $m(j)$ in period t; and $\epsilon^U_{i,j,t}$ is an idiosyncratic preference shock that is distributed i.i.d across programs and periods and follows a type I extreme value distribution with a scale parameter κ. We specify a location normalization and set the systematic value of the outside option to $U_{i,\emptyset,t} = 0$ for all $t \in \mathcal{T}$ and set the distance coefficient to minus one, allowing the scale of the shocks to vary.

4.3. *Labor Market*

Upon graduating from program $j \in \mathcal{J} \cup \{\emptyset\}$ in period $t' \in \mathcal{T}$, we assume that student i enters the work force and obtains a utility $V_{i,j,t'}$ consisting of two elements: (i) a non-pecuniary utility $V^{NP}_{i,j,t'}$, capturing the utility derived from their preferences during their time in college; and (ii) a pecuniary utility $V^P_{i,j,t'}$, capturing the present value of their lifetime wages. Given a retirement period T and a common discount factor δ, we model this as:

$$V_{i,j,t'} = \underbrace{\beta^V_{m(j)} + \beta^V_\alpha \alpha_{i,m(j),t'} + \beta^V_A A^o_{i,j} + \beta^V_{PU} \cdot \left(A^p_{i,j} + A^u_{i,j}\right)}_{:=V^{NP}_{i,j,t'}} + \underbrace{\beta^V_W \log\left(\mathbb{E}\left[\sum_{t=t'}^{T} \delta^{t-t'} \cdot P^W_{m(j),t} \cdot W_{i,j,t}\right]\right)}_{:=V^P_{i,j,t'}}. \tag{4}$$

The first term captures the non-pecuniary utility, which depends on student i's preferences for major $m(j)$ encoded in the random coefficient $\alpha_{i,m(j),t'}$ and on their abilities. The second term captures the pecuniary utility, which is separable over time and depends on the (major-specific) employment probability $P^W_{m(j),t}$ and the wage $W_{i,j,t}$, which we model as

$$W_{i,j,t} = W\left(X^W_{i,j}, G_{i,j,t'}, \Lambda_{j,t-t'}\right), \tag{5}$$

where $X^W_{i,j}$ is a set of time-invariant observable characteristics of student i; program j and their interaction; $G_{i,j,t'}$ is student i's GPA at the time of graduation; and $\Lambda_{j,t-t'} = \Lambda(X^\Lambda_{i,j}, t-t')$ captures the (potentially heterogeneous across observable characteristics in $X^\Lambda_{i,j}$) effect of tenure on program j's wage. As we discuss later, we assume that the effect of grades on wages is heterogeneous across students, and we model it as a random coefficient that is normally distributed with mean $\mu_{W\sim G}$ and variance $\sigma^2_{W\sim G}$. Finally, note that student i receives this continuation value only if they graduate from their program. If, instead, student i drops out, we assume that their continuation value is equal to $V_{i,\emptyset,t'}$.

4.4. *Implementation*

Timing. For estimation, we consider a three-period model. Periods 1 and 2 correspond to the first and second years of college, while period 3 begins in the third year and aggregates all subsequent years until graduation, including discounted labor market payoffs. In period 1, students apply to colleges, receive their assignment, decide whether to retake the PSU, obtain first-year grades, and update their beliefs. In period 2, they may reapply and, given their current enrollment, new assignment, and what they learned earlier, decide whether to remain, switch programs, or drop out. In period 3, students face dropout and graduation probabilities and enter the labor market.

Decisions. Throughout their first two years in college, the key decisions we model are (i) enrollment (in period $t=2$), i.e., conditional on the assignment, whether to enroll in the newly assigned program, stay in the current one, or dropout; (ii) exam re-taking (in period $t=1$), i.e., whether or not to re-take the standardized test; and (iii) application (in periods $t\in\{1,2\}$), i.e., conditional on the scores received after (re)taking the test, the current beliefs about abilities and the updated preferences, which programs to include in the ROL (if any). For the latter, we assume that students differ in their level of sophistication (Pathak and Sönmez, 2008), with each student's application strategy being exogenously given and following one of two types: (i) weak truth-telling (w.p. ρ), i.e., students report their true preferences; and (ii) strategic (w.p. $1-\rho$), i.e., students submit a ROL that maximizes their expected value (as in Chade and Smith (2006)) assuming correct beliefs about their marginal admission probabilities. Specifically, a strategic student i solves:

$$R_i^\star \in \underset{R \subseteq \mathcal{R}_i}{\operatorname{argmax}} \left\{ \sum_{l=1}^{|R|} p_{i,R(l),t} \cdot \bar{p}_{i,R(l),t} \cdot \Psi^\star_{i,R(l)} - C^A_{i,t}(R) \right\}, \tag{6}$$

where $p_{i,j,t}$ is the probability of getting assigned to program j in period t; $\bar{p}_{i,R(l),t} = \prod_{l'=1}^{l-1}(1 - p_{i,R(l'),t})$ is the probability of not being assigned to any of the top $l-1$ programs listed in R, with $\bar{p}_{i,R(1),t} = 1$; $C^A_{i,t}(R) = c\,|R|$ is the cost of applying to ROL R; and $\mathcal{R}_i := \left\{ R \in \mathcal{R}(\mathcal{J}) : |R| \leq K,\ p_{i,j',1} > 0,\ \forall j' \in R \right\}$ is the set of ROLs involving at most K programs with positive admission probability. Note that students include programs in their ROL only if it is strictly profitable, so they will skip programs for which their admission chances are zero. This assumption is in line with the findings in (Larroucau et al., 2025), and rules out multiplicity of optimal ROLs involving zero-chance programs (He, 2012).

The key element in (6) is the term $\Psi^\star_{i,R(l)}$, which represents the expected continuation value associated with being assigned to program $j = R(l)$ in period t. This value integrates all subsequent dynamic decisions that follow, including whether to enroll, retake the admissions exam, re-apply to new programs, switch programs or dropout and join the labor force (see Appendix C.2). It also embeds the evolution of beliefs about abilities and the learning processes that unfold conditional on enrollment—students update their expectations based on observed outcomes and experiences—and also capture the probabilities of graduation, dropout, which we model as functions of observable characteristics from period three (see Appendix C.3), and labor market outcomes upon graduation.

Computing these continuation values is particularly challenging, as it requires taking expectations over a high-dimensional state space that encompasses the multiple sources of uncertainty previously mentioned. To address this complexity, we assume a large matching market, so students take the distribution of cutoffs as given (Azevedo and Leshno, 2016), and we adapt the results of Fack et al. (2019) to our dynamic setting, showing that students' continuation values over their optimal portfolios can be represented by their most preferred ex-post feasible programs, integrated over the empirical distribution of future cutoffs. Exploiting the assumption that preference shocks are i.i.d. Type-I extreme value, we then approximate the corresponding expectations using a log-sum expression (see Appendix C.4). This approach allows us to evaluate continuation values without integrating over all preference shocks, greatly reducing computational burden.

Finally, these continuation values can be directly influenced by policy interventions—for example, by altering the information environment, which shapes the learning process, or through market design changes that modify the incentives and feasibility of re-application or switching. We discuss such interventions in greater detail in Section 7.

Heterogeneity. The learning process described above incorporates heterogeneity at the student level, allowing the learning parameters—related to private signals ($\varsigma_{i,k}$ for $k \in \mathcal{M} \cup \mathcal{S}$), unknown abilities ($\sigma_{i,k}$, $\tilde{\mu}_{i,k}$, $\tilde{\sigma}_{i,k}$ for $k \in \mathcal{M} \cup \mathcal{S}$), and preferences ($\mu^{\alpha}_{i,m}$, $\sigma^{\alpha}_{i,m}$, and $\zeta^{\alpha}_{i,m}$ for $m \in \mathcal{M}$)—to vary across students and majors/subjects. To ease estimation, we group majors into four *broad major*—Science (Science, Farming, and Technology), Social Sciences (Social Sciences, Business and Management, Art and Architecture, and Law), Education and Humanities, and Health. Major-specific structural parameters vary at the broad major level, while random variables (e.g., major-specific abilities) vary at the *major* level. We also classify programs as *math intensive* (*verbal intensive*) if their math weight exceeds the verbal one (otherwise) and use the average math weight within each type, i.e., $\frac{\sum_{j \in \mathcal{J}} \omega_{j,s_m} \cdot \mathbb{1}_{\{s(j)=s_m\}}}{\sum_{j \in \mathcal{J}} \mathbb{1}_{\{s(j)=s_m\}}}$, where $s(j)$ indicates the type of program j. Then, the unobserved ability of student i in program j is $A^u_{ij} = A^u_{i,m(j)} + \bar{\omega}_{s(j)} A^u_{i,s_m} + (1 - \bar{\omega}_{s(j)}) A^u_{i,s_v}$.

Furthermore, we capture heterogeneity at the student-level using a small set of observable characteristics that are likely to influence the learning process and biases in students' beliefs. Specifically, we model each of the mean parameters mentioned above (i.e., $\tilde{\mu}_{i,k}$ for $k \in \mathcal{M} \cup \mathcal{S}$ and $\mu^{\alpha}_{i,m}$ for $m \in \mathcal{M}$) as linear functions of gender, income level, and their average ability, measured as the normalized average between the Math and Verbal entrance exams. We apply a similar approach to the logarithm of the variance parameters ($\varsigma_{i,k}$, $\sigma_{i,k}$ for $k \in \mathcal{M} \cup \mathcal{S}$ and $\sigma^{\alpha}_{i,m}$, $\zeta^{\alpha}_{i,m}$ for $m \in \mathcal{M}$). For instance, for the true variance of unknown abilities, we assume $\log\left(\sigma^2_{i,k}\right) = \gamma^{\sigma}_k + \gamma^{\sigma}_F \cdot \text{female}_i + \gamma^{\sigma}_{LI} \cdot \text{low-income}_i + \gamma^{\sigma}_A \cdot \bar{A}_i$, where

female_i, low-income_i are dummy variables for whether student i is female or their income below the median, respectively, and $\bar{A}_i = \frac{1}{|\mathcal{S}|}\sum_{s\in\mathcal{S}} A^o_{i,s}$ is i's average observable ability.[5]

Functional forms. We assume that the grade, flow utility, and log-wage functions (in (2), (3) and (5), respectively) are linear in their arguments, i.e.,

$$G\left(X^G_{i,j}, A^p_{i,j}, A^u_{i,j}, \epsilon^G_{i,j,t}\right) = \beta^G_X X^G_{i,j} + A^p_{i,j} + A^u_{i,j} + \epsilon^G_{i,j,t}, \tag{7a}$$

$$U\left(X^U_{i,j}, \alpha_{i,m(j),t}, \epsilon^U_{i,j,t}\right) = \beta^U_X X^U_{i,j} + \alpha_{i,m(j),t} + \epsilon^U_{i,j,t}, \tag{7b}$$

$$\log W\left(X^W_{i,j}, G_{i,j,t'}, \Lambda_{j,t-t'}\right) = \beta^W_X X^W_{i,j} + \beta^W_{G,i} G_{i,j,t'} + \Lambda_{m(j),t-t'} + \epsilon^W_{i,j,t}, \tag{7c}$$

where $\Lambda_{m(j),t-t'} = \beta^\Lambda_{m(j),1}(t-t') + \beta^\Lambda_{m(j),2}(t-t')^2$. Furthermore, following our survey evidence, we assume that the effect of grades on wages, $\beta^W_{G,i}$, are heterogeneous across students and can be modeled as a random coefficient that is log-normally distributed with mean $\log(\mu_{W\sim G})$ and variance $\sigma^2_{W\sim G}$. Finally, note that for estimation, we use beliefs on expected wages and grades as opposed to their actual values, so we estimate (7c) considering $\tilde{W}_{i,j,t}$, $\tilde{G}_{i,j,t}$ and adding a measurement error term. This approach, made possible by our rich survey data, allows us to relax the rational expectations assumption and enable us to more accurately capture the (potentially biased) information that students actually used when making their decisions (Wiswall and Zafar, 2015a,b). Furthermore, it allows us to understand the role of biases (in both mean and variances) in affecting the dynamics of college admissions and how different interventions can mitigate these.

Controls. As shown in (7), each function controls for a set of time-invariant observable characteristics—encoded in $X^G_{i,j}, X^U_{i,j}$ and $X^W_{i,j}$—that include characteristics of the student, the program, and their interaction. In the grade equation, we control for broad major fixed effects and students' observed ability. The flow-utility equation includes program fixed effects—parametrized as a combination of major fixed effects, college fixed effects, a quality index and dummies for the top 20 programs in terms of market shares—-, observed

[5] We use the average instead of the program-specific observed ability (multiplying each $A^o_{i,s}$ by the corresponding weight $\omega_{j,s}$) to avoid program-specific heterogeneity in defining students' types and simplify the estimation.

ability, distance to the program's municipality, distance to the program's mean quality, and net costs (tuition minus scholarships). Finally, the wage equation includes broad major fixed effects, expected grades at graduation, average posted wages at enrollment, and fixed effects for gender and low-income status. See Appendix C.5 for additional details.

5. IDENTIFICATION

We outline our identification strategy by focusing on the sources of variation that identify the main parameters. While full identification relies on all moment conditions, several parameters can be isolated using targeted moments. For clarity, we present the arguments constructively and conditional on observable characteristics.

Preferences. We follow Agarwal and Somaini (2018), who show that students' preferences can be non-parametrically identified from their reported ROLs assuming correct beliefs about their (marginal) admission probabilities and exploiting distance as a special regressor.

Application. To estimate the probability of being weak truth-teller in the baseline (ρ), we use (i) the share of applications for which the top-reported choice has zero admission probability, and (ii) the share of applicants who include their top-true preference (elicited in the survey) as their top-reported program. The first moment helps to identify strategic behavior by capturing applications to programs with no admission chance. The second moment directly measures truth-telling alignment.

Tuition and welfare effects. We follow Kapor et al. (2024) and exploit the discontinuous change in tuition generated by the scholarship *Beca Vocación de Profesor* and its heterogeneous effect on enrollment to identify the effect of the net cost that students pay.[6]

Private abilities and preferences. Students know their private abilities and the draw of their initial preferences when making their initial choices (Arcidiacono, 2005). Thus, we

[6] Students with average scores above 600 can enroll in Education programs tuition-free, creating a discontinuity in enrollment around this cutoff (see Figure D.1 in Appendix D.2) that we exploit for identification.

exploit the correlation in students' initial reported preferences—i.e., before learning about their unknown abilities—across majors and subjects to identify the role of private information on students' choices, and rely on the correlation between students' ROL composition and first-year grades to identify the effect of private abilities on grades using a control function approach (Heckman, 1979, Kline and Walters, 2016). If private abilities largely shape preferences through grades, we should observe a strong correlation between the two.

Variance of abilities and grades. Since variation in students' grades, can only be explained by their private abilities, their unknown abilities, and the grade shock—the former already identified—, we separately identify the variance of unknown abilities from the variance of the grade noise by constructing a system of equations using the variance of the first-year and second-year signals among non-switchers (see Appendix D.1).

Biased unknown abilities. Assuming that beliefs about grades follow a similar structure to that in Equation (7a), considering students' (potentially biased) prior about their unknown ability and an additional measurement error term $\epsilon^{\tilde{G}}_{i,j,t} \sim N(0, \sigma^2_{\tilde{G}})$, we identify the initial mean bias in students' beliefs about their first-year grade, $\tilde{\mu}_{i,j}$, by averaging the difference between students' beliefs about their first-year grades and their actual grades.

Learning about preferences and abilities. We directly identify subjective prior variances exploiting the probability that students assign to their grades falling between their expected grades plus or minus 10 percentiles conditional on attending their *top-reported program*, which reflect students' uncertainty about their GPA given their subjective beliefs. We also use the correlation between the signal and the choices that students make after the first period (including their decisions to switch and dropout) to further separate the effect of subjective beliefs from grade noise. Intuitively, a high signal-to-noise ratio implies a low variance of the grade shock relative to the subjective variance of unknown abilities, which means students' choices respond more strongly to their observed grades. This translates into stronger correlations between first-year grades and switching decisions. Moreover, if students are more likely to switch majors than subjects when they receive negative signals (lower grades than expected), then the variances related to majors outweigh those for subjects. Finally, we identify the standard deviation of the evolution of random coefficients, $\zeta^{\alpha}_{i,m}$, by exploiting variation in preference changes across majors among students who reapply after receiving a non-informative signal, allowing us to isolate the effect of changing preferences from learning about unknown abilities.

Labor Market. To identify the (subjective) wage equation, we rely on the correlation between students beliefs about their wages and grades at the individual level and across programs, including students' reported preferences, their top-true program, and also a random program outside their reported ROL—ruling out potential selection concerns. This type of variation allows us to estimate the marginal effect of expected grades on wages. To separately identify the wage equation parameters from measurement error, we exploit the fact that we elicited students' beliefs about their admission chances to their top-reported program in two ways: (i) directly, and (ii) indirectly, by asking them to report the probability that their application score will be above the cutoff. The difference between these two responses should be independent of students unknown ability, providing a proxy of the measurement error in the survey. Then, to separate the variance of the effect of grades on wages from the measurement error—i.e., $\sigma^2_{W\sim G}$ from $\sigma^2_{W\sim M}$—we leverage the elicited beliefs on wages for different programs (both included and excluded from the reported preferences)—same as we do for beliefs in grades—to construct a system of equations involving the residuals of the subjective wage regression, as discussed in Appendix D. Finally, we identify the parameter scaling the wage component by leveraging the correlation between students' subjective wage beliefs and their preferences—specifically, by comparing top true and random programs.

Initial mismatch and Counterfactual outcomes. To identify the strength of the initial mismatch channel and its relationship to parameters governing unobserved heterogeneity and first-time enrollment costs, we exploit variation in initial assignments generated by the RDDs described in Section 3.2. In addition, these variations are key to identify the distribution of outcomes in the counterfactual analysis.[7] Indeed, because the primary effect of our

[7] See Agarwal et al. (2024a) for a thorough discussion of this identification challenge.

counterfactuals is to reassign students around admission cutoffs, these local variations provide a credible basis for predicting counterfactual outcomes. We then use the full model's structure to extrapolate beyond the cutoffs and account for equilibrium effects that may alter application behavior and initial assignments.

6. ESTIMATION

We estimate the model using a random sample of 4,000 students drawn from the population described in Section 3.1. We next describe the estimation procedure (Section 6.1), present the estimated parameters (Section 6.2), and assess the model fit (Section 6.3).

6.1. *Estimation Procedure*

Estimation of Exogenous Components. As discussed in Section 4.4, we estimate the distributions of admission probabilities $\hat{p}_{j,t}$ using the bootstrap procedure detailed in Appendix C.3.1. Similarly, we estimate the probabilities of enrollment $\hat{P}^{e}_{i,t}$, graduation $\hat{P}^{g}_{i,j,t}$ and dropout $\hat{P}^{d}_{i,j,t}$ as discussed in Appendices C.3.2 and C.3.3, respectively. All these probabilities are estimated outside the main estimation procedure, serving as inputs.

Indirect Inference. We estimate the model parameters via Indirect Inference (II), which relies on building a statistical model that yields a rich description of the data patterns to identify the model parameters (Bruins et al., 2018). This statistical model—also known as the *auxiliary* model—is estimated on both the data and on simulated data from the structural model. The II estimator—also known as Wald approach to II—minimizes a function of the distance between the estimated data parameters and those estimated from the simulated data. In this sense, the Simulated Method of Moments is a particular case of II, in which the *auxiliary* model is a vector of moments. In Appendices E.1, E.2 and E.3, we formalize the estimator, describe the estimation algorithm, and discuss the moment conditions, auxiliary models and target parameters, respectively.

Marginal Improvement Algorithm. An important part of the estimation procedure is to solve for the optimal portfolio that a strategic student would choose, given the estimated parameters. To do this, we use the Marginal Improvement Algorithm (MIA) introduced

TABLE IV

ESTIMATION RESULTS: SUMMARY OF KEY PARAMETERS

	Unknown Abilities							
	Beliefs biases				True		Preferences	
	Mean ($\tilde{\mu}$)		Log-Var ($\tilde{\sigma}^2$)		Log-Var (σ^2)		Log-Var (ζ_m^2)	
Parameter	Value	S.E.	Value	S.E.	Value	S.E.	Value	S.E.
Science	0.63	(-)	-1.28	0.21	-0.27	0.19	0.57	0.09
Social Science	0.66	(-)	-2.03	0.13	-0.20	0.22	1.14	0.01
Education and Humanities	0.36	(-)	-0.74	0.16	-0.58	0.26	0.48	< 0.01
Health	0.84	(-)	-1.99	0.09	-1.44	0.18	2.69	0.02
Low income	0.08	(-)	-0.19	0.21	0.12	0.09	2.47	0.02
Female	-0.19	(-)	0.20	0.25	-0.37	0.10	-0.04	0.02
Avg. obs. ability	-0.21	(-)	-1.90	0.08	-0.80	0.16	1.20	0.03

Note: Standard errors (S.E.) are computed using the asymptotic variance-covariance of the estimator. Parameters from external regressions are indicated by (-).

by Chade and Smith (2006), which sequentially adds the program that yields the highest marginal utility gain, stopping when the list-length constraint binds or no further improvement is possible. We then assume that a student's reported ROL corresponds to the MIA-selected programs, ordered by decreasing utility. See Appendix E.4 for details.

6.2. *Estimated parameters*

Table IV presents a summary of the key estimated parameters. Full estimation results are presented in Tables E.III, E.IV, E.V, E.VI, and E.VII in Appendix E.6.

Students' beliefs about their academic abilities display systematic biases. Prior means are upward-biased across all major categories, with subjective mean biases ($\tilde{\mu}_m$) ranging from 0.36 to 0.84, consistent with pervasive overconfidence. Students also underestimate the uncertainty in their abilities, as indicated by negative biases in log-variances reflecting overprecision in their beliefs, i.e., $\log(\tilde{\sigma}_m^2) < \log(\sigma_m^2)$. These distortions have important behavioral implications: overconfidence leads students to overestimate their chances of success in competitive programs, generating ex post mismatches once grade signals reveal true abilities; overprecision dampens belief updating in response to new information, slowing learning and discouraging optimal switching. Biases also vary systematically across demographics: low-income students tend to be more overconfident and overprecise than high-income students; female students are generally less overconfident and overprecise than males; and higher-ability students are less overconfident but more overprecise.

Beyond learning about their abilities, students also learn about their preferences for majors through direct experience. The estimated log-variance parameters (ζ^2) reveal substantial heterogeneity across majors and student characteristics. Low-income students, for example, display higher variances and thus greater initial uncertainty about their preferences.

A few additional parameters are worth highlighting (see Appendix E.6 for details). The estimated share of strategic students is 0.88, suggesting that most applicants account for admission probabilities when choosing their rank-ordered lists, underscoring the importance of explicitly modeling strategic behavior. Regarding preference heterogeneity, the estimated scale of idiosyncratic shocks ($\kappa = 17.78$) and the standard deviation of major-specific random coefficients ($\sigma_\alpha = 33.7$) indicate substantial variation across students, with systematic differences in major preferences dominating idiosyncratic tastes. Finally, the variance of the grades shock ($\sigma_g^2 = 0.11$) and the cross-sectional variance of the realized grade signal ($\mathbb{V}(a_{i,j,1}) \approx 0.26$) imply that the grade signals are highly informative and, thus, students can meaningfully learn about their abilities from their performance in college.

Overall, these findings reveal that students exhibit substantial and heterogeneous belief biases—typically overestimating their abilities and underestimating uncertainty—yet the grade signal remains highly informative about their unknown abilities. These patterns can have far-reaching implications for application strategies, learning dynamics, and the welfare effects of policy interventions, highlighting the importance of accounting for both bias and heterogeneity in students' beliefs and learning parameters.

6.3. *Model fit*

To illustrate the key identifying variations discussed in Section 5, Table V summarizes the moments that help identify the two main channels: (i) initial mismatches and (ii) learning about abilities and preferences. Panel Va reports the estimated causal effects from the

TABLE V

GOODNESS OF FIT: ADDITIONAL MOMENTS

(a) RDD Moments

Targets	Model	Data
RDD Switching		
Constant	0.120	0.167
Discontinuity	-0.088	-0.051
RDD Re-Application		
Constant	0.276	0.234
Discontinuity	-0.144	-0.087
RDD Switch Up/Out		
Constant	0.067	0.048
Discontinuity	-0.063	-0.034

(b) Beliefs on Prior Log-Variance

	Belief Variance				Learning Pref.	
	Subjective		Real			
Coefficient	Model	Data	Model	Data	Model	Data
Science	-1.602	-1.768	-0.252	-0.191	0.117	0.089
Social Science	-1.709	-1.746	-0.227	-0.139	0.053	0.029
Educ. and Human.	-1.522	-1.770	-0.480	-0.548	0.080	0.048
Health	-1.836	-1.749	-0.957	-0.825	0.058	0.070
Female	-0.031	-0.070	-0.249	-0.245		
Low-income	0.001	0.003	0.081	0.068		
Observed ability	-0.307	-0.302	-0.524	-0.503		
Verbal weight²	-0.069	-0.175	0.092	0.351		
Signal					-0.082	-0.120

(c) Signal Correlations

Targets	Model	Data
Grades × Switch		
Dropout	-0.220	-0.281
Program	-0.158	-0.218
Broad Major	-0.134	-0.158
Major	-0.166	-0.213
Math Type	-0.117	-0.139
Switch Up	-0.027	0.065
Switch Down	-0.048	-0.133
Out (feas.)	-0.153	-0.225
Out (unfeas.)	-0.032	-0.011

Notes: Goodness of fit for key moments used in estimation. Panel (a) shows RDD discontinuity coefficients. Panel (b) shows regression coefficients of log total variance of prior beliefs (Belief Variance columns show both Model and Data for Subjective and Real beliefs) and learning preference coefficients (showing both Model and Data). Panel (c) shows correlations between grades and switching behaviors.

RDD models around admission cutoffs for students' top-reported preferences. These variations are crucial for identifying initial mismatches and accurately predicting switching rates in counterfactuals, since students may shift around the cutoffs in equilibrium. The model matches these moments reasonably well, though it slightly overpredicts the magnitudes of the discontinuities. Panel Vb shows the fit of regressions of the log total variance of students' prior beliefs on demographics and ability measures. Along with the correlation structure of grade signals, these moments identify the degree of learning about unknown abilities and their effects on outcomes. The model replicates well the observed heterogeneity across broad majors and demographic groups (see Table C.II in Appendix E.6). The same panel also reports the fit for regressions of changes in the norm of reported preferences on students' grade signals and major. While preference changes and switching are negatively correlated with grade signals (slope), there remains residual switching among students whose grades convey no information—evidence that students also learn about their match-quality (preferences) outside the grade production function. See Appendix E.3 for detailed specifications of these auxiliary models. Finally, Panel Vc presents the correlations between grade signals and switching outcomes. The model reproduces these patterns closely: switching *Up* is nearly uncorrelated with grades, while switching *Out* to ex ante feasible programs is strongly negatively correlated, consistent with both channels.

7. COUNTERFACTUALS

We now turn to our counterfactual analysis. As outlined in Section 1, we consider two families of counterfactual policies: (i) changes to the information and learning environment, aimed at reducing belief biases and front-loading learning about abilities and preferences; (ii) market design interventions aimed to reduce initial mismatches and capture students' dynamic incentives. We consider two types of market design interventions: (ii-a) modifications to the assignment mechanism designed to elicit cardinal preferences and mitigate initial mismatches; and (ii-b) adjustments to the reapplication rules that encourage students to internalize the systemic effects of their dynamic decisions. We evaluate each of these policies in equilibrium, accounting for the strategic incentives they may generate—that is, we assume that strategic students respond to changes in the environment by updating their beliefs about the distribution of cutoffs and adjusting their behavior accordingly. Algorithm 2 in Appendix F.3—which extends Kapor et al. (2020) to account for the mixture between first-time and re-applicants—describes how we estimate students' equilibrium beliefs over the stationary cutoff distributions.

Before proceeding to our primary counterfactual analysis, we assess how switching and dropout decisions are shaped by the two behavioral channels (see Table F.I in Appendix F.5). First, eliminating biases about mean unknown abilities—by setting $\tilde{\mu} = 0$—decreases switching by roughly 15%, while increasing unassignment by close to 19%. Second, eliminating biases in the mean and variance of unknown abilities—by additionally setting $\tilde{\sigma}_k^2 = \sigma_k^2$ for $k \in \mathcal{M} \cup \mathcal{S}$—increases the number of switches and dropouts by close to 81% and 23%, respectively. The intuition behind this result is that students are generally overoptimistic (i.e., positive mean bias) and underestimate variability (i.e., their prior variance is lower than the true), so eliminating these biases increases the informativeness of the signals received from grades, thereby increasing the value of choosing other options (either switching or dropping out). Third, removing the learning about abilities channel—by setting $\sigma_k^2 = 0$ for $k \in \mathcal{M} \cup \mathcal{S}$—reduces program switching by roughly 23% relative to the baseline. Beyond this, eliminating the learning about preferences channel—setting $\zeta^{\alpha}_{i,m} = 0$ for all $i \in \mathcal{I}$ and $m \in \mathcal{M}$—yields an additional 11% drop in switching and 12% in dropout. Finally, eliminating initial mismatches—by assigning every student to their top true choice—reduces switching by over 94%. Overall, the results indicate that initial mismatches and learning about abilities are the primary drivers of students' dynamic decisions.

7.1. *Information and Learning Environment*

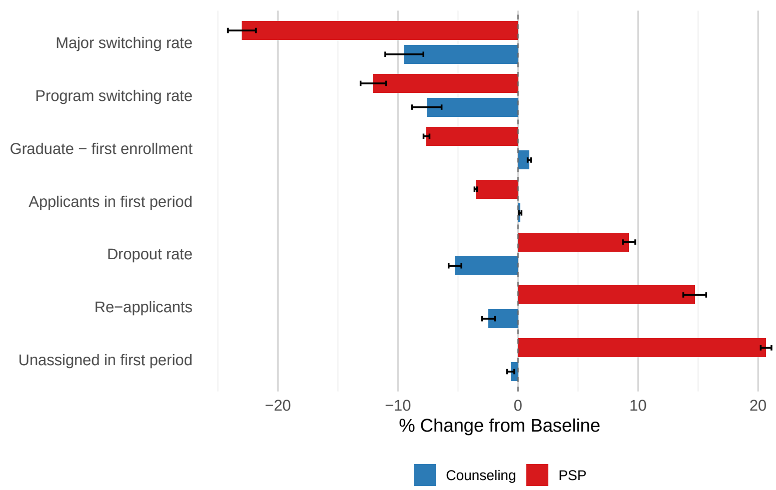
We first evaluate the effects of modifying the information environment to reduce belief biases and front-loading the learning process. The counterfactuals in this section are best interpreted as conceptual exercises designed—intentionally stylized to isolate the specific channels—to help us understand the mechanisms through which information and learning shape student outcomes in equilibrium, and to decompose the relative importance of learning about abilities versus learning about preferences. In particular, we consider two interventions:

1. Counseling: Students often receive (free) guidance and counseling during their last year of high school to help them make informed decisions about their future studies. To obtain an upper bound of their potential effect, we implement this counterfactual assuming that counseling eliminates the variance of the learning process about preferences, i.e., $\zeta^{\alpha}_{i,m} = 0$ for all $i \in \mathcal{I}$ and $m \in \mathcal{M}$.
2. Mandatory year of Post Secondary Program (PSP): In some higher education systems (e.g., Canada and France), students are required to complete one- or multi-year preparatory programs before applying to a specific set of degree programs. This structure allows students to learn about their abilities prior to specialization without generating any crowd-out externality. We model this counterfactual as a one-year intervention implemented by: (i) debiasing the prior mean, i.e., setting $\tilde{\mu}_{i,k} = 0$ for all $i \in \mathcal{I}$ and $k \in \mathcal{M} \cup \mathcal{S}$; (ii) transfer all the uncertainty from unknown abilities to private abilities, i.e., set $\tilde{\sigma}^2_{i,k} = 0$ and $\hat{\varsigma}^2_{i,k} = \varsigma^2_{i,k} + \sigma^2_{i,k}$ for all $i \in \mathcal{I}$ and $k \in \mathcal{M} \cup \mathcal{S}$; and (iii) decreasing the terminal period by one year.

Both interventions seek to reduce inefficient switching by helping students make more informed application decisions, though they target different sources of learning. Counseling facilitates learning about students' preferences, allowing them to identify programs that align with their true interests before submitting their applications. In contrast, PSP enables students to learn about their abilities before applying to a specific program through the centralized system.[8]

Results. In Figure 4, we present the estimated effects—measured as percentage changes relative to the baseline—on several outcomes, including switching (program and major), dropout rates, unassigned and re-application rates, and graduation rates from the first program of enrollment. In Table F.II in Appendix F.5, we present the detailed results.

FIGURE 4.—Counterfactuals: Information and Learning Environment



We observe that both interventions lead to a reduction in switching rates. However, the magnitude of some effects and the mechanism through which they operate differ significantly. First, we observe that the reduction in major switchings is significantly higher in the PSP intervention. The reason behind this result is that PSP allows students to fully learn

[8] Bordon and Fu (2015) analyze a shift from a joint system—where students apply directly to degree programs—to a sequential system, where students first choose a college and later declare a major. In contrast, our analysis modifies the learning environment leaving both the college structure and the assignment process unchanged.

about their major- and subject-specific unknown abilities, which play a more relevant role in driving students' dynamic decisions compared to learning about preferences. Second, we observe that PSP leads to a significantly higher number of unassigned students in the first period, which explains the higher incidence of re-applications compared to the counseling intervention. The reason behind this result is that debiasing prior means and eliminating the uncertainty about unknown abilities leads students to submit more "focused" rank-ordered lists, thereby increasing the risk of being unassigned. Finally, we observe that counseling leads to a reduction in first year and overall dropout rates, significantly improving the system's yield.

Overall, these results suggest that front-loading learning can help to reduce switching and dropout rates. However, the two interventions operate through different channels and have different implications. Counseling primarily helps students learn about their preferences, leading to a reduction in dropout rates, while a mandatory year of college allows students to learn about abilities—which have productive meaning—resulting in a more substantial decrease in switching rates, but also an increase in unassigned students.

7.2. *Market Design Interventions*

We now focus on market design interventions aimed at eliciting the intensity of students' preferences and accounting for their dynamic incentives. We consider two broad types of policy changes. The first targets the assignment mechanism, modifying how preferences are expressed and processed to better capture their intensity—either by restricting the length of applicants' rank-ordered lists (Constrained Deferred Acceptance, CDA) or enabling students to signal one of their preferences to receive a score bonus (Choice-Augmented Deferred Acceptance, CADA (Abdulkadiroğlu et al., 2015)). The second set of interventions alters the reapplication rules, seeking to address dynamic incentives and improve the system's yield by adjusting how the mechanism treats first-time versus returning applicants, as in the Finnish and Turkish reapplication policies. In addition to the outcomes discussed in the previous section, we also evaluate the welfare implications of these policies. We add two measures of students' welfare: (i) interim and (ii) ex-post, which capture welfare before and after students learn about their abilities, respectively.[9] Both welfare measures are translated into millions of Chilean pesos as of 2014. Unlike before, we can now meaningfully compare welfare across counterfactuals because these interventions are implemented conditional on the existing information environment—that is, incorporating students' biases. As such, any differences relative to the baseline reflect changes in the mechanism or reapplication rules rather than differences in the information available to students.

Assignment Mechanisms. We evaluate the effects of eliciting the intensity of students' preferences through changes in the assignment mechanism using two alternative approaches (see Appendix F.1):

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 an ROL with length less than or equal to 3.
2. Choice-Augmented Deferred Acceptance (CADA): Students can signal one program on their ROL and receive a score bonus (score multiplier $\psi > 1$) applied to their high school GPA component. We implement this mechanism only for first-period applicants; consequently, students who apply in the second period do not receive the bonus.

Both mechanisms elicit the intensity of students' preferences, since they introduce opportunity costs that students must take into account when submitting their applications. In the case of CDA, constraining the length of applicants' lists limits students in including other programs on their ROLs, so they must account for the opportunity cost of including each program. In the case of CADA, students can signal only one program, so they must carefully decide which program to target to get the bonus.

Notice that eliciting the intensity of students' preferences may not necessarily lead to higher yield. On the one hand, if eliciting this information decreases initial mismatches, we would expect to reduce inefficient switching. On the other hand, if the assignment mechanism also elicits the intensity of preferences of students who reapply to the system and

[9] 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 of their choices in the model.

these change considerably due to learning, we would see an increase in efficient switching due to a higher value of reapplication. In this sense, we expect that under CADA—which provides a score bonus only in the first period—switching would decrease more than in the case in which the score bonus is applied in both periods, because the policy also gives a comparative advantage to first-period applicants, which increases switching costs through the higher equilibrium cutoff scores produced by the bonus.

Reapplication Rules. Another policy to reduce the incentives to switch is to provide bonuses to students applying for the first time or to penalize students who reapply and try to switch programs; these policies have been implemented in Finland and Turkey, respectively. We implement these policies in our counterfactual analysis as follows, where ψ represents a score multiplier applied to the high school GPA component:

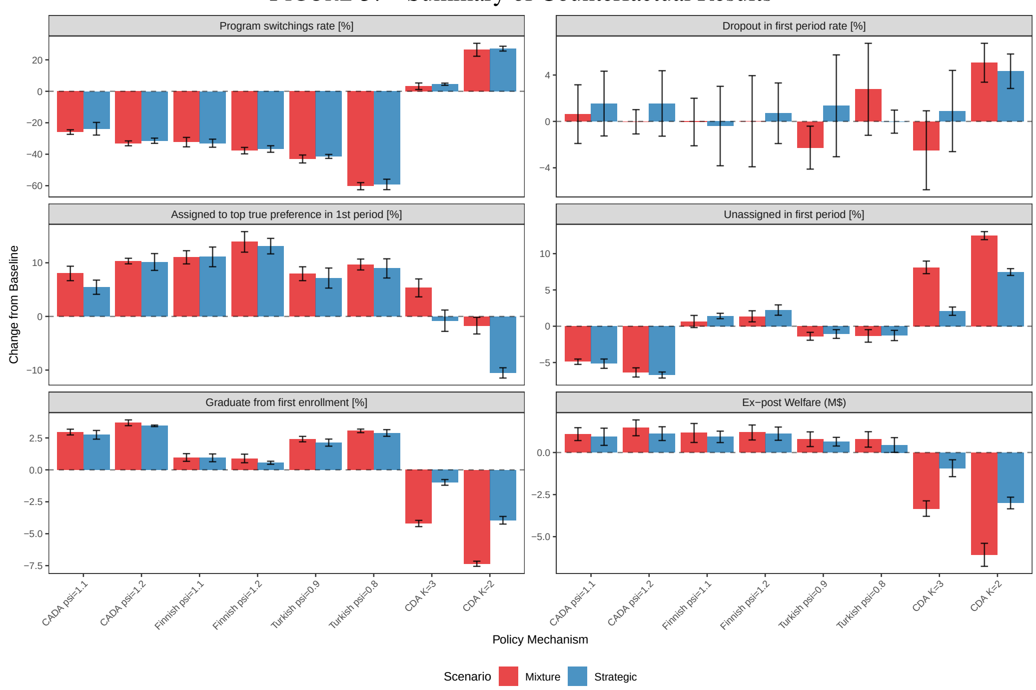
(i) Turkish rule: Applicants currently enrolled in the centralized system receive a score penalty (score multiplier $\psi < 1$).

(ii) Finnish rule: First-time applicants receive a score bonus (score multiplier $\psi > 1$).

Even though both policies aim to reduce the incentives for switching, they affect students' applications and reapplications in different ways. On the one hand, the Finnish policy directly reduces the incentives to reapply to the system, regardless of the programs 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 who wait an extra year to submit their first application. On the other hand, the Turkish policy reduces the incentives to reapply if students were 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 who enroll in the first period in less preferred programs. Despite these differences, we expect that both policies would decrease the frequency of reapplications and switches. In contrast, the welfare effects of these policies is unclear. Students may benefit from these policies, since 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 to switching, the benefits of learning become lower, so students' welfare may decrease.

Results. In Figure 5, we present a summary of the counterfactual results as percentage changes relative to the baseline. The detailed results are reported in Tables F.III and F.IV in Appendix F. Since students' application behavior—i.e., whether they act strategically or as weak truth-tellers—might not be invariant to counterfactuals that heighten strategic incentives, we present bounds on policy effects by considering two benchmark scenarios: (i) all students acting strategically ($\rho = 0$), and (ii) students behaving according to a mixture with probability of strategic behavior given by that estimated in the baseline model ($\rho = \hat{\rho}$).

FIGURE 5.—Summary of Counterfactual Results



On the one hand, we find that implementing CDA increases the share of students who remain unassigned, switch programs, and drop out in the first period. This pattern is intuitive: reducing the maximum length of students' ROLs raises the risk of remaining unassigned and generates more initial mismatches, thereby increasing the incentives to switch or drop out. While the program switching rate (conditional on assignment) increases monotonically as K decreases, the absolute number of program switchers peaks at $K = 2$ and then declines at $K = 1$ due to a sharp rise in unassigned students, which shrinks the pool of potential switchers. Moreover, CDA significantly decreases the share of students assigned to their top true preference—particularly when students behave strategically, as they may skip preferred programs to improve their chances of admission elsewhere. Consequently, CDA leads to significantly lower graduation rates and lower ex-post welfare. Consistent with this, ex-post welfare declines as K decreases.

On the other hand, CADA and the policies involving changes to the re-application rules can improve student outcomes. From the top panel—showing the percentage change in switching and dropout rates relative to the baseline—we observe that all these policies lead to a sizable reduction in program switching, with the largest effect under the Turkish policy.

From the middle panel—reporting the percentage change in the probability of assignment to the top true preference and the probability of being unassigned in the first period relative to the baseline—we find that CADA and the re-application rule changes significantly increase the share of students assigned to their top true preference, with the Finnish and CADA policies delivering comparable improvements. In the fully strategic case, we also observe that changes to the re-application rules cause a small increase in the share of unassigned students in the first period, whereas CADA produces a slight decrease.

From the bottom panel—showing the percentage change in the probability of graduating from the first enrollment and the change in ex-post welfare relative to the baseline—we observe that these policies lead to a significant increase in the probability of graduation from the first enrollment, with the largest gains under CADA and the Turkish policy. When considering the mixture of student types, ex-post welfare increases under all these policies, with CADA generating the highest welfare improvement, and the Finish policy being the best at eliciting preference intensity—assigning students to their top-true preference.

Finally, regarding student types, Tables F.III and F.IV in Appendix F reveal that males generally benefit more than females across the different policies. However, the effects across income groups depend on the policy's intensity, underscoring that policy design may entail trade-offs in equity and distributional fairness.

Overall, these results indicate that restricting the length of applicants' lists can have detrimental effects on both switching behavior and student welfare. In contrast, CADA and the re-application rule policies appear promising for increasing the system's yield, mitigating the congestion externalities caused by initial mismatches, and enhancing welfare. However, the optimal policy depends on the policymaker's objective: if the goal is to reduce switching and increase system yield, the Turkish policy performs best, whereas if the aim is to maximize student welfare, CADA delivers superior outcomes.

8. CONCLUSIONS

In this paper, we analyze the determinants of students' college progression in the presence of a centralized assignment mechanism. We aim to understand why students reapply, switch, or drop out, and how the incentives generated by the mechanism and the learning environment shape these outcomes and students' welfare. We therefore study admissions as a dynamic environment with information frictions, belief updating about abilities and preferences, strategic incentives, and continued interaction with the mechanism over time.

Using data from the Chilean college admissions system and multiple nationwide surveys that we designed and conducted, we provide empirical evidence showing that two central behavioral channels explain students' dynamic decisions. The first channel, 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, the *learning* channel, encompasses both learning about abilities and preferences over majors, leading students to switch to programs where expected outcomes—and their preference intensities—are higher.

Given these findings, we introduce a structural model that captures students' decisions during their academic progression, allowing them to learn about their abilities from grades and about their preferences from initial enrollment, while holding heterogeneous and potentially biased beliefs over college outcomes and labor-market returns. We use the estimated model to evaluate (i) changes to the learning environment—such as counseling in high school and a mandatory year of a post-secondary program—and (ii) different market design choices—including changes in the assignment mechanism such as adding further constraints on list length and allowing a single signal that grants a score bonus, or modifying the reapplication rules to match those used in Turkey and Finland. Our results show that the former can increase the system's yield, while reapplication rules and the signaling mechanism can also increase students' welfare. Importantly, these effects are robust to changes in the fraction of participants who behave strategically, unlike other approaches such as constraining list lengths. The best policy ultimately depends on the planner's objective and the trade-off between yield and the distribution of welfare gains across students.

Overall, our results show that incorporating dynamic incentives and eliciting information on participants' cardinal preferences can significantly increase students' welfare and improve downstream outcomes. These insights can be informative for improving the design of many matching markets that exhibit similar features. For instance, in organ transplant systems, patients have private information regarding their health, face dynamic considerations such as when to accept or reject an organ, and even learn about organs' qualities over time (Zhang, 2010), all of which may affect patient survival (Agarwal et al., 2021). In entry-level labor markets, candidates 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 (reapply), exit (dropout), or rematch (switch), which affect employers' retention in turn. Our key insight is that market designers should correctly balance the gains from learning through experience and the crowd-out externality produced by initial mismatches to improve the efficiency of these markets.

REFERENCES

ABDULKADIROĞLU, ATILA, YEON KOO CHE, AND YOSUKE YASUDA (2015): “Expanding "choice" in school choice,” *American Economic Journal: Microeconomics*, 7, 1–42. [36, 86]

AGARWAL, NIKHIL, ITAI ASHLAGI, MICHAEL REES, PAULO SOMAINI, AND DANIEL WALDINGER (2021): “Equilibrium Allocations Under Alternative Waitlist Designs: Evidence from Deceased Donor Kidneys,” *Econometrica*, 89, 37–76. [42]

AGARWAL, NIKHIL, CHARLES HODGSON, AND PAULO SOMAINI (2024a): “Choices and Outcomes in Assignment Mechanisms: The Allocation of Deceased Donor Kidneys,” *Econometrica*, 92, 1213–1251. [6, 28]

AGARWAL, NIKHIL AND PAULO SOMAINI (2018): "Demand Analysis using Strategic Reports: An application to a school choice mechanism," *Econometrica*, 86, 391–444. [6, 26, 66]

——— (2019): "Revealed Preference Analysis of School Choice Models," *Annual Review of Economics*, 12. [6]

AGARWAL, NIKHIL, PAULO SOMAINI, AND DANIEL WALDINGER (2024b): "Equilibrium Allocations under Alternative Waitlist Designs: Evidence from Deceased Donor Kidneys," *Econometrica*, 92, 1–38. [6]

ALTONJI, J. G., P. ARCIDIACONO, AND A. MAUREL (2016): "The Analysis of Field Choice in College and Graduate School. Determinants and Wage Effects," in *Handbook of the Economics of Education*, Elsevier, vol. 5, 305–396. [6]

ALTONJI, JOSEPH G, ERICA BLOM, AND COSTAS MEGHIR (2012): "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers," *Annu. Rev. Econ*, 4, 185–223. [6]

ARCIDIACONO, PETER (2004): "Ability sorting and the returns to college major," *Journal of Econometrics*, 121, 343–375. [17]

——— (2005): "Affirmative action in higher education: How do admission and financial aid rules affect future earnings?" *Econometrica*, 73, 1477–1524. [6, 17, 26]

ARCIDIACONO, PETER, ESTEBAN AUCEJO, ARNAUD MAUREL, AND TYLER RANSOM (2025): "College Attrition and the Dynamics of Information Revelation," *Journal of Political Economy*, 133, 53–110. [6, 17]

AZEVEDO, EDUARDO M AND JACOB D LESHNO (2016): "A supply and demand framework for two-sided matching markets," *Journal of Political Economy*, 124, 1235–1268. [23]

BORDON, PAOLA AND CHAO FU (2015): "College-major choice to college-then-major choice," *Review of Economic Studies*, 82, 1247–1288. [6, 35]

BRUINS, MARIANNE, JAMES A DUFFY, MICHAEL P KEANE, AND ANTHONY A SMITH JR (2018): "Generalized indirect inference for discrete choice models," *Journal of Econometrics*, 205, 177–203. [29, 74]

CALSAMIGLIA, CATERINA, CHAO FU, AND MAIA GÜELL (2020): "Structural estimation of a model of school choices: The Boston mechanism versus its alternatives," *Journal of Political Economy*, 128, 642–680. [6]

CARD, DAVID AND ALEX SOLIS (2020): "Measuring the Effect of Student Loans on College Persistence," *Education Finance and Policy*, 1–60. [10]

CHADE, HECTOR AND LONES SMITH (2006): "Simultaneous Search," *Econometrica*, 74, 1293–1307. [23, 30, 81]

DE GROOTE, OLIVIER, ANAÏS FABRE, MARGAUX LUFLADE, AND ARNAUD MAUREL (2025): "Sequential college admission mechanisms and off-platform options," Tech. rep., National Bureau of Economic Research. [6]

DE GROOTE, OLIVIER AND MINYOUNG RHO (2024): "Matching disadvantaged children to day care: Evidence from a centralized platform," . [6]

DONG, YINGYING (2017): "Regression Discontinuity Designs With Sample Selection," *Journal of Business & Economic Statistics*, 37, 171–186. [55]

EDERER, TIM (2023): "Labor market dynamics and teacher spatial sorting," *Job Market Paper*. [6]

FACK, GABRIELLE, JULIEN GRENET, AND YINGHUA HE (2019): "Beyond truth-telling: Preference estimation with centralized school choice and college admissions," *American Economic Review*, 109, 1486–1529. [23, 69]

FRIEDRICH, BENJAMIN, MARTIN B HACKMANN, ADAM KAPOR, SOFIA J MORONI, AND ANNE BRINK NANDRUP (2024): "Interdependent Values in Matching Markets: Evidence from Medical School Programs in Denmark," Tech. rep., National Bureau of Economic Research. [6]

HAHM, DONG WOO AND MINSEON PARK (2022): "A dynamic framework of school choice: Effects of middle schools on high school choice," in *Proceedings of the 23rd ACM Conference on Economics and Computation*, 292–293. [6]

HE, YINGHUA (2012): "Gaming the Boston School Choice Mechanism in Beijing," *Manuscript, Toulouse School of Economics*, 2012. [6, 23, 82]

HECKMAN, JAMES J (1979): "Sample selection bias as a specification error," *Econometrica: Journal of the econometric society*, 153–161. [27]

HERNANDEZ-CHANTO, ALLAN (2021): "Recovering preferences in college assignment problems under strategic and constrained reports," *Available at SSRN 3784310*. [6]

HUITFELDT, INGRID, VICTORIA MARONE, AND DANIEL WALDINGER (2025): "Designing Dynamic Reassignment Mechanisms: Evidence from GP Allocation," Tech. rep., NBER Working Paper No. 32458. [6]

KAPOR, ADAM, MOHIT KARNANI, AND CHRISTOPHER A. NEILSON (2024): "Aftermarket Frictions and the Cost of Off-Platform Options in Centralized Assignment Mechanisms," *Journal of Political Economy*, 132, 2346–2395. [2, 26]

KAPOR, ADAM J, CHRISTOPHER A NEILSON, AND SETH D ZIMMERMAN (2020): "Heterogeneous beliefs and school choice mechanisms," *American Economic Review*, 110, 1274–1315. [6, 7, 33]

KLINE, PATRICK AND CHRISTOPHER R WALTERS (2016): "Evaluating public programs with close substitutes: The case of Head Start," *The Quarterly Journal of Economics*, 131, 1795–1848. [27]

KUNO, KAN (2023): "Dynamic Incentives in Centralized Matching: The Case of Japanese Daycare," *arXiv preprint arXiv:2311.07920*. [6]

LARROUCAU, TOMAS AND IGNACIO RIOS (2018): "Do "Short-List" Students Report Truthfully? Strategic Behavior in the Chilean College Admissions Problem," . [82]

LARROUCAU, TOMÁS, IGNACIO A. RIOS, ANAÏS FABRE, AND CHRISTOPHER NEILSON (2025): "College Application Mistakes and the Design of Information Policies at Scale," Tech. Rep. 34164, National Bureau of Economic Research. [23, 67]

LOCHNER, LANCE AND ALEXANDER MONGE-NARANJO (2012): "Credit Constraints in Education," *Annual Review of Economics*, 4, 225–256. [10]

——— (2016): "Student loans and repayment: Theory, evidence, and policy," in *Handbook of the Economics of Education*, Elsevier, vol. 5, 397–478. [10]

LUFLADE, MARGAUX (2017): "The value of information in centralized school choice systems," . [6]

MALAMUD, OFER (2011): "Discovering one's talent: Learning from academic specialization," *Industrial and Labor Relations Review*, 64, 375–405. [6]

NARITA, YUSUKE (2018): "Match or Mismatch? Learning and Inertia in School Choice," *SSRN Electronic Journal*. [6]

OLSZEWSKI, WOJCIECH AND RAKESH VOHRA (2016): "Simultaneous selection," *Discrete Applied Mathematics*, 200, 161–169. [81]

OTERO, SEBASTIÁN, NANO BARAHONA, AND CAUÊ DOBBIN (2021): "Affirmative action in centralized college admission systems: Evidence from Brazil," *Unpublished manuscript*. [6]

PATHAK, PARAG A. AND TAYFUN SÖNMEZ (2008): "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," *American Economic Review*, 98, 1636–52. [22]

REELING, CARSON AND VALENTIN VERDIER (2021): "Welfare Effects of Dynamic Matching: An Empirical Analysis," *The Review of Economic Studies*, 89, 1008–1042. [6]

RIOS, IGNACIO, TOMÁS LARROUCAU, GIORGIOGIULIO PARRA, AND ROBERTO COMINETTI (2021): "Improving the Chilean College Admissions System," *Operations Research*, 69, 1186–1205. [8, 52, 67]

SOLIS, ALEX (2017): "Credit Access and College Enrollment," *Journal of Political Economy*, 125, 562–622. [10]

TANAKA, MARI, YUSUKE NARITA, AND CHIAKI MORIGUCHI (2020): "Meritocracy and Its Discontent: Long-run Effects of Repeated School Admission Reforms," . [6]

TINCANI, MICHELA M., FABIAN KOSSE, AND ENRICO MIGLINO (2023): "Beliefs and the Incentive Effects of Preferential College Admissions: Evidence from an Experiment and a Structural Model," Tech. rep., CEPR Discussion Paper No. 18564. [7]

WALDINGER, DANIEL (2021): "Targeting in-kind transfers through market design: A revealed preference analysis of public housing allocation," *American Economic Review*, 111, 2660–2696. [6]

WISWALL, MATTHEW AND BASIT ZAFAR (2015a): "Determinants of college major choice: Identification using an information experiment," *The Review of Economic Studies*, 82, 791–824. [7, 25]

——— (2015b): "How do college students respond to public information about earnings?" *Journal of Human Capital*, 9, 117–169. [25]

——— (2021): "Human capital investments and expectations about career and family," *Journal of Political Economy*, 129, 1361–1424. [7]

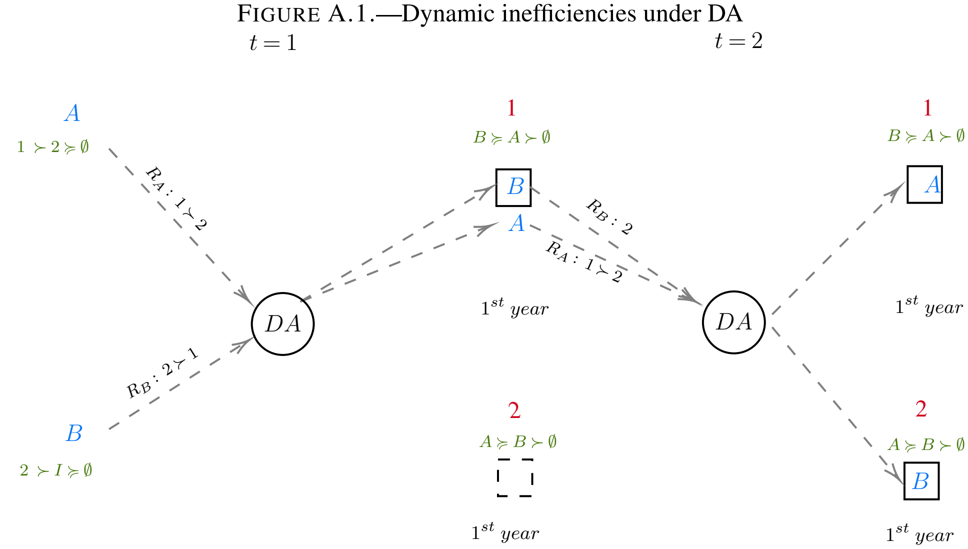
ZHANG, JUANJUAN (2010): "The sound of silence: Observational learning in the US kidney market," *Marketing Science*, 29, 315–335. [42]

APPENDIX A: 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 reapplication rules. Furthermore, we provide intuition on how switching behavior can be affected by the assignment mechanism in a dynamic setting.

Consider a centralized college admissions problem with reapplications and two periods. Let $S = \{A, B\}$ and $C = \{1, 2\}$ be the sets of students and colleges, respectively. We model admissions uncertainty by assuming that colleges announce seat availability only after students submit their applications. For simplicity, each college offers one new seat per period with probability 1/2, and no new seats otherwise. Students already enrolled may choose to retain their seat, in which case they do not consume a newly posted vacancy. As a result, a college may enroll more than one student in a given period if a returning student keeps their seat and a new seat is also offered. In addition, we assume that the preferences of colleges are $B \succ_1 A,\ A \succ_2 B$, i.e., college 1 prefers student B over A, and college 2 prefers A over B. Finally, we assume that colleges care about students' persistence and incur a cost τ per student who does not remain enrolled. This cost captures the idea that colleges make investments in their students and that the vacancy (and corresponding future tuition payments) is lost when students switch. On the other side of the market, 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 about admission probabilities. We assume that the utility of student i in college j is given by $v_{i,j} = u_{i,j} + \xi_{i,j}$, where $u_{i,j}$ is known exante and such that $u_{A,1} > u_{A,2} \gg 0$ and $u_{B,2} > u_{B,1} \gg 0$. $\xi_{i,j}$ is unknown exante but learned after the first year. We assume that this random component $\xi_{i,j}$ is equal to l with probability p, and $-l$ otherwise for each student i and college j, and we assume that this distribution is commonly known exante. We make two additional assumptions regarding the component of the utility that is learned: (i) $u_{A,1} - l < u_{A,2}$ ($u_{B,2} - l < u_{B,1}$)—i.e., if students learn that they may get a higher utility elsewhere they prefer to switch; and (ii) $u_{A,2} - l > 0$ ($u_{B,1} - l > 0$)—i.e., students prefer to enroll in their assigned colleges over the

FIGURE A.1.—Dynamic inefficiencies under DA



outside option. Then, each student i chooses the ROL $R_{i,t}$ that maximizes their expected utility in each period t.

Depending on the mechanism and the reapplication rules, students may submit different ROLs, which in turn may affect their assignment and their academic progression. To illustrate this, we compare the outcomes of two mechanisms: (i) Deferred Acceptance (DA), whereby students can apply to as many colleges as they want; and (ii) DA with no switches (DA-NS); i.e., once students are admitted they cannot reapply and switch.

DA. Under DA, both students apply according to their true preferences, i.e., $R_{A,1} = 1 \succ 2$ and $R_{B,1} = 2 \succ 1$. Then, when only one college opens a seat, one student gets assigned to their second choice while the other remains unassigned. Thus, both students may have incentives to re-apply in the next period. This situation is illustrated in Figure A.1, where we show the case when only college 1 opens a seat.

DA-NS. When no switching is allowed, students still report their true preferences when they apply to the system. However, forbidding switchings introduces a trade-off relative to DA. On the one hand, it reduces switches—eliminating the cost paid by colleges—and increases the probability that unassigned students in the first period will be admitted to their top preference in the second period. On the other hand, preventing students from switching imposes a higher cost if their utility in their current assignment is lower. Hence, it is unclear which mechanism leads to higher aggregate welfare. We formalize this result in Proposition 2.

PROPOSITION 2: *The difference between the aggregated exante welfare generated by DA-NS relative to DA is given by*

$$\Delta^{DA-NS} = \underbrace{\frac{3\cdot\beta\cdot(1-p)}{8}\cdot\tau}_{\textit{Higher Retention}} + \underbrace{\frac{\beta\cdot(1-p)}{4}\cdot\left[\frac{(u_{A,1}-u_{A,2})+(u_{B,2}-u_{B,1})}{4}\right]}_{\textit{Improvement for first-year unassigned students}} - \underbrace{\frac{3\cdot\beta\cdot(1-p)}{8}\cdot l}_{\textit{Less switches after learning}}. \tag{8}$$

The first term in (8) captures the lower cost for universities, since switches disappear under DA-NS. The second term captures the increase in students' welfare due to the higher chances of assignment to their top preference. Finally, the last term captures the negative effect of not switching when students learn that their match quality with the college is poor. Then, depending on the relative magnitude of these three components, either reapplication rule may be better.

PROOF: Proof of Proposition 2

DA. Under DA, students apply to all schools. Then, in the first period, the expected utility is

$$\frac{1}{4}\cdot 0+\frac{1}{4}\cdot\left(u_{A,1}+u_{B,2}\right)+\frac{1}{4}\cdot u_{A,2}+\frac{1}{4}\cdot u_{B,1}.$$

In the second period, the value depends on the realized assignment in the first period, all of which happens with probability $1/4$:

1. If $\mu=((A,\emptyset),(B,\emptyset))$, then the second-period expected utility is

$$\frac{1}{4}\cdot 0+\frac{1}{4}\cdot\left(u_{A,1}+u_{B,2}\right)+\frac{1}{4}\cdot u_{A,2}+\frac{1}{4}\cdot u_{B,1}.$$

2. If $\mu=((A,1),(B,2))$, there are four scenarios depending on the signals observed by the two students. More specifically, let $\xi=(\xi_A,\xi_B)$ be the signals observed by the end of period 1. Then,

- If $\xi = (l, l)$, which happens with probability p^2, both students remain enrolled. Then the expected utility in the second period is $u_{A,1} + u_{B,2} + 2l$.
- If $\xi = (l, -l)$, which happens with probability $p \cdot (1-p)$, B reapplies and switches with probability 1/2 and remains in 2 otherwise. Then the expected utility in the second period is $u_{A,1} + l + \frac{1}{2} \cdot \big(u_{B,2} - l + u_{B,1}\big) - \frac{\tau}{2}$.
- If $\xi = (-l, l)$, which happens with probability $p \cdot (1-p)$, A reapplies and switches with probability 1/2 and remains in 1 otherwise. Then the expected utility in the second period is $u_{B,2} + l + \frac{1}{2} \cdot \big(u_{A,1} - l + u_{A,2}\big) - \frac{\tau}{2}$.
- If $\xi = (-l, -l)$, which happens with probability $(1-p)^2$, both students reapply and switch only if they prefer their newly assigned program (while they remain in their current program otherwise). Then the expected utility in the second period is $\frac{1}{4} \cdot \big(u_{A,1} + u_{B,2} - 2l\big) + \frac{1}{4} \cdot \big(u_{A,2} + u_{B,2} - l\big) + \frac{1}{4} \cdot \big(u_{A,1} + u_{B,1} - l\big) + \frac{1}{4} \cdot \big(u_{A,2} + u_{B,1}\big) - \tau$.

3. If $\mu = ((A, 2), (B, \emptyset))$, only A learns, and thus there are two scenarios:
   - If $\xi_A = l$, which happens with probability p, then A stays and B reapplies. Thus, the expected utility is $u_{A,2} + l + \frac{1}{2} \cdot u_{B,2} + \frac{1}{4} \cdot u_{B,1}$.
   - If $\xi_A = -l$, which happens with probability $1-p$, then A and B reapply. Then the expected utility in the second period is $\frac{1}{4} \cdot \big(u_{A,2} - l\big) + \frac{1}{4} \cdot \big(u_{A,2} + u_{B,1} - l\big) + \frac{1}{4} \cdot \big(u_{A,2} + u_{B,2} - l\big) + \frac{1}{4} \cdot \big(u_{A,1} + u_{B,2}\big) - \frac{1}{4} \cdot \tau$.
4. If $\mu = ((A, \emptyset), (B, 1))$, only B learns, and thus there are two scenarios:
   - If $\xi_B = l$, which happens with probability p, then A reapplies and B stays. Thus, the expected utility is $u_{B,1} + l + \frac{1}{2} \cdot u_{A,1} + \frac{1}{4} \cdot u_{A,2}$.
   - If $\xi_B = -l$, which happens with probability $1-p$, then A and B reapply. Then the expected utility in the second period is $\frac{1}{4} \cdot \big(u_{B,1} - l\big) + \frac{1}{4} \cdot \big(u_{B,1} + u_{A,1} - l\big) + \frac{1}{4} \cdot \big(u_{B,1} + u_{A,2} - l\big) + \frac{1}{4} \cdot \big(u_{A,1} + u_{B,2}\big) - \frac{1}{4} \cdot \tau$.

DA-NS. Under DA-NS, the assignment is performed using DA but students are not allowed to switch in the second period. Then,

$$\frac{1}{4} \cdot 0 + \frac{1}{4} \cdot \big(u_{A,1} + u_{B,2}\big) + \frac{1}{4} \cdot u_{A,2} + \frac{1}{4} \cdot u_{B,1}.$$

In the second period, the value depends on the realized assignment in the first period, all of which happens with probability $1/4$:

1. If $\mu = ((A,\emptyset),(B,\emptyset))$, then the second-period expected utility is

$$\frac{1}{4}\cdot 0 + \frac{1}{4}\cdot\big(u_{A,1}+u_{B,2}\big) + \frac{1}{4}\cdot u_{A,2} + \frac{1}{4}\cdot u_{B,1}.$$

2. If $\mu = ((A,1),(B,2))$, there are four scenarios, as in the case for DA:
   - If $\xi = (l,l)$, which happens with probability p^2, the expected utility in the second period is $u_{A,1}+u_{B,2}+2l$.
   - If $\xi = (l,-l)$, which happens with probability $p\cdot(1-p)$, the expected utility in the second period is $u_{A,1}+u_{B,2}$.
   - If $\xi = (-l,l)$, which happens with probability $p\cdot(1-p)$, the expected utility in the second period is $u_{A,1}+u_{B,2}$.
   - If $\xi = (-l,-l)$, which happens with probability $(1-p)^2$, the expected utility in the second period is $u_{A,1}+u_{B,2}-2l$.
3. If $\mu = ((A,2),(B,\emptyset))$, only A learns, and thus there are two scenarios:
   - If $\xi_A = l$, which happens with probability p, the expected utility is $u_{A,2}+l+\frac{1}{2}\cdot u_{B,2}+\frac{1}{4}\cdot u_{B,1}$.
   - If $\xi_A = -l$, which happens with probability $1-p$, the expected utility in the second period is $u_{A,2}-l+\frac{1}{2}\cdot u_{B,2}+\frac{1}{4}\cdot u_{B,1}$.
4. If $\mu = ((A,\emptyset),(B,1))$, only B learns, and thus there are two scenarios:
   - If $\xi_B = l$, which happens with probability p, the expected utility is $u_{B,1}+l+\frac{1}{2}\cdot u_{A,1}+\frac{1}{4}\cdot u_{A,2}$.
   - If $\xi_B = -l$, which happens with probability $1-p$, the expected utility in the second period is $u_{B,1}-l+\frac{1}{2}\cdot u_{A,1}+\frac{1}{4}\cdot u_{A,2}$.

Then, taking the difference for each scenario, we obtain no differences in the expected utility in the first period. For the second period we obtain that:

1. If $\mu = ((A,\emptyset),(B,\emptyset))$, there is no difference between DA and DA-NS.
2. If $\mu = ((A,1),(B,2))$, there are four scenarios as in the case for DA:
   - If $\xi = (l,l)$, then the difference in expected utility is zero.

- If $\xi = (l, -l)$, then the difference in expected utility is

$$u_{A,1} + l + \frac{1}{2} \cdot \big(u_{B,2} - l + u_{B,1}\big) - \frac{\tau}{2} - \big(u_{A,1} + u_{B,2}\big)$$
$$= \frac{u_{B,2} - u_{B,1}}{2} + \frac{l - \tau}{2}.$$

- If $\xi = (-l, l)$, then the difference in expected utility is

$$u_{B,2} + l + \frac{1}{2} \cdot \big(u_{A,1} - l + u_{A,2}\big) - \frac{\tau}{2} - \big(u_{A,1} + u_{B,2}\big)$$
$$= \frac{u_{A,1} - u_{A,2}}{2} + \frac{l - \tau}{2}.$$

- If $\xi = (-l, -l)$, then the difference in expected utility is

$$\frac{1}{4} \cdot \big(u_{A,1} + u_{B,2} - 2l\big) + \frac{1}{4} \cdot \big(u_{A,2} + u_{B,2} - l\big) + \frac{1}{4} \cdot \big(u_{A,1} + u_{B,1} - l\big)$$
$$+ \frac{1}{4} \cdot \big(u_{A,2} + u_{B,1}\big) - \tau - \big(u_{A,1} + u_{B,2} - 2l\big)$$
$$= \frac{u_{A,1} - u_{A,2}}{2} + \frac{u_{B,2} - u_{B,1}}{2} + l - \tau.$$

3. If $\mu = ((A, 2), (B, \emptyset))$, only A learns, and thus there are two scenarios:
   - If $\xi_A = l$, then the difference in expected utility is zero.
   - If $\xi_A = -l$, then the difference in expected utility is

$$\frac{1}{4} \cdot \big(u_{A,2} - l\big) + \frac{1}{4} \cdot \big(u_{A,2} + u_{B,1} - l\big) + \frac{1}{4} \cdot \big(u_{A,2} + u_{B,2} - l\big) + \frac{1}{4} \cdot \big(u_{A,1} + u_{B,2}\big) - \frac{1}{4} \cdot \tau$$
$$- \left(u_{A,2} - l + \frac{1}{2} \cdot u_{B,2} + \frac{1}{4} \cdot u_{B,1}\right)$$
$$= \frac{u_{A,1} - u_{A,2}}{4} + \frac{l - \tau}{4}.$$

4. If $\mu = ((A, \emptyset), (B, 1))$, only B learns, and thus there are two scenarios:
   - If $\xi_B = l$, then the difference in expected utility is zero.

- If $\xi_B = -l$, then the difference in expected utility is

$$\frac{1}{4}\cdot\left(u_{B,1}-l\right)+\frac{1}{4}\cdot\left(u_{B,1}+u_{A,1}-l\right)+\frac{1}{4}\cdot\left(u_{B,1}+u_{A,2}-l\right)+\frac{1}{4}\cdot\left(u_{A,1}+u_{B,2}\right)-\frac{1}{4}\cdot\tau$$
$$-\left(u_{B,1}-l+\frac{1}{2}\cdot u_{A,1}+\frac{1}{4}\cdot u_{A,2}\right)$$
$$=\frac{u_{B,2}-u_{B,1}}{4}+\frac{l-\tau}{4}.$$

Finally, multiplying by the corresponding probabilities and adding up terms, we obtain that:

$$DA-DA-NS=\frac{\beta\cdot(1-p)}{4}\cdot\left(\frac{u_{A,1}-u_{A,2}}{4}+\frac{u_{B,2}-u_{B,1}}{4}+\frac{3}{2}\cdot(l-\tau)\right).$$

Therefore,

$$\Delta^{DA-NS}=DA-NS-DA=\frac{\beta\cdot(1-p)}{4}\cdot\left(\frac{u_{A,2}-u_{A,1}}{4}+\frac{u_{B,1}-u_{B,2}}{4}+\frac{3}{2}\cdot(\tau-l)\right).$$

Q.E.D.

Notice that these theoretical examples assume 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.2 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.

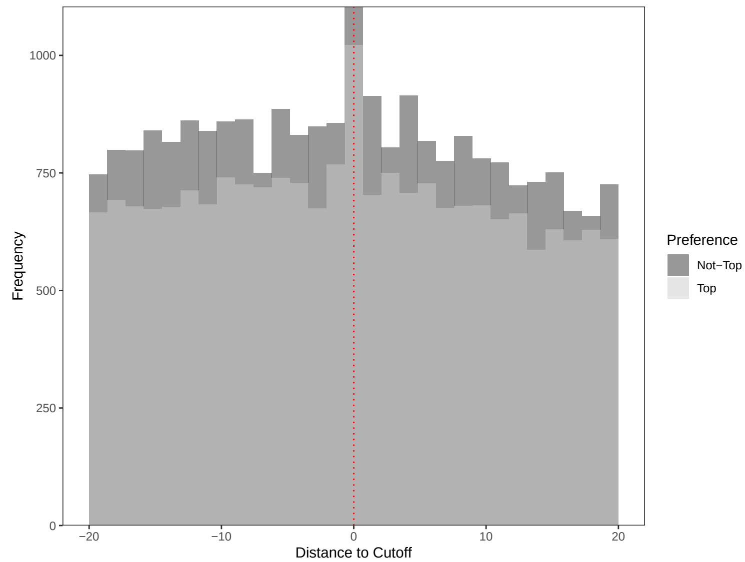
APPENDIX B: APPENDIX TO SECTION 3

B.1. *The Chilean Mechanism*

The Chilean mechanism is a variant of the student-proposing Deferred Acceptance algorithm[10] in which students who tie for the last seat of a program are not rejected if vacancies are exceeded. More formally, the allocation rule can be described as follows:

[10]Before 2014; the algorithm used was the university-proposing version. The assignment differences between both implementations of the algorithm are negligible Rios et al. (2021).

FIGURE A.2.—Distribution of preference of assignment around admission cutoffs



Step 1**.** Each student proposes to their 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 on their ROL that did not reject them. As a side outcome of this, the algorithm generates a set of cutoffs $\{\bar{s}_j\}_{j\in\mathcal{J}}$, where $\bar{s}_j$ is the minimum application score among students matched to program $j \in \mathcal{J}$. Hence, student i with ROL R_i and set of scores $\{s_{i,j}\}_{j\in\mathcal{J}}$ gets allocated to program j if and only if (i) $s_{i,j} \geq \bar{s}_j$ and (ii) $s_{i,j'} < \bar{s}_{j'}$ for all $j' \in R_i$ such that $j' \succ_i j$, i.e., student i's score is less than the cutoff in all programs they prefer over j. The fact that the allocation can be described in terms of cutoffs is known as the *cutoff structure* of the algorithm, and it is a key feature of DA that we exploit in our empirical analysis and estimation.

B.2. *Surveys*

All surveys discussed in Section 3.1were sent to all students who participated in the entrance exam (more than 150,000 each year). In 2019 and 2020, MINEDUC sent the survey right at the end of the application time window; in 2023, we sent the pre-application survey a few days before the application window opened and the post-application survey right after it closed (similar to those in previous years).

To identify model parameters for the 2014 admissions cohort using preference and belief moments from these surveys, we assume that the distributions of preferences and beliefs are stationary across cohorts, conditional on key observables such as gender, income, test scores, and broad major. This stationarity assumption allows us to map belief and preference data from later cohorts onto the 2014 sample.

B.3. *Regression discontinuities*

We use a regression discontinuity design that exploits the algorithm's cutoff structure to estimate the causal effect of being assigned to a higher preference on students' outcomes. In particular, we estimate the following specification:

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

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. Similar results are obtained running these models at the student-preference level.

Notice that many of the outcomes we consider—e.g., switches and dropouts, among others—rely on students enrolling in the centralized system. If there are significant differences in the enrollment rates between students right above and below the cutoff, then the two samples would not be directly comparable, leading to a selection problem (Dong, 2017). To show that this is not the case, in Figure B.1a we show the binned means of the probability of enrolling in the centralized system as a function of distance to the cutoff. In addition, the line represents the predicted values obtained from estimating the regression discontinuity model described in (B.3), considering as dependent variable the probability of enrolling in the centralized system. As Figure B.1a and Column (1) in Table B.I show, there are no significant differences in the enrollment probabilities between students above and below the cutoff, so we conclude that selection problem is not a concern.

Figure B.1 also displays binned means for different outcomes as a function of the distance between the cutoffs in their top preference and the students' scores, and Table B.I reports the corresponding estimation results. In all these analyses, we consider students who applied to at least two programs in the centralized system and focus on their top preference of each student for simplicity. Similar results follow from considering other preferences. As shown in Figure B.1b and Column (2) in Table B.I, 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.1c and B.1d are discussed in Section 2, and show that being above the cutoff significantly reduces the probability of reapplying and switching programs within the system. These results are confirmed in Columns (3) and (4) in Table B.I. Figure B.1f and Column (5) in Table B.I exhibit a similar pattern, as it shows that the probability of major-switching also decreases among students above the cutoff. Similarly, Column (6) shows that university switching is also significantly reduced. Finally, in Figure B.1g, we show that there is no effect of exceeding the cutoff on dropout rates.

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 preference on different outcomes. Using the 2019 cohort 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 preference on their outcomes. In Figures B.2a and B.2b,

FIGURE B.1.—Effect of Cutoff Crossing

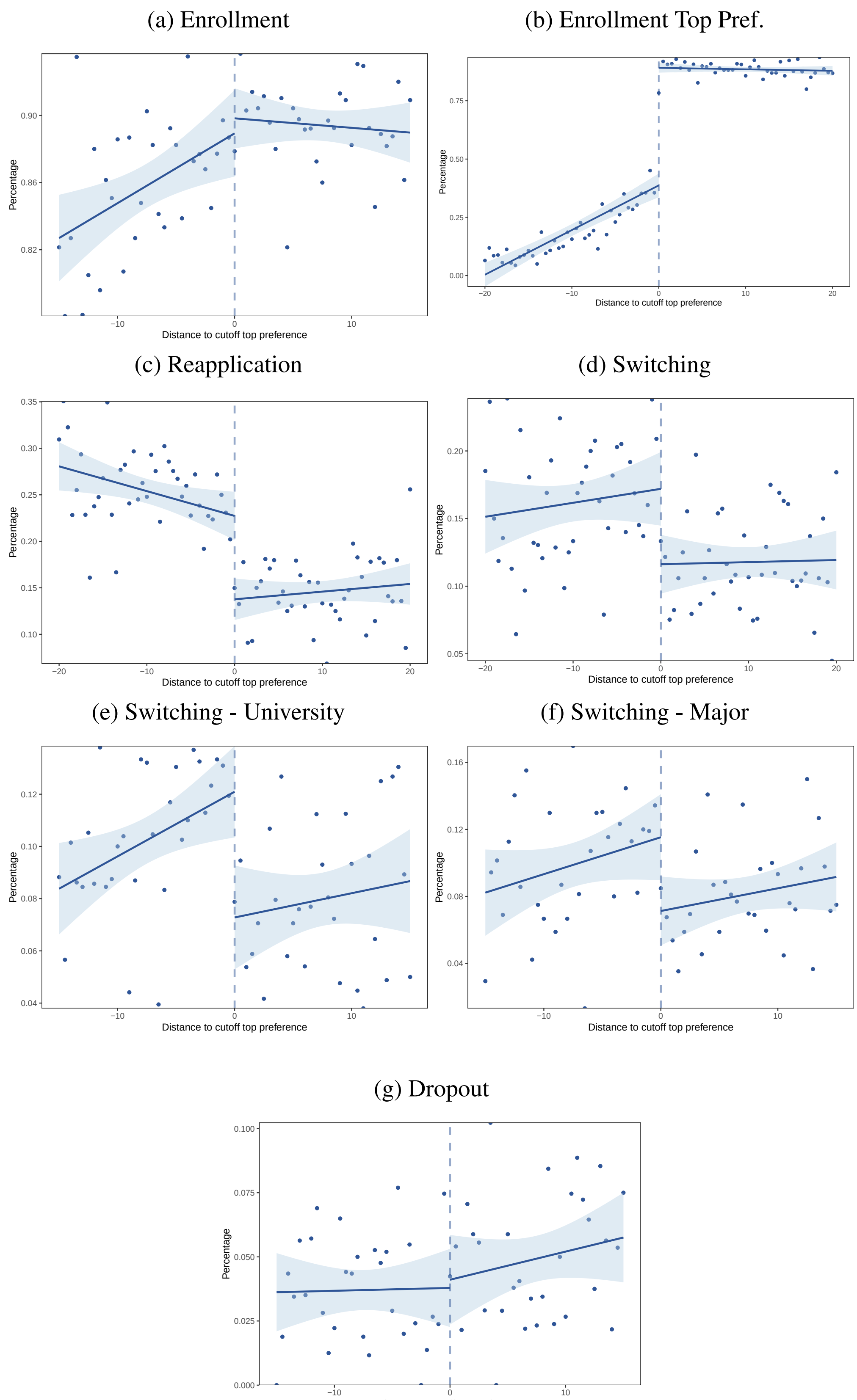


TABLE B.I

CAUSAL EFFECT OF CROSSING THE CUTOFF IN TOP REPORTED PREFERENCE

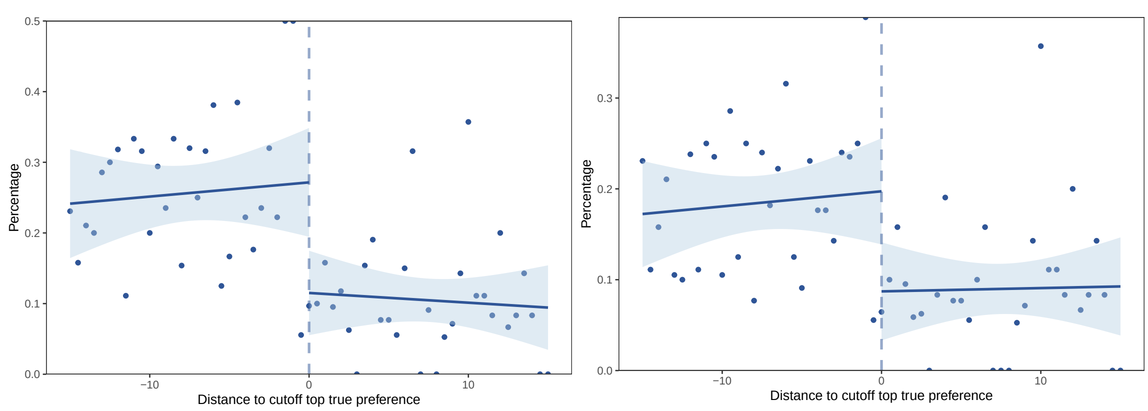
	Enroll (1)	Enroll Top Pref. (2)	Reapp. (3)	Switch (4)	Switch Major (5)	Switch University (6)	Dropout (7)
Z_{ip}	0.028	0.514^{***}	-0.084^{***}	-0.075^{***}	-0.044^{***}	-0.049^{***}	0.008
	(0.018)	(0.019)	(0.020)	(0.020)	(0.016)	(0.016)	(0.012)
Constant	0.867^{***}	0.370^{***}	0.228^{***}	0.184^{***}	0.115^{***}	0.123^{***}	0.035^{***}
	(0.013)	(0.014)	(0.014)	(0.015)	(0.012)	(0.012)	(0.009)
Mean	0.864	0.565	0.199	0.14	0.09	0.091	0.043
Observations	5,426	5,426	6,050	4,686	4,686	4,686	4,686

Note: Standard errors in parenthesis.

being above the cutoff significantly reduces the probability of reapplying to the system and being assigned to a different program the next year, consistent with those reported in Figure B.1.

FIGURE B.2.—Effect of Cutoff Crossing for Top-True Preference

(a) Reapplication (b) Reassignment



B.4. *Additional Evidence*

B.4.1. *Perceived persistence and preference of assignment*

The regression discontinuity results show a causal effect of preference of assignment on students' persistence with respect to their initial assignment. To show that this is partially

FIGURE B.3.—Students flow across states

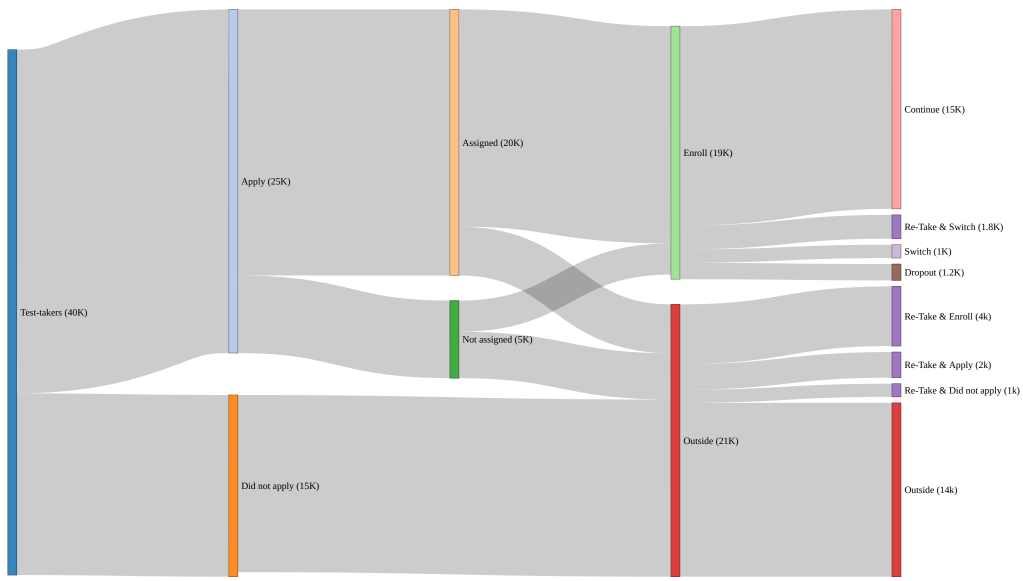
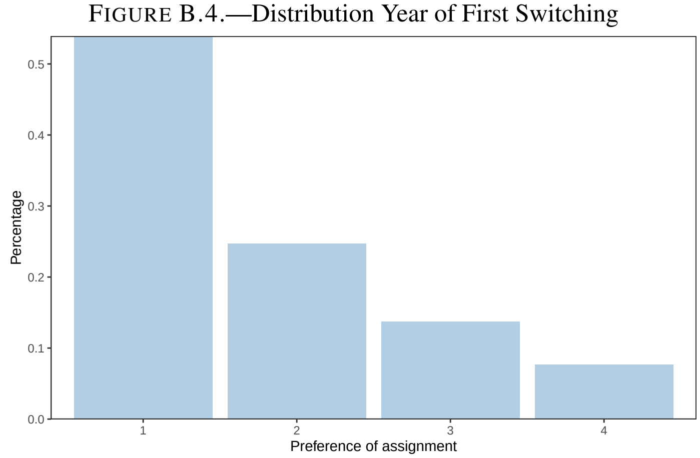


FIGURE B.4.—Distribution Year of First Switching



explained by the mismatch channel, we use the 2020 survey on students' preferences and beliefs. Figure B.5 shows that the average "perceived" probability of remaining enrolled in the same program after the first year is a significantly lower for lower-ranked preferences.

On average, students believe that there is an 85% probability of staying enrolled in their top reported preference, whereas it is close to 65% for programs ranked below the fourth choice. Futhermore, comparing Figure B.5 with Figure B.6, which provides actual probabilities of remaining enrolled, shows evidence of forward-looking behavior and suggests that—on aggregate—students' subjective beliefs are close to being correct.

FIGURE B.5.—Average "perceived" probability of remaining enrolled in the same program, by preference of enrollment

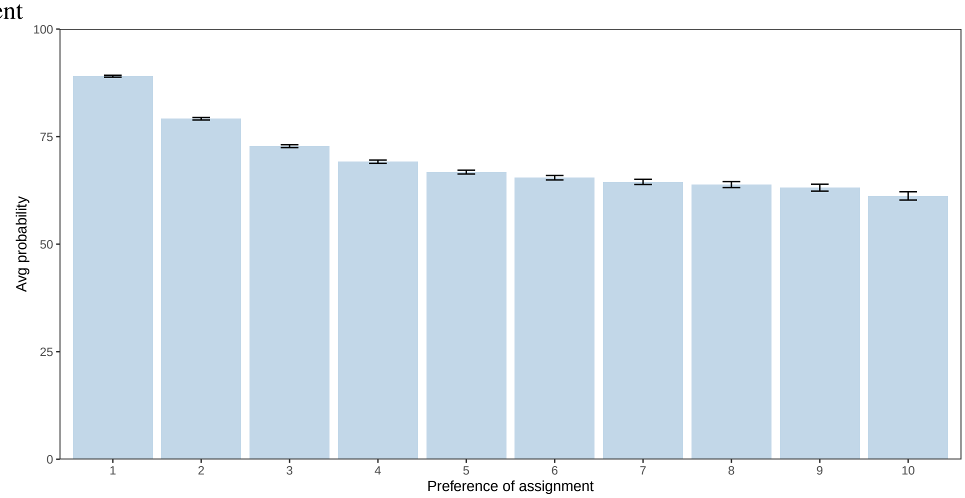
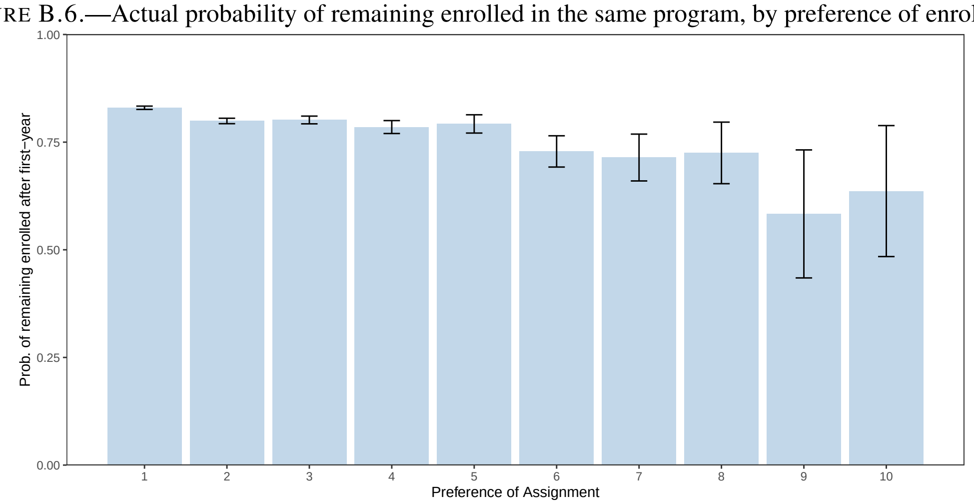


FIGURE B.6.—Actual probability of remaining enrolled in the same program, by preference of enrollment



The previous evidence does not guarantee that there are match effects that are correlated with college persistence. For instance, a similar pattern would arise if all students had identical preferences over programs—determined solely by program-level characteristics that are highly correlated with persistence—so that lower-ranked programs systematically have lower retention rates regardless of individual student-program matches. To rule out this alternative explanation and provide evidence of match effects, we exploit the panel structure of students' ROLs—we elicited perceived persistence probability for every program listed in the ROL—and estimate the following model:

$$\tilde{P}^r_{i,R_i(l)} = \psi^r_i + \psi^r_l + \psi^r_{R_i(l)} + \psi^r_X X^r_{i,R_i(l)} + \epsilon^r_{i,l},$$

where R_i is student i's ROL and $R_i(l)$ is the program listed in the l-th position; $\tilde{P}^r_{i,R(l)}$ is their perceived probability of remaining enrolled in program $R_i(l)$, with $l \in \{1,\ldots,|R_i|\}$; ψ^r_i is students i's fixed effect; ψ^r_l is position l's fixed effect; ψ^r_j is program j's fixed effect; $X^r_{i,j}$ are time-invariant observable characteristics of student i and program j (third-degree polynomial of the application score of student i in program j); and $\epsilon^r_{i,l}$ is an i.i.d shock.

In Table B.II, we report the estimation results. We observe that the preference of enrollment has a significant and strong effect on the perceived probability of remaining enrolled, so we conclude that there are match effects in the setting, which exhibit a strong correlation with students' college persistence. Overall, these results show that (i) there is a significant effect of the preference of assignment on the switching behavior of students, (ii) a significant fraction of students correctly forecast this, and (iii) that match effects (at least) partially explain these patterns.

APPENDIX C: APPENDIX FOR SECTION 4

C.1. *Proof of Proposition 1*

Let W,X,Y,Z be independent random variables normally distributed with means μ_W,μ_X,μ_Y,μ_Z and variances $\sigma^2_W,\sigma^2_X,\sigma^2_Y,\sigma^2_Z$, and let $A = \alpha_W W + \alpha_X X + \alpha_Y Y + \alpha_Z Z$. Using properties of the normal distribution, we know that A is also normally distributed with mean $\mu_A := \alpha_W\mu_W + \alpha_X\mu_X + \alpha_Y\mu_Y + \alpha_Z\mu_Z$ and variance $\sigma^2_A := \alpha^2_W\sigma^2_W + \alpha^2_X\sigma^2_X +$

TABLE B.II

TWO-WAY FIXED EFFECTS REGRESSION RESULTS

	Dependent variable: Prob. of Persistence	
	Estimate	Std. Error
Preference 2	-9.891	8.3652e-03
Preference 3	-16.844	8.7019e-03
Preference 4	-21.355	9.3909e-03
Preference 5	-24.831	1.0631e-02
Preference 6	-27.148	1.1805e-02
Preference 7	-29.164	1.3202e-02
Preference 8	-30.329	1.5004e-02
Preference 9	-31.995	1.6876e-02
Preference 10	-34.757	1.9483e-02
Constant	89.181	1.0186e-02
Observations	159,894	
R^2	0.095	
Adjusted R^2	0.095	

Notes: Standard errors in parentheses. The table reports estimates from a two-way fixed effects regression of students' perceived probability of remaining enrolled in the same program on preference rank indicators, controlling for student fixed effects, program fixed effects, and a third-degree polynomial of the application score. The dependent variable is the perceived probability of remaining enrolled (0-100 scale) elicited from the 2020 survey.

$\alpha_Y^2\sigma_Y^2 + \alpha_Z^2\sigma_Z^2$. Moreover, we know that

$$\mathbb{E}\left[X \mid A = a\right] = \mu_X + \frac{cov(X, A)}{\sigma_X \cdot \sigma_A} \cdot \frac{\sigma_X}{\sigma_A} \cdot (a - \mu_A)\,.$$

By definition, we know that

$$
\begin{aligned}
cov(X,A) &= \mathbb{E}[X\cdot A] - \mathbb{E}[X]\cdot\mathbb{E}[A] \\
&= \mathbb{E}[X\cdot(\alpha_W W + \alpha_X X + \alpha_Y Y + \alpha_Z Z)] - \mathbb{E}[X]\cdot\mathbb{E}[\alpha_W W + \alpha_X X + \alpha_Y Y + \alpha_Z Z] \\
&= \alpha_X \mathbb{E}\left[X^2\right] + \alpha_W \mathbb{E}[XW] + \alpha_Y \mathbb{E}[XY] + \alpha_Z \mathbb{E}[XZ] \\
&\quad - \alpha_X \mathbb{E}[X]^2 - \alpha_W \mathbb{E}[X]\mathbb{E}[W] - \alpha_Y \mathbb{E}[X]\mathbb{E}[Y] - \alpha_Z \mathbb{E}[X]\mathbb{E}[Z] \\
&= \alpha_X \mathbb{E}\left[X^2\right] - \alpha_X \mathbb{E}[X]^2 \\
&= \alpha_X \sigma_X^2
\end{aligned}
$$

where the first equality is by definition; the second follows from the linearity of the expectation; the third follows from independence; and the fourth follows from the definition of the variance. Hence,

$$
\mathbb{E}[X \mid A = a] = \mu_X + \frac{\alpha_X \sigma_X^2}{\sigma_X \cdot \sigma_A} \cdot \frac{\sigma_X}{\sigma_A} \cdot (a - \mu_A) = \mu_X + \alpha_X \cdot \frac{\sigma_X^2}{\sigma_A^2} \cdot (a - \mu_A). \quad (9)
$$

Next, let $\tilde{\mu}'_{\theta j'}(a) := \mathbb{E}\left[A^u_{\theta j'} \mid a_{ijt} = a\right]$, where we may have that $j \neq j'$ (if the student switches). Then, if $\theta(i) = \theta$,

$$
\begin{aligned}
\tilde{\mu}'_{\theta j'}(a) &= \mathbb{E}\left[\tilde{A}^u_{\theta m_{j'}} + \sum_{k\in\{s_m,s_v\}} \omega_{j'k}\cdot \tilde{A}^u_{\theta k} + \varepsilon_{ij't+1} \mid a_{ijt} = a\right] \\
&= \mathbb{E}\left[\tilde{A}^u_{\theta m_{j'}} + \sum_{k\in\{s_m,s_v\}} \omega_{j'k}\cdot \tilde{A}^u_{\theta k} \mid a_{ijt} = a\right] \\
&= \mathbb{E}\left[\tilde{A}^u_{\theta m_{j'}} \mid a_{ijt} = a\right] + \sum_{k\in\{s_m,s_v\}} \omega_{j'k}\cdot \mathbb{E}\left[\tilde{A}^u_{\theta k} \mid a_{ijt} = a\right]
\end{aligned}
$$

Note that, since the student knows their observable characteristics and their private abilities, the signal is such that

$$\begin{aligned}
a_{ijt} &= G_{ijt} - \gamma_{m_j} - \gamma_2 \cdot A_{ij} - A^p_{ij} \\
&= \tilde{A}^u_{\theta m_j} + \sum_{k \in \{s_m, s_v\}} \omega_{jk} \cdot \tilde{A}^u_{\theta k} + \varepsilon_{ijt} \\
&\sim N\left(\tilde{\mu}_{\theta m_j} + \sum_{k \in \{s_m, s_v\}} \omega_{jk} \cdot \tilde{\mu}_{\theta k}, \quad \tilde{\sigma}^2_{\theta m_j} + \sum_{k \in \{s_m, s_v\}} \omega^2_{jk} \cdot \tilde{\sigma}^2_{\theta k} + \sigma^2_g\right) \\
&\sim N\left(\mu_{\theta j}, \quad \tilde{\sigma}^2_{\theta j} + \sigma^2_g\right)
\end{aligned}$$

where $\tilde{\mu}_{\theta j} := \tilde{\mu}_{\theta m_j} + \sum_{k \in \{s_m, s_v\}} \omega_{jk} \cdot \tilde{\mu}_{\theta k}$ and $\tilde{\sigma}^2_{\theta j} := \tilde{\sigma}^2_{\theta m_j} + \sum_{k \in \{s_m, s_v\}} \omega^2_{jk} \cdot \tilde{\sigma}^2_{\theta k}$.

Note that we have four independent random variables $\tilde{A}^u_{\theta m_j}, \tilde{A}^u_{\theta s_m}, \tilde{A}^u_{\theta s_v}$ and ε_{ijt}, and we have another random variable a_{ijt} given by the weighted sum of these four ones considering as weights $\{1, \omega_{js_m}, \omega_{js_v}, 1\}$. Hence, we can use (9) to compute $\tilde{\mu}'_{\theta j'}$. Specifically, noticing that $\tilde{A}^u_{\theta m_{j'}} \perp \tilde{A}^u_{\theta m_j}, \tilde{A}^u_{\theta s_m}, \tilde{A}^u_{\theta s_v}, \varepsilon_{ijt}$ if $m_{j'} \neq m_j$ and $\tilde{A}^u_{\theta m_{j'}} \perp \tilde{A}^u_{\theta s_m}, \tilde{A}^u_{\theta s_v}, \varepsilon_{ijt}$ otherwise, we know that

$$\mathbb{E}\left[\tilde{A}^u_{\theta m_{j'}} \mid a_{ijt} = a\right] = \begin{cases} \tilde{\mu}_{\theta m_{j'}} & \text{if } m_j \neq m_{j'} \\ \tilde{\mu}_{\theta m_{j'}} + \frac{\tilde{\sigma}^2_{\theta m_{j'}}}{\tilde{\sigma}^2_{\theta j} + \sigma^2_g} \cdot \left(a - \tilde{\mu}_{\theta j}\right) & \text{if } m_j = m_{j'} \end{cases}$$

Similarly, for $k \in \{s_m, s_v\}$,

$$\mathbb{E}\left[\tilde{A}^u_{\theta k} \mid a_{ijt} = a\right] = \tilde{\mu}_{\theta k} + \omega_{jk} \cdot \frac{\tilde{\sigma}^2_{\theta k}}{\tilde{\sigma}^2_{\theta j} + \sigma^2_g} \cdot \left(a - \tilde{\mu}_{\theta j}\right)$$

Combining these two equations we obtain that

$$\tilde{\mu}'_{\theta j'}(a) = \mathbb{E}\left[\tilde{A}^u_{\theta m_{j'}} \mid a_{ijt} = a\right] + \sum_{k\in\{s_m,s_v\}} \omega_{j'k} \cdot \mathbb{E}\left[\tilde{A}^u_{\theta k} \mid a_{ijt} = a\right]$$

$$= \begin{cases} \tilde{\mu}_{\theta m_{j'}} + \sum_{k\in\{s_m,s_v\}} \omega_{j'k}\left(\tilde{\mu}_{\theta k} + \omega_{jk} \cdot \frac{\tilde{\sigma}^2_{\theta k}}{\tilde{\sigma}^2_{\theta j}+\sigma^2_g} \cdot \left(a - \tilde{\mu}_{\theta j}\right)\right) & \text{if } m_j \neq m_{j'} \\ \tilde{\mu}_{\theta m_{j'}} + \frac{\tilde{\sigma}^2_{\theta m_{j'}}}{\tilde{\sigma}^2_{\theta j}+\sigma^2_g} \cdot \left(a - \tilde{\mu}_{\theta j}\right) + \sum_{k\in\{s_m,s_v\}} \omega_{j'k}\left(\tilde{\mu}_{\theta k} + \omega_{jk} \cdot \frac{\tilde{\sigma}^2_{\theta k}}{\tilde{\sigma}^2_{\theta j}+\sigma^2_g} \cdot \left(a - \tilde{\mu}_{\theta j}\right)\right) & \text{if } m_j = m_{j'} \end{cases}$$

$$= \begin{cases} \tilde{\mu}_{\theta j'} + \sum_{k\in\{s_m,s_v\}} \omega_{j'k}\omega_{jk} \cdot \frac{\tilde{\sigma}^2_{\theta k}}{\tilde{\sigma}^2_{\theta j}+\sigma^2_g} \cdot \left(a - \tilde{\mu}_{\theta j}\right) & \text{if } m_j \neq m_{j'} \\ \tilde{\mu}_{\theta j'} + \left[\frac{\tilde{\sigma}^2_{\theta m_{j'}}}{\tilde{\sigma}^2_{\theta j}+\sigma^2_g} + \sum_{k\in\{s_m,s_v\}} \omega_{j'k}\omega_{jk} \cdot \frac{\tilde{\sigma}^2_{\theta k}}{\tilde{\sigma}^2_{\theta j}+\sigma^2_g}\right] \cdot \left(a - \tilde{\mu}_{\theta j}\right) & \text{if } m_j = m_{j'} \end{cases}$$

C.2. *Additional Decisions*

In addition to the initial application decision discussed in Section 4.4, students also make the following decisions:

C.2.1. *Enrollment.*

Consider a student i who is enrolled in program j at the beginning of period $t = 2$ and decides to reapply, getting assigned to some program $j' \in \mathcal{J} \cup \{\emptyset\}$ as a result. Then, the expected utility of student i is given by

$$\Psi_{i,j,t}(j') := \left(v_{i,j',t} - C^S_{i,t}\right) \cdot P^E_{i,t} + \max\left\{v_{i,j,t}, v_{i,\emptyset,t}\right\} \cdot \left(1 - P^E_{i,t}\right), \tag{10}$$

where $P^E_{i,t}$ is the enrollment probability of student i in period t; $C^S_{i,t}$ is the cost of starting a new program; and $v_{i,j',t}$ is the expected utility that student i receives from being enrolled in program j' from period t onwards, which is given by

$$v_{i,j,t} = \sum_{t'=t}^{T}\left[\mathbb{E}\left[V_{i,j,t'}\right] \cdot \frac{P^g_{i,j,t'}}{P^g_{i,j,t'} + P^d_{i,j,t'}} + \mathbb{E}\left[V_{i,\emptyset,t'}\right] \cdot \frac{P^d_{i,j,t'}}{P^g_{i,j,t'} + P^d_{i,j,t'}} + \sum_{\tau=t}^{t'} \delta^{\tau-t} \cdot \mathbb{E}\left[U_{i,j,\tau}\right]\right] \cdot P^{g\vee d}_{i,j,t'}, \tag{11}$$

where $P^g_{i,j,t'}$ and $P^d_{i,j,t'}$ are the probabilities of graduating and dropping out from program j in period t' conditional on being enrolled, respectively, and $P^{g\vee d}_{i,j,t'} := (P^g_{i,j,t'} + P^d_{i,j,t'}) \cdot \prod_{\tau=1}^{t'-1}(1 - P^g_{i,j,\tau} - P^d_{i,j,\tau})$ is the unconditional probability of student i graduating or dropping out from program j in period t'. The first term in (11) captures the expected workforce utility of graduating from program j in period t', while the second term captures the expected utility of dropping out in period t'. The third term captures the present value of the flow utility that student i receives from attending program j from period t to t'. Finally, if $j' = \emptyset$ (i.e., the student does not get assigned to a new program), we consider $\Psi_{i,j,t}(\emptyset) = \max\left\{v_{i,j,t+1}, v_{i,\emptyset,t+1}\right\}$, and with a slight abuse of notation, we define

$$\Psi_{i,j,t}(R) := \sum_{l=1}^{|R|} p_{i,R(l),t} \cdot \bar{p}_{i,R(l),t} \cdot \Psi_{i,j,t}(R(l)),$$

where $p_{i,j,t}$ is the probability of getting assigned to program j in period t; $\bar{p}_{i,R(l),t} = \prod_{l'=1}^{l-1}(1 - p_{i,R(l'),t})$ is the probability of not being assigned to any of the top $l-1$ programs listed in R, with $\bar{p}_{i,R(1),t} = 1$.[11] With a slight abuse of notation, we denote by $p_{j,t}$ the function that returns the CDF of the admission probability to program j in period t, i.e., $p_{j,t}(s)$ is the probability of admission given a score s.

C.2.2. *Re-application.*

In period $t = 2$, each student i who is strategic and decides to re-apply solves the following optimal portfolio problem:

$$R^{\star}_{i,j,t} \in \underset{R \subseteq \mathcal{R}_{i,j,t}}{\operatorname{argmax}} \left\{\Psi_{i,j,t}(R) - C^A_{i,t}(R)\right\}$$

where $\mathcal{R}_{i,j,t} := \left\{R \in \mathcal{R}\left(\mathcal{J} \setminus \{j\}\right) : |R| \leq K,\ v_{i,j',t+1} - C^S_{i,t+1} > v_{i,j,t+1},\ p_{i,j',t>0} \forall j' \in R\right\}$ is the set of ROLs that include at most K programs that have positive admission probability and that are strictly prefer than program j starting from period $t+1$, accounting for the cost $C^S_{i,t}$ that student i incurs upon starting a new program; and $C^A_{i,t}(R) = c\,|R|$ is the cost of applying to ROL R. With a slight abuse of notation, we say that $R^{\star}_{i,j,t} = \{\emptyset\}$ if

[11]For simplicity, we assume that enrollment, graduation and dropout occur according to a semi-exogenous statistical process that depends on observable characteristics, as discussed in Appendix C.3. Furthermore, we assume that students neglect potential correlations across cutoff distributions and that strategic students have correct beliefs about the marginal admission probabilities. Finally, note that (11) correctly captures the indirect utility of enrolling in program j, as there are no decisions after period 2.

student i does not re-apply in period t. Note that students include programs in their ROL only if it is strictly profitable, so they will skip programs for which their chances are zero. This assumption is for simplicity, as it prevents from having multiplicity of optimal ROLs involving zero-chance programs.

C.2.3. *Exam retaking.*

We assume that students endogenously decide whether to retake the entrance exam in period $t = 1$, weighing the expected benefits—e.g., increasing their chances of admission into a more preferred program—against the associated costs. This decision shapes the application behavior discussed above, which is conditional on a set of scores determining admission probabilities $\{p_{i,j,t}\}_{j\in\mathcal{J}}$. Formally, student i decides whether to retake the exam in period $t = 1$ by solving:

$$\max_{x\in\{0,1\}} \left\{ \mathbb{E}\left[\Psi_{i,j,t+1}\left(R^{\star}_{i,j,t+1}\right) - C^{A}_{i,t+1}(R^{\star}_{i,j,t+1}) \mid x \right] - x \cdot C^{T}_{i,t} \right\}. \tag{12}$$

The first term is the expected utility of re-applying in period $t+1$, while the second is the cost of taking the exam. Importantly, the expectation is taken over the distribution of the signal about the student's unknown ability (learning about abilities), the preferences shocks and random coefficients (learning about preferences), as well as the distribution of test scores that the student receives, which is affected by the decision of taking the exam. Note that the latter affects the admission probabilities considered in determining $R^{\star}_{i,j,t+1}$ (with $R^{\star}_{i,j,t+1} = \{\emptyset\}$ representing the option of not re-applying), evaluating $\Psi_{i,j,t+1}$, and also in the observed ability that student i considers for later outcomes. Finally, with a slight abuse of notation, let $x^{\star}_{i,j}$ be the optimal solution of (12) and $\Psi^{\star}_{i,j}$ the corresponding value of the objective.

C.3. *Exogenous Components*

C.3.1. *Admission Probabilities*

To compute the admission probabilities, we use a bootstrap procedure similar to that in Agarwal and Somaini (2018). Specifically, we perform 1,000 bootstrap simulations, each consisting of the following steps:

1. Sample with replacement a number of applicants equal to the total number of students that applied in the admissions process.
2. Run the assignment mechanism used in the Chilean system (see Rios et al. (2021)).
3. Compute the cutoff of each program for both the regular and BEA admission tracks.

As a result of this procedure, we obtain two matrices (for the regular and BEA processes) with 1,000 cutoffs for each program. The next step is to estimate the distribution of the cutoff of each program in each admission track. To accomplish this, we estimate the parameters of a truncated normal distribution for each program and admission track via maximum likelihood. Then, using the estimated distributions, we evaluate the CDF on the application score of the student to obtain an estimate of the admission probability, taking into account whether the student participates only in the regular process or also in the BEA track.

As discussed in Section 4.4, these estimated admission probabilities are considered by strategic students when making their application decisions. In contrast, we assume that truth-tellers consider an admission probability equal to one for all programs and continuation values obtained assuming pairwise stability. Although there is evidence that students have biased beliefs and make applications mistakes (Larroucau et al., 2025), our counterfactuals do not aim to correct these and, instead, focus on the dynamic aspects of the mechanism, re-application rules, and the information environment over future outcomes. For this reason, we abstract from this for simplicity and assume correct beliefs.

C.3.2. *Enrollment*

As discussed in Section C.2.1, we assume for simplicity that the decision to enroll in a newly assigned program is exogenously given, while the decision to remain in the current program or take the outside option conditional on not enrolling in the newly assigned program is endogenous. Specifically, our model assumes that students enroll in their newly assigned program with some probability $P^e_{i,t}$ that depends on a vector of time-invariant observable characteristics X^e_i as follows:

$$P^e_{i,t} = \frac{\exp\left(X^e_i \psi^e\right)}{1 + \exp\left(X^e_i \psi^e\right)}.$$

For estimation, we include in X_i^e 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.

Note that we only incorporate students' observable characteristics in the estimation of enrollment probabilities. This is because these probabilities multiply the continuation values as shown in Equation 10, so making them program specific would not allow us to remove those from the expectation and, more importantly, it would break the type I extreme value distribution of shocks and, consequently, the inclusive value formulas would not hold, preventing us from using the approach discussed in Section C.4.

C.3.3. *Dropout and Graduation*

We assume that graduation and dropout after the period 2 are exogenously. Specifically, we assume that conditional on being enrolled in program j in for the last τ periods, student i's probability of graduating or dropping out depends on vectors of time-invariant observable characteristics $X_{i,j}^{g\vee d}$, which may include characteristics of student i, program j, and the interaction between their characteristics. Furthermore, we estimate these probabilities using a multinomial logit model, as follows:

$$P_{i,j,\tau}^{g} = \frac{\exp\left(X_{i,j}^{g\vee d}\psi_\tau^g\right)}{1+\sum\limits_{a\in\{g,d\}}\exp\left(X_{i,j}^{g\vee d}\psi_\tau^a\right)}, \quad \text{and} \quad P_{i,j,\tau}^{d} = \frac{\exp\left(X_{i,j}^{g\vee d}\psi_\tau^d\right)}{1+\sum\limits_{a\in\{g,d\}}\exp\left(X_{i,j}^{g\vee d}\psi_\tau^a\right)}$$

where $P_{i,j,\tau}^{g}$ and $P_{i,j,\tau}^{d}$ represent the probability of student i graduating or dropping out from program j in after τ periods of enrollment, respectively; and ψ_τ^g, ψ_τ^d are vectors of parameters that need to be estimated. For estimation, we include in $X_{i,j}^{g\vee d}$ a constant, student's gender, a dummy variable that identifies whether the student's family income is below the median of the income distribution, and student's high school GPA. Note that we do not include students' GPA during college for two reasons: (i) we only have data for their first two years, so it may not be representative of their grades upon graduation / dropout; and (ii) to keep consistency in the observable characteristics considered with those in the enrollment probabilities.

C.3.4. *Evolution of scores.*

Since students can retake the test before re-applying, we need to model the evolution of (i) their scores, and (ii) their beliefs about the weights that programs will use in the future. For simplicity, we assume that students have correct beliefs about these weights,[12] while we assume that student i's scores in period $t+1$ (conditional on re-taking the test), $\vec{s}_{i,t+1}$, are exogenously given and evolve depending on their initial scores and a vector of time-invariant observable characteristics, X_i^s, according to the following process:

$$\vec{s}_{i,k,t+1} = X_i^s \psi_X^s + \psi_k^s (1+\nu_{i,t+1}) \cdot s_{i,k,t} + \mathbb{1}_{\{s_{i,k,t}=0\}} \psi_{k,0}^s \cdot (1+\nu_{i,t+1}) \bar{s}_{i,t},$$

where $s_{i,k,t}$ is the score of student i in subject k in period t; $\bar{s}_{i,t}$ is the average Math-Verbal score of student i in period t; $\nu_{i,t+1} \sim N(0,\sigma_s^2)$ is random shock; and $\left\{\psi_k,\, \psi_{k,0}\right\}_k$ and σ_s^2 are parameters to be estimated. For estimation, we include in X_i^s a constant, student's gender, a dummy variable that identifies whether the student's family income is below the median of the income distribution, and student's high school GPA.

C.4. *Continuation values.*

The model discussed above involves computing the expectation of the continuation value that students receive after solving a series of complex decision problems—e.g., deciding to take the exam or solving an optimal portfolio problem—for each possible state of the world in each period t—summarized in $\omega_{i,t} := \{\vec{s}_{i,t}, a_{i,j,t}, \vec{\epsilon}_{i,j,t}, \vec{\alpha}_{i,t}\} \in \Omega$, which includes the scores obtained $\vec{s}_{i,t}$ conditional on taking the exam; the signal $a_{i,j,t}$ received; the random shocks $\vec{\epsilon}_{i,t} = \left(\epsilon^G_{i,j,t}, \left\{\epsilon^U_{i,j',t}\right\}_{j' \in \mathcal{J}\cup\{\emptyset\}}\right)$; and random coefficients $\vec{\alpha}_{i,t} = \{\alpha_{i,m,t}\}_{m\in\mathcal{M}}$. Such computation imposes several challenges that we address in the following way.

First, to avoid taking expectation over all preference shocks $\vec{\epsilon}^{\,U}_{i,t} = \left\{\epsilon^U_{i,j',t}\right\}_{j'\in\mathcal{J}\cup\{\emptyset\}}$, we adapt the results in Fack et al. (2019) to our dynamic setting. Specifically, Fack et al. (2019) show that, under mild assumptions, the outcome of constrained DA satisfies pairwise stability relative to students' true preferences and, consequently, the optimal portfolio for student

[12] This assumption is likely to hold in practice, since admission weights are relatively stable over time.

i in period t is equivalent to choosing the most preferred program $r^{\star}_{i,j,t}(\omega)$ among those that are ex-post feasible, i.e., those for which their application score $s_{i,j,t}(\omega)$ is above the program's cutoff $\bar{s}_{j,t}(\omega)$. Combining this result with (10), we know that

$$\begin{aligned}\mathbb{E}_{\omega}\left[\Psi_{i,j,t}\left(R^{\star}_{i,j,t}\left(\omega\right),\omega\right) \mid x\right] = &\left(\mathbb{E}_{\vec{\epsilon}^{U}}\left[\bar{v}_{i,r^{\star}(\omega),t+1} + \epsilon_{i,r^{\star}(\omega),t}\right] - C^{S}_{i,t}\right) \cdot P^{E}_{i,t+1} \\ &+ \mathbb{E}_{\vec{\epsilon}^{U}}\left[\max\left\{\bar{v}_{i,j,t+1} + \epsilon_{i,j,t}, \bar{v}_{i,\emptyset,t+1} + \epsilon_{i,\emptyset,t}\right\}\right] \cdot \left(1 - P^{E}_{i,t+1}\right)\end{aligned}$$

where $\bar{v}_{i,j',t+1} = \mathbb{E}_{\omega\setminus\vec{\epsilon}^{U}}\left[v_{i,j',t+1} \mid x\right]$ is the expectation of the continuation values taken over all sources of randomness except of the preference shocks, and

$$r^{\star}(\omega) = r^{\star}_{i,j,t}(\omega) \in \underset{\substack{j'\in\mathcal{J}\setminus\{j\} \\ s_{i,j',t}(\omega)\geq\bar{s}_{j',t}(\omega)}}{\operatorname{argmax}} \left\{\Psi_{i,j,t}\left(j',\omega\right)\right\}.$$

Leveraging the fact that the preference shocks ϵ are i.i.d. type-I extreme value, the expression in (C.4) can be approximated as

$$\begin{aligned}\mathbb{E}_{\omega\setminus\vec{\epsilon}^{U}}\left[\Psi_{i,j,t}\left(R^{\star}_{i,j,t}(\omega),\omega\right) \mid x\right] \approx &\left(\log\left(\exp\left(\bar{v}_{i,r^{\star}(\omega),t+1}\right)\right) - C^{S}_{i,t}\right) \cdot P^{E}_{i,t+1} \\ &+ \log\left(\exp\left(\bar{v}_{i,j,t+1}\right) + \exp\left(\bar{v}_{i,\emptyset,t+1}\right)\right) + \gamma\end{aligned}$$

where γ is Euler's constant. This approximation allows us to compute the expected utility of student i from program j in period t without having to integrate over all preference shocks, which would be computationally infeasible.

TABLE C.I

TIME-INVARIANT OBSERVABLE CHARACTERISTICS INCLUDED IN THE MODEL

Equation	Coefficient	Description
Grades (β_X^G)	$\beta_{m(j)}^G$	Broad major-fixed effects
	β_A^G	Observed-ability ($A_{i,j}^o$)
Flow-utility (β_X^U)	β_j^U	Program-fixed effects
	β_A^U	Observed-ability ($A_{i,j}^o$)
	β_D^U	Distance between student i; and program j's municipality ($D_{i,j}$)
	β_Q^U	Distance to program mean quality ($\frac{A_{ij}^o - \bar{A}_j^o}{\sigma_j}$)
	β_C^U	Net cost (i.e., tuition - scholarships) that student i pays for program j ($C_{i,j}$)
Wage (β_X^W)	$\beta_{m(j)}^W$	Broad major-fixed effects
	β_W^W	Average wage posted at the moment of enrollment
	β_F^W	Female fixed effect
	β_{LI}^W	Low-income fixed effect
	β_G^W	Expected grades at graduation

Note: Broad majors include Science, Social Sciences, Education and Humanities and Health.

C.5. *Observable Characteristics*

C.6. *Evidence Supporting Implementation Choices*

APPENDIX D: APPENDIX FOR SECTION 5

D.1. *Identifying Equations*

D.1.1. *Grades.*

The variance of the first-year grades is given by

$$\mathbb{V}\left(G_{i,j,1} \mid \theta(i) = \theta\right) = \underbrace{\varsigma^2_{\theta,m(j)} + \sum_{s\in\mathcal{S}} \omega_{j,s} \cdot \varsigma^2_{\theta,s}}_{:=\mathbb{V}\left(A^p_{ij}\mid\theta(i)=\theta\right)} + \underbrace{\sigma^2_{\theta,m(j)} + \sum_{s\in\mathcal{S}} \omega_{j,s} \cdot \sigma^2_{\theta,s}}_{:=\mathbb{V}\left(A^u_{ij}\mid\theta(i)=\theta\right)} + \underbrace{\sigma^2_G}_{:=\mathbb{V}\left(\epsilon^G_{ij1\cdot}\right)}$$

We use variation on the variance of the noisy signals that students who stayed enrolled in the same program received across periods. Specifically, if we denote by $j_{i,1}$ and $j_{i,2}$ the programs of enrollment of student i in the first and second periods, respectively, note that

TABLE C.II

GRADES AND WAGES: ACTUAL VS. BELIEFS

	Grades				Wages
	Mean		Log-Variance		Mean
	Actual	Belief	Actual	Belief	Belief
Social Sciences	-0.012^{**}	0.011	0.081^{***}	0.015	0.024
	(0.006)	(0.007)	(0.016)	(0.039)	(0.016)
Education and Humanities	0.009	0.011	-0.187^{***}	-0.004	-0.085^{***}
	(0.009)	(0.009)	(0.026)	(0.054)	(0.028)
Health	-0.025^{***}	0.022^{***}	-0.404^{***}	0.006	-0.006
	(0.007)	(0.006)	(0.021)	(0.036)	(0.016)
Female	0.061^{***}	0.009^{*}	-0.251^{***}	-0.071^{***}	0.009
	(0.005)	(0.005)	(0.013)	(0.026)	(0.012)
Low-Income	0.000	0.009^{**}	0.155^{***}	-0.006	-0.034^{***}
	(0.005)	(0.005)	(0.013)	(0.026)	(0.011)
Observed Ability	0.083^{***}	0.019^{***}	-0.050^{***}	-0.301^{***}	0.078^{***}
	(0.004)	(0.003)	(0.001)	(0.016)	(0.007)
Omega Verbal	0.051	-0.000			-0.144
	(0.033)	(0.041)			(0.097)
Omega Verbal Squared			0.652^{***}	-0.151	
			(0.067)	(0.154)	
Avg. Log-Income					0.624^{***}
					(0.021)
Belief on Grades					0.038^{***}
					(0.007)
Constant	0.378^{***}	0.618^{***}	-0.345^{***}	-1.777^{***}	0.795^{***}
	(0.023)	(0.029)	(0.035)	(0.084)	(0.104)
Mean	0.523	0.64	-0.649	-1.885	2.668
Observations	15,710	10,423	11,024	7,198	13,351

Notes: Standard errors in parentheses. The "Actual" columns report coefficients from regressions of realized outcomes on student characteristics estimated on the sample of students who enrolled in their top-reported preference. The "Belief" columns report coefficients from regressions of students' stated beliefs (from survey data) on the same characteristics estimated on the sample of students who reported beliefs over their top-reported preference. Science is the omitted broad major category. Significance levels: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

the variances of signals, conditional on staying in the same program, are given by:

$$\begin{aligned}
\mathbb{V}\left(a_{i,j,1} \mid j = j_{i,1} = j_{i,2}\right) &= \mathbb{V}\left(A^u_{ij} \mid j = j_{i,1} = j_{i,2}\right) + \mathbb{V}\left(\epsilon^G_{ij1} \mid j = j_{i,1} = j_{i,2}\right) \\
\mathbb{V}\left(a_{i,j,2} \mid j = j_{i,1} = j_{i,2}\right) &= \mathbb{V}\left(A^u_{ij} \mid j = j_{i,1} = j_{i,2}\right) + \mathbb{V}\left(\epsilon^G_{i,j,2} \mid j = j_{i,1} = j_{i,2}\right) \qquad (13) \\
&= \mathbb{V}\left(A^u_{ij} \mid j = j_{i,1} = j_{i,2}\right) + \sigma^2_G.
\end{aligned}$$

The last equality follows because the noise term is i.i.d. across periods and selection on staying only occurs in the first one. Then, using properties of the normal distribution and denoting $\sigma^2_{A^u_{i,j}} := \mathbb{V}\left(A^u_{ij} \mid j = j_{i,1} = j_{i,2}\right)$, we can write the following system of equations to jointly identify the variances of the true unknown abilities (conditional on staying) and the grade noise variance:

$$\mathbb{V}\left(a_{i,j,2} \mid j = j_{i,1} = j_{i,2}\right) = \sigma^2_{A^u_{i,j}} + \sigma^2_G$$

$$\mathbb{V}\left(a_{i,j,2} \mid j = j_{i,1} = j_{i,2}\right) - \mathbb{V}\left(a_{i,j,1} \mid j = j_{i,1} = j_{i,2}\right) = \frac{1}{\sigma^2_{A^u_{i,j}} + \sigma^2_G} \cdot \left[\frac{\sigma^2_G \cdot \mathbb{V}\left(a_{i,j,1} \mid j = j_{i,1} = j_{i,2}\right)}{\sigma^2_{A^u_{i,j}} + \sigma^2_G} - \sigma^2_{A^u_{i,j}}\right]. \tag{14}$$

D.1.2. *Wages.*

Since $\log\left(\tilde{W}_{i,j,t}\right) = \beta^W_{G,i} \cdot \tilde{G}_{i,j,t'} + \beta^W_X X^W_{i,j} + \Lambda_{m(j),t-t'} + \epsilon^{W\sim M}_{i,j,t}$—where the latter term is the measurement error of wages—, and we have modeled $\beta^W_{G,i} = \mu_{WG} \times \exp\left(\epsilon^{W\sim G}_i\right)$, were $\epsilon^{W\sim G}_i \sim N\left(0, \sigma^2_{W\sim G}\right)$; the residual of this regression is given by

$$e_{i,j,t} = \left(\beta^W_{G,i} - \bar{\beta}_{W\sim G}\right)\tilde{G}_{i,j,t'} + \varepsilon^{W\sim M}_{i,j,t},$$

where the pooled slope is given by

$$\bar{\beta}_{W\sim G} \equiv \mathbb{E}[\beta^W_{G,i}] = \mu_{W\sim G}\exp\left(\tfrac{1}{2}\sigma^2_{W\sim G}\right).$$

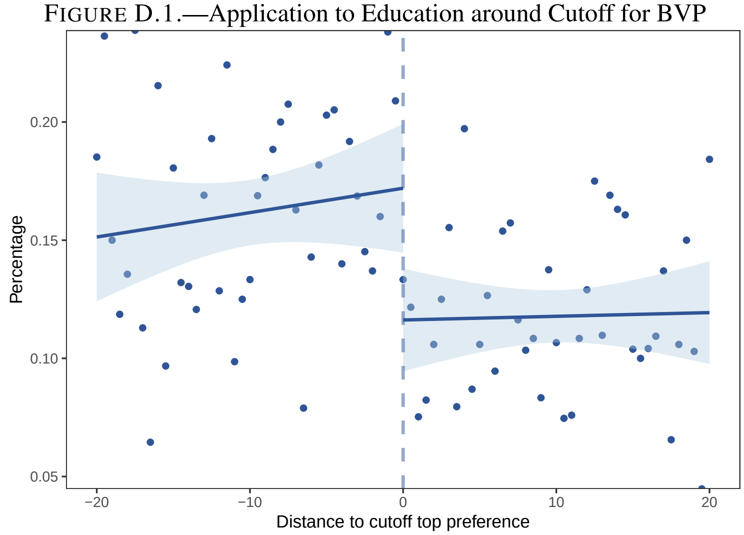
Define $v_\beta \equiv \mathrm{Var}(\beta^W_{G,i}) = \mu^2_{W\sim G}\left(e^{2\sigma^2_{W\sim G}} - e^{\sigma^2_{W\sim G}}\right)$. Assuming $\varepsilon^{W\sim M}_{i,j,t}$ is mean-zero, independent of $(\beta^W_{G,i}, \tilde{G}_{i,j,t'})$, and uncorrelated across programs, we obtain the system of moments:

$$\mathbb{E}\left[e^2_{i,j,t}\right] = v_\beta \cdot \mathbb{E}\left[\tilde{G}^2_{i,j,t'}\right] + \sigma^2_{W\sim M}, \tag{15}$$

$$\mathbb{E}\left[e_{i,j,t}\, e_{i,j',t'}\right] = v_\beta \cdot \mathbb{E}\left[\tilde{G}_{i,j,t'}\, \tilde{G}_{i,j',t'}\right], \qquad (j \neq j'). \tag{16}$$

Thus, after identifying the pooled slope by OLS, we have three moments and three unknowns.

FIGURE D.1.—Application to Education around Cutoff for BVP



D.2. *Additional Evidence*

APPENDIX E: APPENDIX FOR SECTION 6

E.1. *Estimator*

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

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

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 independent of 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:

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

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:

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

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.[13] Solving the model is computationally expensive and especially computing the continuation value terms, since they depend on realization of the random coefficients $\{\alpha_i\}_{i=1}^N$ (which are known to 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. The current set of estimation results uses 50 draws for shocks that do not affect the backward-induction procedure (preference shocks, enrollment shocks, etc.) and 5 draws for random coefficient shocks per student.[14] In Algorithm 1, we describe in detail how we perform the estimation and discuss some related technical considerations.

[13] In which we have suppressed the dependency on the first-stage estimators for readability.

[14] We have run robustness checks estimating the model with up to 10 realizations of random coefficient shocks per student, and estimation results are relatively similar.

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

E.2. *Estimation Moments Summary*

Tables E.I and E.II provide a comprehensive summary of all moment conditions used in the estimation procedure, grouped by outcome type following the organization of our goodness-of-fit validation exercises.

E.3. *Auxiliary models*

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, linear probability models of switching and dropout, and RDD models—and a vector of moment conditions. The parameters of the model are identified jointly by the moment conditions generated by the auxiliary models. We describe now in detail the regressions and moment conditions we use in the estimation and the sets of parameters that explain most of each moment's variation.

TABLE E.I

ESTIMATION MOMENTS SUMMARY (PANEL A)

Group	Moment description
Application and Preference Moments	Top true preference reported first in R_1
	Top true preference reported first in R_2
	Share of students who apply in period 1
	Share of students who apply in period 2
	Share changing broad major preference
	Share changing broad major from original order
Additional Application and Assignment Characteristics	Share unassigned in period 1
	Share assigned to first reported preference
	Share assigned to second reported preference
	Share assigned to third reported preference
	Share assigned to top true preference
	Share of reapplicants assigned to top true preference
	Share with top true preference changed
	Student observed ability (top reported preference)
	Program observed ability (top reported preference)
	Tuition at top reported preference
	Distance to top reported preference
	Relative position at top reported preference
	Portfolio risk in R_1
	Share of broad majors within R_1 (enrolled)
	Correlation between preference changes and grades (major)
	Correlation between preference changes and grades (broad major)
	Correlation between preference changes and grades (math type)
Switching and Outcome Moments	Share switching programs
	Share switching majors
	Share switching major within math (period 1)
	Share switching math within major (period 1)
	Share switching broad majors
	Share switching universities
	Share switching math types
	Share switching up (period 1)
	Share switching down (period 1)
	Share switching out (feasible)
	Share switching out (unfeasible)
	Share retaking the PSU
	Share dropouts
	Share of reapplicants
	Share of reapplicants assigned top reported preference
	Share enrolling for the first time in period 2
	Share dropout at the end of period 1
	Share first-year students in period 2
	Share second-year students in period 2
Market Structure Moments	Major shares in R_1 (10 majors)
	Major shares in R_2 (10 majors)
	Major inclusion indicators in R_1
	Major inclusion indicators in R_2
	Broad major shares in R_1
	Broad major shares in R_2
	Market shares by major (period 1)
	Market shares by major (period 2)
	Market shares by broad major (period 1)
	Market shares by broad major (period 2)

Notes: Moments are grouped following the comprehensive goodness-of-fit tables used in the estimation and validation exercises.

TABLE E.II

ESTIMATION MOMENTS SUMMARY (PANEL B)

Group	Moment description
Portfolio and Reapplication Structure	Norm difference of college-type shares for reapplicants Norm difference of major shares for reapplicants Norm difference of math-type shares for reapplicants Norm difference of ω shares for reapplicants Applications by major and college type (gender groups, R_1) Applications by major and college type (gender groups, R_2)
Grade and Academic Performance	Grade equation coefficients (year 1) Grade equation coefficients (year 2) Grade time-series coefficients (no switchers) Grade time-series coefficients (switchers) Grade SSR (year 1) Grade SSR (year 2) Residual variance of grades for stayers (years 1-2)
Signal and Belief Moments	Signal correlations with graduation and dropout Signal correlations with switching decisions Difference between subjective and realized grades (broad majors & demographics) Log variance of ability beliefs (realized) Log variance of ability beliefs (subjective) Learning regression on change in broad major norm
Wage Regression and Expectations	Wage regression coefficients (broad majors, ability, demographics) Wage regression residual diagnostics Expected log wage regression (top true vs random) Belief-based wage dispersion (dependent mass within quintile) Expected log wages: top true vs random assignment
PSU Score Evolution	Mean and variance of log score ratios (Language, Math) Mean and variance of log score ratios (History, Science takers) Mean and variance of log score ratios (History, Science non-takers)
RDD Auxiliary Moments	RDD switching intercepts RDD switching discontinuities RDD reapplication intercepts RDD reapplication discontinuities RDD switch up/out intercepts RDD switch up/out discontinuities RDD BVP intercepts RDD BVP discontinuities

Notes: Moments are grouped following the comprehensive goodness-of-fit tables used in the estimation and validation exercises.

Grades. The auxiliary models targeting the grade equations are:

$$G_{i,j,1} = \beta^G_{m(j),1} + \beta^G_{A,1} A^o_{i,j} + \beta^G_{TP,1} \mathbb{1}_{\{j=R^1_i(1)\}} + \beta^G_{M,1} M_{i,m(j),1} + \epsilon^G_{i,j,1}$$

$$G_{i,j,2} = \beta^G_{m(j),2} + \beta^G_{A,2} A^o_{i,j} + \beta^G_{TP,2} \mathbb{1}_{\{j=R^2_i(1)\}} + \beta^G_{M,2} M_{i,m(j),2} + \beta^G_{CP,2} \mathbb{1}_{\{j^1_i=j^2_i\}} + \epsilon^G_{i,j,2}$$

$$\begin{aligned} G_{i,j,2} = {} & \beta^G_{m(j),3} + \beta^G_{A,3} A^o_{i,j} + \beta^G_{TP,3} \mathbb{1}_{\{j=R^2_i(1)\}} + \beta^G_{M,3} M_{i,m(j),3} + \beta^G_{CP,3} \mathbb{1}_{\{j^1_i=j^2_i\}} \\ & + \beta^G_{a,3} a_{i,j^1_i,1} + \left(\beta^G_{CM,3} \mathbb{1}_{\{m(j^1_i)=m(j^2_i)\}} + \beta^G_{CT,3} \mathbb{1}_{\{s(j^1_i)=s(j^2_i)\}} \right) a_{i,j^1_i,1} + \epsilon^G_{i,j,3} \end{aligned}$$

where R_i^t is the ROLs of student i in periods $t \in \{1,2\}$ (with $R_i^2 = R_i^1$ if student i does not reapply); j_i^t is the program of enrollment of student i in period t; $M_{i,m(j),t} = \frac{\sum_{l=1}^{\left|R_i^t\right|} \mathbb{1}_{\left\{m\left(R_i^t(l)\right)=m(j)\right\}}}{\left|R_i^t\right|}$ is the fraction of programs in student i's ROL in period t that share the same major as program j $(m(j))$; $a_{i,j_i^1,1}$ is the first-year grade signal of student i in program j_i^1; and $s(j)$ captures the type of program j (i.e., whether or not it is math intensive). We also estimate autoregressions of second-period grades on first-period grades separately for students who switch programs and those who do not, to capture grade persistence dynamics. Additionally, we use the sum of squared residuals (SSR) and residual variances from the grade equations as auxiliary moments.

Wages. The auxiliary models targeting the initial wage equation is:

$$\begin{aligned}\log\left(\tilde{W}_{i,j}\right) = {} & \beta^W_{m(j),1} + \beta^W_{A,1} A^o_{i,j} + \beta^W_{TP,1} \mathbb{1}_{\{j=R_i^1(1)\}} \\ & + \beta^W_{G,1} \tilde{G}_{i,j} + \beta^W_{m(j),1} \bar{W}_{m(j)} + \beta^W_{F,1} fem_i + \beta^W_{LI,1} li_i + \epsilon^W_{i,j}\end{aligned}$$

where $\tilde{W}_{i,j}$ and $\tilde{G}_{i,j}$ are the subjective wage and grade reported by student i for program j (both elicited from the survey); $\bar{W}_{m(j)}$ is the average wage of graduates from major $m(j)$ after four years of graduation (same horizon as that considered in the survey); fem_i is a dummy variable that equals one if student i is a female student; and li_i is a dummy variable that equals one if student i's family income is below the median of the income distribution.

To target the parameters related to the evolution of wages with tenure, we use the following auxiliary model:

$$\bar{W}_{j,\tau} = \beta^\Lambda_{m(j),1} + \beta^\Lambda_{m(j),2}\tau + \beta^\Lambda_{m(j),3}\tau^2 + \epsilon^\Lambda_{j,\tau},$$

where $\tau > 4$ is the number of years after graduation; and $\bar{W}_{j,\tau}$ is the average wage. We use the latter instead of beliefs on wages because we only elicited beliefs at the fourth year after graduation.

Nonpecuniary utilities. The auxiliary model targeting the non-pecuniary payoffs is:

$$grad_{i,j} = \beta^U_{m(j)} + \beta^U_A A^o_{i,j} + \beta^U_{TP} \mathbb{1}_{\{j=R_i^1(1)\}} + \beta^U_{M,1} M_{i,m(j)} + \beta^U_G G_{i,j,1} + \epsilon^U_{i,j},$$

where $grad_{i,j} = \mathbb{1}_{\{j_i^g = j_i^1\}}$, with $j_i^g \in \mathcal{J} \cup \{\emptyset\}$ being the program of graduation of student i and j_i^1 being the program of enrollment of student i in their first year.

Learning. The auxiliary models targeting students' learning process are:

$$y_{i,j}^o = \beta_{m(j),1}^o + \beta_{A,1}^o A_{i,j}^o + \beta_{TP,1}^o \mathbb{1}_{\{j=R_i^1(1)\}} + \beta_{M,1}^o M_{i,m(j),1} + \beta_{a,1}^o a_{i,j_i^1,1} + \epsilon_{i,j,1}^o$$

where $y_{i,j}^o$ equals 1 if student i is enrolled in program j in their first year and outcome o holds at the end of the first year, where o is switching program, dropout, among others; and $a_{i,j_i^1,1}$ is the first-year grade signal of student i in program j_i^1.

To separately identify learning about abilities from learning about preferences, we estimate regressions of preference changes on first-year grade signals. These regressions help distinguish belief updating based on grade information (ability learning) from shifts in tastes (preference learning):

$$\mathbb{1}\{\Delta \text{Pref}_i > \tau\} = \beta_{m(i)}^{\Delta p} + \beta_{A,1}^{\Delta p} A_{i,j}^o + \beta_{F,1}^{\Delta p} fem_i + \beta_{LI,1}^{\Delta p} li_i + \beta_a^{\Delta p} a_{i,j_i^1,1} + \varepsilon_i^{\Delta p},$$

where $\mathbb{1}\{\Delta \text{Pref}_i > \tau\}$ equals one if the change in the norm of student i's reported preferences between the initial application and reapplication exceeds threshold τ, and zero otherwise; $a_{i,j_i^1,1}$ is the first-year grade signal of student i in the initial program of enrollment j_i^1; $A_{i,j}^o$ is student i's ability in program j; fem_i and li_i are indicators for female and low-income students. The coefficient $\beta_a^{\Delta p}$ captures how preference changes correlate with grade signals, while the major fixed effects $\beta_{m(i)}^{\Delta p}$ capture baseline preference drift that is independent of grade information.

We use the following auxiliary models to separately target the parameters related to subjective and actual variance of the unknown ability:

$$\log(\tilde{\sigma}_{i,j}^2) = \beta_{m(i)}^{\tilde{\sigma}} + \beta_{A,1}^{\tilde{\sigma}} A_{i,j}^o + \beta_{F,1}^{\tilde{\sigma}} fem_i + \beta_{LI,1}^{\tilde{\sigma}} li_i + \omega_{j,s_v}^2 \beta_{s_v}^{\tilde{\sigma}} + \epsilon_{i,j}^{\tilde{\sigma}},$$

$$\log(\sigma_{i,j}^2) = \beta_{m(i)}^{\sigma} + \beta_{A,1}^{\sigma} A_{i,j}^o + \beta_{F,1}^{\sigma} fem_i + \beta_{LI,1}^{\sigma} li_i + \omega_{j,s_v}^2 \beta_{s_v}^{\sigma} + \epsilon_{i,j}^{\sigma},$$

where $\log(\tilde{\sigma}_{i,j}^2)$ is the log of the subjective variance of student i's prior beliefs about unknown abilities in program j, $\log(\sigma_{i,j}^2)$ is the log of the actual variance, and ω_{j,s_v}^2 is the squared weight that program j places in the verbal test.

Bias. The auxiliary models targeting the mean bias are:

$$\begin{aligned} G_{i,j,t} &= \left(\beta^{\tilde{\mu}}_{m(i),1} + \beta^{\tilde{\mu}}_{A,1} A^o_{i,j} + \beta^{\tilde{\mu}}_{F,1} fem_i + \beta^{\tilde{\mu}}_{LI,1} li_i\right) \\ &\quad + \omega_{j,s_m}\left(\beta^{\tilde{\mu}}_{s_m,1} + \beta^{\tilde{\mu}}_{A,2} A^o_{i,j} + \beta^{\tilde{\mu}}_{F,2} fem_i + \beta^{\tilde{\mu}}_{LI,2} li_i\right) + \epsilon^{\tilde{\mu}}_{i,j,1}, \\ \tilde{G}_{i,j,t} &= \left(\beta^{\tilde{\mu}}_{m(i),2} + \beta^{\tilde{\mu}}_{A,3} A^o_{i,j} + \beta^{\tilde{\mu}}_{F,3} fem_i + \beta^{\tilde{\mu}}_{LI,3} li_i\right) \\ &\quad + \omega_{j,s_m}\left(\beta^{\tilde{\mu}}_{s_m,2} + \beta^{\tilde{\mu}}_{A,4} A^o_{i,j} + \beta^{\tilde{\mu}}_{F,4} fem_i + \beta^{\tilde{\mu}}_{LI,4} li_i\right) + \epsilon^{\tilde{\mu}}_{i,j,2}. \end{aligned}$$

Unobserved preferences. The auxiliary models targeting unobserved and persistent preferences are:

$$y^o_{i,j} = \beta^o_{1,2}\left(s_{i,j,1} - \bar{s}_{j,1}\right) + \beta^o_{2,2}\mathbb{1}_{\left\{\left(s_{i,j,1} - \bar{s}_{j,1}\right) > 0\right\}} + \beta^o_{3,2}\left(s_{i,j,1} - \bar{s}_{j,1}\right)\mathbb{1}_{\left\{\left(s_{i,j,1} - \bar{s}_{j,1}\right) > 0\right\}} + \epsilon^o_{i,j,2}$$

where $s_{i,j,1}$ is the application score of student i in program j in the first period; and $\bar{s}_{j,1}$ is the cutoff of program j in the first period. Note that we estimate these equations (for each outcome) only considering students whose application score is around the cutoff of their top reported preference.

E.4. *MIA*

Chade and Smith (2006) show that, when admission probabilities are independent and the cost of submitting ROL R only depends on its cardinality—i.e., $C^A_i(R) = c(|R|)$ for some function c—the unconstrained problem is Downward Recursive, and the optimal solution is given by the MIA algorithm presented next.

MIA: Marginal Improvement Algorithm

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

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. Note that, if some programs have zero-admission chance, the strict inequality in MIA's stopping criteria becomes a weak inequality, which could lead to multiplicity of best responses (He (2012)). We discuss this potential identification threat in Larroucau and Rios (2018).

E.5. *Weighting Matrix and Standard Errors*

E.5.1. *Weighting Matrix*

The weighting matrix W is a diagonal matrix with the inverse of the variance of each data moment. We weight up moments by outcome type, ensuring that moments from different data sources and with different scales contribute appropriately to the objective function.[15] 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.

E.5.2. *Standard Errors*

We compute standard errors for 145 out of 165 structural parameters using a simulated method of moments (SMM) framework with finite-simulation correction. The estimation proceeds in three steps. First, we construct the Jacobian matrix G by computing numerical derivatives of all moment conditions with respect to free parameters at the estimated optimum $\hat{\theta}$, using parallel finite-difference approximations with common random numbers to reduce simulation noise. Second, we build the asymptotic variance of moments Σ by combining the data moment covariance matrix Ω (recovered from the inverse of our diagonal weighting matrix W) with a finite-simulation correction factor. Third, we compute standard errors from the sandwich formula $\widehat{\text{Var}}(\hat{\theta}) = N_{\text{data}}^{-1}(G'WG)^{-1}G'W\Sigma WG(G'WG)^{-1}$, which properly accounts for both sampling variation in the data and simulation error. This formula does not account for the effect of first-stage estimation error.

TABLE E.III

ESTIMATION RESULTS - MAIN STRUCTURAL PARAMETERS

Parameter	Value	Std. Error
Application behavior and Dropout		
Share of strategic ROLs $(1-\rho)$	0.88	0.04
Cost of retaking PSU (C^{psu})	21.84	< 0.01
Dropout flow-utility for females $(\alpha^{dropout}_{female})$	38.64	0.10
Dropout flow-utility for males $(\alpha^{dropout}_{male})$	-12.73	0.05
Dropout flow-utility for low-income $(\alpha^{dropout}_{low\text{-}income})$	34.64	0.20
Dropout observed ability effect $(\alpha^{dropout}_{obs})$	9.40	< 0.01
First-time enrollment cost (C^{e})	206.55	< 0.01
Flow-utility		
Tuition (β^U_3)	-3.28	0.16
Relative position (β^U_2)	1.35	0.17
Distance (β^U_1)	-1.00	(-)
Student observed ability (β^U_4)	13.46	0.08
Female effect by major (Δ_m female)	(13.85, 15.71, 10.45, 9.68)	(0.11, 0.12, < 0.01, 0.07)
Low-income effect by major (Δ_m low-income)	(-19.02, -31.31, 0.69, 31.90)	(0.01, 0.01, < 0.01, 0.14)
Avg. observed ability effect by major (Δ_m avg. obs. ability)	(10.42, 20.98, -1.68, 19.48)	(0.05, 0.10, < 0.01, 0.09)
Variance major random coefficient (σ^2_α)	1134.98	< 0.01
Scale gumbel parameter (κ)	17.78	0.21
Grade equations		
Constant by major $(\beta^G_{m_j})$	(4.03, 4.54, 4.41, 4.00)	(0.25, 0.22, 0.25, 0.24)
Student observed ability (β^G_1)	0.90	0.14
Variance grade shock (σ^2_g)	0.11	0.02
Evolution of scores		
Std. of ν (σ_{psu})	0.08	0.23
Mean prop. change $(\{\alpha_l\}_l)$	(1.04, 1.02, 1.03, 1.03)	(0.13, 0.18, 0.26, 0.20)
Mean prop. change from zero $(\{\alpha_{0l}\}_l)$	(1.05, 1.02)	(0.05, 0.04)

Notes: Standard errors (S.E.) are computed using the asymptotic variance-covariance of the estimator. Parameters from external regressions are indicated by (-).

TABLE E.IV

ESTIMATION RESULTS - WORK UTILITY AND WAGE PARAMETERS

Parameter	Value	Std. Error
Non-pecuniary work utility		
Random coefficient by major (β_1^V)	6.43	0.12
Observed ability (β_2^V)	311.54	$<$ 0.01
Unknown and private ability by major (β_3^V)	(553.65, 571.55, 741.39, 691.94)	(0.03, $<$ 0.01, $<$ 0.01, 0.02)
Major fixed effects ($\beta_{m_j}^V$)	(-3.40, 44.45, 25.84, -9.23)	($<$ 0.01, 0.01, $<$ 0.01, $<$ 0.01)
Wage parameters		
Constant by major ($\beta_{m_j}^W$)	(-0.08, 0.08, -0.17, -0.10)	(0.20, 0.20, 0.19, 0.20)
Grades (β_G^W, mean)	0.08	0.07
Gender effects (β_1^W)	-0.01	0.21
Wage low income effect (β_2^W)	-0.02	0.24
Wage parameter (β_3^W)	0.56	0.13
Variance of wage grades effect ($\sigma^2_{W\sim G}$)	0.43	0.07
Measurement error log wages variance ($\sigma^2_{\tilde{W}}$)	0.27	0.19
Variance wage shock (σ^2_w)	0.09	0.02
Wage growth		
Linear term by major ($\beta^\Lambda_{m_j,1}$)	(0.23, 0.16, 0.15, 0.27)	(-)
Quadratic term by major ($\beta^\Lambda_{m_j,2}$)	(-0.01, -0.01, -0.01, -0.02)	(-)

Notes: Standard errors (S.E.) are computed using the asymptotic variance-covariance of the estimator. Wage growth parameters are fixed from auxiliary regressions (indicated by (-)).

TABLE E.V

ESTIMATION RESULTS - SUBJECTIVE PRIOR MEANS BIAS

Major Unknown Ability			**Subject Unknown Ability**		
Parameter	Value	Std. Error	Parameter	Value	Std. Error
By major and subject type					
Social Sciences ($\tilde{\mu}_{m,0}$)	0.63	(-)	Verbal type ($\tilde{\mu}_{s,v}$)	-0.05	(-)
Science ($\tilde{\mu}_{m,1}$)	0.66	(-)			
Education and Humanities ($\tilde{\mu}_{m,2}$)	0.36	(-)			
Health ($\tilde{\mu}_{m,3}$)	0.84	(-)			
Covariates					
Low income ($\tilde{\mu}_{m,low}$)	0.08	(-)	Low income ($\tilde{\mu}_{s,low}$)	0.00	(-)
Female ($\tilde{\mu}_{m,f}$)	-0.19	(-)	Female ($\tilde{\mu}_{s,f}$)	0.00	(-)
Avg. obs. ability ($\tilde{\mu}_{m,\bar{a}}$)	-0.21	(-)	Avg. obs. ability ($\tilde{\mu}_{s,\bar{a}}$)	0.00	(-)

Notes: Standard errors (S.E.) are computed using the asymptotic variance-covariance of the estimator. Parameters from external regressions are indicated by (-).

TABLE E.VI

ESTIMATION RESULTS - LOG-VARIANCES: TRUE VS SUBJECTIVE BIAS

True Log-Variances			Subjective Bias Log-Variances		
Parameter	Value	Std. Error	Parameter	Value	Std. Error
Major unknown ability (by major)					
Social Sciences ($\log \sigma^2_{m,0}$)	-0.27	0.19	Social Sciences ($\log \tilde{\sigma}^2_{m,0}$)	-1.28	0.21
Science ($\log \sigma^2_{m,1}$)	-0.20	0.22	Science ($\log \tilde{\sigma}^2_{m,1}$)	-2.03	0.13
Education and Humanities ($\log \sigma^2_{m,2}$)	-0.58	0.26	Education and Humanities ($\log \tilde{\sigma}^2_{m,2}$)	-0.74	0.16
Health ($\log \sigma^2_{m,3}$)	-1.44	0.18	Health ($\log \tilde{\sigma}^2_{m,3}$)	-1.99	0.09
Major unknown ability (covariates)					
Low income ($\log \sigma^2_{m,low}$)	0.12	0.09	Low income ($\log \tilde{\sigma}^2_{m,low}$)	-0.19	0.21
Female ($\log \sigma^2_{m,f}$)	-0.37	0.10	Female ($\log \tilde{\sigma}^2_{m,f}$)	0.20	0.25
Avg. obs. ability ($\log \sigma^2_{m,\bar{a}}$)	-0.80	0.16	Avg. obs. ability ($\log \tilde{\sigma}^2_{m,\bar{a}}$)	-1.90	0.08
Subject unknown ability					
Verbal type ($\log \sigma^2_{s,v}$)	-3.86	0.04	Verbal type ($\log \tilde{\sigma}^2_{s,v}$)	-3.61	0.03
Low income ($\log \sigma^2_{s,low}$)	0.00	(-)	Low income ($\log \tilde{\sigma}^2_{s,low}$)	0.00	(-)
Female ($\log \sigma^2_{s,f}$)	0.00	(-)	Female ($\log \tilde{\sigma}^2_{s,f}$)	0.00	(-)
Avg. obs. ability ($\log \sigma^2_{s,\bar{a}}$)	0.00	(-)	Avg. obs. ability ($\log \tilde{\sigma}^2_{s,\bar{a}}$)	0.00	(-)
Learning about preferences - Major (by major)					
Social Sciences ($\log((\zeta^\alpha_{m,0})^2)$)	0.57	0.09			
Science ($\log((\zeta^\alpha_{m,1})^2)$)	1.14	0.01			
Education and Humanities ($\log((\zeta^\alpha_{m,2})^2)$)	0.48	$<$ 0.01			
Health ($\log((\zeta^\alpha_{m,3})^2)$)	2.69	0.02			
Learning about preferences - Major (covariates)					
Low income ($\log((\zeta^\alpha_{m,low})^2)$)	2.47	0.02			
Female ($\log((\zeta^\alpha_{m,f})^2)$)	-0.04	0.02			
Avg. obs. ability ($\log((\zeta^\alpha_{m,\bar{a}})^2)$)	1.20	0.03			

Notes: Standard errors (S.E.) are computed using the asymptotic variance-covariance of the estimator. Parameters from external regressions are indicated by (-).

E.6. *Results*

APPENDIX F: APPENDIX FOR SECTION 7

F.1. *Additional Details*

F.1.0.1. *CADA.* 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 computa-

[15] The exact weighting scheme is available upon request.

TABLE E.VII

ESTIMATION RESULTS - PRIVATE ABILITIES

Log-Variances			Means		
Parameter	Value	Std. Error	Parameter	Value	Std. Error
Private ability - Subjects					
Intercept ($\log \varsigma^2_{s,0}$)	-0.75	0.11	Verbal type ($\mu^p_{s,v}$)	-1.16	0.03
Low income ($\log \varsigma^2_{s,low}$)	0.00	(-)	Low income ($\mu^p_{s,low}$)	0.35	0.03
Female ($\log \varsigma^2_{s,f}$)	0.00	(-)	Female ($\mu^p_{s,f}$)	1.79	0.01
Avg. observed ability ($\log \varsigma^2_{s,\bar{a}}$)	0.00	(-)	Avg. observed ability ($\mu^p_{s,\bar{a}}$)	-0.02	0.11
Private ability - Major (by major)					
Social Sciences ($\log \varsigma^2_{m,0}$)	-2.04	0.08	Social Sciences ($\mu^p_{m,0}$)	0.00	(-)
Science ($\log \varsigma^2_{m,1}$)	-2.02	0.08	Science ($\mu^p_{m,1}$)	0.00	(-)
Education and Humanities ($\log \varsigma^2_{m,2}$)	-0.97	0.19	Education and Humanities ($\mu^p_{m,2}$)	0.00	(-)
Health ($\log \varsigma^2_{m,3}$)	0.05	0.14	Health ($\mu^p_{m,3}$)	0.00	(-)
Private ability - Major (covariates)					
Low income ($\log \varsigma^2_{m,low}$)	-0.27	0.06	Low income ($\mu^p_{m,low}$)	-0.36	0.00
Female ($\log \varsigma^2_{m,f}$)	-0.72	0.01	Female ($\mu^p_{m,f}$)	-0.20	0.01
Avg. observed ability ($\log \varsigma^2_{m,\bar{a}}$)	-0.72	0.15	Avg. observed ability ($\mu^p_{m,\bar{a}}$)	-0.35	0.01

Notes: Standard errors (S.E.) are computed using the asymptotic variance-covariance of the estimator. Parameters from external regressions are indicated by (-).

tional 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 3 in Appendix F describes a procedure to accomplish this. See Abdulkadiroğlu et al. (2015) for details.

F.2. *Market Structure and Capacities*

To construct the counterfactual market, we scale from our estimation sample of 4,000 students to the full market of approximately 40,000 students. We first back out program capacities that are consistent with the empirical distribution of cutoffs observed in the data, given our sample composition. After recomputing these capacities for the entire market, we solve for the baseline equilibrium by running Algorithm 2 instead of using the empirical distribution of cutoffs for assignments. This yields an equilibrium cutoff distribution that serves as the baseline for all counterfactual experiments. We fix this vector of capacities throughout all counterfactual analyses, allowing us to isolate the effects of policy changes on student behavior and outcomes while holding market structure constant.

F.3. *Finding equilibrium beliefs*

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 $\left(R_1^0, R_2^0, s_1^0, s_2^0\right)$

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, $\hat{\theta}$, $\hat{P}^e$, $\hat{P}^d$, $\hat{P}^g$, $\hat{P}^w$, and $\hat{p}^{k-1}$

Step 3.b. Forward simulate first period ROL R_{i1}^k given those objects

Step 3.c. For each program j, estimate the mean and standard deviation $\hat{\delta}_j^k \equiv \left(\hat{\mu}_j^k, \hat{\sigma}_j^k\right)$

Step 3.d. Given R_1^k, R_2^{k-1}, s_1^k, and s_2^{k-1}, run the Chilean matching mechanism and obtain μ^k

Step 3.e. Given μ^k and τ, simulate second-period ROLs R_{i2}^k

Step 3.f. Bootstrap to estimate $\tilde{p}^k$ and update $\hat{p}^k = \rho^k \hat{p}^{k-1} + (1 - \rho^k)\tilde{p}^k$

Step 3.g. Compute $\delta_{diff} = ||\hat{\delta}^k - \hat{\delta}^{k-1}||$

$\hat{p}(\tau) = \hat{p}^{k-1}$; $k++$

Algorithm 3 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 $R\,(v, s, P, \psi_\tau)$

Step 1. For each program j

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

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

Step 2. Compute optimal signal

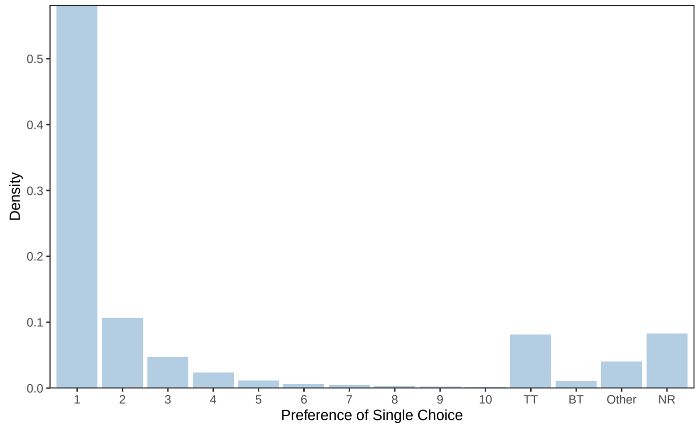
$$s_j^* = \operatorname*{argmax}_j \{R(v, \tilde{p}_j)\}$$

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

F.4. *Hypothetical scenario*

In 2022, we conducted a survey to elicit preferences and beliefs, which is similar to the 2019-2021' versions. In this version, we included a question to elicit information about how students would change their application lists if they could apply to a single program—i.e., applying under CDA with $K = 1$. Figure F.1 shows the distribution of the chosen program in the hypothetical scenario relative to their submitted ROL. Labels from *1-10* identify the share of students whose hypothetical program coincides with their k-th reported preference; *TT*, *BT*, and *Other* identify the share of students who report in the hypothetical scenario a program outside their submitted ROL; finally, *NR* identifies the share of students who did not respond to the survey question. We observe that under CDA with $K = 1$, a significant fraction of students would choose a program that is not their top-reported program on their current list (close to 40%). This suggests that a significant fraction of students would react strategically and take into account their admission probabilities when facing a binding constraint on the length of the list, consistent with our modeling assumptions.

FIGURE F.1.—Distribution of applications under CDA with $K = 1$



F.5. *Additional Results*

APPENDIX G: APPENDIX FOR SECTION 6.3

Co-editor [Name Surname; will be inserted later] handled this manuscript.

TABLE F.I

UNDERSTANDING BEHAVIORAL CHANNELS - PERCENTAGE CHANGES (EXOGENOUS VERSION)

Outcome	Baseline (exogenous)	No bias mean	No bias	No learning abilities	No learning prefs	No mismatch
Re-applicants [%]	18.80	-5.20	19.35	-7.99	-10.78	-54.86
	(0.041)	(0.258)	(0.196)	(0.240)	(0.200)	(0.203)
Program switchings [%]	4.21	-15.62	80.80	-22.80	-34.51	-93.97
	(0.055)	(0.505)	(1.452)	(0.904)	(1.118)	(0.225)
Program switching rate [%]	6.69	-8.35	89.02	-20.25	-31.42	-95.35
	(0.093)	(0.614)	(1.518)	(0.987)	(1.214)	(0.170)
Major switchings [%]	3.44	-20.35	93.42	-26.41	-39.68	-93.29
	(0.043)	(0.427)	(1.798)	(0.765)	(0.985)	(0.326)
Major switching rate [%]	5.47	-13.48	102.22	-23.98	-36.83	-94.83
	(0.072)	(0.532)	(1.877)	(0.823)	(1.056)	(0.248)
Retakes PSU [%]	7.65	-8.51	-8.03	5.34	4.09	-69.77
	(0.020)	(0.394)	(0.205)	(0.585)	(0.590)	(0.485)
Dropouts [%]	7.45	3.82	23.29	-0.99	-12.31	0.20
	(0.119)	(0.893)	(0.806)	(0.767)	(0.813)	(1.905)
Dropout rate [%]	11.83	12.78	28.90	2.29	-8.15	-22.71
	(0.178)	(0.966)	(0.829)	(0.687)	(0.748)	(1.393)
Dropout in first period rate [%]	5.27	23.99	55.97	-2.71	-27.70	-7.01
	(0.131)	(2.459)	(2.251)	(2.095)	(2.050)	(2.988)
Applicants in first period [%]	76.58	-7.74	-4.86	-3.82	-4.94	19.89
	(0.089)	(0.033)	(0.025)	(0.094)	(0.104)	(0.125)
Unassigned in first period [%]	29.79	18.90	10.27	7.61	10.72	-72.53
	(0.116)	(0.232)	(0.046)	(0.339)	(0.380)	(0.215)
First stated preference assignment [%]	1.59	-1.74	-1.83	-5.60	-5.58	-37.11
	(0.009)	(0.114)	(0.079)	(0.151)	(0.166)	(0.341)
Assigned to top true preference in period 1 [%]	11.60	19.67	16.79	62.02	64.74	761.77
	(0.137)	(0.550)	(0.464)	(1.360)	(1.382)	(10.175)
Graduate from 1st enrollment [%]	45.67	-8.50	-13.73	-1.59	-0.67	45.79
	(0.046)	(0.113)	(0.093)	(0.092)	(0.086)	(0.151)

Note: Baseline column shows levels; other columns show percentage changes relative to the baseline. Standard errors computed using the delta method for percentage changes, accounting for correlation between baseline and counterfactual simulations (common seeds across 5 simulations). Switching and dropout rates are computed relative to first-period enrollees. This table uses the exogenous version of the counterfactuals (empirical distribution of cut-offs rather than solving for equilibrium cut-offs under the policy change).

TABLE F.II

COUNSELING AND PSP POLICIES: % CHANGE FROM BASELINE

Outcome	Baseline	Counseling	PSP
Re-applicants [%]	15.06	-2.46	14.72
	(0.096)	(0.298)	(0.290)
Program switchings [%]	2.68	-7.34	-19.23
	(0.020)	(1.602)	(0.743)
Program switching rate [%]	4.19	-7.59	-12.05
	(0.037)	(1.658)	(0.786)
Major switchings [%]	2.09	-9.22	-29.28
	(0.018)	(1.614)	(0.873)
Major switching rate [%]	3.27	-9.47	-23.00
	(0.033)	(1.699)	(0.930)
Retakes PSU [%]	5.56	-2.80	9.64
	(0.062)	(0.975)	(0.421)
Dropouts [%]	6.48	-4.99	0.32
	(0.075)	(0.903)	(1.122)
Dropout rate [%]	10.12	-5.25	9.25
	(0.100)	(0.832)	(1.071)
Dropout in first period rate [%]	3.72	-11.09	14.61
	(0.053)	(1.533)	(1.065)
Applicants in first period [%]	75.21	0.18	-3.52
	(0.094)	(0.170)	(0.133)
Unassigned in first period [%]	28.60	-0.62	20.66
	(0.094)	(0.354)	(0.152)
Graduate - first enrollment [%]	48.43	0.94	-7.62
	(0.065)	(0.225)	(0.143)

Note: First column shows baseline levels. Columns 2-3 show percentage changes relative to baseline. Standard errors computed using the delta method, accounting for correlation between baseline and counterfactual simulations (common seeds). Counseling and PSP (Post-Secondary Program) policies represent full information environments. Switching and dropout rates are computed relative to first-period enrollees.

TABLE F.III

COUNTERFACTUAL POLICIES: CDA AND CADA MECHANISMS (% CHANGE FROM BASELINE)

		CDA Rules			CADA Rules		
Outcome	Baseline	$K=3$	$K=2$	$K=1$	$\psi=1.1$	$\psi=1.2$	$\psi=1.3$
Re-applicants [%]	15.06	3.23	11.67	56.12	-10.37	-13.85	-15.70
	(0.096)	(0.494)	(0.633)	(0.390)	(0.417)	(0.571)	(0.510)
Program switchings [%]	2.68	3.65	23.44	11.50	-22.31	-29.72	-34.15
	(0.020)	(0.572)	(1.576)	(0.864)	(1.199)	(1.265)	(1.741)
Program switching rate [%]	4.19	4.50	27.12	28.68	-23.80	-31.42	-36.04
	(0.037)	(0.583)	(1.590)	(0.907)	(1.151)	(1.260)	(1.766)
Major switchings [%]	2.09	2.11	17.18	2.57	-20.38	-27.03	-31.66
	(0.018)	(0.719)	(1.845)	(1.016)	(1.695)	(1.683)	(2.343)
Major switching rate [%]	3.27	2.95	20.66	18.38	-21.91	-28.79	-33.63
	(0.033)	(0.736)	(1.866)	(1.045)	(1.633)	(1.679)	(2.375)
Retakes PSU [%]	5.56	2.12	9.70	47.14	-16.62	-20.43	-22.66
	(0.062)	(1.383)	(1.058)	(0.224)	(1.018)	(0.657)	(1.279)
Dropouts [%]	6.48	-1.95	-3.30	-9.94	-1.07	-1.69	-2.28
	(0.075)	(1.298)	(1.523)	(1.545)	(2.016)	(1.864)	(0.868)
Dropout rate [%]	10.12	-1.13	-0.41	3.96	-2.96	-4.05	-5.08
	(0.100)	(1.207)	(1.390)	(1.449)	(1.849)	(1.771)	(0.871)
Dropout in first period rate [%]	3.72	0.90	4.33	14.62	1.54	1.55	2.39
	(0.053)	(1.142)	(1.195)	(2.815)	(2.356)	(2.717)	(0.621)
Applicants in first period [%]	75.21	-0.01	-0.20	0.25	1.00	1.71	2.25
	(0.094)	(0.108)	(0.292)	(0.151)	(0.116)	(0.153)	(0.134)
Unassigned in first period [%]	28.60	2.07	7.46	34.55	-5.16	-6.72	-7.77
	(0.094)	(0.163)	(0.324)	(0.070)	(0.377)	(0.303)	(0.131)
Assigned to top true preference in first period [%]	11.03	-0.79	-10.53	-26.16	5.45	10.15	13.80
	(0.105)	(0.286)	(1.212)	(0.866)	(1.184)	(0.513)	(0.552)
Graduate - first enrollment [%]	48.43	-0.98	-3.95	-13.99	2.74	3.45	4.27
	(0.065)	(0.156)	(0.165)	(0.140)	(0.079)	(0.096)	(0.035)
Difference in Interim Welfare Relative to Baseline							
Overall	–	-0.63	-1.80	-4.51	0.90	1.30	1.97
		(0.204)	(0.190)	(0.107)	(0.204)	(0.087)	(0.124)
Difference in Ex-Post Welfare Relative to Baseline							
Overall	–	-0.93	-3.00	-11.03	0.94	1.12	1.74
		(0.150)	(0.161)	(0.112)	(0.104)	(0.135)	(0.112)
Males	–	-0.90	-3.14	-11.66	1.33	1.93	2.15
		(0.254)	(0.138)	(0.258)	(0.238)	(0.134)	(0.187)
Females	–	-0.96	-2.86	-10.38	0.54	0.30	1.33
		(0.258)	(0.300)	(0.306)	(0.285)	(0.363)	(0.283)
Low income	–	-0.60	-1.77	-6.55	0.99	1.19	1.48
		(0.332)	(0.214)	(0.299)	(0.166)	(0.213)	(0.278)
High income	–	-1.28	-4.29	-15.74	0.88	1.05	2.01
		(0.504)	(0.341)	(0.316)	(0.281)	(0.301)	(0.455)

Note: Baseline column shows levels; other columns show percentage changes relative to the baseline for outcomes, and absolute differences for welfare measures. Standard errors computed using the delta method for percentage changes and error propagation for welfare differences, accounting for correlation between baseline and counterfactual simulations (common seeds across 5 simulations). Switching and dropout rates are computed relative to first-period enrollees. Welfare measures are in millions of Chilean pesos (2014) and are given subjective beliefs.

TABLE F.IV

COUNTERFACTUAL POLICIES: TURKISH AND FINNISH MECHANISMS (% CHANGE FROM BASELINE)

		Turkish Rules			Finnish Rules		
Outcome	Baseline	$\psi = 0.9$	$\psi = 0.8$	$\psi = 0.7$	$\psi = 1.1$	$\psi = 1.2$	$\psi = 1.3$
Re-applicants [%]	15.06	-21.62	-31.98	-35.46	-23.15	-25.24	-26.24
	(0.096)	(0.494)	(0.633)	(0.390)	(0.417)	(0.571)	(0.510)
Program switchings [%]	2.68	-41.20	-59.03	-71.30	-33.35	-37.25	-40.92
	(0.020)	(0.572)	(1.576)	(0.864)	(1.199)	(1.265)	(1.741)
Program switching rate [%]	4.19	-41.39	-59.19	-71.45	-32.99	-36.67	-40.11
	(0.037)	(0.583)	(1.590)	(0.907)	(1.151)	(1.260)	(1.766)
Major switchings [%]	2.09	-38.09	-56.23	-68.61	-31.08	-33.82	-39.28
	(0.018)	(0.719)	(1.845)	(1.016)	(1.695)	(1.683)	(2.343)
Major switching rate [%]	3.27	-38.29	-56.41	-68.78	-30.70	-33.22	-38.46
	(0.033)	(0.736)	(1.866)	(1.045)	(1.633)	(1.679)	(2.375)
Retakes PSU [%]	5.56	-12.53	-19.44	-22.20	-5.58	-2.56	1.30
	(0.062)	(1.383)	(1.058)	(0.224)	(1.018)	(0.657)	(1.279)
Dropouts [%]	6.48	-1.81	-1.68	-2.69	-0.68	0.07	-0.52
	(0.075)	(1.298)	(1.523)	(1.545)	(2.016)	(1.864)	(0.868)
Dropout rate [%]	10.12	-2.12	-2.08	-3.20	-0.12	1.01	0.85
	(0.100)	(1.207)	(1.390)	(1.449)	(1.849)	(1.771)	(0.871)
Dropout in first period rate [%]	3.72	1.35	-0.02	0.93	-0.40	0.70	2.50
	(0.053)	(1.142)	(1.195)	(2.815)	(2.356)	(2.717)	(0.621)
Applicants in first period [%]	75.21	0.06	-0.26	-0.20	-4.13	-4.33	-4.50
	(0.094)	(0.108)	(0.292)	(0.151)	(0.116)	(0.153)	(0.134)
Unassigned in first period [%]	28.60	-1.07	-1.27	-1.32	1.41	2.23	3.17
	(0.094)	(0.163)	(0.324)	(0.070)	(0.377)	(0.303)	(0.131)
Assigned to top true preference in first period [%]	11.03	7.15	8.95	9.87	11.10	13.10	13.36
	(0.105)	(0.286)	(1.212)	(0.866)	(1.184)	(0.513)	(0.552)
Graduate - first enrollment [%]	48.43	2.13	2.89	3.61	0.94	0.55	0.29
	(0.065)	(0.156)	(0.165)	(0.140)	(0.079)	(0.096)	(0.035)
Difference in Interim Welfare Relative to Baseline							
Overall	–	0.52	0.47	0.60	1.27	1.89	2.22
		(0.204)	(0.190)	(0.107)	(0.204)	(0.087)	(0.124)
Difference in Ex-Post Welfare Relative to Baseline							
Overall	–	0.64	0.45	0.74	0.93	1.12	1.01
		(0.150)	(0.161)	(0.112)	(0.104)	(0.135)	(0.112)
Males	–	0.77	0.87	1.01	1.41	1.43	1.41
		(0.254)	(0.138)	(0.258)	(0.238)	(0.134)	(0.187)
Females	–	0.51	0.01	0.46	0.44	0.81	0.62
		(0.258)	(0.300)	(0.306)	(0.285)	(0.363)	(0.283)
Low income	–	0.52	0.02	0.78	0.73	0.99	0.77
		(0.332)	(0.214)	(0.299)	(0.166)	(0.213)	(0.278)
High income	–	0.77	0.90	0.69	1.15	1.26	1.27
		(0.504)	(0.341)	(0.316)	(0.281)	(0.301)	(0.455)

Note: Baseline column shows levels; other columns show percentage changes relative to the baseline for outcomes, and absolute differences for welfare measures. Standard errors computed using the delta method for percentage changes and error propagation for welfare differences, accounting for correlation between baseline and counterfactual simulations (common seeds across 5 simulations). Switching and dropout rates are computed relative to first-period enrollees. Welfare measures are in millions of Chilean pesos (2014) and are given subjective beliefs.

TABLE G.I

GOODNESS OF FIT: APPLICATION AND PREFERENCE MOMENTS

	Overall		Female		Low-income		High-ability	
Moment	Model	Data	Model	Data	Model	Data	Model	Data
Top true pref. is reported first (Period 1)	0.24	0.44	0.28	0.40	0.27	0.38	0.26	0.52
Top true pref. is reported first (Period 2)	0.44	0.45	0.43	0.45	0.42	0.40	0.51	0.51
Share that applies (Period 1)	0.77	0.66	0.74	0.68	0.58	0.56	0.89	0.77
Share that applies (Period 2)	0.21	0.22	0.21	0.25	0.20	0.20	0.18	0.24
Share changed broad major preference	0.09	0.06	0.13	0.06	0.10	0.07	0.10	0.05
Share changed broad major from original order	0.10	0.04	0.13	0.03	0.09	0.09	0.07	0.00

Notes: This table presents moments related to application behavior and reported preferences. Moments are shown across four demographic groups: Overall population, Female students, Low-income students, and High-ability students.

TABLE G.II

GOODNESS OF FIT: SWITCHING AND OUTCOME MOMENTS

	Overall		Female		Low-income		High-ability	
Moment	Model	Data	Model	Data	Model	Data	Model	Data
Share switching math type	0.01	0.02	0.01	0.02	0.02	0.01	0.02	0.02
Share switching major within math (Period 1)	0.02	0.03	0.02	0.03	0.02	0.02	0.02	0.04
Share switching math within major (Period 1)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Share switching up (Period 1)	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.01
Share switching down (Period 1)	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.02
Share switching out feasible (Period 1)	0.03	0.03	0.02	0.03	0.03	0.02	0.03	0.04
Share switching out unfeasible (Period 1)	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01
Share switching major	0.04	0.05	0.03	0.04	0.04	0.03	0.03	0.06
Share switching broad major	0.03	0.03	0.02	0.03	0.03	0.02	0.03	0.04
Share switching university	0.03	0.04	0.02	0.04	0.03	0.03	0.04	0.06
Share retakes	0.08	0.21	0.08	0.23	0.08	0.20	0.07	0.21
Share dropouts	0.07	0.06	0.06	0.06	0.10	0.07	0.07	0.07
Share program switching	0.04	0.08	0.03	0.07	0.04	0.05	0.05	0.10
Share reapplicants assigned top reported preference	0.43	0.25	0.47	0.23	0.41	0.26	0.52	0.26
Share reapplicants	0.19	0.15	0.17	0.16	0.16	0.13	0.17	0.18
Share major switching	0.04	0.05	0.03	0.04	0.04	0.03	0.03	0.06
Share enrolls in first period (Period 2)	0.09	0.17	0.11	0.17	0.08	0.14	0.07	0.19
Share dropout end of period 1	0.03	0.03	0.03	0.03	0.06	0.03	0.04	0.03
Share first year students (Period 2)	0.13	0.22	0.14	0.23	0.12	0.18	0.12	0.26
Share second year students (Period 2)	0.56	0.39	0.54	0.39	0.34	0.30	0.70	0.49

Notes: This table presents moments related to switching behavior and enrollment dynamics across the four demographic groups.

TABLE G.III

GOODNESS OF FIT: GRADE MOMENTS

	Grade Yr 1		Grade Yr 2			Time Series No Switchers		Time Series Switchers	
Coefficient	Model	Data	Model	Data	Coefficient	Model	Data	Model	Data
Observed ability	0.423	0.456	0.408	0.408	Intercept	0.664	0.854	3.731	3.560
Top preference	-0.000	0.076	0.003	0.045	Year 1 grade	0.865	0.808	0.282	0.279
Science	3.720	3.691	3.806	3.820					
Social Science	4.434	4.058	4.520	4.234					
Education and Humanities	4.582	4.272	4.692	4.453					
Health	4.283	4.261	4.386	4.605					
ROL share same major	0.269	0.170	0.209	0.130					
Signal period 1			0.820	0.702					
Switch major × Signal 1			-0.725	-0.556					
Switch math type × Signal 1			0.015	-0.127					

Notes: This table presents coefficients from grade equations (year 1 and year 2 for students enrolled both periods) and time series for grades equations (for no switchers and switchers). Each coefficient is labeled in the first and middle columns.

TABLE G.IV

GOODNESS OF FIT: REGRESSION DISCONTINUITY DESIGN (RDD) MOMENTS

	RDD Switching		RDD Switch Up/Out		RDD Reapplications		RDD BVP	
Coefficient	Model	Data	Model	Data	Model	Data	Model	Data
Intercept	0.120	0.167	0.067	0.048	0.276	0.234	0.059	0.046
Causal effect	-0.088	-0.051	-0.063	-0.034	-0.144	-0.087	0.042	0.025

Notes: This table presents coefficients from four regression discontinuity designs (RDD). RDD Switching examines the effect of admission to your top-reported preference on switching. RDD Switch Up/Out examines switching up or out to unfeasible programs. RDD Reapplications examines reapplication decisions. RDD BVP examines the effect of scoring above the 600 threshold on applying to BVP-eligible programs (Beca Vocación de Profesor, a teacher scholarship). Each RDD shows the intercept and the causal effect (treatment effect of being above the cutoff).

TABLE G.V

GOODNESS OF FIT: MAJOR MARKET SHARES - MAJORS IN ROL

Moment	Overall		Female		Low-income		High-ability	
	Model	Data	Model	Data	Model	Data	Model	Data
Average Share - Majors in ROL 1								
Business & Management	0.06	0.13	0.03	0.11	0.02	0.11	0.09	0.13
Agriculture	0.03	0.02	0.05	0.02	0.04	0.02	0.03	0.02
Arts & Architecture	0.11	0.07	0.12	0.08	0.07	0.05	0.10	0.07
Basic Sciences	0.06	0.05	0.04	0.04	0.05	0.04	0.06	0.05
Social Sciences	0.10	0.12	0.16	0.14	0.06	0.12	0.13	0.12
Law	0.03	0.05	0.05	0.05	0.02	0.04	0.03	0.05
Education	0.09	0.07	0.10	0.09	0.11	0.09	0.06	0.07
Humanities	0.01	0.02	0.02	0.03	0.02	0.02	0.01	0.02
Health	0.23	0.21	0.23	0.30	0.37	0.23	0.21	0.21
Technology	0.28	0.24	0.21	0.12	0.25	0.27	0.29	0.25
Average Share - Majors in ROL 2								
Business & Management	0.03	0.11	0.01	0.08	0.01	0.09	0.05	0.11
Agriculture	0.02	0.03	0.04	0.03	0.03	0.03	0.01	0.03
Arts & Architecture	0.06	0.06	0.08	0.06	0.04	0.05	0.05	0.06
Basic Sciences	0.03	0.04	0.03	0.04	0.02	0.04	0.04	0.05
Social Sciences	0.06	0.14	0.09	0.16	0.04	0.14	0.08	0.14
Law	0.02	0.04	0.04	0.04	0.02	0.04	0.02	0.04
Education	0.14	0.09	0.19	0.11	0.16	0.12	0.06	0.08
Humanities	0.02	0.02	0.04	0.03	0.03	0.02	0.01	0.02
Health	0.38	0.26	0.29	0.36	0.49	0.28	0.35	0.26
Technology	0.24	0.19	0.19	0.09	0.16	0.20	0.34	0.20
Average Dummy - Majors in ROL 1								
Business & Management	0.09	0.22	0.05	0.19	0.03	0.19	0.13	0.22
Agriculture	0.05	0.04	0.07	0.05	0.05	0.04	0.03	0.04
Arts & Architecture	0.17	0.12	0.19	0.14	0.10	0.10	0.15	0.12
Basic Sciences	0.13	0.15	0.10	0.14	0.11	0.13	0.12	0.16
Social Sciences	0.17	0.22	0.23	0.25	0.10	0.20	0.19	0.22
Law	0.06	0.09	0.11	0.09	0.04	0.07	0.06	0.09
Education	0.11	0.13	0.13	0.16	0.14	0.16	0.07	0.12
Humanities	0.03	0.06	0.04	0.08	0.04	0.05	0.01	0.06
Health	0.27	0.27	0.26	0.38	0.42	0.30	0.25	0.27
Technology	0.37	0.35	0.28	0.21	0.33	0.38	0.36	0.36
Average Dummy - Majors in ROL 2								
Business & Management	0.05	0.18	0.03	0.14	0.02	0.16	0.07	0.18
Agriculture	0.04	0.06	0.06	0.06	0.04	0.06	0.02	0.06
Arts & Architecture	0.10	0.11	0.14	0.11	0.07	0.09	0.08	0.11
Basic Sciences	0.09	0.12	0.07	0.11	0.06	0.13	0.09	0.13
Social Sciences	0.10	0.23	0.15	0.25	0.06	0.23	0.12	0.23
Law	0.06	0.07	0.10	0.06	0.05	0.06	0.05	0.07
Education	0.17	0.16	0.23	0.18	0.20	0.20	0.07	0.15
Humanities	0.06	0.06	0.10	0.07	0.07	0.07	0.01	0.06
Health	0.43	0.33	0.33	0.43	0.54	0.35	0.39	0.33
Technology	0.31	0.29	0.25	0.17	0.21	0.30	0.41	0.29

Notes: This table shows major shares and dummies in rank-ordered lists (ROL) in periods 1 and 2, across four demographic groups. “Average Share” is the average within-student fraction of a ROL in a given major. “Average Dummy” is the proportion of students who list at least one program from that major.

TABLE G.VI

GOODNESS OF FIT: BROAD MAJOR AND MATH TYPE SHARES IN ROL

Moment	Overall		Female		Low-income		High-ability	
	Model	Data	Model	Data	Model	Data	Model	Data
Average Share - Broad Majors in ROL 1								
Science	0.37	0.37	0.30	0.39	0.34	0.32	0.37	0.37
Social Science	0.30	0.31	0.35	0.19	0.16	0.33	0.35	0.32
Education and Humanities	0.10	0.09	0.12	0.12	0.13	0.11	0.07	0.09
Health	0.23	0.21	0.23	0.30	0.37	0.23	0.21	0.21
Average Share - Broad Majors in ROL 2								
Science	0.30	0.35	0.26	0.34	0.21	0.31	0.39	0.35
Social Science	0.16	0.26	0.22	0.16	0.10	0.27	0.19	0.27
Education and Humanities	0.16	0.12	0.22	0.14	0.20	0.14	0.07	0.11
Health	0.38	0.26	0.29	0.36	0.49	0.28	0.35	0.26
Average Dummy - Broad Majors in ROL 1								
Science	0.43	0.50	0.35	0.51	0.40	0.45	0.42	0.50
Social Science	0.33	0.45	0.38	0.32	0.19	0.46	0.39	0.46
Education and Humanities	0.12	0.17	0.14	0.21	0.16	0.19	0.08	0.17
Health	0.27	0.27	0.26	0.38	0.42	0.30	0.25	0.27
Average Dummy - Broad Majors in ROL 2								
Science	0.35	0.47	0.30	0.45	0.25	0.43	0.45	0.47
Social Science	0.19	0.39	0.25	0.28	0.12	0.40	0.22	0.40
Education and Humanities	0.18	0.20	0.25	0.23	0.21	0.24	0.07	0.19
Health	0.43	0.33	0.33	0.43	0.54	0.35	0.39	0.33
Average Share - Math Types in ROL 1								
Low math intensity	0.40	0.39	0.69	0.47	0.47	0.41	0.37	0.38
High math intensity	0.60	0.61	0.31	0.53	0.53	0.59	0.63	0.62
Average Share - Math Types in ROL 2								
Low math intensity	0.38	0.45	0.69	0.53	0.48	0.49	0.28	0.43
High math intensity	0.62	0.55	0.31	0.47	0.52	0.51	0.72	0.57
Average Dummy - Math Types in ROL 1								
Low math intensity	0.46	0.36	0.75	0.45	0.54	0.37	0.42	0.36
High math intensity	0.66	0.64	0.39	0.55	0.59	0.63	0.67	0.64
Average Dummy - Math Types in ROL 2								
Low math intensity	0.45	0.53	0.75	0.61	0.55	0.59	0.31	0.50
High math intensity	0.68	0.47	0.39	0.39	0.58	0.41	0.75	0.50

Notes: This table shows broad major and math type shares and dummies in rank-ordered lists (ROL) in periods 1 and 2, across four demographic groups.

TABLE G.VII

GOODNESS OF FIT: MARKET SHARES IN PERIODS 1 AND 2

Moment	Overall		Female		Low-income		High-ability	
	Model	Data	Model	Data	Model	Data	Model	Data
Major Market Shares - Period 1								
No College	0.36	0.26	0.39	0.28	0.55	0.30	0.21	0.22
Business & Management	0.04	0.12	0.02	0.11	0.01	0.11	0.08	0.11
Agriculture	0.02	0.02	0.03	0.02	0.02	0.01	0.02	0.02
Arts & Architecture	0.08	0.06	0.08	0.07	0.03	0.04	0.08	0.06
Basic Sciences	0.04	0.04	0.03	0.04	0.03	0.04	0.05	0.06
Social Sciences	0.07	0.07	0.10	0.09	0.03	0.06	0.11	0.08
Law	0.02	0.04	0.03	0.04	0.01	0.02	0.02	0.04
Education	0.06	0.05	0.06	0.06	0.05	0.06	0.05	0.05
Humanities	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.02
Health	0.13	0.12	0.13	0.18	0.15	0.12	0.15	0.13
Technology	0.18	0.20	0.12	0.09	0.11	0.22	0.22	0.21
Major Market Shares - Period 2								
No College	0.31	0.12	0.32	0.11	0.54	0.15	0.18	0.09
Business & Management	0.04	0.15	0.02	0.14	0.01	0.14	0.08	0.14
Agriculture	0.02	0.02	0.04	0.02	0.02	0.02	0.02	0.02
Arts & Architecture	0.08	0.07	0.09	0.08	0.03	0.05	0.08	0.07
Basic Sciences	0.04	0.04	0.03	0.04	0.02	0.04	0.05	0.05
Social Sciences	0.07	0.09	0.11	0.12	0.03	0.08	0.11	0.10
Law	0.02	0.04	0.03	0.04	0.01	0.03	0.03	0.05
Education	0.07	0.07	0.08	0.09	0.06	0.08	0.05	0.07
Humanities	0.01	0.02	0.02	0.02	0.02	0.02	0.00	0.02
Health	0.16	0.16	0.16	0.24	0.19	0.17	0.17	0.17
Technology	0.18	0.23	0.11	0.10	0.08	0.24	0.23	0.23
Broad Major Market Shares - Period 1								
No College	0.36	0.26	0.39	0.28	0.55	0.30	0.21	0.22
Science	0.24	0.26	0.18	0.15	0.16	0.27	0.30	0.28
Social Science	0.20	0.28	0.22	0.31	0.08	0.24	0.29	0.30
Education and Humanities	0.07	0.06	0.08	0.08	0.07	0.07	0.06	0.07
Health	0.13	0.12	0.13	0.18	0.15	0.12	0.15	0.13
Broad Major Market Shares - Period 2								
No College	0.31	0.12	0.32	0.11	0.54	0.15	0.18	0.09
Science	0.24	0.29	0.18	0.16	0.12	0.29	0.30	0.30
Social Science	0.21	0.35	0.24	0.38	0.07	0.30	0.30	0.36
Education and Humanities	0.08	0.08	0.10	0.11	0.08	0.09	0.06	0.08
Health	0.16	0.16	0.16	0.24	0.19	0.17	0.17	0.17

Notes: This table shows market shares across majors and broad majors in periods 1 and 2 across four demographic groups. Market shares represent the proportion of students enrolled in each major category. Although the model tends to overpredict non-enrollment, we hold the outside option value flow fixed in counterfactuals, which attenuates issues with our welfare comparisons.

TABLE G.VIII

GOODNESS OF FIT: ADDITIONAL APPLICATION AND ASSIGNMENT CHARACTERISTICS

Moment	Overall		Female		Low-income		High-ability	
	Model	Data	Model	Data	Model	Data	Model	Data
Share unassigned in 1st period	0.292	0.414	0.309	0.418	0.496	0.522	0.134	0.296
Share assigned to 1st preference	0.526	0.297	0.522	0.293	0.385	0.243	0.681	0.353
Share assigned to 2nd preference	0.100	0.145	0.090	0.142	0.067	0.109	0.100	0.176
Share assigned to 3rd preference	0.033	0.079	0.033	0.079	0.021	0.062	0.033	0.095
Student observed ability score (top reported pref.)	0.963	1.111	0.900	1.108	0.763	0.948	1.413	1.216
Program observed ability score (top reported pref.)	1.162	1.467	1.091	1.464	1.091	1.349	1.295	1.543
Tuition (top reported pref.)	3.601	3.780	3.432	3.708	3.000	3.345	3.581	3.821
Distance (top reported pref.)	12.126	11.714	12.187	11.684	13.271	12.416	11.462	11.563
Relative position (top reported pref.)	-0.352	-1.849	-0.388	-1.928	-0.883	-2.099	0.780	-1.750
Correlation: Normalized major change and grades	-0.203	-0.098	-0.133	-0.089	-0.169	-0.109	-0.064	-0.106
Correlation: Normalized broad major change and grades	-0.116	-0.063	0.030	-0.093	-0.089	-0.116	-0.087	-0.064
Correlation: Normalized math type change and grades	-0.038	0.006	-0.089	0.010	-0.028	0.007	-0.094	-0.010
Risk in ROL 1 (portfolio diversity)	0.140	0.325	0.133	0.349	0.219	0.363	0.050	0.303
Share of broad majors within ROL 1 (enrolled)	0.949	0.858	0.958	0.850	0.944	0.854	0.959	0.857
Share assigned to top true preference	0.116	0.230	0.152	0.046	0.127	0.168	0.152	0.282
Share of reapplicants assigned to top true pref.	0.046	0.090	0.080	0.046	0.055	0.069	0.068	0.107
Share with top true preference changed	0.462	0.665	0.607	0.617	0.509	0.718	0.448	0.582

Note: This table shows additional application and assignment characteristics from the model and data across four demographic groups. The moments include assignment status, characteristics of top reported preferences, correlations between preference changes and academic performance, portfolio risk measures, and top true preference assignment outcomes.

TABLE G.IX

GOODNESS OF FIT: RESIDUAL VARIANCES AND EXPECTED WAGES

Moment	Overall		Female		Low-income		High-ability	
	Model	Data	Model	Data	Model	Data	Model	Data
SSR: Grades equation (Year 1)	0.672	0.713	0.570	0.638	0.744	0.785	0.568	0.698
SSR: Grades equation (Year 2)	0.608	0.646	0.541	0.588	0.633	0.681	0.517	0.626
Variance of residuals: Stayers (Year 1)	0.603	0.401	0.529	0.364	0.632	0.407	0.524	0.396
Variance of residuals: Stayers (Year 2)	0.607	0.636	0.538	0.575	0.647	0.675	0.521	0.613
Expected log wages: Top true preference	2.784	2.775	2.722	2.765	2.710	2.692	2.830	2.863
Expected log wages: Random assignment	2.521	2.548	2.491	2.540	2.478	2.472	2.567	2.638

Note: This table shows the sum of squared residuals (SSR) from grade equations, variance of residuals for students who stay in their initial assignment, and expected log wages under different assignment scenarios across four demographic groups. Expected wages compare outcomes when students are assigned to their top true preference versus random assignment in the surveys.

TABLE G.X

GOODNESS OF FIT: CORRELATIONS BETWEEN SIGNALS AND OUTCOMES

Moment	Overall		Female		Low-income		High-ability	
	Model	Data	Model	Data	Model	Data	Model	Data
Correlation: Signal and graduation from 1st enrollment	0.163	0.318	0.140	0.283	0.222	0.338	0.158	0.307
Correlation: Signal and dropout	-0.220	-0.281	-0.213	-0.301	-0.273	-0.344	-0.225	-0.259
Correlation: Signal and program switch	-0.158	-0.218	-0.125	-0.182	-0.196	-0.246	-0.111	-0.216
Correlation: Signal and major switch	-0.166	-0.213	-0.132	-0.181	-0.193	-0.227	-0.111	-0.208
Correlation: Signal and broad major switch	-0.134	-0.158	-0.076	-0.138	-0.157	-0.208	-0.114	-0.159
Correlation: Signal and math type switch	-0.117	-0.139	-0.069	-0.140	-0.138	-0.121	-0.117	-0.152
Correlation: Signal and major switch within math type	-0.120	-0.171	-0.117	-0.134	-0.139	-0.206	-0.057	-0.157
Correlation: Signal and math type switch within major	-0.036	-0.099	-0.022	-0.097	-0.050	-0.066	-0.053	-0.111
Correlation: Signal and switch up (higher preference)	-0.027	0.065	-0.016	0.062	-0.039	0.016	-0.018	0.069
Correlation: Signal and switch down (lower preference)	-0.048	-0.133	-0.050	-0.122	-0.043	-0.129	-0.019	-0.130
Correlation: Signal and switch out (feasible)	-0.153	-0.225	-0.121	-0.211	-0.181	-0.216	-0.101	-0.239
Correlation: Signal and switch out (unfeasible)	-0.032	-0.011	-0.025	-0.011	-0.044	-0.044	-0.045	-0.011

Note: This table shows correlations between students' ability signals and various outcomes across four demographic groups.

TABLE G.XI

GOODNESS OF FIT: PSU SCORE EVOLUTION FOR RETAKERS

Moment	Model	Data
Mean log ratio (future/actual): Language	0.026	0.040
Variance of log ratio: Language	0.006	0.010
Mean log ratio (future/actual): Math	0.024	0.043
Variance of log ratio: Math	0.006	0.007
Mean log ratio (future/actual): History (takers)	0.034	0.055
Variance of log ratio: History (takers)	0.005	0.010
Mean log ratio (future/actual): Science (takers)	0.016	0.070
Variance of log ratio: Science (takers)	0.006	0.012
Mean log ratio (future/actual): History (non-takers)	0.045	0.067
Variance of log ratio: History (non-takers)	0.005	0.011
Mean log ratio (future/actual): Science (non-takers)	0.014	0.023
Variance of log ratio: Science (non-takers)	0.006	0.013

Note: This table shows the mean and variance of the logarithm of the ratio between future and actual PSU (Prueba de Selección Universitaria) scores for students who retake the exam. For each test, we compute $\log(\text{future score}/\text{actual score})$ and report the mean and variance of these log ratios. Language and Math are mandatory tests taken by all retakers. For History and Science, which are optional, we separate students into "takers" (those who had a score > 0 in the initial attempt) and "non-takers" (those who had a score of 0 initially). These moments capture the distribution of score improvements when retaking the PSU, which is important for understanding reapplication and admission dynamics.

TABLE G.XII

GOODNESS OF FIT: WAGE REGRESSION MOMENTS

Moment	Model	Data
Wage Regression Coefficients		
Intercept (log avg wage)	0.574	0.647
Science (broad major)	0.714	0.610
Social Science (broad major)	0.793	0.634
Educ. & Humanities (broad major)	0.436	0.509
Health (broad major)	0.747	0.629
Dummy: Top reported preference	0.001	0.140
Grades	0.050	0.026
Observed ability	0.004	0.067
Female	-0.005	-0.000
Low-income	-0.030	-0.057
Regression Residuals		
Sum of squared residuals (SSR)	0.414	0.404
Average product of residuals	0.145	0.314
Expected Log Wages Regression		
Coeff: Top true vs. Random (intercept)	0.051	0.063
Beliefs over Wage Uncertainty		
Dep. mass within quintile (top rep. pref.)	0.493	0.496

Note: This table shows wage regression moments. The wage regression estimates the relationship between log individual wages and program characteristics (broad major), student characteristics (grades, observed ability, demographics), and an indicator for being assigned to the top reported preference. The regression pools observations from students assigned to their top true preference and random programs. SSR (Sum of Squared Residuals) and average product of residuals assess regression fit quality. The expected log wages regression compares predicted wages between top true preference and random assignments, with the intercept capturing the average wage difference. The dependent mass within quintile measures the subjective probability that students report of their wage falling within $\pm 10\%$ of the expected wage reported in the survey.

TABLE G.XIII

GOODNESS OF FIT: LEARNING AND SUBJECTIVE BELIEFS

Moment	Model	Data
Panel A: Difference in Grade Expectations (Subjective - Realized)		
Science (broad major)	0.630	0.630
Social Science (broad major)	0.664	0.664
Educ. & Humanities (broad major)	0.361	0.361
Health (broad major)	0.843	0.843
Female	-0.185	-0.185
Low-income	0.084	0.084
Avg. obs. ability	-0.207	-0.207
Verbal type	-0.045	-0.045
Panel B: Log Variance of Ability Beliefs (Realized)		
Science (broad major)	-0.252	-0.191
Social Science (broad major)	-0.227	-0.139
Educ. & Humanities (broad major)	-0.480	-0.548
Health (broad major)	-0.957	-0.825
Female	-0.249	-0.245
Low-income	0.081	0.068
Avg. obs. ability	-0.524	-0.503
Verbal type	0.092	0.351
Panel C: Log Variance of Ability Beliefs (Subjective)		
Science (broad major)	-1.602	-1.768
Social Science (broad major)	-1.709	-1.746
Educ. & Humanities (broad major)	-1.522	-1.770
Health (broad major)	-1.836	-1.749
Female	-0.031	-0.070
Low-income	0.001	0.003
Avg. obs. ability	-0.307	-0.302
Verbal type	-0.069	-0.175
Panel D: Change in Norm of Broad Majors (Top Quartile)		
Science (broad major)	0.117	0.089
Social Science (broad major)	0.053	0.029
Educ. & Humanities (broad major)	0.080	0.048
Health (broad major)	0.058	0.070
Signal	-0.082	-0.120

Note: This table shows moments related to learning and subjective beliefs. Panel A reports the difference between subjective grade expectations and realized grades at the top reported preference. Panels B and C show how the log variance of ability beliefs varies by demographics and program characteristics, comparing realized beliefs (based on actual enrollment) and subjective beliefs (stated expectations). Panel D shows the regression of the change in the norm of broad majors for students in the top quartile, with coefficients on broad major dummies and the signal about ability.