College Application Mistakes and the Design of Information Policies at Scale

Tomás Larroucau
Arizona State University

Ignacio Rios
University of Texas at Dallas

Anaïs Fabre
Institute for Fiscal Studies

Christopher Neilson
Yale University and NBER

February 27, 2026

Abstract

We study whether large-scale information interventions can improve college applicant outcomes. Using nationwide surveys, we document widespread information frictions and the prevalence of application mistakes. To address these frictions, we conducted a field experiment that provided applicants with personalized information on admission chances and program characteristics. The intervention increased assignment probabilities for previously unmatched students by 44% and increased placement into higher-ranked programs by 20%. Following these results and with the collaboration of policymakers, we successfully scaled up the policy nationwide. Our findings suggest that personalized information policies implemented at scale can effectively reduce application mistakes and improve student outcomes.

The authors wish to thank *Ministerio de Educación* (MINEDUC) of the government of Chile and DEMRE for their support and collaboration throughout the research project. This project would not have been possible without the support of JPAL and ConsiliumBots. The authors would like to disclose the following: Tomás Larroucau and Ignacio Rios have previously collaborated with the centralized clearinghouse and the Ministry of Education in designing the assignment process. Christopher Neilson is the founder and CEO of ConsiliumBots, which currently holds active contracts with the Ministry of Education related to the school assignment process.

1 Introduction

The choice of college and major is a key driver of earnings inequality over the life cycle (Altonji et al., 2014). Beyond financial constraints, information frictions may hinder students from making optimal education decisions. While evidence from small-scale college counseling programs suggest they can effectively mitigate these barriers, the increasing popularity of centralized admission systems has further facilitated the market-wide implementation of interventions providing targeted, personalized information to students. Yet, it remains unclear whether such information policies can successfully scale.

Many small-scale programs have failed to replicate when implemented at a broader level, undermining the value of empirical research in the eyes of policymakers. Although several explanations for the so-called "scale-up problem" have been proposed—ranging from statistical inference issues to unrepresentativeness of the experimental population and situation (Al-Ubaydli et al., 2017)—these challenges remain largely unresolved. Information-based policies in education markets are not immune to such concerns. When implemented broadly, these may alter application behavior in ways that generate congestion effects, particularly in the presence of limited capacities and correlated preferences, thereby diminishing their effectiveness. Nevertheless, the ability to leverage college admission platforms to automate the delivery of personalized information offers a promising path forward, as such automation can help circumvent two common barriers to scale: quality drops and diseconomies of scale in implementation costs (Davis et al., 2017).

This paper investigates the unique scale-up opportunities and challenges offered by information policies in college admissions, leveraging a multi-year collaboration with policy-makers. First, we develop a model of students' college application decisions and propose a taxonomy of application mistakes. Second, in collaboration with Chile's Ministry of Education (MINEDUC), we design and administer large-scale surveys to measure the prevalence of these mistakes. Third, to assess whether information policies can mitigate them, we implement a natural field experiment that directly addresses the scale-up challenges identified in Al-Ubaydli et al. (2017). Finally, we use the nationwide rollout of the intervention to assess its effectiveness at scale.

To establish a conceptual framework, we begin by developing a taxonomy of application mistakes, defined as discrepancies between a student's optimal and submitted application list. Using the standard model of students' college application decisions, we identify four sources of mistakes that information policies could alleviate at scale: (i) biased beliefs about admission probabilities, leading to over- and under-confidence mistakes; (ii) incorrect information about program characteristics; (iii) unawareness of available options; and (iv) misunderstanding of the mechanism.

We then turn to measuring the prevalence of the application mistakes identified in our taxonomy and the relevance of their underlying drivers. To do so, we analyze large-scale surveys designed and implemented over a five-year collaboration with MINEDUC, summarized in Figure 1. Initial surveys show that a significant fraction of students misreport their preferences even when restrictions on the length of their lists are not binding, resulting in payoff-relevant application mistakes. In particular, close to 40% of applicants do not list their most preferred program, and 11% of them would have benefited from doing so as they had a strictly positive admission probability. Similarly, approximately 28% of students do not apply to programs they prefer over being unassigned. Among them, at least 4% would have been strictly better off by including such programs in their application. By analyzing elicited beliefs, we find that these mistakes are mainly driven by biased beliefs about admission chances: pessimistic students omit their most preferred programs, while optimistic students neglect to list safety options. These biases are more pronounced among students from public schools and with lower test scores. Addressing these mistakes could therefore play a crucial role in reducing inequalities in educational opportunities.

Figure 1: Timeline of Interventions

To assess whether information policies can help mitigate these mistakes, we continued our collaboration with MINEDUC and designed a natural field experiment. In Chile, as in many countries, the college admissions process is centralized on a single application platform, which we used to gather students' initial rank-ordered lists (ROLs) and create personalized websites accessible throughout the application window. These websites provided relevant information over different margins, including general details about programs on each student's list, personalized admission probabilities and risk assessments, and customized recommendations for other potentially suitable majors. The information shown to students was randomized across four treatment arms, enabling us to evaluate the impact of reducing information frictions on various decision-making margins.

Our results show that providing personalized information about admission probabilities and the risk of remaining unassigned significantly influences students' application decisions. Specifically, students who received this information were 10% more likely to modify their rank-ordered lists compared to those in the control group. These adjustments led to notable improvements in students' admission outcomes. For instance, initially unassigned students were 44% more likely to gain admission to a program after receiving the information. Students who were initially assigned also benefit from the information, experiencing a 20% increase in the probability of being matched to a program ranked higher than the one where they would have been assigned to given their initial list. Furthermore, we find that these effects persist in the longer run, as the intervention increases the likelihood that students remain enrolled in the same program two years later by 34%. These effects suggest that our intervention effectively decreased the incidence of application mistakes, particularly *over-confidence* mistakes, improving students' outcomes.

This large-scale intervention was crafted to carefully address the three sources of the "scale-up" problem highlighted by Al-Ubaydli et al. (2017)—unrepresentativeness of the experimental population, unrepresentativeness of the situation, and statistical inference issues. First, we designed the experiment to be as informative as possible for future scale-up efforts by closely mirroring the target population and real-world conditions. The information provided was directly randomized within the target population, i.e. the sample of early-applicants, ensuring that the experimental sample was representative of the relevant population with respect to both treatment effects and participation costs. The opportunity to run the experiment directly within the actual, high-stake, college admissions procedure also allowed us to avoid introducing any artificial margins of choice. Additionally, by leveraging the centralized admission mechanism to automatize the personalized information we provide, the intervention we crafted is not threatened by quality drops and display economies of scale, ensuring its scalability. On top of these ex-ante steps, and recognizing the risks of scaling up a policy based on false positive findings, we went beyond assessing results with respect to their statistical significance. In particular, we evaluate their Post-Study Probability (Maniadis et al., 2014), which measures the likelihood that a finding is valid once it reaches statistical significance. Our results support the decision to scale up the policy. Finally, acknowledging that the experimental results may still fail to scale due to potential congestion effects,[1] we also

[1]Specifically, if a large number of initially unmatched students respond to the information by shifting their applica-

investigate the underlying drivers of the observed changes in student assignment. We find that the combination of ample seat availability and heterogeneity in student preferences allowed the intervention to improve assignment outcomes. We further exploit the properties of the centralized assignment mechanism and the experimentally collected data to simulate the effect of scaling-up the policy and assess the magnitude of such congestion effects. Our findings indicate that the information policy would still yield benefits if implemented system-wide, suggesting that congestion effects are not large enough to offset its benefits.

In light of these findings, MINEDUC implemented the information policy nationwide in 2023. The evaluation of the scale-up policy, performed using an encouragement design through *WhatsApp* messages, confirms most of the aforementioned results. We find that providing real-time personalized information about students' admission probabilities, alongside warning messages and cutoff scores for all programs in the centralized system—resembling a one-shot implementation of the iterative Deferred Acceptance (DA) algorithm—causally improves students' outcomes, in line with the results of our field experiment. Specifically, by reducing primarily *over-confidence* and *under-confidence* mistakes, we estimate that the policy roughly doubles the probability that initially unmatched students get admitted to the centralized system and the probability to improve their assignment relative to their initial match. This translates into a substantial decrease in unmatched and undermatched students. Improvements in assignment outcomes persist at the enrollment stage, as students who get assigned thanks to the policy are three times more likely to enroll in their assigned program. Furthermore, by observing students' preferences and beliefs before and after the policy implementation, we find that these improvements are primarily driven by changes in beliefs concerning admission probabilities at the bottom of their submitted list rather than at the top, reducing the incidence of biased beliefs on students' application decisions.

Overall, our work demonstrates that well-designed information policies—those that equip students with personalized insights into their admission prospects—can significantly reduce application mistakes at scale, lowering the number of unmatched and undermatched students. By seamlessly integrating with the centralized admissions mechanism, our intervention showcases how market design can drive meaningful improvements in student outcomes at minimal cost. Moreover, because the marginal cost of the intervention is near zero and the induced assignment changes are associated with positive predicted long-run earnings gains, the policy is highly cost-

tions to the same program, the latter might become oversubscribed. As a result, despite modifying their rank-ordered list in response to the information provided, some students would still remain unmatched.

effective. These findings underscore the power of leveraging existing structures to enact scalable, cost-effective, and transformative policy solutions in education.

Related Literature. Our paper is related to three strands of literature. First, we contribute to the literature emphasizing the importance of information policies in alleviating barriers to higher education access. Interest in such interventions grew after evidence that financial constraints alone do not fully explain why high-achieving, low-income students do not apply to selective universities (Hoxby and Avery, 2012). In response, many interventions have focused on small-scale mentoring and counseling programs, which have been shown to significantly improve college enrollment outcomes (Oreopoulos et al., 2017; Carrell and Sacerdote, 2017; Castleman et al., 2014; Castleman and Goodman, 2018; Oreopoulos and Ford, 2019). However, such "boots-on-the-ground" approaches are costly and may be difficult to scale. In contrast, the results of light-touch interventions that provide information about options' characteristics, financial aid, and costs are much more mixed (Hoxby et al., 2013; Gurantz et al., 2021; Hyman, 2020), despite evidence that students update their beliefs after receiving information (Baker et al., 2018; Wiswall and Zafar, 2015b; Bleemer and Zafar, 2018).

These results may suggest that information needs to be provided through in-person interactions to be successful, which lowers the scalability of such policies. However, an alternative explanation for their mixed results is the lack of personalization, as these light-touch interventions often provide general information about college options and broad application recommendations, failing to account for individual students' circumstances and preferences. To address this limitation, we design and evaluate a unique intervention that leverages data from centralized college admission platforms to deliver highly personalized information and targeted alerts, ensuring that students receive personalized guidance based on their individual circumstances and increasing the likelihood of making better application decisions.

Second, we contribute to an emerging literature studying whether and how education interventions can successfully scale by addressing the recently highlighted "scale-up problem" (List, 2022). For example, Davis et al. (2017) develop a method that uses experimental variation to estimate how quality and cost evolve as programs such as tutoring scale up. Similarly, Agostinelli et al. (2025) evaluate a national mentoring program in Mexico and find that enhanced mentor training significantly improves the program's effectiveness at scale.

Although tutoring and mentoring programs could yield sizable benefits, information-based interventions—due to their low marginal costs and scalability—are especially attractive for large-scale implementation. Yet, recent work underscores important congestion externalities that can attenuate their gains. Bobba et al. (2023) implement an RCT providing academic performance feedback to prospective high school students in Mexico City and demonstrate that, under full roll-out, oversubscribed school capacities would partially offset the policy's positive effects. Similarly, Allende et al. (2019) combine a small-scale RCT with a structural model to evaluate an information intervention targeted at parents, finding that capacity constraints would significantly diminish the welfare gains of a citywide expansion.

We contribute to this literature by providing model-free evidence from both a randomized information intervention targeted at college applicants and its nationwide scale-up. Unlike previous studies, we find that our policy remains highly effective even at scale. This success lies in three key elements of its design. First, by leveraging centralized admission mechanisms—which streamline data collection and enable the automated distribution of personalized information—we ensure the scalability of the intervention while avoiding two of the most common pitfalls of large-scale interventions: quality deterioration and diseconomies of scale (Davis et al., 2017). Second, by conducting a thorough analysis of the randomized control trial—identifying the key drivers of its effects—we carefully refined the information policy to maximize its effectiveness at scale. Finally, contrasting with Allende et al. (2019) who provide vertical information about schools' quality, we provide personalized information that is unlikely to lead students to shift their applications in the same direction, minimizing congestion effects. Analyzing the changes in admission triggered by the intervention, we highlight the role of both programs' sufficient capacities and heterogeneity in students' preferences in the successful scale-up of the program.

Finally, our paper contributes to the growing literature on behavioral economics in education market design; see Rees-Jones and Shorrer (2023) for an excellent review. Several studies in this literature have documented application mistakes in settings that use variants of DA. Two prominent types are *obvious mistakes*, where students only list the full-fee version of a program despite being eligible for a reduced-fee alternative (Artemov et al., 2017; Shorrer and Sóvágó, 2021; Hassidim et al., 2020), and *skipping*, where students omit their favorite programs from their list (Hakimov et al., 2023; Chrisander and Bjerre-Nielsen, 2023). These mistakes are typically payoff-irrelevant and stem from underconfidence or misunderstanding of the mechanism's rules. At the same time, other studies have emphasized the role of overconfidence and limited awareness in shaping application behavior. For instance, Arteaga et al. (2022) show that, in the context of school choice, families often submit overly risky applications by not listing enough options, driven by overconfidence about admission chances, lack of awareness of alternatives and search costs. Consistent with our findings, the authors show that providing information on available programs can improve outcomes by helping families make more informed decisions.

We make several contributions to this literature. First, we provide a unifying taxonomy of application mistakes, expanding their sources beyond those typically studied in the literature. Second, we use large-scale survey data to provide lower-bounds on their prevalence and relevance. Third, we identify that strategic behavior coupled with information frictions and biased beliefs are the main drivers of these mistakes. Finally, we highlight how a highly-scalable information policy can mitigate these application mistakes.

Overview. The remainder of this paper is organized as follows. Section 2 introduces the Chilean college admissions system, the data, and our taxonomy of mistakes. Section 3 describes the prevalence and relevance of application mistakes and sheds light on their potential drivers. Section 4 presents our field experiment and the evaluation of its impact. Section 5 reports the results of the scaled-up policy implementation. Finally, Section 7 concludes.

2 Background and Taxonomy of Application Mistakes

In this section, we first present the most relevant institutional details of the Chilean college admissions system. We then describe the different data sources used throughout the paper. Finally, we provide a taxonomy of application mistakes.

2.1 Institutional Details

We focus on the centralized part of Chile's tertiary education system, which encompasses the country's most selective universities.[2] Hereafter, we refer to this as the admissions system. To participate, students are required to take a series of standardized tests: the *Prueba de Selección Universitaria* (PSU) until 2020, the *Prueba de Transición* (PDT) from 2021 to 2022, and the *Prueba de Acceso a la Educación Superior* (PAES) starting in 2023. While Math and Language are mandatory, students can

[2]See Kapor et al. (2020) and Larroucau and Rios (2023) for a more general description of tertiary education in Chile and additional institutional details. As of 2023, 45 out of the 58 universities in Chile participated in the centralized admission system.

select either Sciences or History & Social Sciences as their elective.[3] In addition to these test scores, students receive two additional scores based on their high school performance: one derived from their average grades during high school (*Notas de Enseñanza Media* (NEM)) and another reflecting their relative position within their cohort (*Ranking de Notas* (Rank)). Each program offered by the institutions participating in the centralized system announces the weights assigned to these components, which are then used to calculate a program-specific application score for each student.

After scores are published, students can access an online platform to submit their applications in the form of a Rank-Ordered List (ROL) of up to ten programs.[4] DEMRE, the Chilean equivalent of the College Board, collects these applications, checks students' eligibility in each of their listed programs and, if eligible, computes their program-specific application scores. Applicants to each program are then ranked based on these scores. The final assignment is determined by an algorithm that takes into account students' ROL as well as programs' rankings and capacities and relies on a variant of the DA algorithm, whereby ties on students' scores are not broken (Rios et al., 2021). As a result, programs may exceed their capacities only if there are ties for their last seat. We refer to the application score of the last admitted student as the *cutoff* of each program.

2.2 Data

We combine a panel of administrative data on the admissions process with survey data that we collected to analyze students' mistakes. The administrative data includes information about students (socio-economic characteristics, scores, and applications), programs (weights, capacities, and admission requirements), and admission outcomes (i.e., whether each application was valid and whether the student was admitted or wait-listed) from 2020 to 2023.

To complement this administrative data, we conducted three nationwide surveys in 2020, 2022, and 2023 in collaboration with MINEDUC and DEMRE. These surveys were designed to elicit students' preferences, beliefs, and to characterize the drivers of application mistakes. To this end, these surveys included three main modules: (i) preferences, (ii) beliefs, and (iii) understanding of the admission process. We describe each module and include its most relevant questions in Appendix A.1.2.[5] These surveys were also carefully timed to capture students' beliefs and preferences before key moments in the admissions process. In 2020, 2022 and 2023, MINEDUC and

[3]In 2023, an additional, more advanced, Math test was added to the admission process. See Appendix C.1 for more details on the changes implemented in 2023.

[4]Students apply directly to programs, i.e. university-major pairs. In 2023, MINEDUC increased the number of programs students can list to 20.

[5]The complete translations of the surveys are available upon request.

DEMRE emailed the surveys to students shortly after the application deadline but before the publication of assignment results, so students knew their scores and submitted ROLs but had not yet learned their admission outcomes when completing the surveys. In 2023, we also implemented a baseline survey before the release of the national test results. Since students had not yet applied to the centralized system, this survey elicited their preferences and beliefs regarding hypothetical programs.

2.3 Application Behavior and Taxonomy of Mistakes

In this section, we introduce a framework of application behavior under incomplete information and use it to provide a taxonomy of application mistakes that we will study throughout the paper.

We follow the model of students' applications under strategic reporting introduced by Agarwal and Somaini (2018). Let $\mathcal{J} = \{1, ..., J\}$ be the set of programs, $\mathcal{I} = \{1, ..., I\}$ be the set of students, and $\boldsymbol{X}_i = \{X_{ij}\}_{j \in \mathcal{J}}$ be a matrix that gathers characteristics of student i, program $j \in \mathcal{J}$, and their interaction. For instance, X_{ij} may include student i's gender, program j's tuition, and the distance between student i's home and program j's campus. Furthermore, let $v_i(\boldsymbol{X}_i) = \{v_{ij}(X_{ij})\}_{j \in \mathcal{J}} \in \mathbb{R}^J$ be the vector of indirect utilities that student i derives from attending each program.

As previously discussed, when applying to college, students must choose an ordered subset of programs to apply to. To formalize this, let $\mathcal{R}$ denote the set of all possible rank-ordered lists (ROLs) of programs in $\mathcal{J}$. With a slight abuse of notation, we represent each ROL $R \in \mathcal{R}$ as $R = \left\{r_1, \ldots, r_{|R|}\right\}$, with $r_k \in \mathcal{J}$ for all $k \in \{1, \ldots, |R|\}$ and $k < l$ implying that student i ranks program r_k higher than program r_l. Moreover, let $\mathcal{L}_i : \mathcal{R} \to [0, 1]^J$ be the function mapping each ROL to student i's vector of beliefs about their admission chances at each program in $\mathcal{J}$ under rational expectations. The set of optimal rank-ordered lists for student i corresponds to those maximizing their expected utility, i.e., :[6]

$$\mathcal{R}_i^*(\boldsymbol{X}_i, \mathcal{L}_i, \mathcal{J}) = \left\{ R \in \mathcal{R} \ : \ R \in \underset{\substack{R' \in \mathcal{R}, \\ |R'| \leq K, \, R' \subseteq \mathcal{J}}}{\operatorname{argmax}} \ v_i(\boldsymbol{X}_i) \cdot \mathcal{L}_i(R') - c_i\left(R'\right) \right\} \tag{1}$$

where $\cdot$ represents the dot product between two vectors, K is the maximum number of programs that students are allowed to rank, and $c_i(R)$ is the cost faced by student i from submit-

[6] We define optimal behavior as the one corresponding to expected utility maximization. Thus, we abstract away from recent work exploring non-standard preferences such as risk aversion, expected loss aversion, among others. See Rees-Jones and Shorrer (2023) for a comprehensive review.

ting a rank-ordered list R. The set $\mathcal{R}_i^*(\boldsymbol{X}_i, \mathcal{L}_i, \mathcal{J}) \subseteq \mathcal{R}$ in (1) includes all ROLs that maximize student i's expected utility, assuming accurate knowledge about program characteristics, correct beliefs about admission chances, and awareness of all the available programs. Hence, we say that student i makes a *payoff-relevant* application mistake upon submitting a rank-ordered list $\tilde{R}$ if $\tilde{R} \notin \mathcal{R}_i^*(\boldsymbol{X}_i, \mathcal{L}_i, \mathcal{J})$ and, consequently, there exists an alternative ROL $R \in \mathcal{R}_i^*(\boldsymbol{X}_i, \mathcal{L}_i, \mathcal{J}) \subseteq \mathcal{R}$ such that

$$v_i(\boldsymbol{X}_i) \cdot \mathcal{L}_i(\tilde{R}) - c_i(\tilde{R}) < v_i(\boldsymbol{X}_i) \cdot \mathcal{L}_i(R) - c_i(R). \tag{2}$$

Based on this definition, application mistakes can be classified into two categories based on students' strategic behavior. On the one hand, a student i who behaves strategically—i.e., who submits a list that maximizes their expected utility, as in Equation (1)—may still make an application mistake if they (i) have incorrect information about program characteristics, $\boldsymbol{X}_i$; (ii) miscalculate the admission probabilities associated with their list, $\mathcal{L}_i(\cdot)$; or (iii) are unaware of certain programs in $\mathcal{J}$.[7] While these mistakes are not mutually exclusive, we analyze them separately. On the other hand, a student i who acts non-strategically may also make an application mistake if they submit a ROL based on a criterion that differs from expected utility maximization, leading to a different report.

2.3.1 Mistakes by Strategic Students

Mistakes on Assignment Probabilities. As defined in Agarwal and Somaini (2018), students hold *rational expectations* over their admission probabilities if they (i) know their vector of program-specific admission scores, (ii) accurately estimate the distribution of other students' scores, (iii) correctly infer others' application strategies, and (iv) understand how the assignment mechanism translates these scores and strategies into outcomes. Consequently, they form *correct beliefs* about their admission chances for each program in their rank-ordered list. Since this is a cognitively demanding task, students may hold *biased beliefs* about their admission probabilities. Formally, we say that student i makes a *mistake on assignment probabilities* if they submit a rank-ordered list $\tilde{R} \in \mathcal{R}_i^*(\boldsymbol{X}_i, \tilde{\mathcal{L}}_i, \mathcal{J})$ that is strictly dominated by a ROL $R \in \mathcal{R}_i^*(\boldsymbol{X}_i, \mathcal{L}_i, \mathcal{J})$ relative to the rational expectation probabilities $\mathcal{L}_i$ (i.e., Equation (2) holds), where $\mathcal{R}_i^*(\boldsymbol{X}_i, \tilde{\mathcal{L}}_i, \mathcal{J})$ is the set of optimal ROLs that arises from solving Equation (1) considering student i's biased beliefs about their ad-

[7]Note that, while determining the optimal ROL may be cognitively demanding, Chade and Smith (2006) provide a simple and intuitive algorithm to solve for the portfolio that maximizes students' expected utility when the cost is a function of the number of programs included in the ROL.

mission chances $\tilde{\mathcal{L}}_i$.

Biases in students' beliefs can lead to mistakes in either direction. On the one hand, students may be overly optimistic, exhibiting a positive bias in their assessment of their admission chances. As a result, they may wrongly believe that they will be admitted to one of the programs they have already ranked, perceiving no benefit in adding a safety program to their list. This may potentially lead them to an *over-confidence mistake*, increasing students' risk of resulting *unmatched*. On the other hand, students may be overly pessimistic, exhibiting a negative bias in their perceived admission chances. In this case, they might wrongly believe they have no chance of being admitted to a program they prefer, leading them to omit it from their ROL, even when they are not constrained by its length. This behavior could result in an *under-confidence mistake*, potentially causing students to forgo options that are preferable and attainable.

Beyond misestimating their admission probabilities, students may also misunderstand how the assignment mechanism operates. For instance, they may fail to recognize that DA sequentially processes their rank-ordered list and that they should rank programs in decreasing order of utility. This misunderstanding can result in *ordering mistakes*, where less-preferred programs are ranked above more desirable ones where students might have been admitted. Both *under-confidence* and *ordering mistakes* can lead to *undermatching*, where students receive placements in programs they prefer less than others they could have realistically secured.

Mistakes on Valuations. Information frictions may also lead students to have incorrect information about programs' characteristics such as their costs, expected earnings, and employment rates (Wiswall and Zafar, 2015a; Bleemer and Zafar, 2018). As a result, students may misperceive the indirect utility associated to each program, which may generate sub-optimal applications. Formally, we say that student i makes a *mistake on valuations* if they submit a rank-ordered list $\tilde{R} \in \mathcal{R}_i^*(\tilde{\boldsymbol{X}}_i, \mathcal{L}_i, \mathcal{J})$ that is strictly dominated by a ROL $R \in \mathcal{R}_i^*(\boldsymbol{X}_i, \mathcal{L}_i, \mathcal{J})$ (i.e., Equation (2) holds), where $\mathcal{R}_i^*(\tilde{\boldsymbol{X}}_i, \mathcal{L}_i, , \mathcal{J})$ is the set of optimal ROLs that arises from solving Equation (1) while holding incorrect information about programs characteristics, $\tilde{\boldsymbol{X}}_i$.

Mistakes on Awareness. The mistakes described above assume that students are aware of all programs in the centralized system and choose their rank-ordered list from the full set of programs, $\mathcal{J}$. In practice, however, students may only know a subset of programs, potentially leading them to suboptimal choices. Formally, we say that student i makes a *mistake on awareness* if they submit a rank-ordered list $\tilde{R} \in \mathcal{R}_i^*(\boldsymbol{X}_i, \mathcal{L}_i, \tilde{\mathcal{J}}_i)$ that is strictly dominated by some $R \in \mathcal{R}_i^*(\boldsymbol{X}_i, \mathcal{L}_i, \mathcal{J})$, where $\tilde{\mathcal{J}}_i \subset \mathcal{J}$. Note that such mistakes could also be driven by biased beliefs about admission probabilities. For instance, if search is costly, students may decide to stop their search if they believe that they would be admitted with high enough probability to at least one of the programs they are currently aware of (Arteaga et al., 2022).

2.3.2 Mistakes by Non-Strategic Students.

Regardless of how students perceive $(\boldsymbol{X}_i, \mathcal{L}_i, \mathcal{J}_i)$, they may submit a rank-ordered list that maximizes an objective other than their expected utility (as defined in Equation (1)). A prominent example of this behavior is *truth-telling*, where a student ranks programs solely based on their preferences, without accounting for admission probabilities or application costs. Formally, we say that student i makes a *truth-telling mistake* if they submit a rank-ordered list $\tilde{R} \in \tilde{\mathcal{R}}(\boldsymbol{X}_i, \mathcal{J})$ that is strictly dominated by some $R \in \mathcal{R}_i^*(\boldsymbol{X}_i, \mathcal{L}_i, \mathcal{J})$, where

$$\tilde{\mathcal{R}}(\boldsymbol{X}_i, \mathcal{J}) = \left\{ R \in \mathcal{R} \;:\; R \in \underset{\substack{R' \in \mathcal{R}, \\ |R'| \leq K,\, R' \subseteq \mathcal{J}}}{\operatorname{argmax}} \quad v_i(\boldsymbol{X}_i) \cdot \vec{\mathbf{1}}(R') \right\}, \tag{3}$$

and $\vec{\mathbf{1}}(R') \in \{0,1\}^J$ is a vector indicating the programs included in R', with entries equal to one for programs in R' and zero otherwise.

Note that for students who prefer strictly fewer than K programs over being unassigned (i.e., those for whom $\tilde{R} \in \tilde{\mathcal{R}}(\boldsymbol{X}_i, \mathcal{J})$ satisfies $\left|\tilde{R}\right| < K$), submitting such a rank-ordered list $\tilde{R} \subseteq \mathcal{R}_i^*(\boldsymbol{X}_i, \mathcal{L}_i, \mathcal{J})$ is undominated under the student-proposing Deferred Acceptance algorithm (Haeringer and Klijn, 2009). However, students who prefer more than K programs may benefit from strategically selecting them based on their admission probabilities to maximize their expected utility.

2.3.3 Discussion

Our taxonomy provides a structured micro-foundation for the occurrence of the *obvious mistakes* previously studied in the literature (Artemov et al., 2017; Shorrer and Sóvágó, 2021), linking them to specific information frictions and deviations from expected utility maximization. For instance, our model can rationalize *obvious mistakes* as resulting from students being unaware of programs they should clearly prefer (i.e., *mistakes on awareness*), having biased beliefs about their chances of admission (i.e., *mistakes on admission probabilities*), or misperceiving key characteristics of these programs (i.e., *mistakes on valuations*).

We also want to highlight that having incorrect information about program characteristics and biased beliefs about admission probabilities are necessary but not sufficient conditions for having *application* mistakes, since these deviations need to be (i) large enough, (ii) in the relevant programs (programs the students prefer to their outside option), and (iii) over characteristics with a high enough preference weight, such that correcting beliefs could lead to strict improvements in their expected utility. For instance, if students care mostly about non-pecuniary elements, informing them about returns associated to different majors might not impact their choices (Wiswall and Zafar (2015a)). In addition, note that we define mistakes from a static perspective given students' current preferences and beliefs at the time of the application. Although there is evidence that students' valuations over programs and beliefs about their admission chances might change over time due to learning,[8] we consider interventions that are implemented during the application process, making it unlikely that dynamic effects are relevant for our analysis.

3 Evidence of Mistakes

A key challenge in identifying the application mistakes described above is that they are not directly observable in administrative data. To address this, we conducted several large-scale nationwide surveys designed to elicit students' preferences, beliefs about program characteristics, and perceptions of their admission chances.[9] In this section, we use these survey data to present empirical evidence on the prevalence and underlying drivers of the different types of mistakes made by strategic students.[10]

3.1 Mistakes on Admission Probabilities

As discussed in Section 2.3, mistakes on admission probabilities fall into three categories: *under-confidence*, *ordering*, and *over-confidence* mistakes. Fully characterizing these mistakes would require eliciting students' valuations, their beliefs about admission chances across all programs, and the costs associated with submitting applications. Since collecting such comprehensive data is infeasible, we designed surveys to capture students' perceptions of a carefully selected subset of

[8] See for instance Narita (2018) for school choice settings and Arcidiacono et al. (2016), Fu (2014), and Larroucau and Rios (2023) for college admissions.

[9] Details of these surveys, including translations of key questions, are provided in Appendix A.1.

[10] We also explore *mistakes by non-strategic students* and, specifically, mistakes by truth-tellers. We defer the results from this analysis to Appendix A.4 for completeness.

programs. These survey data, combined with the assumption that application costs are small—which reflects the Chilean setting, where applications are free but may involve a small cognitive cost—enable us to establish lower bounds on the prevalence of each type of mistake, which we summarize in Table 1.

Table 1: Prevalence of Mistakes on Admission Probabilities

	N	%
Panel A: Top-True Preferences (N=10,390)		
Misreport - Exclude	3264	31.41
Ex-Ante Under-confidence Mistake	265	8.12
Ex-Post Under-confidence Mistake	220	6.74
Misreport - Order	891	8.58
Ex-Ante Ordering Mistake	197	22.11
Ex-Post Ordering Mistake	162	18.18
Panel B: Bottom-True Preferences (N=7,311)		
Misreport - Exclude	2033	27.81
Ex-Ante Over-confidence Mistake	85	4.18
Ex-Post Over-confidence Mistake	55	2.71

NOTES. This table provides lower-bounds on the share of students making different types of mistakes. Panel A is restricted to the sample of short-list students who were asked about their top-true preference (2020 survey), while panel B focuses on students who were asked about their bottom-true preference (2022 survey). In both cases, we include all short-list survey respondents who completed the survey, provided consistent answers regarding their true top preference, and are not PACE.

Under-confidence and Ordering Mistakes. A student makes an *under-confidence mistake* if (i) they skip a program they prefer over those they listed, (ii) they have a positive probability of admission to that program, and (iii) the constraint on the length of their list is not binding. Similarly, a student makes an *ordering mistake* if conditions (ii) and (iii) hold, but instead of omitting the program, they rank it below a less-preferred option. Our survey allows us to identify one such program: the student's top-true preference. If a short-list student has a strictly positive chance of admission at their top-true preference, their expected utility would be strictly higher if they placed this program at the top of their ranked-ordered list. Thus, skipping it constitutes an under-confidence mistake, while ranking it below less-preferred options constitutes an ordering mistake. Since students may exclude or misrank multiple programs that meet conditions (i) and (ii)—and we only elicit their most-preferred one—our estimates provide a lower bound on the prevalence of these mistakes.[11]

[11] A challenge to measure mistakes using students' top-true preference is that they may incorrectly interpret the hypothetical scenario posited by our survey question. Indeed, when asked why they skipped their reported top true preference, some students declare that they did so because of their low chances of graduation or of the program's cost. These reasons are inconsistent with the definition of top-true preference because these characteristics should enter the indirect utility and be captured in the students' preference order. Thus, we take a conservative approach to analyze those

Panel A in Table 1 provides a characterization of under-confidence and ordering mistakes. First, we observe that 39.99% of short-list students misreport their top-true preference: 78.54% of them exclude it from their rank-ordered list, while the remainder includes it, but do not rank it first.[12] Second, among students who skipped their top-true preference, 8.12% had a positive admission probability, and 6.74% would have been admitted, making ex-ante and ex-post under-confidence mistakes, respectively. Finally, among students who misranked their top-true preference, 22.11% (18.18%, resp.) had a positive admission probability (an application score above the cutoff, resp.) but were assigned to a program they ranked higher, thus making an *ex-ante* (*ex-post*, resp.) *ordering mistake*. Overall, we estimate that at least 11.1% of short-list students who misreport their top-true preference made a mistake. Since this estimate only provides a lower bound, it suggests that a substantial fraction of students make *under-confidence* or *ordering* mistakes, indicating that these mistakes are sizable in our setting.

Over-confidence. A student makes an over-confidence mistake if (i) they skip a program that they prefer to being unassigned and for which they have a positive admission probability, (ii) they face a positive risk of being unassigned, and (iii) the constraint on the length of their list is not binding. As the set of programs students prefer to the outside option is likely to be large, it is not feasible to gather an exhaustive list. To address this, we ask students about their *bottom-true preference*, defined as any program not included in their list that they would prefer over being unassigned. Consequently, this analysis only provides a lower-bound on the prevalence of *over-confidence mistakes*, similar to the mistakes discussed above.

Panel B in Table 1 presents the results for *over-confidence* mistakes. We find that 27.81% of short-list students exclude their bottom-true preference from their rank-ordered list. Among them, 4.18% face a positive risk of being unassigned and have a strictly positive probability of admission to this program, making an ex-ante over-confidence mistake. Additionally, 2.71% have an application score that exceeds the admission cutoff for their bottom-true program while still facing a positive risk of being unassigned, resulting in an ex-post over-confidence mistake.[13] These mistakes are particularly concerning because they are highly consequential. While under-confidence and or-

mistakes, restricting the sample to students who either did not skip their top-true preference or listed "low chances of admission" as the reason for skipping it. They represent 67% of survey respondents.

[12]Note that students for whom the constraint on the length of their list is binding may also make an *ordering* mistake. However, since we focus on providing lower bounds for these mistakes, and nearly 90% of students are short-list students, we omit these cases for simplicity.

[13]Note that, as students may be affected by the field experiment we discuss in Section 4, we exclude students who received one of our information treatments from this analysis.

dering mistakes primarily affect the quality of assigned programs, over-confidence mistakes can result in students being unmatched by the centralized system. Given the substantial monetary returns associated with university enrollment (Hastings et al., 2013), failing to secure a spot can have lasting educational and economic consequences, making these mistakes especially impactful.

Main Driver: Biased Beliefs. Understanding the drivers of mistakes is essential to design effective information policies and provide students with tools to improve their applications. To this end, we analyze the extent to which biases in beliefs on admission probabilities explain *mistakes on assignment probabilities*.[14] Specifically, we used our surveys to elicit students' beliefs about their admission probabilities for a subset of programs, including their top-true and bottom-true preferences, as well as their belief about their overall probability of being assigned. Using these reported beliefs, we compute biases in students' perceptions by taking the difference between their elicited beliefs and their rational-expectation probabilities. Consequently, positive bias values indicate optimism (over-confidence), while negative bias values reflect pessimism (under-confidence).[15]

In Table 2, we report the results of linear probability models in which the dependent variable is a dummy equal to one if the student made a specific type of mistake, and zero otherwise. The main variables of interest are the bias η at the top- and bottom-true preferences and the bias in the overall assignment probability $\bar{\eta}$. To capture the differential effects of optimism and pessimism on mistakes, we consider the bias separately depending on whether it is positive or negative.[16] We additionally control for students' demographic characteristics such as their gender, score, and region of residence and, consistent with the above analyses, we focus on short-list students and exclude survey respondents whose top-/bottom-true preference is not valid.

Columns (1) and (2) show that students who are more optimistic about their chances of admission to their top true preference are significantly less likely to make under-confidence and ordering mistakes. Conversely, students with greater pessimism are significantly more likely to make these mistakes. Notably, the effect of pessimism is substantially larger than that of optimism, suggesting that pessimism is the primary driver of under-confidence and ordering mistakes. Column

[14]In Appendix A.2, we further analyze the drivers of these biases.

[15]We describe the computation of their rational-expectation probabilities in Appendix B.3). Formally, we denote by $\tilde{p}_{ij}$ and p_{ij} the elicited belief and the rational-expectation admission probability of student i in program j, respectively, and we denote the bias as $\eta_{ij} = \tilde{p}_{ij} - p_{ij}$. Similarly, $\tilde{\rho}_i = 1 - \prod_{r \in R}(1 - \tilde{p}_r)$ and $\rho_i = 1 - \prod_{r \in R}(1 - p_r)$ denote the elicited and rational-expectation overall probability of student i getting assigned in any preference of their ROL R, and $\bar{\eta}_i = \tilde{\rho}_i - \rho_i$ captures the bias on the overall admission probability.

[16]We define the positive and negative parts of the bias as $\bar{\eta}^+_{ij} = \bar{\eta}_{ij}\mathbb{1}_{\{\bar{\eta}_{ij}>0\}}$ and $\bar{\eta}^-_{ij} = -\bar{\eta}_{ij}\mathbb{1}_{\{\bar{\eta}_{ij}<0\}}$. As a result, $\bar{\eta}_{ij} = \bar{\eta}^+_{ij} - \bar{\eta}^-_{ij}$. Finally, we define $\bar{\rho}^+_i$ and $\bar{\rho}^-_i$ accordingly.

Table 2: Effect of Biased Beliefs on Mistakes on Admission Probabilities

	Underconfidence	Ordering	Overconfidence	
	(1)	(2)	(3)	(4)
Optimism Bias - Top True	-0.014^{*}	-0.028^{***}		
	(0.008)	(0.007)		
Pessimism Bias - Top True	0.231^{***}	0.113^{***}		
	(0.011)	(0.010)		
Optimism Bias - Bottom True			-0.056^{***}	
			(0.021)	
Pessimism Bias - Bottom True			0.030	
			(0.023)	
Optimism Bias - Overall				0.772^{***}
				(0.024)
Pessimism Bias - Overall				-0.011
				(0.022)
Mean	0.03	0.022	0.038	0.038
Demographics	Yes	Yes	Yes	Yes
Observations	7,411	7,411	1,458	1,458

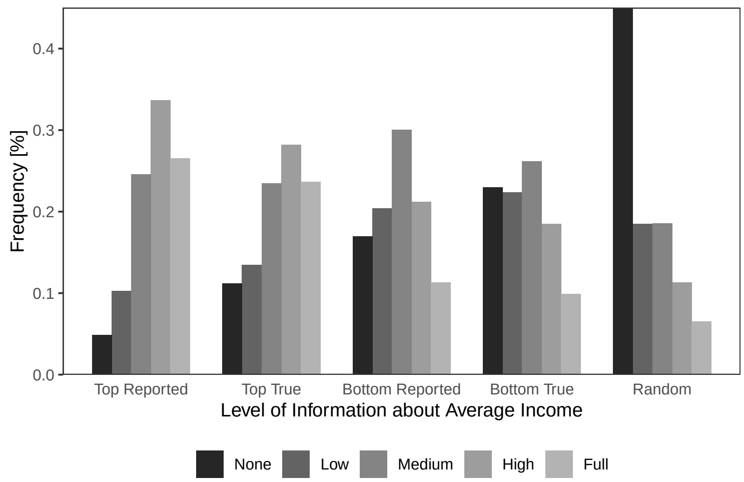
NOTES. Columns (1) and (2) use data from the 2020 survey, while Columns (3) and (4) use data from the 2022 survey. In both cases, the sample includes all survey respondents who are not PACE, are short-list, provided consistent answers, and reported a top/bottom true preference for which they satisfy the application requirements. Significance reported: $^{*}p < 0.1;^{**} p < 0.05;^{***} p < 0.01$.

(3) reveals a similar pattern for biases at the bottom true preference, where lower optimism and higher pessimism about admission chances significantly increase the likelihood of over-confidence mistakes. Finally, Column (4) highlights the role of overall biases in admission assessments, showing that optimistic students are significantly more prone to over-confidence mistakes. Altogether, these results indicate that biased beliefs are relevant drivers of application mistakes, underlining the potential for policies correcting such biases to improve students' outcomes.

3.2 Mistakes on Valuations and Awareness

Measuring mistakes in valuations and awareness presents two key challenges. First, it requires assessing students' misperception of program characteristics that significantly affect their indirect utilities. Second, measuring awareness requires identifying programs—including their attributes and admission chances—that students may be willing to attend but are not aware of. We address the former by focusing on expected earnings, which are one of the primary drivers of college choices (Arcidiacono et al., 2012). For the latter, we leverage our survey design, which elicits students' beliefs about program characteristics and admission chances for a randomly selected program based on their reported preferences.[17] This randomization increases the likelihood of capturing a program the student might reasonably apply to but has not yet considered, allowing us to better assess the role of limited awareness in shaping application decisions.

Figure 2: Knowledge about Average Income



Mistakes on Valuations. In Figure 2, we report the distribution of knowledge about expected earnings for the different elicited programs, separating by level of knowledge (ranging from none to full). We observe that a substantial fraction of students reports being poorly informed, even for programs they included in their ROL. For instance, close to 15% of students report being poorly or not informed at all about their expected earnings for their top reported program, while over 37% do so for their bottom reported preference. These results suggest that information frictions regarding key program characteristics are widespread, which is likely to lead to *mistakes on valuations*.[18]

Mistakes on Awareness. Figure 2 shows that close to 45% of students declare having no information regarding their expected earnings at the random program, which is substantially higher than the reported level of misinformation for other programs known to be in the student's awareness set (i.e., their reported and true preferences). This result suggests that a significant fraction of students might not be aware of all the programs available, potentially leading to *mistakes on awareness*.

[17]For each student, this random program is drawn from a distribution of second-reported programs, conditional on their top-reported program, excluding those already listed in the ROL.

[18]In Appendix A.3, we analyze students' beliefs about their expected earnings upon graduation and analyze their biases. Table A.3 shows that, on average, students significantly overestimate the average expected earnings upon graduation. This result holds even for programs for which we expect students to be well informed, such as their top true preference.

4 Randomized Information Intervention

In collaboration with MINEDUC, we designed and implemented an intervention to provide students with personalized information during the 2022 college admissions process. To do so, we created a customized website for each student who submitted an application within the first two days of the five-day application window. Each website was crafted based on the last rank-ordered list that students submitted during that period—referred to as their initial rank-ordered list. At the start of the third day, MINEDUC emailed these students a link to their customized website, where they could (i) access personalized information based on their characteristics and ROL, designed to address various information frictions; and (ii) revise their application accordingly, as many times as they wished before the deadline. By comparing application and admission outcomes before and after the intervention—and leveraging the random assignment of students to one of four information treatments, which varied the content shown on the websites—we assess the impact of different types of information on students' application behavior and outcomes. The reminder of this section is organized as follows. Section 4.1 outlines the experimental design, including the information provided, treatment conditions, and implementation details. In Section 4.2, we examine the effects of the intervention on application and admission outcomes, and in Section 4.3, we explore the key drivers behind these effects. Finally, Section 4.4 discusses the steps and analyses conducted to ensure the successful scale-up of the intervention, with the implementation and results discussed in detail in Section 5.

4.1 Experimental Design

4.1.1 Information Modules.

The information included in the personalized websites was carefully tailored to address the information frictions and causes of mistakes outlined in the above sections. Specifically, the intervention had four main modules:[19]

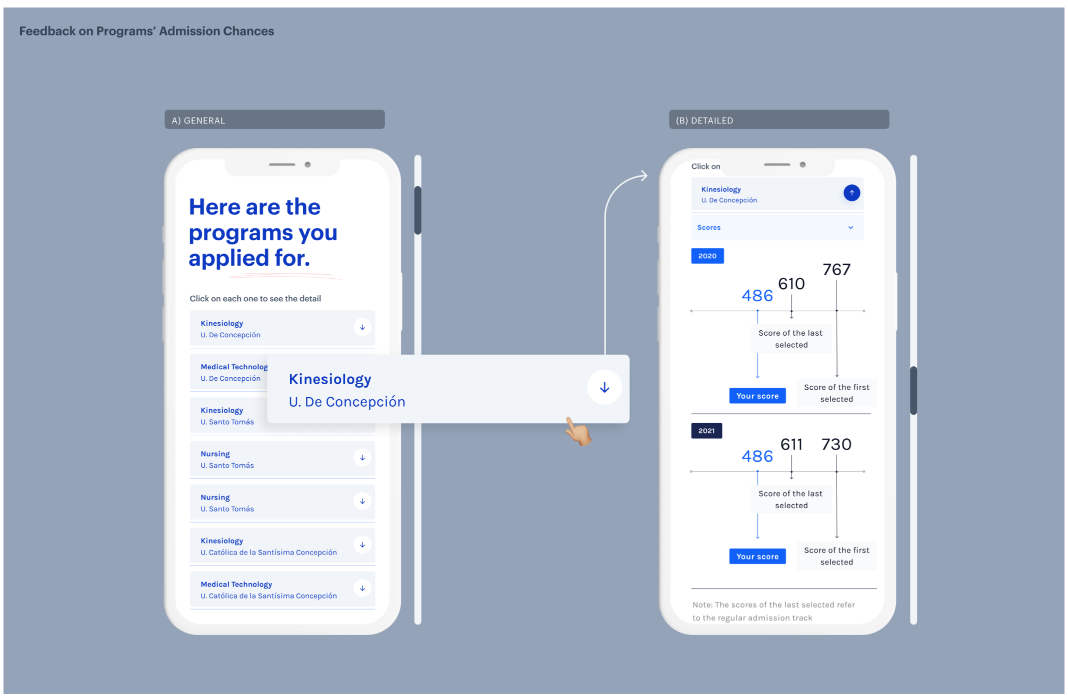
M1 - General information. This module served as a baseline, providing basic information and allowing us to isolate the impact of the additional information introduced in other modules. Specifically, M1 displayed the student's initial application list and, when clicking on a particular program,

[19] Appendix B provides additional details on these modules. The images shown in this section are modified versions of the actual intervention, adapted to improve clarity and exposition. Screenshots of the actual implementation are included in the Appendix B, and the full set of screenshots is available upon request.

presented detailed information that included the program's address, the number of years the institution is accredited for,[20] the benefits and types of financial aid the student would be eligible for if enrolled in the program, the program's official duration (measured in semesters), and the annual tuition fee in Chilean pesos (see Figure B.1 in Appendix B).

M2 - Personalized information about scores for programs included in the rank-ordered list. This module was designed to reduce *mistakes on admission probabilities*. To that end, it displayed the admission scores of the first and last students admitted to each program in the student's initial rank-ordered list during the 2020 and 2021 admissions processes and visually compared them to the student's own application score. Additionally, it flagged programs for which the student did not meet the admission requirements, providing program-specific alerts (see Figure 3).

Figure 3: M3 - Feedback on Programs' Cutoffs



M3 - Personalized alerts depending on student's admission probabilities. Like Module M2, this module aimed to reduce mistakes on admission probabilities by addressing students' misperceptions about their chances of admission. However, M3 presented this information in a more salient and actionable way, introducing alerts at both the program and portfolio levels that flagged poten-

[20] Accreditation length serves as a signal of institution quality. Institutions without accreditation are not eligible to offer public student aid. See details in `https://www.cnachile.cl/`.

tial issues and provided personalized recommendations to help students improve their application strategy. Specifically, this module included:

1. **Program-level alerts:** If the student's estimated admission probability for a given program was below 1%, M3 displayed a red alert (see Figure 4) warning about the low likelihood of admission and encouraging the student to add more programs to their list (see Figure B.3 in Appendix B).

2. **Portfolio-level alerts:** Based on the student's overall admission probability and their chances of getting into their top-ranked program, M3 displayed one of three customized messages (see Figure 5):

   (a) If the overall probability of being assigned to any program is below 99%, M3 recommended adding *safety* programs—i.e., programs for which the student has a positive admission probability—to help prevent potential *overconfidence* mistakes.

   (b) If the probability of admission to the top-ranked program exceeds 99%, M3 encouraged students to consider adding *reach* programs—i.e., programs that they may be interested in and for which they face a positive admission probability—to help prevent potential *underconfidence* mistakes.[21]

   (c) Otherwise, M3 invited students to *explore* and learn more about additional programs, encouraging broader consideration beyond their current list. Since these students were very likely to be assigned to some program on their list (although not necessarily their top choice), this message aimed to prevent potential *underconfidence* mistakes.

M4 - Personalized recommendations of majors of potential interest. This module aimed to mitigate mistakes on awareness and valuations by recommending four majors, carefully selected based on each student's scores and their submitted rank-ordered list.[22] For each recommended major, the module provided information on average program duration, historical cutoff scores (minimum and maximum among programs within the major), and labor market outcomes, including average employment rates and wages four years after graduation (see Figure B.4 in Appendix B).

[21]To identify relevant *reach* programs, we use 2021 survey data on students' top true preferences, computing transition matrices based on typical program pairs.

[22]Section B.4 details the recommendation algorithm. MINEDUC did not allow program-specific recommendations to avoid favoring particular universities.

Figure 4: Feedback on Programs' Admission Chances

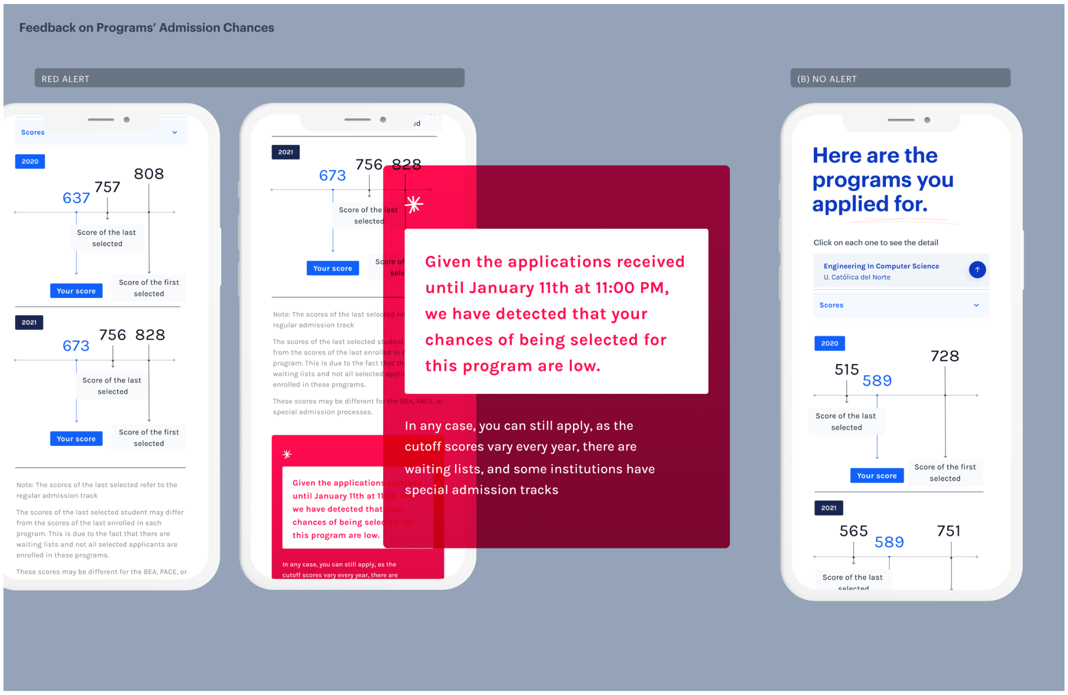
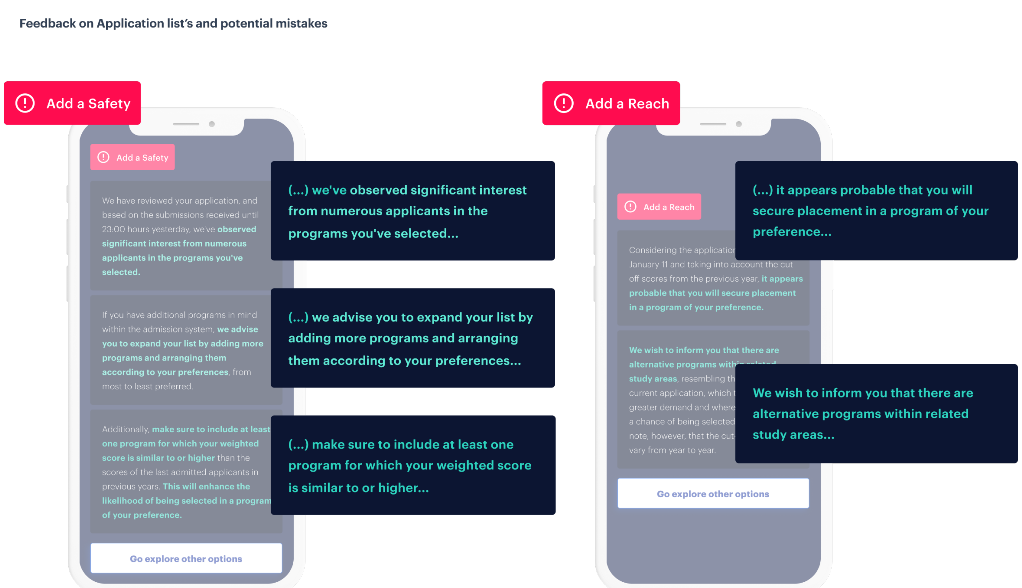


Figure 5: Feedback on Application and Potential Mistakes



4.1.2 Treatments.

To evaluate the impact of the different modules, we defined four treatment groups (T1–T4) and randomly assigned each student who applied before January 12, 2023 at 8 pm CT (i.e., within the first half of the application time window) to one of these groups in a stratified way,[23] with each treatment involving a different combination of modules in addition to M1.[24]

T1: Only M1 is displayed.

T2: M1 and M2 are displayed.

T3: M1, M2 and M3 are displayed.

T4: M1 and M4 are displayed.

Given the prevalence of mistakes on admission probabilities documented in Section 3, the remainder of this section focuses on comparing applicants who received information designed to address such mistakes—i.e., information about cutoff scores, estimated admission probabilities, and personalized alerts about potential mistakes (included in treatments T2 and T3)—to those who did not receive this type of information.[25]

4.2 Impact on Application and Admission Outcomes

A total of 107,663 students (approximately 72.36% of applicants) applied during the first half of the application window, received an email with a link to their personalized website, and ultimately submitted an application by the end of the application window (i.e., they did not withdraw from the process). Among these students, 30,709 (approximately 28.52%) opened their personalized website and thus received the intervention. The analysis in this section focuses on this group, excluding stratification misfits, which results in a final sample of 29,905 students.[26] Descriptive statistics by treatment group are provided in Table B.1 in Appendix B. As expected (given that random treatment assignment), we observe no significant differences in take-up rates or observable characteristics across groups. Therefore, comparing outcomes among openers assigned to different treatments allows us to estimate the effects of the intervention among those who engaged with the website.

[23] Appendix B.2 describes the variables used for stratification and reports the results of several balance checks.

[24] MINEDUC requested that the general information in M1 not be randomized, as it is publicly available on the government website `https://mifuturo.cl/`.

[25] Neither M1 nor M4 provided students with information on these margins, making it unlikely that they affected our outcomes of interest. In fact, institutional constraints made T4 broadly similar to T1, as MINEDUC did not authorize program-specific recommendations. Consistent with this, Table B.6 shows no significant differences in application and assignment outcomes between students assigned to T1 and T4.

[26] We require students not to withdraw their application to compute the outcomes of interests. The results are similar when including misfits.

Table 3: Regression Results among Openers

	Applications		Assignment		
	Modified	Incr. Prob.	Improved	Entered	Benefited&Persisted
	(1)	(2)	(3)	(4)	(5)
Treatment 2	0.0133	0.0079	0.0032	0.0031	0.0015
	(0.0049)	(0.0060)	(0.0029)	(0.0052)	(0.0021)
	[0.0495] {0.0180}	[0.5149] {0.2670}	[0.5248] {0.3370}	[0.7426] {0.5500}	[0.7426] {0.5330}
Treatment 3	0.0139	0.0173	0.0064	0.0122	0.0071
	(0.0050)	(0.0063)	(0.0030)	(0.0056)	(0.0022)
	[0.0396] {0.0180}	[0.0495] {0.0180}	[0.1089] {0.0600}	[0.0990] {0.0600}	[0.0099] {0.0160}
Mean (Control)	0.1348	0.0444	0.0328	0.0279	0.0210
Observations	29,905	7,626	23,371	6,277	29,648
Stratas FE	Yes	Yes	Yes	Yes	Yes

NOTES. Modified is a binary variable equal to 1 if the student modified her application after the information was sent, 0 otherwise. Incr. Prob. is a binary variable equal to 1 if the admission probability associated to the initial rank-ordered list submitted is lower than the one associated to the final rank-ordered list, 0 otherwise. This variable is defined only for students with a positive admission risk given their initial rank-ordered list. Improved is a dummy variable equal to 1 if the student was assigned to a program ranked above the one where they would have been assigned given their initial rank-ordered list, 0 otherwise. It is only defined for the sample of students who would have been matched to a program given their initial rank-ordered list and who did not remove this program from their list. Entered is defined for students who would not have been assigned to any program given the initial rank-ordered list they submitted. It is a binary variable equal to 1 if the student is assigned given their rank-ordered list submitted after the information was sent, 0 otherwise. Benefited&Persisted is a binary variable equal to 1 if the student either entered or improved and persisted for two years in the same program, 0 otherwise. Robust standard errors are reported in parentheses. We report in brackets the p-values adjusted for multiple hypothesis testing following the procedure described in Romano and Wolf (2005) and in braces the q-values computed following Anderson (2008a).

Table 3 presents the results of linear probability models capturing the impact of the information intervention on students' application behavior and assignment outcomes. We observe that students assigned to treatment groups T2 and T3 were 10% more likely to modify their rank-ordered list after opening their personalized website (Column (1)). While this did not translate into significant improvements on other application and assignment outcomes for students who only received information regarding the cutoffs of the programs included in their initial rank-ordered list (T2), students who also received personalized alerts (T3) experienced significant benefits.

Specifically, the intervention induced these students (in T3) to submit a final rank-ordered list associated with a higher overall admission probability compared to their initial application. To show this, we focus on students who faced a positive risk of remaining unassigned based on their initial rank-ordered list and define the outcome *Increased Probability* to be equal to one if the overall admission probability of the final list exceeded that of the initial one, and zero otherwise.[27] We find that the share of students who increased their overall admission probability after opening their personalized website is 39% higher for students assigned to T3, compared to those who received no information about their admission chances (Column (2)). As we show next, these adjustments

[27]To compute these probabilities, we follow a bootstrapping procedure similar to the one described in Appendix B.3.

translated into substantial improvements in assignment outcomes.

First, we find that the alerts students received were effective at decreasing the share of under-matched students. To illustrate this, we restrict attention to students who would have been assigned given their initial rank-ordered list and define the outcome *Improved* to be equal to one if the student obtained a better assignment (according to their final ROL) compared to their initial one, and zero otherwise.[28] We find that students assigned to T3 were 20% more likely to be assigned to a program they ranked higher than the one where they would have been assigned given their initial ROL (Column (3)), consistent with a decrease in ex-post *under-confidence mistakes*.

Second, we find that the intervention was particularly effective at reducing the share of unmatched students. To assess this, we restrict attention to students who would not have been assigned to any program based on their initial rank-ordered list, and define the outcome *Entered* to be equal to one if the student was assigned to a program included in their final list, and zero otherwise. We find that students in T3 were 44% more likely to secure an assignment based on their final ROL compared to those who received no information about their admission risk (Column (4)). This effect is consistent with a reduction in ex-post *over-confidence mistakes*, as the alerts prompted students to adjust their choices in ways that ultimately increased their admission chances.

Third, we find that the positive effects on admissions also translated to downstream outcomes, including enrollment and persistence. To capture this, we define the outcome *Benefited & Persisted* to be equal to one if the student entered or improved, enrolled and remained in the same program in the next two years, and zero otherwise. Although the baseline is low—only 2% of the students in T1 and T4 manage to both gain or improve their admission outcome given their final applications and remain enrolled after two years—being assigned to T3 increases this probability by 34%.

Overall, these results suggest that alerts at both the program and rank-ordered-list levels can effectively induce changes in students' application behavior, leading to improved outcomes. Importantly, this finding is robust to a series of checks reported in Appendix B.5. We conduct two placebo tests: one among students who never opened the intervention and another restricted to students who opened the intervention but were assigned to treatments that did not provide cutoff information. In both cases, we find no effect in any of the outcomes considered (see Tables B.4 and B.6, respectively). We also find no evidence of adverse behaviors such as shortening application lists, exiting the system, or worsening the preference rank of the assigned program. Finally, we

[28]To assess this, we compare the application scores in the programs included in the initial (final) rank-ordered list against the cutoffs that resulted from the admissions process, assigning each student to their most desired program among those for which they were eligible and their application score was greater than or equal to the cutoff.

Table 4: Impact on Biases in Admission Probabilities

	All Applicants			Reach Group	Safety Group
	Top-True	Bottom-True	Overall Prob.	Top-True	Overall Prob.
	(1)	(2)	(3)	(4)	(5)
Treatment 2	-0.0058	-0.0652	0.0069	-0.0608	-0.0021
	(0.0104)	(0.0204)	(0.0058)	(0.0310)	(0.0173)
Treatment 3	-0.0058	-0.0471	-0.0087	-0.0604	-0.0386
	(0.0101)	(0.0195)	(0.0056)	(0.0337)	(0.0169)
Mean (Control)	0.3556	0.3550	0.2181	0.2460	0.4620
Observations	5,321	1,206	9,634	409	1,676
Stratas FE	Yes	Yes	Yes	Yes	Yes

NOTES. This table reports results from the OLS estimation of linear regression models where the dependent variable is the absolute value of students' subjective bias in their admission probabilities for different types of programs. The sample is limited to students who responded to the survey and opened the intervention. Robust standard errors reported in parentheses.

examine application withdrawal and find no treatment effects; withdrawals are negligible overall, with fewer than 0.1% of students dropping their application (see Table B.5).

4.3 Drivers

Correcting Biases in Beliefs over Admission Chances. Since we elicited students' beliefs about their admission chances and their knowledge about program characteristics (as discussed in Section 3), we can directly test whether the reduction in application mistakes discussed above results from correcting students' biases about the former. In particular, we link students' survey responses, which we collected after they submitted their final rank-ordered list, to their applications and treatment group. We compute the absolute bias in students' beliefs regarding their admission chances for their top-true and bottom-true programs, as well as for their overall admission probability, and regress these beliefs on students' treatment status, controlling for strata fixed effects.

Table 4 summarizes these results. On the one hand, we find that being exposed to past cutoff scores (T2 and T3) is effective at correcting students' beliefs regarding their bottom-true program, with a reduction of 13% to 18% in students' biases over their admission chances to these programs (Column (2)). Additionally, we observe that students in the Safety group assigned to T3—who received targeted alerts about both specific programs and their overall admission risk—present significantly lower biases in beliefs regarding their overall admission probability (Column (5)). These patterns align with the intervention's effectiveness at enhancing admission outcomes among students at risk of being unassigned, i.e., those making *over-confidence* mistakes. On the other hand, we find that cutoff information also helped to correct beliefs among students in the Reach group, reducing their bias regarding their top-true program by 25% (Column (4)). This suggests that the intervention lead pessimistic students to correct their beliefs regarding their admission chances to programs they really like, in line with the reduction in *under-confidence mistakes* discussed earlier.

Spillover Effects between Channels. By correcting students' misperceptions about their admission chances and prompting them to revise their rank-ordered lists, the information provided to students in T2 and T3 may also encourage them to seek additional information about program characteristics (beyond admission probabilities). Consequently, these treatments could indirectly reduce valuation and awareness mistakes. To test this possibility, we examine whether the absolute bias in students' elicited beliefs about average earnings differs across treatment groups, focusing on the top-true, top-reported, bottom-true, bottom-reported, and a randomly selected program. As shown in Table B.7, we find no evidence supporting this indirect channel.

4.4 Scale-Up Considerations

Since the first pilot version of the randomized information policy in 2021, a nationwide scale-up was envisioned as a central goal, conditional on positive results from the intervention. Accordingly, scalability considerations were embedded in the design from the outset. To this end, we implemented a series of ex-ante steps to facilitate future scale-up and mitigate common threats to large-scale implementation, as highlighted by Al-Ubaydli et al. (2017). Furthermore, we performed several analyses to assess the generalizability and robustness of our findings (List, 2020).

4.4.1 Maximizing Experiment Scalability: Ex-Ante Considerations.

To ensure that our experiment would yield not only internally valid estimates but also insights relevant for large-scale implementation, we took several ex-ante steps to address potential sources of the "scale-up problem" and enhance external validity. We now describe these in detail.

Representativeness of the Population. One of the main sources of the "scale-up problem" is that the experimental population may differ from the broader policy population, either in their treatment effects or in participation costs. Our collaboration with MINEDUC helped mitigate this concern by allowing us to conduct the experiment directly within the relevant policy population:

college applicants in Chile. Specifically, we randomized all students who submitted their applications within the first two days of the five-day application window—representing over 72% of the full applicant pool—into one of the four treatment groups described above. While early applicants differ from later ones along several observable characteristics (see Table B.3), they constitute the relevant target group for our intervention since (i) personalized information on admission probabilities can only be provided *after* students submit an initial rank-ordered list, and (ii) sufficient time is needed for them to process and act on this information. Moreover, the randomization procedure did not introduce any differential motivation or incentives across experimental and non-experimental students. By embedding the experiment within the actual policy population, we thus address concerns about heterogeneity in treatment effects and participation costs that often limit the external validity of experimental findings.

Naturalness of the Experimental Task. Another common source of the "scale-up problem" is the disconnect between experimental conditions and real-world policy environments. For instance, providing information about the returns to different majors may influence survey responses, yet have little effect on actual enrollment decisions due to the low stakes and hypothetical nature of the task. Our collaboration with MINEDUC allowed us to closely align the experimental setting with real-world decision-making. We conducted a natural field experiment in a high-stakes environment, embedded within the official admissions process and implemented in partnership with policymakers. Crucially, we did not introduce any artificial margins of choice or hypothetical decisions. As a result, both the choices students faced and the outcomes we measured were identical to those in the policy setting, enhancing the external validity of our findings.

Cost Considerations. "Boot-on-the-ground" education interventions, such as one-on-one tutoring, are likely to face diseconomies of scale due to increasing logistical costs. In contrast, our intervention was designed to scale at minimal cost, leveraging the cost-effectiveness of information policies (Kremer et al., 2013; Angrist et al., 2020). The core technological components—assessing students' admission risk and generating personalized websites—involved primarily fixed costs, rendering the marginal cost of reaching an additional student virtually zero. In addition to being inexpensive, the policy also generates fiscal revenue by increasing enrollment in selective universities—an outcome associated with substantial earnings gains and long-term tax revenue for the government (Hastings et al., 2013). While a detailed quantification of the fiscal returns is beyond the scope of this paper, the combination of low costs and induced tax revenue suggests that the intervention is likely to pay for itself over a short horizon (Hendren and Sprung-Keyser, 2020). We formalize this intuition in Section 6.1, where we assess the Marginal Value of Public Funds and estimate the long-run monetary benefits of the intervention using administrative labor-market data.

Quality Considerations. Many small-scale programs experience quality drops once implemented on a larger scale (Davis et al., 2017; Al-Ubaydli et al., 2017). This challenge is particularly acute for in-person educational interventions, where the supply of qualified tutors or mentors may not grow proportionally with demand. In contrast, our intervention was designed to minimize the risk of quality degradation by leveraging the infrastructure of the centralized admission system. The platform already collects all the data required to compute students' admission probabilities, and the generation and delivery of personalized information can be fully automated once message types and the probability thresholds are defined. As a result, the quality and accuracy of the recommendations can be maintained at scale.[29]

Generalizability to Other Settings. We believe our findings have important implications beyond the Chilean context. First, according to Kapor et al. (2024), at least 46 countries use a centralized system to organize their admissions to college, including Turkey, Taiwan, Tunisia, Hungary, and Chile. Among them, the most common assignment mechanism is the Deferred Acceptance algorithm (Neilson (2024)). Second, our theoretical framework captures behavioral mistakes that are likely to arise in many centralized college admissions systems, formalizing patterns already documented in other settings (Artemov et al., 2017; Shorrer and Sóvágó, 2021; Chrisander and Bjerre-Nielsen, 2023). Together, these elements make our study highly relevant for policymakers seeking to enhance the efficiency and equity of college admission systems worldwide.

4.4.2 Statistical Inference

Another key challenge for scalability is the risk of false positive findings—that is, statistically significant estimates that may not replicate at scale. To mitigate this concern, we adopted a conservative inference strategy in our analysis. Specifically, all regressions include fixed effects for the

[29] Another quality consideration is that mistakes need to be non-negligible and actionable (i.e., that students respond appropriately to the intervention) for information policies to scale appropriately. This is not a concern in our setting, since (i) the primary driver of mistakes are biased in admission probabilities that can be corrected (as the results in Table C.7 show), and (ii) the policy was accompanied with an outreach effort that minimized the risk of inattention.

Table 5: Post-Study Probability

	Applications		Assignment		
	Modified	Incr. Prob.	Improved	Entered	Benefited&Persisted
Prior	(1)	(2)	(3)	(4)	(5)
0.10	0.642	0.639	0.562	0.574	0.667
0.20	0.801	0.799	0.743	0.752	0.819
0.30	0.874	0.872	0.832	0.838	0.886
0.40	0.915	0.914	0.885	0.890	0.923
0.50	0.942	0.941	0.920	0.924	0.947
0.60	0.960	0.960	0.945	0.948	0.964
0.70	0.974	0.974	0.964	0.966	0.977
0.80	0.985	0.985	0.979	0.980	0.986
0.90	0.993	0.993	0.990	0.991	0.994

NOTES. This table shows the Post-Study Probability of the experimental results presented in Table 3. Computations follow Equation (4) for different levels of prior probability of each hypothesis.

strata used in the randomization, combined with the standard heteroskedasticity-consistent estimator of the asymptotic variance. As shown by (Bugni et al., 2019), this approach is conservative: the probability of incorrectly rejecting the null hypothesis is no greater than, and typically strictly less than, the nominal significance level.

While this conservative approach to inference mitigates concerns about false positives, the risk remains due to the multiplicity of outcomes and treatment comparisons. To further limit this risk, we follow List et al. (2019) and adjust for multiple hypothesis testing. Specifically, we report p-values corrected using the stepwise procedure of Romano and Wolf (2005), which controls the family-wise error rate, as well as q-values based on the method developed by Anderson (2008b) to control the false discovery rate.

Finally, we complement our inference strategy by going beyond statistical significance alone. Following the recommendations of Al-Ubaydli et al. (2017) and Maniadis et al. (2014), we assess the post-study probability (PSP) of our main findings—that is, the probability that a result is truly non-zero given that it achieves statistical significance. Table 5 presents PSP estimates for each of our outcomes of interest: whether students modify their list, increase their admission probability, improve their assignment outcome, and persist in their program. We focus on the effect of Treatment 3, our main treatment of interest, and compute the PSP as follows:

$$PSP = \frac{(1-\beta)\pi}{(1-\beta)\pi + \alpha(1-\pi)}, \tag{4}$$

where $(1-\beta)$ denotes the statistical power of our experiment—the probability of detecting a true effect—while α is the significance level (set at 0.05), and π is the prior probability that the hypothesis is true. Even under a conservative prior of $\pi = 0.1$, which reflects strong initial skepticism, we obtain PSP values above 0.56 across all key outcomes. This indicates a meaningful update from the prior belief. Furthermore, the PSP estimates quickly surpass 90% as the prior increases, suggesting that our findings can reliably be used by policymakers.

4.4.3 Congestion Effects and Scale-Up.

A final scalability concern relates to potential congestion effects. Since program seats are limited, providing information at scale could significantly change application patterns and alter the market equilibrium, thus reducing the effectiveness of the intervention. In particular, if many initially unmatched or undermatched students respond to the information by applying to the same program, that program may become oversubscribed. As a result, even though students adjust their rank-ordered lists in response to the information, some may still remain unmatched or undermatched.

We first investigate whether the intervention triggered any displacement effects within the experimental sample. Although the scope for displacement is inherently limited—only marginally admitted students in oversubscribed programs are at risk of being displaced—, this is a key concern since such effects would violate the Stable Unit Treatment Value Assumption (SUTVA).

Leveraging the fact that we observe all application changes, we can identify potential displacement effects. To do so, we simulate a counterfactual assignment by running the assignment algorithm under the assumption that students in treatment groups T2 and T3 did not alter their rank-ordered lists in response to the information intervention.[30] We then compare this counterfactual assignment to that obtained using students' final rank-ordered lists. Specifically, we compute the number of students who receive a worse outcome in the final assignment relative to the counterfactual, either because they are admitted to a lower-ranked program or lose admission altogether. We find that such displacement effects on control group students are negligible: fewer than 0.5% of students in T1 and T4 receive a worse outcome due to changes in the rank-ordered lists of T2 and T3 students.[31] We thus conclude that displacement effects do not compromise the integrity of

[30]Only a small fraction of control students are even at risk of being displaced. Only 7.7% of programs meet the necessary conditions: (1) the last admitted student is from the control group, and (2) the program is oversubscribed. Displacement would then require that a treated student gains admission due to the information intervention and that no higher-priority student vacates a seat in the program.

[31]This estimate is likely an upper bound, as some students in T2/T3 may have changed their preferences even without the intervention.

our experimental results.

Having ruled out within-experiment displacement effects as a driver of the improvements generated by T2 and T3, we now turn to investigate the underlying sources of these admission gains. One possibility is that treated students are displacing late applicants who fall outside the experimental sample. Alternatively, the gains may result from students gaining admission to programs that would otherwise have remained undersubscribed, indicating that the intervention generates net improvements in vacancy utilization.

Applying the same counterfactual procedure discussed above, we find limited support for the displacement of late-applicants. Specifically, only 206 students outside the experimental sample are displaced as a result of the changes in T2 and T3 students' applications, representing 0.5% of the non-experimental population. In contrast, we find substantial evidence backing net gains in vacancy utilization. First, many programs fail to fill all available seats: 55% of the 1,993 programs on the system still have vacancies at the end of the admission process, and 20% of the 147,056 total seats remain unassigned. Among the programs to which T2 and T3 students are ultimately admitted, the average number of unfilled seats is 10.8. Second, we observe substantial heterogeneity in students' preferences, which reduces the risk of congestion. Students in T2 and T3 who benefited were assigned to 910 different programs and, among these, the median number of treated students gaining admission was one. This dispersion indicates that students respond to the intervention in diverse ways, avoiding potentially negative congestion effects. Taken together, the combination of ample seat availability and heterogeneity in student preferences enables the intervention to generate meaningful improvements in student assignment outcomes.

While our findings suggest that congestion is limited within the experimental setting, it may become more pronounced as the intervention scales. To assess the potential magnitude of such congestion effects under broader implementation, we leverage the properties of the centralized assignment mechanism and our experimental data to simulate a scaled version of the policy using a bootstrap procedure (Karnani, 2023). Specifically, we simulate two counterfactual assignments: (i) No Intervention, representing the assignment that would arise without the policy, and (ii) Scale-Up, representing the assignment that would result if all students applying during the first half of the application window received the full information treatment (as in T3). Comparing outcomes across these two counterfactuals allows us to estimate the potential effects of delivering the T3 intervention to all early applicants. To construct each simulated market, we replace every student who received the intervention with a student randomly drawn from either T1 (for No Intervention)

Table 6: Scale-Up Simulations

	Entered [%]	Improved [%]
No Intervention	3.194	3.698
Scale-Up	3.732	4.058

NOTES. This table reports counterfactual assignment results obtained by averaging over 1,000 bootstrap simulations, where each student who participated in the assignment is replaced by a student assigned to T1 for the *No Intervention* counterfactual and by a student assigned to T3 for the *Scale-Up* counterfactual.

or T3 (for Scale-Up), while keeping all other applicants unchanged. This approach ensures that the simulated markets involve the same total number of students, comprising those eligible for the intervention (bootstrapped from the appropriate group) and those applying in the second half of the application window. For each counterfactual, we compute the resulting assignments considering students' submitted rank-ordered lists both before and after the information treatment and compare the distribution of assignment outcomes, as detailed in Section 4.2.[32]

Table 6 presents the results. While we find that the estimated effects of the intervention are smaller under the Scale-Up scenario than in the experimental setting—indicating that congestion effects are not negligible—we still observe meaningful gains. In particular, students who would have remained unassigned based on their initial rank-ordered lists are 0.54 percentage points more likely to be assigned in the Scale-Up counterfactual compared to the No Intervention scenario. This suggests that, even in the presence of increased competition, the intervention continues to deliver improvements in assignment outcomes—especially at the extensive margin—implying that these gains are not purely redistributive. Moreover, it is important to note that the simulation procedure—based on sampling with replacement from a limited subset of students—may understate the heterogeneity in submitted rank-ordered lists, potentially overstating congestion effects. These results therefore indicate that while some crowding may occur, a broader rollout of the policy is likely to improve students' outcomes.

[32]This approach relies on a large-market assumption: students are assumed to be "cutoff takers" whose individual applications do not influence admission thresholds.

5 Scaled-Up Information Intervention

5.1 Policy Implementation and Empirical Evaluation Strategy

Given the positive results discussed above, MINEDUC decided to implement the information intervention nationwide in 2023, adapting its design to accommodate important changes in the application process. While the core structure of the intervention remained consistent with the randomized experiment—students received personalized websites created based on their rank-ordered lists submitted within the first half of the application period—the national roll-out included all informational modules and was delivered to the entire target group without randomization. Furthermore, due to the higher uncertainty and potential impact, MINEDUC promoted the intervention through social media and sent multiple reminder emails to students, resulting in higher take-up compared to the experimental intervention.

As detailed in Appendix C.1, the 2023 admissions process was among the most challenging and uncertain since the inception of the centralized system. Not only was an additional entrance exam added, but more significantly, the scoring scale for all admissions factors was overhauled—shifting from a $[210, 850]$ range to a $[100, 1000]$ scale. This change rendered previous years' cutoffs uninformative, leaving students without reliable benchmarks to assess their chances of admission. To address this uncertainty, the most relevant change we implemented in 2023 was replacing previous years' cutoffs with the actual score of the last student admitted to each program based on early applications. Figure 6 provides an example of the personalized information students received, with additional illustrations available in Appendix C.1.

Motivated by the limitations of the recommendation module discussed in Section 4—which provided major- rather than program-specific recommendations—, we updated Module M4 to include a search engine that allowed students to find programs based on various filters (e.g., location, major, university). For each program, this search engine displayed the current admission cutoff given the received applications, aiming to reduce *mistakes on admission probabilities*. In addition, it provided key additional information such as tuition fees, program duration, and available benefits, with the goal of mitigating *mistakes on valuations* and *mistakes on awareness* (see Figure 7). When implemented at scale, this updated design closely resembles a one-shot version of the *iterative* Deferred Acceptance algorithm, allowing students to access "real-time" cutoff information and adjust their applications accordingly.

To evaluate the impact of the policy in the absence of an untreated group, we use an encour-

Figure 6: General

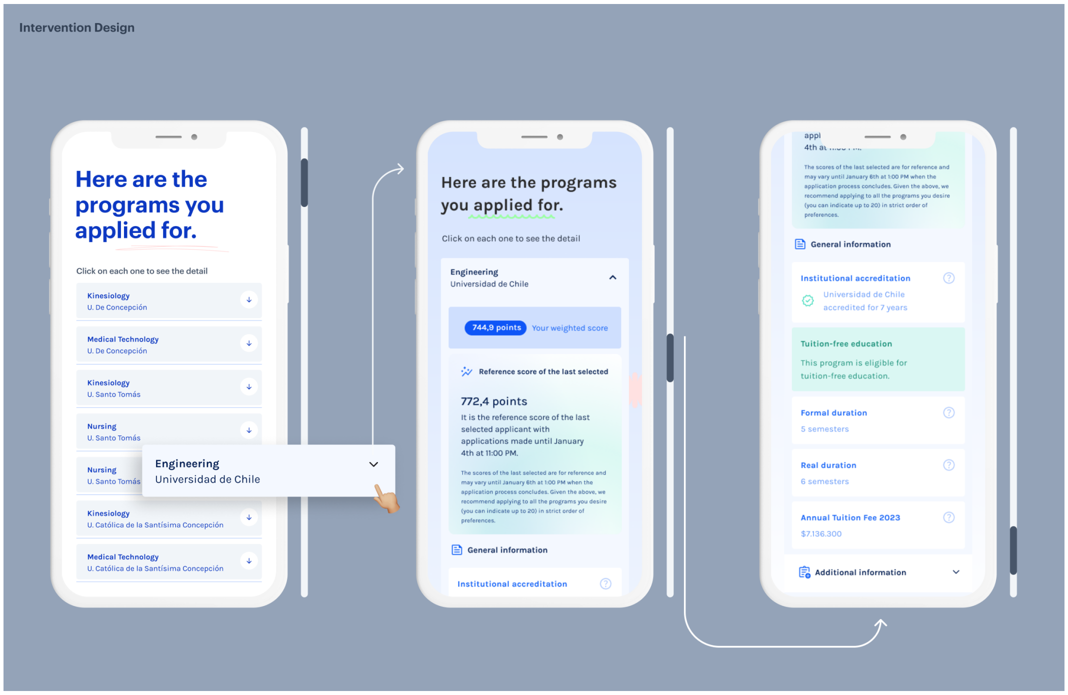


Figure 7: Search module

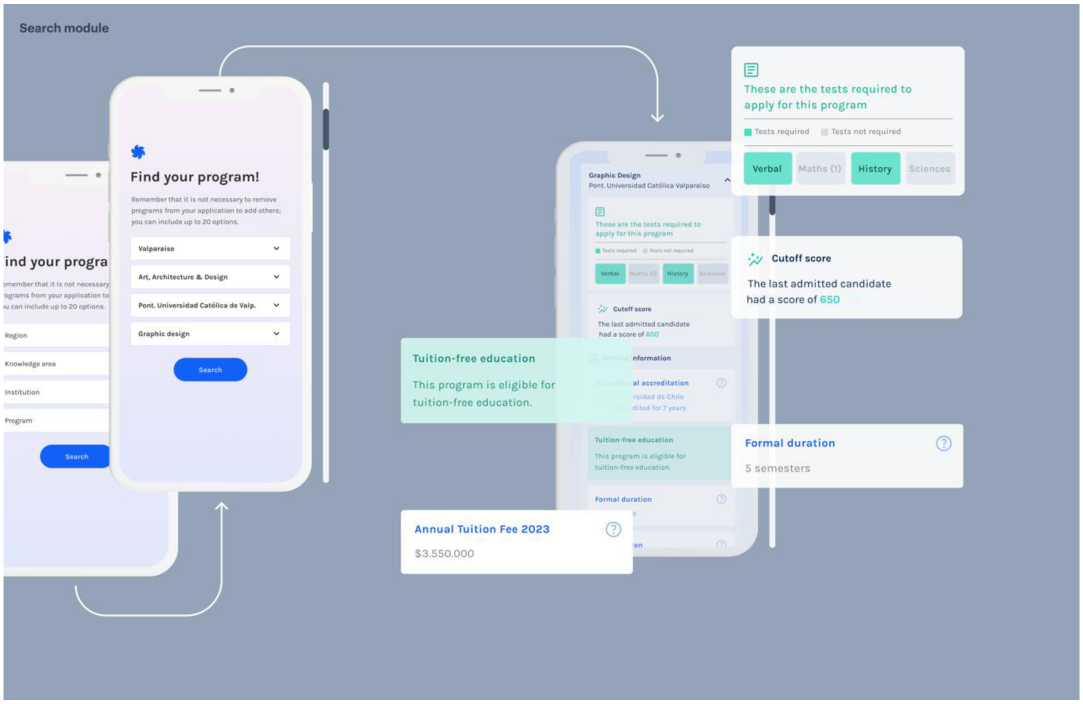


Table 7: Regression Results: Instrumental Variables

	Applications		Assignment		
	Modified	Incr. Prob.	Improved	Entered	Benefited & Enrolled
	(1)	(2)	(3)	(4)	(5)
Open	0.1177	0.0523	0.0364	0.0347	0.0308
	(0.0148)	(0.0152)	(0.0110)	(0.0149)	(0.0077)
	[0.0099] {0.0010}	[0.0099] {0.0010}	[0.0198] {0.0020}	[0.0297] {0.0200}	[0.0099] {0.0010}
Mean (No Text)	0.2322	0.0764	0.0680	0.0602	0.0486
Observations	132,893	41,659	95,739	35,317	131,056
Risk Group	Yes	Yes	Yes	Yes	Yes

NOTES. Modified is a binary variable equal to 1 if the student modified her application after the information was sent, 0 otherwise. Incr. Prob. is a binary variable equal to 1 if the admission probability associated to the initial rank-ordered list submitted is lower than the one associated to the final rank-ordered list, 0 otherwise. This variable is defined only for students with a positive admission risk given their initial rank-ordered list. Improved is a dummy variable equal to 1 if the student was assigned to a program ranked above the one where they would have been assigned given their initial rank-ordered list, 0 otherwise. It is only defined for the sample of students who would have been matched to a program given their initial rank-ordered list and who did not remove this program from their list. Entered is defined for students who would not have been assigned to any program given the initial rank-ordered list they submitted. Entered is a binary variable equal to 1 if the student is assigned given their rank-ordered list submitted after the information was sent, 0 otherwise. Benefited&Enrolled is a binary variable equal to 1 if the student either entered or improved and enrolled, 0 otherwise. Robust standard errors are reported in parentheses. We report in brackets the p-values adjusted for multiple hypothesis testing following the procedure described in Romano and Wolf (2005) and in braces the q-values computed following Anderson (2008a).

agement design. Specifically, we randomly selected a group of students to receive a WhatsApp message encouraging them to access their personalized website. We then use the random assignment of the message as an instrument to identify the causal effect of opening the personalized website. This approach allows us to estimate the Local Average Treatment Effect (LATE)—that is, the average effect for students who are induced to open their personalized website as a result of receiving the Whatsapp message. In Appendix C.2, we provide details of the encouragement design and a series of checks that confirm the validity of our instrument (see Tables C.1 and C.2).

5.2 Results

Table 7 presents the second-stage results of our two-stage least squares estimation strategy, using the same outcome measures as in Table 3. Since we only observe enrollment data for 2023, we replace the outcome *Benefited & Persisted* with *Benefited & Enrolled*, defined as an indicator equal to one if the student benefited—i.e., entered or improved—and ultimately enrolled.[33]

First, consistent with the results reported in the previous section, we find that the information intervention prompts students to revise their applications. Specifically, opening the personalized

[33]In Table C.4 in Appendix C.4, we report summary statistics for several outcomes of interest, separating by whether the student open the email and their risk group.

website increases the probability of updating their rank-ordered list by 11.77 percentage points among compliers, relative to a baseline update of 23% among students who did not receive the WhatsApp reminder (Column (1)). These changes translate into significant improvements in students' applications. Among students who faced a positive risk of being unassigned based on their initial rank-ordered list, accessing the personalized website increases the likelihood of submitting a final rank-ordered list with a higher probability of admission by 5.2 percentage points (Column (2)). This represents a 68% increase compared to students who did not receive the reminder.

In line with our experimental findings, we also find that the personalized information improves students' assignment outcomes for both students who would have been assigned and those at risk of being unassigned based on their initial rank-ordered list. Among the former, students who accessed their personalized website due to the WhatsApp reminder are 3.64 percentage points more likely to improve, i.e., to be admitted to a program ranked higher than the one they would have otherwise been assigned to—a 54% increase relative to those who did not receive the reminder.[34] For students who would not have been admitted to any program based on their initial list, the intervention leads to a 3.47 percentage points increase in the probability to be admitted to a program at the end of the assignment procedure, compared to a 6% baseline for those not receiving the reminder (Column (4)). These gains in assignment outcomes also translate to improvements in enrollment, as shown in Column (5).

To gain some insights on the mechanisms behind these improvements, we analyze students' use of the search engine embedded in their personalized websites. Table C.6 shows that among students who opened their website, there is a large and significant correlation between the use of the search engine and all of the outcomes analyzed in Table 7. These results remain true even after controlling for the number of programs included in the list, suggesting that the search engine affected outcomes through channels other than simply expanding the length of rank-ordered lists. In addition to these behavioral responses, we also find that the intervention helped reduce biases in students' beliefs about program cutoffs and their own admission probabilities (Table C.7).

Overall, these results suggest that the scaled-up information intervention was highly effective in improving students' application and assignment-related outcomes. The estimated effects are not only consistent with the experimental results for T3 but notably larger in magnitude. This amplification may stem from three factors. First, it may reflect differences in the composition

[34]Table C.5 shows that this effect is concentrated among applicants with a very low risk of non-admission given their initial rank-ordered list.

of compliers in the encouragement design relative to RCT openers. Table C.3 in Appendix C.3—which applies the methodology of Abadie (2003) to characterize compliers and compare them with RCT openers—documents meaningful differences across these groups. Compliers are less likely to be female or from the metropolitan region, have lower high-school GPAs, higher Math–Verbal scores, and are more likely to come from public and voucher schools. Second, the amplified effects may reflect the heightened uncertainty surrounding the 2023 admissions process, which likely increased students' reliance on timely, personalized information (see Appendix C.1). Finally, the nationwide roll-out expanded the scope of the intervention by allowing students to explore real-time cutoffs for all programs in the system—not only those on their initial rank-ordered list. This broader access plausibly improved students' ability to assess their admission chances, discover additional relevant options, and ultimately make more informed application decisions.

6 Discussion

6.1 Cost-Effectiveness and Scalability of the Intervention

A central advantage of our intervention is its ability to generate meaningful improvements in students' higher-education choices at extremely low cost. In this subsection, we formalize this intuition by drawing on two complementary approaches commonly used in the cost-effectiveness and comparative welfare literature.

Marginal Value of Public Funds (MVPF). We begin by computing the Marginal Value of Public Funds (MVPF), following the unified framework of Hendren and Sprung-Keyser (2020). The MVPF is defined as the ratio between individuals' aggregate willingness to pay (WTP) for the intervention and the net fiscal cost to the government. Information interventions are well known to be highly cost-effective because their marginal cost is negligible and their fixed costs do not scale with the number of beneficiaries (Kremer et al., 2013; Angrist et al., 2020). In our setting, the direct implementation cost—which includes developing the algorithm, building the platform, and designing the communication materials—is approximately $20,000–$50,000 and does not increase with scale.

In addition, the intervention increases enrollment in selective programs, which has been shown to generate substantial long-run earnings gains and corresponding tax revenue (Hastings et al., 2013). When an intervention has low (or even negative) net fiscal cost and generates positive willingness to pay, Hendren and Sprung-Keyser (2020) show that the MVPF is infinite. Our setting satisfies these conditions, implying that the intervention is not only cost-effective but potentially self-financing from a public finance perspective. Full details of the MVPF calculation are provided in Appendix D.1.

Long-Run Monetary Benefits. To complement the MVPF analysis, we also quantify the long-run monetary benefits generated by the intervention. Using administrative labor-market data from the Chilean social security system, we estimate degree-program-specific earnings returns for four historical cohorts and use these estimates to predict long-run income for students in our experimental sample. We then apply our IV strategy to estimate the causal effect of being induced to open the platform on predicted earnings thirteen years after high school graduation.

The results, reported in Appendix D.2, show that the intervention generates economically meaningful gains. For the full sample, the IV estimate implies an increase of approximately $87 in predicted annual income. Among students who experienced a change in assignment following the intervention, the estimated increase is substantially larger, at $1,936.7 US dollars.[35] Importantly, a natural concern is that students induced to switch programs might move into substantially more expensive ones, offsetting the earnings gains. To address this, we restrict attention to students who were initially assigned to a degree program and estimate the causal effect of the intervention on the difference in annual tuition (in CLP) between their final and initial first-choice programs using the same 2SLS strategy. As reported in Table D.2, the estimated effect on tuition differences is not statistically significant, indicating that the intervention does not induce students to move into more expensive programs. This pattern also holds among students who experienced a change in their assigned program. Given that the marginal cost of sending a text message is essentially zero, these long-run benefits underscore the cost-effectiveness and scalability of the intervention.

Taken together, both approaches highlight that our intervention delivers substantial welfare gains at minimal cost, positioning it among the class of highly scalable information treatments emphasized in the cost-effectiveness literature.

6.2 Learning from the Scale-Up and the Last-Mile Problems.

Over our five-year journey, we gathered valuable insights at every stage, which informed the design and implementation of each subsequent step. Theory helped us identify the mechanisms

[35] This comparison conditions on experiencing a change in assignment, which is itself affected by the intervention. This estimate should therefore be interpreted as descriptive rather than causal.

linking information frictions to application mistakes, offering a framework to measure these distortions and design targeted interventions. Surveys and administrative data confirmed the presence of frictions along multiple dimensions and allowed us to assess both the prevalence and the impact of application mistakes. The field experiment demonstrated that light-touch personalized interventions can meaningfully reduce these mistakes. Finally, the policy implementation gave us a unique opportunity to evaluate the intervention at scale—under real-world constraints and amidst considerable uncertainty.

While we view the overall process as a success, particularly given the significant improvements observed in student outcomes, the scale-up also faced several last-mile challenges that offer important lessons for policymakers and market designers. Although the policy was carefully co-designed with MINEDUC, other key stakeholders were not directly involved and, in some cases, opposed the intervention—especially given the concurrent changes to the admissions process. As a result, MINEDUC had to make a number of concessions to ensure the intervention could move forward. One key compromise was the inability to provide students with forecasted ranges for cutoff scores—an important feature we had initially envisioned. The lack of full coordination among stakeholders also created unnecessary confusion. For example, universities, which were not informed about how the displayed cutoff scores were calculated, still had to respond to inquiries from applicants seeking advice about their admission chances.[36]

Overall, we believe that the lack of alignment among key stakeholders—particularly during the design phase—led to a suboptimal version of the policy and introduced avoidable challenges during the implementation. For practitioners, policymakers, and market designers looking to scale similar interventions, our experience underscores the importance of engaging all relevant stakeholders early in the process to ensure their support and alignment. Early coordination not only contributes to a more coherent policy design but also facilitates smoother implementation. Stronger institutional collaboration could have enhanced both the clarity and the effectiveness of the intervention, suggesting that future iterations would benefit from a more integrated and participatory approach.

[36]We believe MINEDUC efficiently addressed these concerns, alleviating a significant part of the implementation issues.

6.3 Implications for Market Design.

Our results indicate that information frictions significantly alter the performance of centralized admission systems, even when students have no clear strategic incentive to misreport their preferences. These results have two key implications for the design of matching markets. First, the presence of application mistakes suggests that strategy-proof mechanisms are not immune to inefficiencies caused by information frictions, highlighting the need for more robust mechanisms. This is particularly relevant as the adoption of centralized assignment systems based on strategy-proof mechanisms is expanding globally, with many countries using them for student allocations. As highlighted by Rees-Jones and Shorrer (2023), sequential assignment procedures—such as dynamic implementations of the Deferred Acceptance algorithm (Bó and Hakimov, 2022)—can improve outcomes in contexts where students face incomplete preferences (Grenet et al., 2022), behavioral biases (Meisner and von Wangenheim, 2021; Dreyfuss et al., 2022), off-platform options (De Groote et al., 2025), or costly information acquisition (Immorlica et al., 2020). The results of our scaled-up intervention, resembling a one-shot version of the *iterative* Deferred Acceptance algorithm, support the use of such sequential mechanisms at scale.

Second, our results highlight the potential of information policies to improve the performance of centralized admissions systems. In cases where policymakers prefer to retain static mechanisms like standard DA over sequential or iterative implementations, our results show that light-touch information interventions can serve as an effective and scalable complement. The personalized websites we implemented in Chile—–providing real-time and personalized information on admission chances—offer a replicable model that can be adapted to other settings. As such, they represent a practical tool for supporting students in the application process and mitigating the effects of information frictions.

7 Conclusion

In this paper, we present the results of a multi-year collaboration with policymakers in Chile aimed at evaluating whether information interventions can effectively reduce application mistakes and improve student outcomes. We introduce a new taxonomy that characterizes different types of application mistakes arising from information frictions across several dimensions, including admission probabilities, program characteristics, unawareness of available options, and misunderstanding of the admission mechanism. Nationwide surveys allow us to quantify the prevalence of these application mistakes and the underlying mechanisms, with biases in students' beliefs about their admission chances being the primary driver of these mistakes.

Based on the surveys' insights, we collaborated with policymakers to design and implement a multi-year outreach policy to reduce information frictions and application mistakes. Using a field experiment where we vary the information provided to students, we find that showing personalized information about admission probabilities for listed programs and customized messages to guide students depending on their overall admission probability has a causal effect on improving students' outcomes, significantly reducing the risk of being unmatched to the centralized system and the incidence of both *over-confidence* and *under-confidence* mistakes. We provide a thorough discussion of how scalability was embedded in the intervention's design from the beginning and implement additional analyses to assess the generalizability and robustness of our main findings.

Given the success of the field experiment, we continued collaborating with MINEDUC to improve the design of the intervention and implement it at scale. By exploiting an encouragement design, we find that showing personalized application advice, information about programs' characteristics, and current cutoff scores for all programs in the centralized system—similar to sequential implementations of DA—significantly improved students' outcomes.

Overall, our work demonstrates that information frictions and application mistakes are prevalent, even in a high-stakes environment like college admissions. However, personalized information policies implemented at scale can effectively alleviate these frictions. Importantly, we show that these improvements are achieved at near-zero marginal cost and are associated with positive long-run earnings gains, making the intervention not only scalable but also highly cost-effective. We believe our work provides a model for how researchers and policymakers can collaborate closely to design, test, refine, and implement at scale education policies to improve students' outcomes. As the availability of education data grows, major opportunities exist for policies addressing information frictions to enhance efficiency and equity. We hope our work inspires future efforts in this direction.

References

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics*, 113(2):231–263.

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

Agostinelli, F., Avitabile, C., and Bobba, M. (2025). Enhancing human capital in children: A case study on scaling. *Journal of Political Economy*, 133(2):000–000.

Al-Ubaydli, O., List, J. A., and Suskind, D. L. (2017). What can we learn from experiments? understanding the threats to the scalability of experimental results. *American Economic Review*, 107(5):282–286.

Allende, C., Gallego, F., and Neilson, C. (2019). Approximating the equilibrium effects of informed school choice. Technical report.

Altonji, J. G., Kahn, L. B., and Speer, J. D. (2014). Trends in earnings differentials across college majors and the changing task composition of jobs. *American Economic Review*, 104(5):387–393.

Anderson, M. L. (2008a). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association*, 103(484):1481–1495.

Anderson, M. L. (2008b). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, 103(484):1481–1495.

Angrist, N., Evans, D. K., Filmer, D., Glennerster, R., Rogers, F. H., and Sabarwal, S. (2020). How to improve education outcomes most efficiently?: A comparison of 150 interventions using the new learning-adjusted years of schooling metric. Technical report, World Bank.

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

Arcidiacono, P., Hotz, V. J., and Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, 166(1):3–16.

Arteaga, F., Kapor, A. J., Neilson, C. A., and Zimmerman, S. D. (2022). Smart matching platforms and heterogeneous beliefs in centralized school choice. *The Quarterly Journal of Economics*, 137(3):1791–1848.

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

Baker, R., Bettinger, E., Jacob, B., and Marinescu, I. (2018). The effect of labor market information on community college students' major choice. *Economics of Education Review*, 65:18–30.

Bleemer, Z. and Zafar, B. (2018). Intended college attendance: Evidence from an experiment on college returns and costs. *Journal of Public Economics*, 157:184–211.

Bó, I. and Hakimov, R. (2022). The iterative deferred acceptance mechanism. *Games and Economic Behavior*, 135:411–433.

Bobba, M., Frisancho, V., and Pariguana, M. (2023). Perceived ability and school choices: Experimental evidence and scale-up effects.

Bugni, F. A., Canay, I. A., and Shaikh, A. M. (2019). Inference under covariate-adaptive randomization with multiple treatments. *Quantitative Economics*, 10(4):1747–1785.

Carrell, S. and Sacerdote, B. (2017). Why do college-going interventions work? *American Economic Journal: Applied Economics*, 9(3):124–151.

Castleman, B. and Goodman, J. (2018). Intensive college counseling and the enrollment and persistence of low-income students. *Education Finance and Policy*, 13(1):19–41.

Castleman, B. L., Page, L. C., and Schooley, K. (2014). The forgotten summer: Does the offer of college counseling after high school mitigate summer melt among college-intending, low-income high school graduates? *Journal of Policy Analysis and Management*, 33(2):320–344.

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

Chrisander, E. and Bjerre-Nielsen, A. (2023). Why do students lie and should we worry? an analysis of non-truthful reporting. *arXiv preprint arXiv:2302.13718*.

Davis, J. M., Guryan, J., Hallberg, K., and Ludwig, J. (2017). The economics of scale-up. Technical report, National Bureau of Economic Research.

De Groote, O., Fabre, A., Luflade, M., and Maurel, A. (2025). Sequential college admission mechanisms and off-platform options.

Dreyfuss, B., Glicksohn, O., Heffetz, O., and Romm, A. (2022). Deferred acceptance with news utility. Technical report, National Bureau of Economic Research.

Fu, C. (2014). Equilibrium Tuition, Applications, Admissions, and Enrollment in the College Market. *Journal of Political Economy*, 122(2):225–281.

Grenet, J., He, Y., and Kübler, D. (2022). Preference discovery in university admissions: The case for dynamic multioffer mechanisms. *Journal of Political Economy*, 130(6):1427–1476.

Gurantz, O., Howell, J., Hurwitz, M., Larson, C., Pender, M., and White, B. (2021). A national-level informational experiment to promote enrollment in selective colleges. *Journal of Policy Analysis and Management*, 40(2):453–479.

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

Hakimov, R., Schmacker, R., and Terrier, C. (2023). Confidence and college applications: Evidence from a randomized intervention. Technical report, Working Paper.

Hassidim, A., Romm, A., and Shorrer, R. I. (2020). The Limits of Incentives in Economic Matching Procedures. *Management Science*, (October):1–13.

Hastings, J. S., Neilson, C. A., and Zimmerman, S. D. (2013). Are some degrees worth more than others? evidence from college admission cutoffs in chile. Technical report, National Bureau of Economic Research.

Hendren, N. and Sprung-Keyser, B. (2020). A unified welfare analysis of government policies. *The Quarterly journal of economics*, 135(3):1209–1318.

Hoxby, C., Turner, S., et al. (2013). Expanding college opportunities for high-achieving, low income students. *Stanford Institute for Economic Policy Research Discussion Paper*, 12(014):7.

Hoxby, C. M. and Avery, C. (2012). The missing" one-offs": The hidden supply of high-achieving, low income students. Technical report, National Bureau of Economic Research.

Hyman, J. (2020). Can light-touch college-going interventions make a difference? evidence from a statewide experiment in michigan. *Journal of Policy Analysis and Management*, 39(1):159–190.

Immorlica, N., Leshno, J., Lo, I., and Lucier, B. (2020). Information acquisition in matching markets: The role of price discovery. *SSRN Electronic Journal*.

Kapor, A., Karnani, M., and Neilson, C. (2024). Aftermarket frictions and the cost of off-platform options in centralized assignment mechanisms. *Journal of Political Economy*, 132(7):2346–2395.

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

Karnani, M. (2023). Rcts with interference in matching systems. Technical report.

Kremer, M., Brannen, C., and Glennerster, R. (2013). The challenge of education and learning in the developing world. *Science*, 340(6130):297–300.

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

Larroucau, T. and Rios, I. (2023). Dynamic college admissions.

List, J. A. (2020). Non est disputandum de generalizability? a glimpse into the external validity trial. Technical report, National Bureau of Economic Research.

List, J. A. (2022). *The voltage effect: How to make good ideas great and great ideas scale*. Currency.

List, J. A., Shaikh, A. M., and Xu, Y. (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics*, 22:773–793.

Maniadis, Z., Tufano, F., and List, J. A. (2014). One swallow doesn't make a summer: New evidence on anchoring effects. *American Economic Review*, 104(1):277–290.

Meisner, V. and von Wangenheim, J. (2021). School choice and loss aversion.

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

Neilson, C. (2024). The rise of coordinated choice and assignment systems in education markets around the world. *Background paper to the World Development Report*.

Oreopoulos, P., Brown, R. S., and Lavecchia, A. M. (2017). Pathways to education: An integrated approach to helping at-risk high school students. *Journal of political economy*, 125(4):947–984.

Oreopoulos, P. and Ford, R. (2019). Keeping college options open: A field experiment to help all high school seniors through the college application process. *Journal of Policy Analysis and Management*, 38(2):426–454.

Rees-Jones, A. and Shorrer, R. (2023). Behavioral economics in education market design: A forward-looking review. Technical report.

Rios, I., Larroucau, T., Parra, G., and Cominetti, R. (2021). Improving the Chilean College Admissions System. *Operations Research*, 69:1186–1205.

Romano, J. P. and Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4):1237–1282.

Shorrer, R. I. and Sóvágó, S. (2021). Dominated Choices in a Strategically Simple College Admissions Environment : The Effect of Admission Selectivity.

Wiswall, M. and Zafar, B. (2015a). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, 82(2):791–824.

Wiswall, M. and Zafar, B. (2015b). How do college students respond to public information about earnings? *Journal of Human Capital*, 9(2):117–169.

Appendices

A Appendix to Section 3

A.1 Survey Details

A.1.1 Survey Design

As exposed in Section 2, we designed and conducted several nationwide surveys between 2020 and 2023 in collaboration with MINEDUC and DEMRE. These surveys included three main modules: (i) preferences, (ii) beliefs, and (iii) understanding of the admission process.[37] In this section, we describe each of these modules and the information they provide to characterize the different types of mistakes.

(i) *Preferences module.* This module aimed to elicit students' true preferences. Towards this end, we asked students about their true top preference, i.e., the program they prefer the most among all the programs in the system, regardless of their admission chances. In addition, we asked students about their true bottom preference, i.e., any program they did not list in their ROL and that they would prefer compared to being unassigned, regardless of their admission chances. These questions allow us to know whether students misreport their true top and bottom preferences, excluding them from their ROL.

(ii) *Beliefs module.* This module aimed to elicit students' beliefs about several relevant factors affecting their application, including admission probabilities, expected earnings, chances of retention and graduation, expected cutoffs, etc. We elicited this information for programs included in the students' preference list and others outside their ROL, including their true top and bottom preference and some random programs. These questions allow us to understand the role of biased beliefs on application mistakes.

(iii) *Knowledge of the mechanism.* This module aimed to measure students' knowledge and understanding of the system's rules, the requirements of the programs they applied to, their awareness of potential mistakes, and also to learn the reasons behind some of their decisions such as excluding their true top or true bottom programs. These questions allow us to further understand the drivers of application mistakes.

[37] In Appendix A.1.3, we report summary statistics for all surveys.

A.1.2 Surveys questions

In this subsection, we describe the main questions used in the analysis.[38]

- ***Current cutoff:*** *We show you now a list of the programs you applied to, in strict order of preference. For each of them, please tell us which do you think will be the value of the cutoff score for the CURRENT Admission Process and how likely do you think your application score will be above the cutoff score. We remind you that this is only a survey, and it DOES NOT affect in any way your application nor your admission probabilities. What do you think will be the value of the cutoff score for the current Admission Process for each of these programs?*

- ***Admission probability to a program:*** *How likely do you think your application score for the following programs will be above the current admission process's cutoff score?*
  On a scale from 0 to 100, where 0 is "completely sure that your application score WILL NOT be above the cutoff score for this program" and 100 is "completely sure that your application score WILL BE above the cutoff score for this program".

- ***Admission probability:*** *Regardless of the admission track. How likely you think that you will be admitted in some preference of your application?*
  On a scale of 0 to 100, where 0 is "completely sure that you WILL NOT be admitted in any of your preferences" and 100 is "completely sure that WILL BE admitted in one of your preferences".

- ***Knowledge about previous cutoffs:*** *It is referred to a cutoff score as the application score of the last admitted students to a given program. Each student is assigned to the highest reported preference for which her application score is greater than or equal to the cutoff score that realizes in the current Admission Process. Do you know which was the cutoff score for the PREVIOUS YEAR for each of the programs you applied to?*

- ***Knowledge about requirements:*** *Do you know the requirements and vacancies for each program in the following list?*
  - *Restricts preference order?*
  - *Requires Science test?*
  - *Number of vacancies?*
  - *Minimum weighted score?*
  - *Minimum math-verbal average?*
  - *Requires HYCS test?*

[38] The complete set of instruments is available upon request.

- ***Knows someone in the program:*** *Among the programs you applied to, do you know someone close to you who is currently studying there (friends, relatives, etc.)?*
- ***True-top:*** *This question aims to know where you would have applied to in the hypothetical case in which your admission did not depend on your scores. We remind you that this is only a hypothetical question and will not affect your application or admission probabilities. If the Admissions Process did not depend on your PSU scores, nor your NEM or Ranking scores. To which program would you have applied?*
- ***True-bottom:*** *Imagine a HYPOTHETICAL scenario in which you were NOT admitted to any program in your application list. Is there any program in the centralized system that you have NOT included in your application but you would prefer than being unassigned?*
- ***Knowledge about income and employment:*** *Regarding the graduation process of higher education and considering your knowledge about characteristics like the average income of the graduates and employment rates, how informed do you think you are about the following programs?*
  - *I don't think I am informed*
  - *Slightly informed*
  - *Moderately informed*
  - *Quite informed*
  - *Completely informed*
- ***Average income:*** *The objective of this question is to know about your expectations of FUTURE INCOME in some of the programs that you applied to and in some that you did not. What do you think is the average monthly income of graduates at the fourth year of graduation from the following programs? In a scale of $0 to $3.000.000*

A.1.3 Survey Summary Statistics

Table A.1: Summary Statistics Applicants and Survey

Year	Group	N	Demographics and Scores			High-School Type		
			Female	Low-Income	Avg. Math-Verbal	Public	Voucher	Private
2020	Applied	146465	0.572 (0.495)	0.621 (0.485)	507.439 (171.069)	0.281 (0.449)	0.537 (0.499)	0.179 (0.384)
	Surveyed	38093	0.613 (0.487)	0.625 (0.484)	530.32 (154.798)	0.281 (0.45)	0.543 (0.498)	0.172 (0.377)
2022	Applied	148819	0.577 (0.494)	0.491 (0.5)	543.032 (88.567)	0.288 (0.453)	0.549 (0.498)	0.161 (0.367)
	Surveyed	31791	0.64 (0.48)	0.498 (0.5)	549.103 (92.367)	0.289 (0.453)	0.562 (0.496)	0.145 (0.353)
2023	Applied	187225	0.58 (0.494)	0.487 (0.5)	624.698 (117.734)	0.302 (0.459)	0.554 (0.497)	0.143 (0.35)
	Surveyed	49230	0.637 (0.481)	0.553 (0.497)	588.923 (144.089)	0.362 (0.481)	0.552 (0.497)	0.083 (0.276)

Note: Summary statistics for the 2020, 2022 and 2023 applicants and surveyed students. Applicants data considers all students who applied to at least one program in their corresponding year. Survey data considers all students who completed their corresponding survey. For 2023, we consider students who completed both base and endline surveys.

A.2 Understanding Biases in Under-Confidence and Ordering Mistakes

In Table A.2, we report the results of linear regressions examining the absolute value of the bias related to admission probabilities and expected cutoff scores,[39] focusing on students' knowledge and various socioeconomic and demographic factors. We observe that greater bias in admission probabilities is positively correlated with increased biases in cutoff scores (*Bias Norm Cutoff*) and negatively correlated with students' awareness of admission requirements (*Requirements Knowledge Share*) and their connections within the program (*Knows someone in the program*). Furthermore, this analysis reveals that students from more affluent backgrounds, with higher application scores, and those attending private schools exhibit lower bias in admission probabilities and cutoff scores. These results suggest a disparity in information accessibility, implying that students from less advantaged socioeconomic backgrounds may possess less accurate beliefs about their admission probabilities, potentially leading to application mistakes.

[39] As part of our survey, we also elicit students beliefs on the cutoffs of some relevant programs, including their true top preference, their top reported one, among others. Moreover, we ask students whether they know the previous year's cutoff. Indeed, we find that only 58% declare to know the previous year's cutoffs for all their listed programs, and 9% declare to ignore all of them. Although DEMRE does not provide any information about programs' cutoffs during the application process, this information can be typically found on universities' websites. One reason behind the lack of centralized information about cutoff scores is the concern that some students might not understand what a cutoff score exactly means. For instance, they might believe that programs predetermine cutoffs and may not understand that these

Table A.2: OLS regression for Bias Norms over Adm. Probabilities and Cutoffs

	Abs. Bias Adm. Prob		Abs. Bias Cutoff	
	(1)	(2)	(3)	(4)
Avg. Math-Verbal Normalized	-0.022^{***}	-0.016^{***}	-2.329^{***}	-1.788^{***}
	(0.001)	(0.001)	(0.084)	(0.083)
Female	0.024^{***}	0.028^{***}	-0.221	0.581^{***}
	(0.001)	(0.001)	(0.226)	(0.220)
Below median income	0.026^{***}	0.022^{***}	1.887^{***}	1.097^{***}
	(0.002)	(0.002)	(0.251)	(0.245)
Public	0.058^{***}	0.042^{***}	9.351^{***}	6.893^{***}
	(0.003)	(0.002)	(0.380)	(0.372)
Voucher	0.052^{***}	0.036^{***}	9.279^{***}	7.134^{***}
	(0.002)	(0.002)	(0.318)	(0.312)
Requirements Knowledge Share		-0.023^{***}		-4.962^{***}
		(0.002)		(0.337)
Knows cutoff for every program		-0.095^{***}		-21.209^{***}
		(0.003)		(0.427)
Knows cutoff for some programs		-0.059^{***}		-13.143^{***}
		(0.003)		(0.435)
Knows someone in the program		-0.008^{***}		-4.435^{***}
		(0.002)		(0.277)
Absolute bias cutoff		0.001^{***}		
		(0.000)		
Mean	0.297	0.297	41.984	41.984
Observations	148,539	148,539	148,539	148,539

Note: Each observation represents a student-program pair where the student is not PACE and completed the 2020 survey, the program has a cutoff, and the student either (i) included the program in their preference list or (ii) reported it as their top-true preference. Requirement Knowledge Share is the fraction of requirements the student claims to know for the applied program. Knows the cutoffs for every (some) program is a dummy variable indicating whether the student knows the cutoffs for all (some) programs. Knows someone in the program is a dummy for knowing someone in the program. Abs. Bias Adm. Prob. (Abs. Bias Cutoff) is the absolute value of the bias (cutoff) for the program's admission probability. Models (2) and (4) also controls for the position of the program in the student's preference list. Significance reported: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Overall, these results suggest that having information about previous year cutoff scores could help students to reduce their biases over admission probabilities and, thus, reduce the prevalence of *mistakes on assignment probabilities*. However, since students might anchor their beliefs to the previous year's information and misunderstand the meaning of cutoff scores, it is essential also to provide information about the current admission process and educate them on the mechanism's rules.

A.3 Mistakes on Valuations and Awareness

To quantify the magnitude of this lack of knowledge and understand how it correlates with students' characteristics, in Table A.3, we report the results of linear regressions on the absolute percentage of bias on the average income,[40] controlling by application scores, gender, income, and type of high school. We only consider students who did not open the intervention to avoid potential effects driven by the intervention we discuss in Section 4.

Table A.3: Regression Results for Absolute Bias on Average Income

	Percentage of Absolute Bias in Average Income				
	Top Reported (1)	Top True (2)	Bottom Reported (3)	Bottom True (4)	Random (5)
Avg. Math-Verbal normalized	-22.036^{***}	-20.669^{***}	-20.892^{***}	-21.811^{***}	-21.246^{***}
	(0.928)	(1.314)	(1.003)	(1.893)	(0.965)
Female	7.232^{***}	4.328^{**}	6.400^{***}	8.600^{***}	6.884^{***}
	(1.551)	(2.116)	(1.672)	(3.330)	(1.624)
Family income below median	-2.970^{*}	-2.996	-3.237^{*}	-2.747	-4.592^{***}
	(1.564)	(2.073)	(1.687)	(3.377)	(1.638)
Public	-9.218^{***}	-6.760^{*}	-8.839^{***}	-13.978^{**}	-6.996^{**}
	(2.628)	(3.560)	(2.814)	(5.646)	(2.745)
Voucher	-7.392^{***}	-6.554^{**}	-5.147^{**}	-12.071^{**}	-6.661^{***}
	(2.352)	(3.211)	(2.509)	(5.191)	(2.467)
Mean	50.397	49.595	53.163	55.745	54.293
Observations	6,855	3,366	6,146	1,607	6,470

Note: Sample includes all students who completed the survey and did not receive the intervention. Significance reported: $^{*}p < 0.1;^{**} p < 0.05;^{***} p < 0.01$.

First, we observe that the constant is large and significant for all the models considered, suggesting that students have incorrect beliefs about their expected income upon graduation. Second, we observe that biases are negatively correlated with application scores, meaning that students with lower scores have less accurate beliefs. Third, we observe that the variable Female is positive

vary from year to year.

[40]This metric is computed as the absolute value of the difference between students' beliefs and the actual value of the average income, divided by the latter.

and significant, i.e., women are more biased about average expected earnings than men. Overall, these results suggest that students have large biases in their assessment of the average future earnings, potentially leading them to *mistakes on valuations*.

A.4 Mistakes by Truth-tellers

As discussed in Section 2.3, a truth-teller is a student who includes all programs they prefer over being unassigned in their ranked-ordered list, unless they prefer more programs than the maximum allowed (i.e., full-list students). In such cases, Equation 3 implies that the student will truncate their list by excluding the least-preferred options. However, if the student includes any program for which they have zero admission probability, excluding any program that the student may prefer over being unassigned and for which they may have positive admission probability would lead to a strict increase in the student's utility, resulting in a mistake on truth-telling.[41]

When the constraint in the length of their list is binding, students may exclude many programs they prefer over being unassigned. As discussed in Section 3.1, our survey allows us to identify one such program: the bottom-true preference. Then, a full-list truth-teller—i.e., a student for whom the constraint is binding and that included their top-true preference in their list—makes an overconfidence mistake if (i) they have a positive admission chance for their bottom-true program; (ii) they include at least one program in their list with zero admission probability; and (iii) they face a positive overall risk of being unassigned. Together, these conditions imply that adding the bottom-true program to their list would have strictly increased the student's expected utility, implying a mistake by truth-telling,

Table A.4 presents summary statistics on full-list truth-tellers from the 2022 survey. Specifically, it reports the total number of full-list truth-tellers among survey respondents (under *Total*) and, among these, the number and fraction of students who include at least one program with zero admission probability and who exclude their bottom-true preference (under *Any zero* and *Exclude Bottom-True*, respectively). Then, we compute the number and fraction of students who make an ex-ante (ex-post) mistake by truth-telling, i.e., students who have at least one program with zero chance, who excluded their bottom-true preference, and that had a positive admission probability (application score above the cutoff) for the latter.

[41] By definition, a short-list truth-teller cannot make a mistake by truth-telling, since such mistakes require excluding programs preferred over being unassigned, contradicting the premise of truthful reporting.

Table A.4: Mistakes on Truth-Telling

Full-List Truth-Tellers					Mistake by Truth-Telling			
Total	Any zero		Exclude Bottom-True		Ex-Ante		Ex-Post	
N	N	%	N	%	N	%	N	%
887	796	89.741	271	30.552	10	1.127	6	0.676

Note: Sample includes students who completed the survey of 2022, who are not PACE and applied to ten programs including their top true preference (i.e., full-list truth-tellers), and were not part of treatments T2 and T3. The columns under *Any zero* report the number and fraction of full-list truth-tellers that included at least one program for which they had zero admission chance in their final ROL. The columns under *Exclude Bottom-True* report the number and fraction of full-list truth-tellers that excluded their bottom-true program.

Among the 7,311 survey respondents (see Table 1), 12.13% ($N = 887$) are classified as full-list truth-tellers. Of these students, 89.74% include at least one program with zero admission probability in their final ROL, while 30.55% exclude their bottom-true program. Finally, we observe that the number of the full-list truth-tellers that meet the latter two conditions and make a mistake ranges from 6 to 10, which represents between 0.676% and 1.127% of this group. Hence, we conclude that the prevalence of mistakes by truth-tellers is relatively low but not negligible.

B Appendix to Section 4

B.1 Intervention Design

Figure B.1: Module 1 - Information on Programs' Characteristics Included in Application

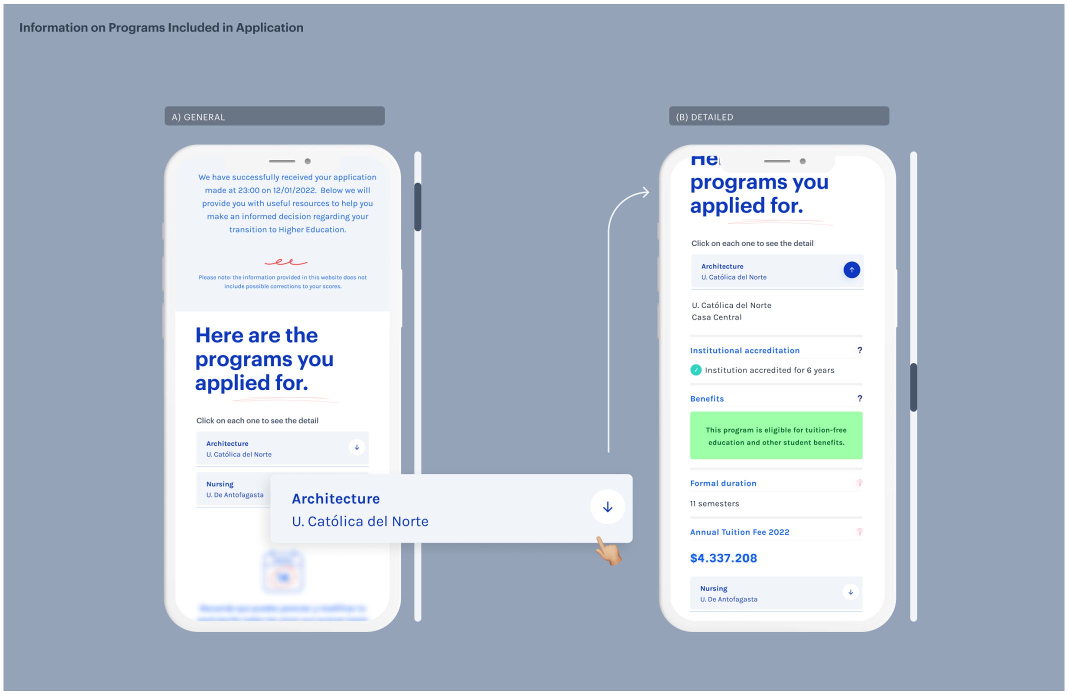


Figure B.2: Detailed zoom

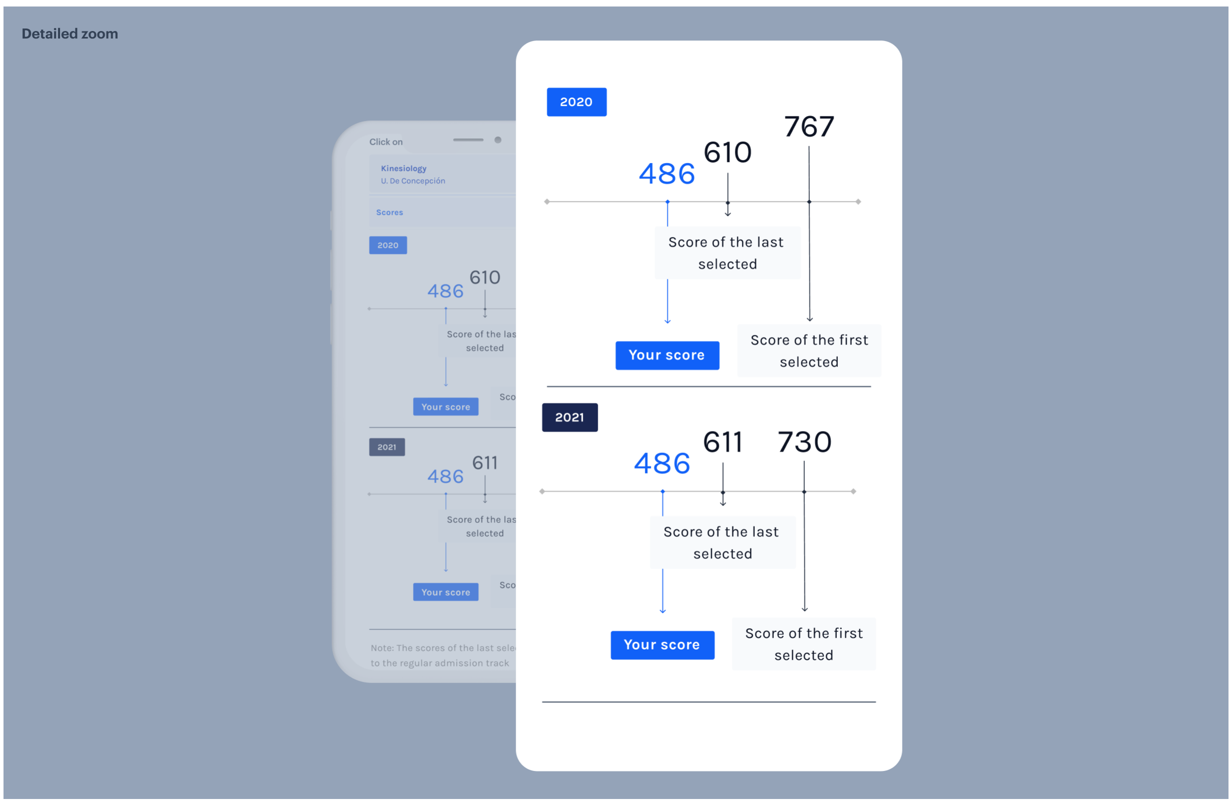


Figure B.3: Feedback on Programs' Admission Chances: red warning

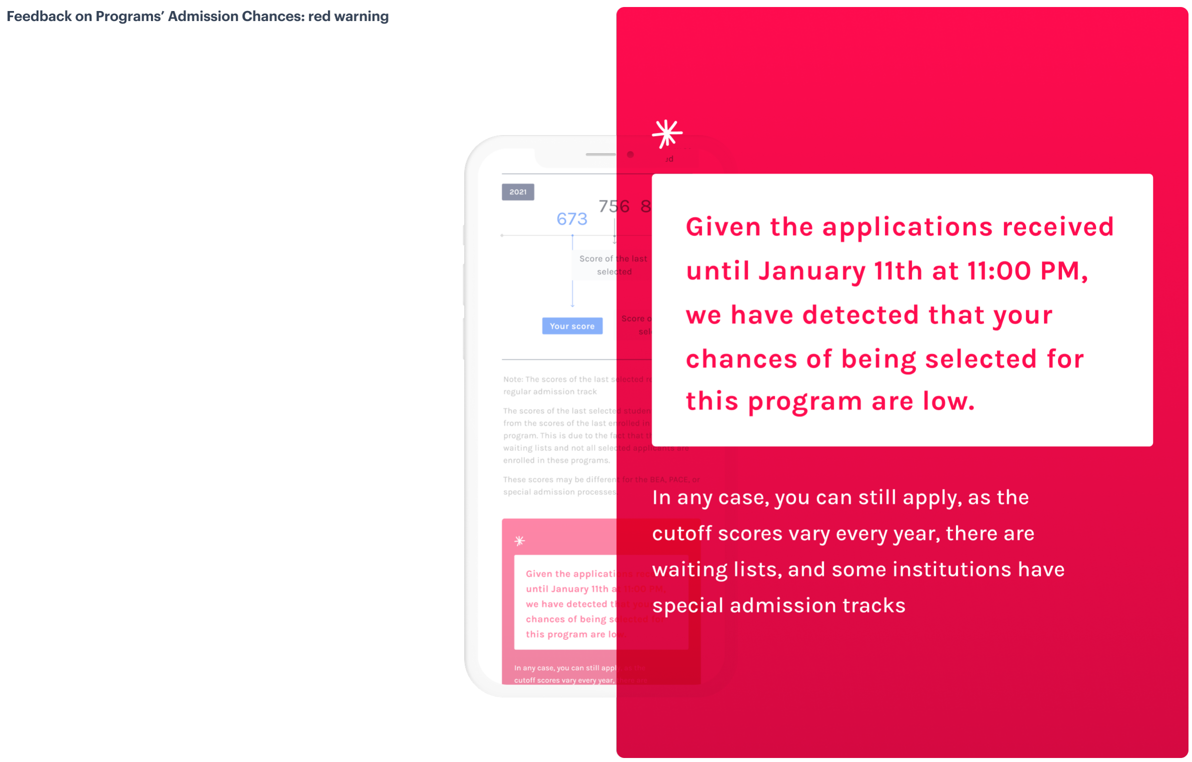
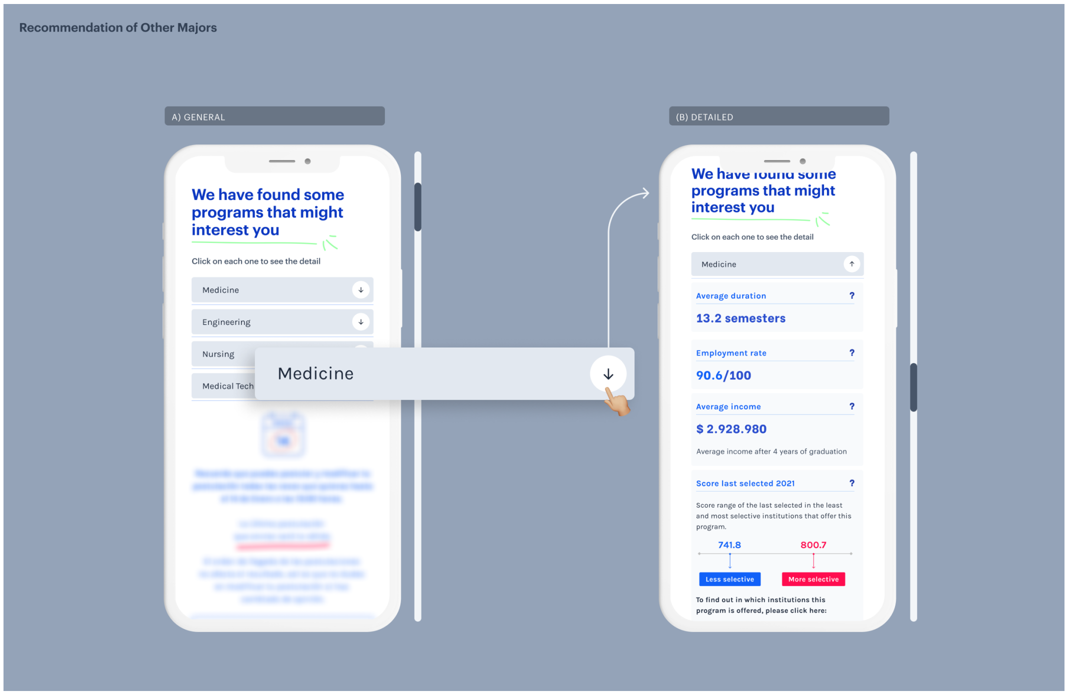


Figure B.4: Major Recommendations



B.2 Treatment Assignment and Stratification

As discussed in Section 4, we assign students to treatments in a stratified way to achieve balance. For the stratification we consider the following characteristics:

- Female: dummy variable equal to 1 if the student is female, and 0 otherwise.
- Region: categorical variable that takes four 3 levels depending on the region where the student graduated from high-school. Specifically, this variable is equal to 1 for students graduating in the north (regions I, II, III, IV and XVII); 2 for students graduating in the center (regions V, XIII, VI, VII); and 3 for students graduating in the south (regions VIII, IX, X, XI, XII, XIV and XVI).
- Score: categorical variable that takes 4 values depending on the average score between the tests in Math and Verbal. Specifically, this variable is equal to 1 for students with average score below 450; 2 for students which average score between 450 and 600; and 3 for students with score above 600.
- Overall alert: there are three types of overall alerts: (i) reach, (ii) safety, and (iii) more information. Each student can be assigned to one of these groups, and thus we also use this assignment as part of the stratification.

In Table B.1, we report the mean of different variables and p-values from tests of differences in means between Treatment 1 and each of the other treatment arm. We observe that we fail to reject equality across all these tests, which confirms that covariates are balanced across treatment arms.

Table B.1: Balance Tests

	Mean				P-Value Difference		
	Treatment 1	Treatment 2	Treatment 3	Treatment 4	T2-T1	T3-T1	T4-T1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Open	0.282	0.282	0.281	0.284	0.985	0.882	0.601
Female	0.598	0.597	0.598	0.597	0.802	0.892	0.778
Low-Income	0.414	0.418	0.420	0.416	0.311	0.186	0.591
Metropolitan Region	0.367	0.367	0.369	0.372	0.919	0.588	0.220
Public High school	0.292	0.293	0.294	0.293	0.725	0.629	0.818
Voucher High school	0.561	0.560	0.561	0.557	0.791	0.895	0.337
GPA	6.046	6.045	6.041	6.046	0.819	0.289	0.972
Math-Verbal	0.337	0.338	0.337	0.340	0.902	0.960	0.623
Assigned Interim	0.779	0.777	0.778	0.777	0.634	0.698	0.483
Overall Ratex Interim	0.770	0.768	0.769	0.768	0.574	0.854	0.632
Observations	26,404	26,546	26,512	26,467			

NOTES. This table reports results from covariates balance checks across treatment arms. Columns (1)–(4) report the mean of the variable for treatment group 1-4, respectively. Columns 5–7 display p-values from tests of differences in means between Treatment 1 and each of the other treatment groups (Treatment 2, Treatment 3, and Treatment 4).

Table B.2: Balance Tests - Open

	Demographics			High-School Type		Scores	
	Female	Low-Income	Metrop. Region	Public	Voucher	GPA	Math-Verbal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Open	0.069	0.006	0.032	-0.002	-0.002	0.056	0.103
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.006)
Constant	0.578	0.415	0.360	0.293	0.560	6.029	0.309
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Observations	105,917	105,929	105,917	105,917	105,230	105,800	105,917

NOTES. This table compares the characteristics of applicants who opened their personalized website against those who did not. Each column reports results from the OLS estimation of a linear regression model, considering different students' characteristics as the outcome variable. Robust standard errors are reported in parentheses.

Table B.3: Balance Tests - Early & Late Applicants

	Demographics			High-School Type		Scores	
	Female	Low-Income	Metrop. Region	Public	Voucher	GPA	Math-Verbal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
RCT Participant	0.071	0.026	-0.067	0.019	0.037	-0.033	-0.199
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)
Constant	0.525	0.391	0.433	0.275	0.522	6.058	0.536
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)
Observations	148,769	148,782	148,769	148,769	147,508	148,769	148,769

NOTES. This table compares the characteristics of early applicants, selected to participate in the RCT, and late applicants, who were not participating in the experiment. Each column reports results from the OLS estimation of a linear regression model, considering different students' characteristics as the outcome variable. Robust standard errors are reported in parentheses.

B.3 Admission Probabilities

To compute the admission probabilities, we use a bootstrap procedure similar to that in Agarwal and Somaini (2018) and Larroucau and Ríos (2018). The main difference is that these approaches use complete information regarding the applications. In our case, we only have the application list of close to 2/3 of the students that ended up applying, so running the bootstrap procedure on this sample would considerably underestimate the cutoffs. For this reason, our first task is to estimate the total number of students that would apply in 2022 based on the applications received so far. To accomplish this, we divide the population into three segments based on their average score between Math and Verbal (the two mandatory exams of the PSU/PDT). Then, using data from 2020 and 2021, we estimate which fraction of all students that take the national exam would apply to at least one program in the centralized system taking the average between these two years. Finally, comparing this number with the actual fraction of students in each score bin that have applied so far, we quantify the number of students that have not applied yet.

Based on the number of applicants missing, we perform 1,000 bootstrap simulations, each consisting of the following steps:

1. Sample with replacement the number of students missing in each bin score, and incorporate the sampled students to the pool of applications received so far.

2. Run the assignment mechanism used in the Chilean system. See Rios et al. (2021) for a detailed description of the mechanism used in Chile to solve the college admissions problem.

3. Compute the cutoff of each program for both the regular and BEA admission processes.

As a result of this procedure, we obtain two matrices (for the regular and BEA processes) with 1,000 cutoffs for each program. Hence, 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.

B.4 Recommendations

The recommendation algorithms works as follows:

1. Find the most and the second most popular majors based on the preferences included in the student's ROL.

2. For each pair of majors, and considering the most and the second most preferred major of each student, compute a transition matrix that returns the probability that a given major is followed by another major as the most preferred ones.

3. For each student, compute the set of feasible majors considering the student's scores and her admission probabilities (obtained as described in the previous section).

4. For students with high scores (i.e., average between Math and Verbal above 600), choose four majors according to the following rule:

   (a) Choose most preferred major according to the student's list of preferences,

   (b) Choose the second most preferred major according to the student's list of preferences,

   (c) Choose the major with the highest average wage[42] among all majors considering the transition matrix previously computed,

   (d) Choose the major with the highest average wage among all feasible majors (i.e., majors for which the student has a positive probability of assignment) considering the transition matrix previously computed.

5. For students with low scores (i.e., average between Math and Verbal below 600), choose four majors according to the following rule:

   (a) Choose the most preferred major according to the student's list of preferences,

   (b) Choose the second most preferred major according to the student's list of preferences,

   (c) Choose the major with the highest expected wage among all majors belonging to IPs or CFTs,

   (d) Choose the major with the highest expected wage among all feasible majors (i.e., majors for which the student has a positive probability of assignment) considering the transition matrix previously computed.

[42] Average wages are measured at the fourth year after graduation. This statistic is computed by SIES and provided to us by MINEDUC.

B.5 Additional Results

Table B.4: Placebo Test

	Applications		Assignment		
	Modified (1)	Incr. Prob. (2)	Improved (3)	Entered (4)	Benefit & Persisted (5)
Treatment 2	0.0005	0.0005	−0.001	0.001	−0.002
	(0.003)	(0.003)	(0.002)	(0.003)	(0.001)
Treatment 3	0.003	0.003	0.001	0.001	0.0002
	(0.003)	(0.003)	(0.002)	(0.003)	(0.001)
Mean (Control)	0.104	0.014	0.028	0.024	0.018
Observations	76,024	21,163	58,216	17,266	75,482
Strata FE	Yes	Yes	Yes	Yes	Yes

Note: Significance reported: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table B.5: Withdraw

	RCT		Scale-Up	
	(1)	(2)	(3)	(4)
Treatment 2	0.0002	0.0002		
	(0.001)	(0.001)		
Treatment 3	−0.0002	−0.0002		
	(0.001)	(0.001)		
Open			0.0004	0.0001
			(0.001)	(0.001)
Mean (Control / No text)	0.001	0.001	0.001	0.001
Observations	29,898	29,898	132,975	132,965
Strata FE	No	Yes	No	Yes

Note: Significance reported: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table B.6: Regression Results among Openers - T1 vs T4

	Applications		Assignment		
	Modified	Incr. Prob.	Improved	Entered	Benefited&Persisted
	(1)	(2)	(3)	(4)	(5)
Treatment 4	0.0068	-0.0064	0.0013	-0.0049	-0.0014
	(0.0056)	(0.0065)	(0.0033)	(0.0058)	(0.0024)
	[0.5842] {0.6730}	[0.6634] {0.6730}	[0.7426] {0.6900}	[0.6634] {0.6730}	[0.7426] {0.6900}
Mean (Control)	0.1316	0.0467	0.0322	0.0299	0.0217
Observations	14,960	3,804	11,724	3,116	14,840
Stratas FE	Yes	Yes	Yes	Yes	Yes

NOTES. Modified is a binary variable equal to 1 if the student modified her application after the information was sent, 0 otherwise. Incr. Prob. is a binary variable equal to 1 if the admission probability associated to the initial rank-ordered list submitted is lower than the one associated to the final rank-ordered list, 0 otherwise. This variable is defined only for students with a positive admission risk given their initial rank-ordered list. Improved is a dummy variable equal to 1 if the student was assigned to a program ranked above the one where they would have been assigned given their initial rank-ordered list, 0 otherwise. It is only defined for the sample of students who would have been matched to a program given their initial rank-ordered list and who did not remove this program from their list. Entered is defined for students who would not have been assigned to any program given the initial rank-ordered list they submitted. It is a binary variable equal to 1 if the student is assigned given their rank-ordered list submitted after the information was sent, 0 otherwise. Benefited&Persisted is a binary variable equal to 1 if the student either entered or improved and persisted for two years in the same program, 0 otherwise. Robust standard errors are reported in parentheses. We report in brackets the p-values adjusted for multiple hypothesis testing following the procedure described in Romano and Wolf (2005) and in braces the q-values computed following Anderson (2008a).

Table B.7: Impact on Absolute Biases in Average Earnings

	Top-True	Top-Reported	Bottom-True	Bottom-Reported	Random
	(1)	(2)	(3)	(4)	(5)
Treatment 2	-7.5128	3.2644	6.1639	-1.6456	-0.4523
	(4.3712)	(2.9052)	(6.9910)	(2.9638)	(3.1996)
Treatment 3	-4.3718	-1.4839	8.3782	-2.3558	-1.6792
	(4.2917)	(2.7023)	(6.2008)	(2.9911)	(3.1411)
Mean (Control)	53.4459	47.4743	48.4287	51.5722	55.3175
Observations	1,353	2,945	673	2,657	2,751
Stratas FE	Yes	Yes	Yes	Yes	Yes

NOTES. This table reports results from the OLS estimation of linear regression models where the dependent variable is the absolute value of students' subjective bias in their admission probabilities for different types of programs. The sample is limited to students who responded to the survey and opened the intervention. Robust standard errors reported in parentheses.

C Appendix to Section 5

C.1 Background

As in the previous year, students participated in a national exam that provided them with test scores that the system uses to compute their application scores in each program they listed in their preference list. However, MINEDUC introduced a series of changes to the admission process. First, they completely redesigned the admission exam by changing its focus (moving from knowledge-based to attitude-based) and adding a math-specific exam. In addition, MINEDUC changed the normalization rules and, more importantly, the range of possible scores, moving from a $[210, 850]$ to a $[100, 1000]$ scale.

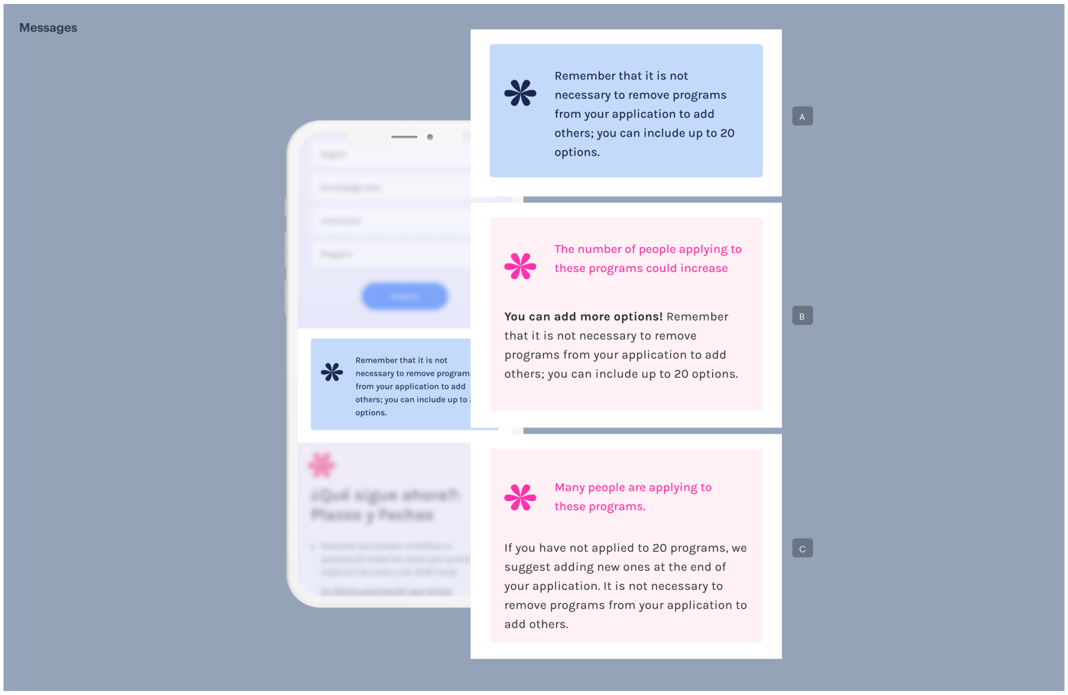
Second, MINEDUC introduced the option to take the national exam twice per year and changed the rules on how to compute application scores for students that took the exam several times and thus have multiple pools of scores.[43] Specifically, they moved from a pool-based approach, in which the application score is computed considering the best pool among all the ones available, to a test-specific approach, in which the application score is computed considering the best score for each specific exam, potentially combining different pools of scores.

Finally, given all the changes mentioned above and the advice from the research team, MINEDUC decided to increase the constraint on the length of preference lists from ten to twenty programs.

A critical consequence of all these changes is that the previous year's cutoffs were not as informative as in previous years. Indeed, many students had no idea how to assess their chances of admission, as they had no reference point, and the uncertainty was considerably higher. As a result, MINEDUC decided it was crucial to provide students with as much guidance as possible, and thus decided to implement our information policy for all students nationwide. Hence, students who opened their personalized websites received the same information fields, so we do not have proper treatment and control groups as described in Section 4. Nevertheless, as we later discuss, we can still estimate the effect of the intervention using an encouragement design.

[43] Moreover, MINEDUC had to introduce conversion tables to transform scores from the previous scale to the new one.

Figure C.1: Feedback on Application Strategy



C.2 Instrument Validity

For an instrument to be valid, we need it to satisfy two conditions: (i) relevance, and (ii) exclusion. The former states that the instrument is correlated with the endogeneous variable of interest. In our case, we need to confirm that receiving a whatsapp is correlated with opening the intervention. We consider the first stage regression

$$O_i \sim W_i + X_i + \epsilon_i$$

where $O_i = 1$ if the student opens the intervention and zero otherwise; $W_i = 1$ if student i receives a Whatsapp encouragement message and zero otherwsie; X_i is a vector of control variables (in this case, the risk level group); and ϵ_i is an error term.

Table C.1 reports the results of the first-stage of our 2SLS estimation. We find that receiving the WhatsApp reminder significantly increases the probability that students access their personalized website, corresponding to a 34% increase (F-stat = 1414.6).

In our context, the instrument is exogeneous by design as we randomize students receiving the encouragement message. Table C.2 confirms this by showing that students who randomly received the reminder are, on average, similar to those who did not. Each column reports results from the

Table C.1: Regression Results: First Stage

	Open Website
	(1)
Receive Whatsapp	0.1782
	(0.0029)
Risk Group	Yes
Mean (No Reminder)	0.5217
Observations	132,893
F-statistic	1414.8112

NOTES. This table shows results from the first-stage regression of the instrumental variable strategy.

OLS estimation of a linear regression model, considering different students' characteristics as the outcome variable. We find that the characteristics of students who received the reminder do not significantly differ compared to those who did not, confirming that the encouragement messages were properly randomized.

Table C.2: Balance Tests - Received Whatsapp Reminder

	Demographics			High-School Type		Scores	
	Female	Low-Income	Metrop. Region	Public	Voucher	GPA	Math-Verbal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Received Whatsapp	0.000	-0.000	-0.003	0.004	-0.001	-0.001	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.006)
Constant	0.603	0.445	0.353	0.306	0.565	6.029	0.143
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)
Observations	132,892	132,893	132,524	132,893	132,060	132,892	132,892

NOTES. This table compares the characteristics of applicants who received the Whatsapp reminder to open their personalized website against those who did not. Each column reports results from the OLS estimation of a linear regression model, considering different students' characteristics as the outcome variable. Robust standard errors are reported in parentheses.

C.3 Characterizing Scale-Up Compliers

To compare the population of RCT openers with the population of compliers in the scale-up, we implement the method proposed by Abadie (2003) to recover the distribution of observable characteristics among compliers. Let D denote whether a student opened the personalized website and Z the instrument indicating the receipt of the WhatsApp message. The procedure proceeds in four steps:

1. **Estimate the first stage.** For each student, we estimate $E[D \mid Z = 1, X]$ and $E[D \mid Z = 0, X]$, where X denotes the vector of observable characteristics.

2. **Compute the complier score:**

$$\pi(X) = E[D \mid Z = 1, X] - E[D \mid Z = 0, X].$$

3. **Construct complier weights:**

$$w(X) = \frac{\pi(X)}{E[\pi(X)]},$$

   which reweight the sample to recover the distribution of X among compliers.

4. **Compute complier means.** Weighted averages of each covariate using $w(X)$ yield estimates of $E[X \mid \text{compliers}]$.

We apply this procedure to the 2023 scale-up data and compare the resulting complier characteristics with those of RCT participants who opened the intervention in 2022. Table C.3 reports the summary statistics and standard errors.

Table C.3: Summary Statistics

		Demographics			High-School Type		Scores	
Group	N	Female	Low-Income	Metro. Region	Public	Voucher	GPA	Math-Verbal
RCT: Openers	30755	0.645 (0.003)	0.501 (0.003)	0.39 (0.003)	0.292 (0.003)	0.557 (0.003)	6.086 (0.003)	544.989 (0.522)
Scale-up: Compliers	132965	0.586 (0.001)	0.502 (0.001)	0.353 (0.001)	0.32 (0.001)	0.57 (0.001)	5.979 (0.002)	596.633 (0.294)

Note: Summary statistics for RCT openers in 2022 and Scale-up compliers in 2023. Summary statistics for Scale-up compliers are computed using the procedure in Abadie (2003). Standard errors reported in parenthesis.

We observe meaningful differences across several observable characteristics. For example, the set of compliers includes a lower fraction of female students and fewer students from the

Metropolitan region. We also find that average GPA is slightly higher among RCT Openers, whereas their average Math–Verbal score is lower compared to Scale-up compliers (the difference is approximately 0.5 standard deviations in the Math exam distribution). Finally, we observe that the fraction of students coming from public and voucher schools is higher among scale-up compliers relative to RCT Openers.

C.4 Additional results

Table C.4: Summary Statistics by Risk Group

	N	Applications		Assignment		
		Modified	Incr. Prob.	Improved	Entered	Benefited & Enrolled
		(1)	(2)	(3)	(4)	(5)
Panel A: Did not open						
High Risk	14,345	0.146	0.047	0.082	0.034	0.017
Medium Risk	6,388	0.163	0.047	0.024	0.053	0.023
Low Risk	36,419	0.193	0.047	0.058	0.045	0.046
Panel A: Opened						
High Risk	15,983	0.291	0.108	0.054	0.083	0.050
Medium Risk	7,462	0.275	0.108	0.044	0.096	0.044
Low Risk	52,296	0.282	0.108	0.083	0.096	0.067

NOTES. The high-risk group (low-risk, resp.) includes students with an overall admission probability below 1% (above 99%, resp.) given their initial rank-ordered list. Students considered as facing a medium risk have an overall admission probability between 1% and 99% given their initial rank-ordered list. Modified is a binary variable equal to 1 if the student modified her application after the information was sent, 0 otherwise. Incr. Prob. is a binary variable equal to 1 if the admission probability associated to the initial rank-ordered list submitted is lower than the one associated to the final rank-ordered list, 0 otherwise. This variable is defined only for students with a positive admission risk given their initial rank-ordered list. Improved is a dummy variable equal to 1 if the student was assigned to a program ranked above the one where they would have been assigned given their initial rank-ordered list, 0 otherwise. It is only defined for the sample of students who would have been matched to a program given their initial rank-ordered list and who did not remove this program from their list. Entered is defined for students who would not have been assigned to any program given the initial rank-ordered list they submitted. Entered is a binary variable equal to 1 if the student is assigned given their rank-ordered list submitted after the information was sent, 0 otherwise. Benefited&Enrolled is a binary variable equal to 1 if the student either entered or improved and enrolled, 0 otherwise.

Table C.5: Regression Results: Instrumental Variables - By Risk Level

	Entered		Improved	
	High	Medium	Medium	Low
	(1)	(2)	(3)	(4)
Open	0.0334	0.0348	0.0104	0.0379
	(0.0157)	(0.0449)	(0.0236)	(0.0119)
Mean (No Reminder)	0.0578	0.0740	0.0342	0.0714
Observations	30,006	5,108	8,547	86,873

NOTES. The high-risk group (low-risk, resp.) includes students with an overall admission probability below 1% (above 99%, resp.) given their initial rank-ordered list. Students considered as facing a medium risk have an overall admission probability between 1% and 99% given their initial rank-ordered list. Improved is a dummy variable equal to 1 if the student was assigned to a program ranked above the one where they would have been assigned given their initial rank-ordered list, 0 otherwise. It is only defined for the sample of students who would have been matched to a program given their initial rank-ordered list and who did not remove this program from their list. Entered is defined for students who would not have been assigned to any program given the initial rank-ordered list they submitted. It is a binary variable equal to 1 if the student is assigned given their rank-ordered list submitted after the information was sent, 0 otherwise. Robust standard errors are reported in parentheses.

Table C.6: Regression Results: Search Engine

	Applications		Assignment		
	Modified	Incr. Prob.	Improved	Entered	Benefited & Enrolled
	(1)	(2)	(3)	(4)	(5)
Panel A:					
Search	0.2037	0.1215	0.0464	0.0831	0.0419
	(0.0045)	(0.0060)	(0.0035)	(0.0057)	(0.0026)
Panel B: Controlling for Number of Applications					
Search	0.2018	0.1144	0.0478	0.0778	0.0417
	(0.0042)	(0.0058)	(0.0035)	(0.0056)	(0.0026)
Mean (No Text)	0.2829	0.1077	0.0801	0.0849	0.0609
Observations	75,741	22,080	55,875	18,650	74,525
Risk Group	Yes	Yes	Yes	Yes	Yes

NOTES. Modified is a binary variable equal to 1 if the student modified her application after the information was sent, 0 otherwise. Incr. Prob. is a binary variable equal to 1 if the admission probability associated to the initial rank-ordered list submitted is lower than the one associated to the final rank-ordered list, 0 otherwise. This variable is defined only for students with a positive admission risk given their initial rank-ordered list. Improved is a dummy variable equal to 1 if the student was assigned to a program ranked above the one where they would have been assigned given their initial rank-ordered list, 0 otherwise. It is only defined for the sample of students who would have been matched to a program given their initial rank-ordered list and who did not remove this program from their list. Entered is defined for students who would not have been assigned to any program given the initial rank-ordered list they submitted. Entered is a binary variable equal to 1 if the student is assigned given their rank-ordered list submitted after the information was sent, 0 otherwise. Benefited&Enrolled is a binary variable equal to 1 if the student either entered or improved and enrolled, 0 otherwise. Robust standard errors are reported in parentheses.

C.5 Drivers

To examine whether the policy influences behavior through changes in beliefs on admission probabilities, we use the panel of respondents from the baseline and endline surveys conducted in 2023. For each student in the panel, we calculate a measure of bias in expected cutoffs and a measure of bias in admission probabilities by taking the difference between students' subjective beliefs (elicited in the baseline and endline surveys) and the rational expectations of expected cutoffs and admission probabilities. We then compute the difference between the absolute value of the bias in beliefs for baseline and end-line measures across the top-reported, bottom-reported, and true top programs declared in the baseline survey, and over their overall admission probability.

Table C.7: Regression Results: Reduction in Absolute Bias

	Cutoffs			Adm. Probs.		
	Top-true (1)	Top-reported (2)	Bottom-reported (3)	Top-true (4)	Top-reported (5)	Bottom-reported (6)
Open	-15.0048	-13.1548	-9.4951	-1.1966	-1.7047	-3.4078
	(5.4299)	(5.3771)	(5.5784)	(1.4676)	(1.4773)	(1.5430)
Mean (Not Opening)	-15.8534	-16.7030	-14.5292	-5.5208	-5.4095	-2.1878
Observations	2,791	2,760	2,701	2,463	2,434	2,434
Risk Group	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

NOTES. The sample considers all students who answered the questions related to the expected cutoffs or admission probabilities in the 2023 baseline and endline surveys. Robust standard errors are reported in parentheses.

D Appendix to Section 6

This appendix provides the full implementation details for the two complementary approaches used to assess the cost-effectiveness of the intervention: (i) the computation of the Marginal Value of Public Funds (MVPF), and (ii) the estimation of long-run monetary benefits using administrative labor-market data.

D.1 MVPF Calculation

Following Hendren and Sprung-Keyser (2020), the Marginal Value of Public Funds (MVPF) is defined as:

$$\text{MVPF} = \frac{\text{WTP}}{G} = \frac{\sum_i WTP_i}{G},$$

where G denotes the net fiscal cost of the intervention and WTP_i is individual i's willingness to pay for the policy.

Cost of the intervention. The intervention involves fixed costs associated with developing the risk-assessment algorithm, building the platform, and designing the communication materials. These costs total approximately \$20,000–\$50,000. Importantly, these costs do not scale with the number of students: the marginal cost of informing an additional student is effectively zero. This places the intervention squarely within the class of highly scalable information treatments emphasized by Kremer et al. (2013) and Angrist et al. (2020).

Fiscal externalities. The intervention increases enrollment in selective programs, which prior work has shown to generate substantial long-run earnings gains and corresponding increases in tax revenue (Hastings et al., 2013). Let τ denote the average tax rate and ΔY_i the predicted earnings gain for individual i. The fiscal externality is:

$$F = \sum_i \tau \cdot \Delta Y_i.$$

Because F is positive and potentially large, the net fiscal cost is:

$$G = \text{Direct Cost} - F,$$

which may be close to zero or even negative.

Implication for MVPF. When $G \leq 0$ and aggregate willingness to pay is positive, Hendren and Sprung-Keyser (2020) show that the MVPF is infinite. Our setting satisfies these conditions, implying that the intervention is not only cost-effective but potentially self-financing.

D.2 Long-Run Monetary Benefits

To quantify the long-run monetary benefits of the intervention, we leverage administrative labor-market data from the Chilean social security system (Unemployment Insurance, AFC, and Pension Funds, AFP). These data provide longitudinal records of formal-sector labor income.

Step 1: Estimating degree-program-specific returns. Using the four oldest available high school graduation cohorts (2006–2009), we estimate the following model for each horizon h:

$$Y_{i,h} = \alpha_0 + \alpha_m + X_i'\beta + \varepsilon_i,$$

where $Y_{i,h}$ is yearly income of student i at horizon h years after high school graduation, α_0 is a constant, α_m is a degree-program fixed effect, and X_i includes gender, region, and test-score strata fixed effects. The estimated $\widehat{\alpha}_m$ capture degree-program-specific earnings returns. For degree programs in the experimental sample that did not exist in the training cohorts, we impute the missing degree-program fixed effects using predictions based on broader program categories, university indicators, and historical test-score stratifications.

Step 2: Predicting long-run income for the experimental sample. For each student in the 2023 experimental cohort, we use the estimated coefficients to predict income under both their *initial* and *final* degree programs. We then define the predicted income change at horizon h as:

$$\Delta Y_{i,h} = \widehat{Y}_{i,h}^{\,\text{fin}} - \widehat{Y}_{i,h}^{\,\text{int}},$$

where $\widehat{Y}_{i,h}^{\,\text{fin}}$ and $\widehat{Y}_{i,h}^{\,\text{int}}$ denote predicted income under the student's final and initial degree-program preferences, respectively. Under this specification, the difference is driven exclusively by the difference in degree-program fixed effects across the two programs.

Step 3: Estimating the causal effect of the intervention. Using $\Delta Y_{i,h}$ as the dependent variable, we estimate the effect of being induced to open the platform via two-stage least squares. The first and second stages are:

$$\text{Open}_i = \alpha_0 + \pi\,\text{WhatsApp}_i + X_i'\delta + \alpha_{r(i)} + u_i, \tag{5}$$

$$\Delta Y_{i,h} = \beta_0 + \beta\,\widehat{\text{Open}}_i + X_i'\gamma + \alpha_{r(i)} + \varepsilon_i, \tag{6}$$

where WhatsApp_i is the text-message assignment indicator used as an instrument for Open_i, X_i includes gender, region, and test-score strata fixed effects, and $\alpha_{r(i)}$ denotes risk-group fixed effects. We focus on horizon $h = 13$, corresponding to predicted income 13 years after high school graduation.

Results. Table D.1 reports the results. For the full sample, the IV estimate implies an increase of \$86.74 in predicted yearly income. For compliers, the effect is \$1,936.7.

Table D.1: IV Estimates: Predicted Income Change: 13 Years after HS Graduation

	All (1)	Changed (2)	Assigned (3)
Open	86.74*	1936.7*	131.8**
	(47.46)	(1061.1)	(59.30)
Mean	23.69	59.7	−5.47
Observations	132,964	6,305	97,614
Risk Group	Yes	Yes	Yes

Note: Assigned is a binary variable equal to 1 if the student was initially assigned to a program, 0 otherwise. Changed is a binary variable equal to 1 if the student changed her program of assignment before and after the intervention, 0 otherwise. Covariates include stratification by gender, test scores, geographic location, and risk group. Robust standard errors are reported in parentheses. Significance reported: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Tuition Differences. A natural concern is whether the intervention induces students to switch to substantially more expensive degree programs, which could offset the predicted earnings gains. To assess this, we restrict the sample to students who were initially assigned to a degree program and estimate the effect of the intervention on the difference in annual tuition (in CLP) between their final and initial first-choice programs, using the same 2SLS strategy. Table D.2 reports the results. In column (1), which includes all initially assigned students, the estimated effect is -\$6.49 and not statistically significant. Column (2) further restricts to students who changed their first-choice program, and the estimate remains statistically insignificant. These results indicate that the intervention does not push students into substantially more expensive programs, and tuition changes do not meaningfully erode the net monetary benefits.

Interpretation. Given that the marginal cost of sending a text message is essentially zero, these long-run predicted income gains highlight the extreme cost-effectiveness of the intervention, consistent with the broader literature on scalable information treatments.

Table D.2: Regression Results: Tuition Differences (Instrumental Variables)

	All	Changed
	(1)	(2)
Open	-6.49	-85.67
	(14.63)	(286.56)
Mean (No Text)	-1.86	-30.85
Observations	92,254	5,733
Strata FE	Yes	Yes

Note: 2SLS estimates of the effect of opening the intervention on the difference in annual tuition (2024 USD) between the student's final and initial programs. The sample is restricted to students who (i) were initially assigned to a program, (ii) were assigned to a program at the end of the process, and (iii) have non-missing tuition data for both programs. The instrument is random assignment to receive the WhatsApp encouragement message. Column (1) includes all students meeting these criteria; column (2) further restricts the sample to students whose initial and final programs differ. Covariates include stratification by gender, test scores, geographic location, and risk group. Robust standard errors are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.