

# Smart Matching Platforms and Heterogeneous Beliefs in Centralized School Choice<sup>†</sup>

Felipe Arteaga, *UC Berkeley*

Adam J. Kapor, *Princeton University*

Christopher A. Neilson, *Princeton University*

Seth D. Zimmerman, *Yale SOM*

June 12, 2021

## Abstract

Many school districts with centralized school choice adopt strategyproof assignment mechanisms to relieve applicants of the need to strategize on the basis of beliefs about their own admissions chances. This paper shows that beliefs about admissions chances shape choice outcomes even when the assignment mechanism is strategyproof by influencing the way applicants search for schools, and that “smart matching platforms” that provide live feedback on admissions chances help applicants search more effectively. Motivated by a model in which applicants engage in costly search for schools and over-optimism can lead to under-search, we use data from a large-scale survey of choice participants in Chile to show that learning about schools is hard, that beliefs about admissions chances guide the decision to stop searching, and that applicants systematically underestimate non-placement risk. We then use RCT and RD research designs to evaluate live feedback policies in the Chilean and New Haven choice systems. 22% of applicants submitting applications where risks of non-placement are high respond to warnings by adding schools to their lists, reducing non-placement risk by 58%. These results replicate across settings and over time. Reducing the strategic burden of school choice requires not just strategyproofness inside the centralized system, but also choice supports for the strategic decisions that inevitably remain outside of it.

---

<sup>†</sup>We thank Joseph Altonji, Francisco Gallego, Paul Goldsmith-Pinkham, John Eric Humphries, Karthik Muralidharan, Nick Ryan, and seminar participants at NYU Economics/Stern, UCSD, Yale, the 2021 AEA meetings, and the NBER Economics of Education workshop for comments and suggestions. We thank Isabel Jacas, Jan Knuf, Manuel Martinez, Cecilia Moreira, and Fernando Ochoa for research assistance. We thank [JPAL-LAC](#) and the implementation team of data scientists and developers at [ConsiliumBots](#) for their help throughout the project. We thank Claudia Allende for her support in survey design and implementation. The authors wish to thank the government partners that made this research possible, in particular the leadership at the *Sistema de Admision Escolar* ([SAE](#)) of the Ministry of Education in Chile and the Office of School Choice and Enrollment at New Haven Public Schools ([NHPS](#)). [Online Appendix](#) and [disclosure statements](#) available at authors’ websites. Arteaga: [fharteaga@berkeley.edu](mailto:fharteaga@berkeley.edu). Kapor: [akapor@princeton.edu](mailto:akapor@princeton.edu). Neilson: [cneilson@princeton.edu](mailto:cneilson@princeton.edu). Zimmerman: [seth.zimmerman@yale.edu](mailto:seth.zimmerman@yale.edu).

# 1 Introduction

Many school systems around the world use centralized mechanisms to assign students to schools. An important contribution economists have made to the design of centralized school choice is to guide policymakers towards centralized mechanisms that are strategically simple for participants to use. In cities including New York, Boston, New Haven, and Santiago, economists have helped design “strategyproof” choice systems where applicants’ dominant strategy is to list schools they like, in the order that they like them.<sup>1</sup> A central point in the case for strategyproof approaches is that knowledge of admissions chances—which may be costly to acquire or unequally distributed—is not required for optimal play.

The conclusion that strategyproof centralized mechanisms relieve choice participants of the need to know about their admissions chances follows from the maintained assumptions of the canonical “school choice problem” ([Abdulkadiroğlu and Sönmez, 2003](#)) that applicants know which schools are available to them and which they like. But what if learning about schools is costly, and families do not know about all of their options? This paper examines how costly search interacts with beliefs about admissions chances to shape what families know about schools, how much this matters for choice outcomes, and what policymakers can do about it. We take an empirical approach, drawing on large-scale surveys and policy variation in the Chilean and New Haven school choice systems.

We make two main points. The first point is that costly search for schools is central to the way families experience choice, and that this places beliefs about assignment chances back in a key role even when the assignment mechanism is strategyproof. Many participants stop searching for schools because they think they will be admitted to a school already on their application. Systematic over-optimism about admissions chances leads participants to submit applications with high non-placement risk.

The second point is that a new kind of intervention, which we call a “smart matching platform,” can help families navigate costly search more effectively. Smart matching platforms aggregate data on the distribution of choice applications to provide live feedback on admissions chances to platform users. We use experimental and quasi-experimental research designs to test smart platforms at scale, and find that they change application behavior and raise placement rates, with consistent effects across years, cities, and countries. We conclude that reducing the strategic burden of choice requires not just strategyproofness inside the centralized system, but also support during the search process that precedes it.

We begin by developing a simple model of school search in a strategyproof assignment mechanism. We draw on models of job search such as [McCall \(1970\)](#), with the key difference being that individuals add schools they find to an application portfolio, rather than making one-time decisions to accept or decline an offer. Applicants engage in costly, sequential search

---

<sup>1</sup>[Abdulkadiroğlu et al. \(2005a,b\)](#); [Abdulkadiroğlu et al. \(2006\)](#); [Correa et al. \(2019\)](#); [Akbarpour et al. \(2020\)](#).

for schools to add to their choice application. Once applicants decide to stop searching, they submit the application to a strategyproof assignment mechanism.

The key insight of the model is that the value of search depends on how likely the applicant thinks she is to be placed in the school she finds. We use the model to derive three results. First, over-optimism about admissions chances can reduce search and increase the risk of non-placement. Second, information interventions implemented after individuals have decided to stop search weakly raise the probability that applicants search for and find additional schools to add to their applications. Applicants who respond to the intervention by conducting additional search are “compliers” with the intervention policy (Angrist et al., 1996). Third, we show that if the choice to enroll in the placed school reflects a preference for that school relative to the outside option, then the individual welfare gains from information interventions rise in proportion to placement rates, except to the extent they are offset by declines in enrollment conditional on placement.

The theory of costly school search has testable predictions. People should not know about all the available schools. People should report that the activities involved in search are challenging, and that one reason they stopped searching is that they thought they would be placed. And, if in addition people tend to be over-optimistic, they should respond to information about admissions chances by searching more and adding schools to their applications.

We test these predictions using data from two school choice systems. The first is the national centralized choice system in Chile. Chile implemented centralized primary and secondary school choice in 2016. All cities in Chile use the same choice platform to implement a strategyproof deferred acceptance (DA) assignment mechanism. We use data from the entire system for the years 2018–2020. Our second setting is the school choice system in New Haven, Connecticut in 2020. New Haven uses a “truncated” DA mechanism in which applicants can list only a limited number of schools. Truncated DA mechanisms are not strategyproof, but they are less manipulable than the common alternative of Immediate Acceptance (Haeringer and Klijn, 2009; Pathak and Sönmez, 2013). Studying the Chilean and New Haven settings together allows us to consider the role of search under different implementations of DA, within different choice platforms, and in different cultural contexts.

We supplement our administrative records with extensive survey data on choice participants in Chile. As part of the 2020 Chilean choice process, the Chilean government surveyed families submitting applications to the choice process about their search for schools, their preferences over schools, and their beliefs about their placement chances. These surveys were administered online, after the submission of applications but before results were known. 48,929 applicants completed the choice survey. The combination of a very large sample size and novel questions about both the choice application and the search process allow us to construct a detailed picture of the way families navigate choice.

Survey findings provide strong evidence that strategic, costly search for schools is one

of the central challenges applicants face in the choice process, and that our stylized model captures important elements of the way students use potentially inaccurate beliefs to build their application portfolios. We have four main survey findings.

Our first survey finding is that search is, in fact, costly, and that applicants have limited information about relevant schools. When asked about what steps they need to take to know a school, large majorities of respondents give a long list of attributes and activities, including academic performance, extracurriculars, and interviews with staff. Obtaining this information would typically require both internet research and in-person visits or phone calls. Only 17% of respondents report that they know a randomly-chosen nearby school well, compared to 73% who report knowing their first choice well.

Our second survey finding is that the choice to terminate search is a strategic one to which beliefs about admissions chances are an important input. When we ask applicants why they did not add more schools to their list, the modal response is that they think they will be placed at one of the schools on the list already. 35% of respondents give this answer, compared to 30% who say they stopped adding schools because there are no more schools around. Applicants reporting higher subjective placement probabilities are much more likely to say they stopped search because they thought they would be placed.

Our third survey finding is that applicants are over-optimistic about their admissions chances. On average, respondents submitting applications with non-zero risk of non-placement believe they have a 75% chance of being placed at a school on their application, 30 percentage points higher than the true mean of 45%. Applicants with true placement chances close to zero report average subjective placement beliefs of nearly 70%. Beliefs matter for search, but they are often wrong.

Our fourth survey finding is that the welfare stakes are large. Only 12% of applicants report that they would be at least somewhat satisfied with an outcome of no placement, compared to 69% who report they would be satisfied with the last-ranked school on their application. Differences in satisfaction manifest in enrollment choices. 99% of applicants placed at a school where they say they would be very satisfied go on to enroll in the school, compared to 65% of students placed at schools where they say they would be unsatisfied.

The survey results suggest that access to information about admissions chances would be helpful to applicants searching for schools. However, providing this information presents a logistical challenge. Placement chances are attributes of applications, not schools, and depend not just on individual submissions but on the distribution of submissions in the market.

We develop a new approach to address these challenges, which we call a “smart matching platform.” The smart matching platform aggregates back-end data on the distribution of submitted applications to produce live, personalized predictions about application risk for platform users on the front end. This approach combines several features that past research has shown to be critical for successful information interventions, including the reduction of

computational burdens, timely provision, and provision from a trusted source (Mani et al., 2013; Fernandes et al., 2014; Fischer and Wagner, 2018; Patterson et al., 2019).

We evaluate smart matching platforms in Chile and New Haven. In both cases, the platform warned applicants submitting applications with high non-placement risk. In Chile, these warnings consisted of a pop-up in the application platform, as well as off-platform text messages. In New Haven, warnings came via email and directed applicants to an application simulator, which they could use to view placement chances for hypothetical applications. To assess risk in advance of application deadlines, policymakers combined data from previous years with data on applications already submitted in the current cycle. These policies were implemented nationwide in Chile starting in 2017, and in New Haven starting in 2020.

Because choice administrators need to choose some cutoff for what makes an application “risky,” risk warnings lend themselves naturally to a regression discontinuity research design. In the face of quantity limits on messaging, choice administrators in Chile also randomized the provision of off-platform messages on the intensive margin. That is, all risky applicants received a text message but some received an additional, earlier message including an image. This allows us to employ RCT research designs as well. These experimental and quasi-experimental approaches allow us to evaluate our theoretical model without restricting access to information or reducing policy efficacy.

Warning applicants about their risky applications leads to lengthened applications, reduced application risk, increased placement rates, and increased rates of enrollment in the placed school. We focus first on Chile, where sample sizes are much larger. Policymakers designated all applications with at least a 30% predicted chance of non-placement as risky. All applications above that cutoff received the live notification on the choice platform. Receiving a warning caused 21.6% of students ( $SE=1.0$ ) to add schools to their applications, corresponding to the complier group in our model. These students reduced their non-placement risk by an average of 15.5 percentage points ( $SE=1.3$ ), or 58% of average ex post application risk at the cutoff. Consistent with model predictions, essentially all of the application changes we observe in response to treatment are additions of new schools.

Results from the text-message RCT show that these findings are not local to the cutoff. 4.4% of students randomly chosen to receive an additional warning message (and 9.4% of students who viewed the message) add at least one school to their application. Treatment compliers reduced their average non-placement risk by 29.7 percentage points, 49% of the mean in the risky population. Because the control group in this experiment also received warning messages, this finding also shows that warnings matter on the intensive margin.

We next turn to welfare. Our theoretical analysis suggests that, from the perspective of an individual, the welfare effects of the intervention are proportional to gains in placement except as offset by declines in enrollment conditional on placement. Applicants receiving the intervention are no less likely to enroll in their placed schools, suggesting that placement

quality does not decline. Most of the additional placements generated by the treatment are at schools with slack capacity, suggesting that the intervention reduces market-level congestion.

Smart matching platforms are effective in a wide variety of choice settings. Splitting the Chilean data by city and year shows that our findings are not driven by particular market-year combinations. Treatment effects do not vary with market-level school choice experience, suggesting a limited role for “learning-by-doing” about admissions chances. We observe treatment effects in markets of all sizes, but applicants with more schooling options nearby tend to add more schools to their applications. Effects are large for both high- and low-SES students.

We find similar results in an analysis of the NGO’s intervention in New Haven, Connecticut. While the broad structure of the New Haven intervention paralleled the approach in Chile, the cultural context, choice institutions, and intervention details differed substantially. Despite these differences, the intervention has similar effects. 11.4% of applicants near the risk cutoff comply with the intervention policy by lengthening their application; these applicants reduce their application risk by 45%.

An additional intervention conducted in New Haven provides evidence that the “smart” component of smart matching platforms is important for their efficacy. Choice administrators randomly assigned a subset of applicants to a “static” intervention that paralleled the smart intervention but provided a suggestion to think about admissions chances rather than a personalized risk assessment. The static intervention had no effect on choice behavior.

We contribute to three strands of literature. Our first contribution is to show that strategyproofness within the school choice problem does not correspond to strategyproofness in the broader choice *process*, and that the divergence between the two places substantial information demands on participants. Many papers consider how students make choices under different assignment mechanisms ([Abdulkadiroğlu et al., 2011](#); [Pathak and Sönmez, 2013](#); [De Haan et al., 2015](#); [Agarwal and Somaini, 2018](#); [Calsamiglia et al., 2020](#); [Kapor et al., 2020](#); [Akbarpour et al., 2020](#)). These papers analyze behavior in the choice problem, and typically ignore deviations from optimal behavior in strategyproof settings. We show that these deviations are empirically important and provide an economic rationale for why they occur.

An emerging literature considers the search aspect of school choice directly. [Immorlica et al. \(2020\)](#) and [Hakimov et al. \(2021\)](#) use theoretical and laboratory approaches to study the equilibrium implications of costly (but rational) search in matching markets, with a focus on education markets. [Ajayi and Sidibe \(2020\)](#) and [Son \(2020\)](#) use application data from centralized choice systems to estimate empirical models that allow for limited consideration sets and belief errors. Our empirical contributions here are to provide survey evidence that the frictions these papers build into their models are important in practice, to provide experimental and quasi-experimental tests of model predictions that shocks to beliefs affect search, and to demonstrate that smart matching platforms are an effective and generalizable policy response. From the theory side, our contribution is to unpack the way systematic

belief errors affect search from the perspective of the individual applicant. Our work fits into a broader set of studies that consider how strategic actions taken prior to participation in centralized mechanisms affect assignments within the mechanism; for example in spectrum auctions (Doraszelski et al., 2017; Milgrom and Segal, 2020).

Our second contribution is to illustrate the importance of information interventions that target search *strategy*, as opposed to fixed product attributes. In both education and product markets there is a large literature exploring the effect of providing consumers with information on the product or school.<sup>2</sup> Findings in this literature are mixed, with some interventions changing choices, and others finding precise zeros. Our intervention, which provides information about *endogenous* admissions chances, is conceptually quite different. However, our findings can help rationalize null results in some attribute-focused studies. If applicants are confident they will be admitted to a school they like, they may not think it is worth it to conduct the additional due diligence required to add a new option to their portfolio, even when prompted with appealing (but limited) information about that option. Another important distinction between interventions on strategies and interventions on attributes such as school quality is that interventions on strategy do not guide people towards a small set of high-performing schools, and may thus reduce congestion. We observe this in our setting.

Our third contribution is to show the power of combining market design principles, which limit the need for strategic sophistication, with “prediction machines” (Agrawal et al., 2018) and “choice engines” (Thaler and Tucker, 2013), which distill complex datasets into the information people need to make the strategic decisions that remain. We bring narrow AI into a matching setting where it aggregates information on market-level outcomes and identifies the part of that information relevant for specific participants. This contrasts with previous work focusing on attribute comparisons in product markets (Gruber et al., 2020).

## 2 Searching for schools

### 2.1 Model overview

We guide our empirical analysis using a model of search for schools with imperfect information about admissions chances. The theoretical analysis has three goals. The first is to highlight how beliefs about admissions chances affect students’ decisions to search for schools to add to their applications. The second is to show how optimism about placement chances affects beliefs about the returns to search, and how interventions that reduce optimism about placement can cause students to increase search and reduce application risk. The third is to show

---

<sup>2</sup>Education markets: Hastings and Weinstein (2008); Jensen (2010); Bettinger et al. (2012); Hoxby and Turner (2013); Hastings et al. (2015); Corcoran et al. (2018); Ainsworth et al. (2020); Neilson et al. (2019); Bergman et al. (2019); Dynarski et al. (2019); Gurantz et al. (2020); Hyman (2020). Product markets: Jin and Leslie (2003); Allcott and Taubinsky (2015); Barahona et al. (2020).

how changes in the placement rates and rates of enrollment conditional on placement provide information about the individual welfare effects of policies that reduce non-placement risk.

Our analysis takes the perspective of an individual student searching for schools to add to her school choice application. The approach is similar to models of job search (McCall, 1970), with the key difference being that agents in our model add schools they find to a multi-school application portfolio, from which placement outcomes are determined by a centralized assignment mechanism. This contrasts with the standard approach to job search models, in which agents decide whether to take jobs as they arrive, and search terminates once the agent accepts an offer. It also contrasts with models of the school choice problem that focus on market equilibria as the main outcomes of interest. Our model highlights the strategic challenges facing individuals even when the centralized assignment mechanism is strategyproof, and allows us to draw out the role of beliefs about admissions chances.

## 2.2 Model setup

Consider an applicant to a strategyproof centralized assignment mechanism with limited information about what schools are available to her. The applicant is endowed at time zero with consideration set  $C_0 \equiv \{1, 2, 3, \dots, K\} \subseteq \mathcal{J}$ , where  $\mathcal{J}$  is the set of all schools. The applicant receives utility  $u_j$  from placement at school  $j$ . Without loss of generality suppose  $u_1 > u_2 > \dots > u_K > 0$ , and that utilities are measured relative to the outside option of non-placement, which yields utility zero. For each  $j \in C_0$ , the individual knows their utility from placement at  $j$ ,  $u_j \in \mathbb{R}$ , and has subjective beliefs about admissions chances  $p_j \in [0, 1]$ , which they believe to be independent across  $j$ .<sup>3</sup> Define  $R_j = 1 - p_j$  as the risk of non-placement at school  $j$ .

The value of the optimal portfolio given consideration set  $C_0$  is given by:

$$V_0 = p_1 u_1 + R_1 p_2 u_2 + \dots + \prod_{j < K} R_j p_K u_K \quad (1)$$

Individuals have accurate beliefs about the distribution of utilities at schools outside their consideration set,  $F_u(u)$ , and potentially inaccurate beliefs about the distribution of admissions chances  $F_p(p|u)$  that may depend on their value of being placed at the school. Individuals may choose to pay a cost  $C$ , known to them, to add a school to their consideration set.<sup>4</sup> These

<sup>3</sup>In the empirical settings that we consider, admissions outcomes are determined by lotteries which are independent across schools. In principle, additional uncertainty about the general number of seats or level of demand might induce correlation in beliefs within a portfolio. For instance, rejection by school  $j$  might indicate that demand for some other school  $k$  was higher than the student had believed. In practice, school choice applicants seem to exhibit “correlation neglect” (Rees-Jones et al., 2020).

<sup>4</sup>Our assumption allows for uncertainty about the amount of effort needed to discover a school. For example, students may pay a flow cost  $c$  in order to discover an additional school with instantaneous probability  $\lambda$ . In this case,  $C$  would denote the expected cost of searching until one additional school is found.



search costs differ across individuals and are distributed according to  $\Phi(C)$ , which we assume is differentiable with pdf  $\phi$ .

This setup captures the idea that students need to know what a school is like before they apply to it. We think of the search cost  $C$  as reflecting the cost of achieving this level of familiarity. We present evidence in section 4 that these costs are fairly large. We assume that students have accurate beliefs about the underlying distribution of utilities. This assumption lets us focus on the novel aspect of our contribution, which is to analyze the effects of erroneous beliefs about admissions chances, even if other aspects of schools can be observed without error.

### 2.3 The value of learning about a school

Suppose that the person learns about an additional school,  $s$ , with utility  $u_s$  and “chance”  $p_s$ . Suppose further that  $u_{r-1} > u_s > u_r$ , for some  $r$ . Letting  $V$  denote the value of the optimal application that includes  $s$ , we have

$$V - V_0 = \prod_{j < r} R_j (p_s(u_s - \Gamma_r)), \quad (2)$$

where

$$\Gamma_r = \sum_{k=r}^K \prod_{r < j < k} R_j p_k u_k.$$

$\Gamma_r$  is the expected value of the application portfolio conditional on being placed at schools ranked above  $r$ .

### 2.4 Optimism and the value of finding a school

We assume a simple, multiplicative structure for belief errors. Let  $R_j = (1 - a)R_j^*$  for all  $r_j$ , where  $R_j^*$  is the true risk. Similarly, let  $p_j^* = 1 - R_j^*$  denote the true chance of being admitted to  $j$ . The parameter  $a$  is then a measure of optimism; as  $a$  grows, people believe risk is smaller. Assume  $a < 1$  so that people do not rule out all application risk. Taking the log of  $V - V_0$  and the taking the derivative with respect to  $a$  yields<sup>5</sup>

$$\frac{d \log(V - V_0)}{da} = \frac{1 - r}{1 - a} + \frac{R_s^*}{1 - R_s^*(1 - a)} + \frac{d\Gamma_r}{da} \frac{1}{\Gamma_r - u_s}.$$

The effect of optimism on the value of adding new schools operates through three channels, corresponding to each term in the sum. The first channel is that more optimism reduces the value of adding school  $s$  by reducing applicants’ subjective beliefs that they will fall all the way to rank  $s$ . This is the first term in the sum. It is equal to zero if  $r = 1$  (i.e., if added

---

<sup>5</sup>See Appendix A for details.

school  $s$  is first-ranked on the new application) and negative for  $r > 1$ . It will tend to be bigger as optimism grows.

Second, increased optimism raises the value of adding a school to the portfolio because applicants think they are more likely to be admitted to that school. The second term of the sum captures this effect. It is positive for all values of  $a$ .

Third, increasing optimism reduces the value of adding school  $s$  by raising the expected value of falling below  $s$  on the application. The third term of the sum is negative whenever  $s$  is not the last school on the application, in which case it is equal to zero.  $\frac{d\Gamma_r}{da} > 0$ , because optimism shifts students towards believing they will be placed at higher-ranked schools given that they have fallen below  $s$ .  $\frac{1}{\Gamma_r - u_s} < 0$  because the value of a placement at  $s$  is always larger than the expected value of placement at schools with lower placement utility than  $s$ .

These channels combine to affect the subjective value of adding school  $s$  to the application.

**Proposition 1.** *Assume  $s$  is the lowest ranked school on an application with at least one higher-ranked school. Then  $1 - \frac{r-1}{rR_s^*} < 1$ , and the value of adding  $s$  to the application is decreasing in  $a$  whenever  $a > 1 - \frac{r-1}{rR_s^*}$ .*

This proposition shows that for sufficiently high levels of baseline optimism, additional increases in optimism reduce the value of adding schools to the bottom of the application. As we discuss below, this case—optimistic students adding schools to the bottom of their applications—is the modal one in our setting. More broadly, this analysis shows that information on admissions chances can be important to choice strategy even if it does not affect the applications students submit given their consideration set.

## 2.5 Information interventions and search behavior

The expected value of search  $U[\text{Search}|C_0, a]$  is given by integrating the value of adding a given school  $s$  over the distribution of utilities and subjective admissions chances:

$$U[\text{Search}|C_0, a] = \int \int_{u_0} (V(u, p) - V_0) dF(u, p)$$

where  $V(u, p)$  is the expected utility from submitting the optimal portfolio that includes a school with utility  $u$  and admissions chance  $p$  in addition to schools in  $C_0$ . At the optimal strategy given applicants' beliefs, we have  $U[\text{Search}|C_0, a] \leq C$ ; otherwise applicants would not have stopped searching.

Taking this optimal portfolio as a starting point, consider how a small change in optimism,  $\Delta a$ , alters search behavior. Individuals for whom this change reduces the value of search cannot “unsearch,” so their search behavior does not change. Individuals for whom changing optimism raises the value of search, such as those identified in Proposition 1, increase search if their decision to stop was marginal. The costs of search are sunk once incurred, so an

applicant induced to search will continue to do so until he finds at least one school to add to his application.

**Proposition 2.** *Once applicants have completed search, the effect of changing optimism by  $\Delta a$  is to weakly raise the probability of search, and to raise the probability of adding at least one school to the choice application by an equal amount.*

*Proof.* See Appendix A. □

Applicants who engage in additional search and add at least one school to their application in response to the information treatment  $\Delta a$  are *compliers* with the intervention policy. This observation is important because we observe application responses to information interventions in our data, but do not directly observe search responses.

Adding schools to the application reduces non-placement risk. We can bound the expected reduction in true non-placement risk from above as the expected change from adding one school. Define non-placement risk prior to the change in  $a$  as  $RISK_0 = \prod_{j < K} R_j^*$ . Then, the change in placement risk after adding a given school  $s$  to the application is

$$RISK - RISK_0 = R_s^* \times \prod_{j < K} R_j^* - \prod_{j < K} R_j^* = -RISK_0 \times p_s^*.$$

Integrating over schools  $s$  that an individual may add to his application, and using Proposition 2's result that search does not terminate after an unacceptable school is discovered, we have

$$\frac{dRISK}{da} \leq \frac{dPr(Search)}{da} \times RISK_0 \times E[p^* | u > 0].$$

In sum, we expect information interventions implemented at the conclusion of search to raise search rates, cause individuals to lengthen their applications, and reduce non-placement risk.

## 2.6 Enrollment choices and the welfare effects of information

We next consider how to use objects we observe in the data to assess the individual welfare effects of changes in optimism. We focus on changes in welfare accrued through the placement process; i.e. excluding search costs. The key insight here is that an applicant's decision to enroll in the school in which they are placed is a measure of how much they prefer that school to the outside option.

We model enrollment as a binary choice between the school where an individual is placed and the outside option. Timing is as follows. At the time of application, individuals observe school- (and person-) specific utilities  $\mu_j$ , with outside the option normalized to zero. Following placement, they receive enrollment shocks  $\epsilon_j$ , iid across schools. Students choose to enroll

in the placed school  $j$  according to the rule

$$Enroll = 1[\mu_j + \epsilon_j > 0].$$

The utilities  $u_j$  defined above capture the expected value of placement at the time of application, so that  $u_j = E[\max(\mu_j + \epsilon_j, 0)]$ . Assume the  $\epsilon_j$  have distribution  $G(\epsilon)$ , which is differentiable with density function  $g$  and has an inverse that is also differentiable.

This framework lets us break down the effect of a change information on  $V$ , the utility value of the application a student submits, into two channels: a placement channel, and a utility conditional on placement channel.

**Proposition 3.** *The individual utility gains from a change in optimism  $\Delta a$  are equal to the sum of two terms:*

- *a term that is proportional to the change in placement chances; and*
- *a term that is, to a first-order approximation, a weighted sum of the resulting changes in enrollment probabilities, with all weights positive.*

*Proof.* See Appendix A. □

Individual utility increases in proportion to placement rate, except to the extent it is offset by declines in enrollment conditional on placement. We use this observation to guide our assessment of welfare effects.

## 3 Setting

### 3.1 Centralized choice in Chile

We study the importance of costly search using nationwide survey and administrative data from Chile and district-level data from New Haven, Connecticut. We focus first on Chile, where sample sizes are several orders of magnitude larger. This section describes school choice institutions in Chile and interactions between policymakers and choice applicants that help us understand the role of search. We return to the New Haven setting in section 5.7.

Chile introduced nationwide, voucher-based school choice in 1981 (Hsieh and Urquiola, 2006). For the first 35 years, school choice in Chile was *decentralized*. Families applied to each school separately. In 2016, policymakers adopted centralized assignment with the goal of making the school choice process more transparent and equitable (Ministerio de Educación, 2017). The centralized choice system was rolled out on a region-by-region basis, with adoption in all cities by 2019 and all grades by 2020. The centralized process includes 93% of primary

school matriculation in the country, covering almost all public schools and private schools that accept school vouchers.<sup>6</sup> 450,000 applicants participated in 2020.

All cities in Chile use the same choice platform, which assigns students to schools using a deferred acceptance (DA) assignment mechanism (Correa et al., 2019). To ration seats in oversubscribed schools, the mechanism combines coarse sibling, school employee, and alumni priorities with lottery-based tiebreakers.<sup>7</sup> Applicants may list as many schools as they want on their choice application.<sup>8</sup> This means that the mechanism is *strategyproof*. The approach Chile takes to centralized assignment is similar to that used in major US districts such as New York and Boston (Abdulkadiroğlu et al., 2005b,a).

The centralized school choice platform opens in August of each year. Applicants have access to the platform for roughly one month, during which time they may view, submit, and edit their applications. The application deadline falls in early September, and students are notified of their placements (or non-placement) in late October. Applicants who receive a placement can choose to turn down that placement if they want. Applicants who reject their placement, who are not placed, or who did not participate in the main round can join a secondary application process in late November that lasts one week. Between early January and the beginning of the school year in March, students who still do not have a placement and placed students who decide they do not want to accept their placement may enroll in undersubscribed schools, outside of the centralized system. We focus our analysis on the first placement round, which accounts for more than 90% of placements over the period we study. See Online Appendix C for further discussion of school choice institutions and enrollment outcomes for unplaced students.

We analyze the choice process using data on all applicants to the centralized platform between 2018 and 2020. We describe the applicant population in Table 1.<sup>9</sup> The platform received just under 1.2 million applications (defined at the student-year level) over this period. 49% of these applications came from students identified by the Chilean Ministry of Education (Mineduc) as “economically vulnerable,” a classification based primarily on income and benefits receipt. 95% of applicants come from urban areas, as defined by the 2017 Census.<sup>10</sup>

Many applicants interact more than once with the application platform between the time it opens and the application deadline. Panels B and C of Table 1 describe these interactions.

---

<sup>6</sup>The remaining 7% of PK-12 students are enrolled in expensive private schools that do not accept government vouchers or in schools where the highest grade offered is Kindergarten. These schools do not participate in the centralized process.

<sup>7</sup>Some schools also use quotas for vulnerable students and, in a very small number of cases, for high-performing students.

<sup>8</sup>Students applying in zones with more than one option who are either entering the schooling system from outside or enrolled in a school that does not offer the next grade must list a *minimum* of two schools.

<sup>9</sup>See Online Appendix D for a discussion of our data sources.

<sup>10</sup>The Census definition of urban areas includes (primarily) all settlements with more than 2000 inhabitants. We define applicants’ geographic zone based on the location of their first-choice school. Individual geocoding is unreliable for a large portion of applicants, while school locations are known precisely.

The first portfolio an applicant submits contains an average of 2.8 schools. Following their initial submission, applicants are free to revisit their submission and change, add, or subtract schools at any time before the deadline. At the deadline, the average portfolio length rises to 3.1 schools. The average applicant submits 1.4 distinct portfolios to the centralized platform before the deadline. 25% of applicants submit a final application that differs from their initial application. The most common change is to add a new school to the application: 21% of applicants have a school on their final application that was not on their initial application. Most people who add schools add them to the bottom of their portfolio— 18% make such an addition— but 3% add a new school to the middle of their application (i.e., above some but not all previously-ranked schools) and 2% add a school to the top (above all previously-ranked schools). We observe a variety of other changes as well. 5% of students change their top-ranked school, and 5% delete a school. Columns 2 and 3 of Table 1 show that lower-income students tend to have shorter applications and are less likely to change their applications.

Most but not all students receive a placement through the centralized process. As reported in Panel D of Table 1, 79% of applicants receive a placement at some school on their first-round application. 54% of students are placed in their first-ranked school, 13% in their second, and 6% in their third. 5% of students place at a school lower than third. Placement rates are *higher* for lower-income students despite their shorter applications. 84% of low-income students receive a placement, compared to 74% of higher-income students. As reported in Panel E of Table 1, 9% of students who participate in the first round go on to participate in the second centralized round, and 7% receive a second-round placement.<sup>11</sup>

Non-placement occurs despite slack capacity. Panel F of Table 1 displays the (applicant-weighted) average share of seats in a market that are unfilled after the first placement round. On average, participants apply in markets where 42% of seats are unfilled; the share of unfilled seats in schools that are free to students is even higher. These values far exceed the share of students placed in the second placement round, indicating that follow-on attempts to fill slack capacity do not fully succeed.

### 3.2 Intervention design

Heading into the 2017 process, non-placement risk was a major concern for education policymakers in Chile. Our research team worked closely with Mineduc to evaluate the causes of non-placement risk and formulate a policy response. Preliminary descriptive and qualitative evidence suggested that some families had inaccurate, overly optimistic beliefs about their chances of being assigned to schools. Based on this evidence, we helped Mineduc design a set of information interventions alerting applicants to non-placement risk. These interventions identified applicants whose submissions placed them at high risk of non-placement, and

---

<sup>11</sup> Applicants who do not participate in the first round are not included in our analysis.

notified them of this risk prior to the close of the application deadline.

The key feature enabling these interventions is the ability to interact with both application data and applicants themselves in real time over the course of the application process, to compute and communicate risk. The technical and logistical demands of implementing live feedback at scale led one member of the research team (Neilson) to found an NGO, ConsiliumBots, specializing in school choice services. The NGO partnered with Mineduc to run the interventions from 2018 on. See our Disclosure Statement for details on the relationship between the research team and the implementing partner NGO.

The remainder of this section summarizes the design of the interventions we study. See Online Appendix C for additional detail.

Mineduc conducted two kinds of information interventions over the period we study. The first was an interactive pop-up embedded in the application platform, which we label the *platform pop-up*. This intervention computed a predicted risk value for each application submitted through the platform. Applications identified as “risky”—defined as having a non-placement risk greater than 30%—received a pop-up warning about their application immediately after they clicked submit. The warning stated that many families were applying to the same schools, and not enough seats were available for all applicants. It encouraged students to add more schools to their applications, while also offering them the option to continue and submit the application as-is. Figure 1 displays the pop-up, with key text translated to English.

Mineduc implemented this intervention throughout the choice system. In 2018 and 2019, Mineduc activated the pop-up functionality one to two days after the date that applications opened. This delay reflected a combination of implementation difficulties and a desire to collect data on early applications for use in demand predictions. Our empirical analysis of pop-up effects in 2018 and 2019 excludes the students who submitted their first application attempt before the pop-up came online. These students made up 39% of all applicants in these years. In 2020, the pop-up was available over the full application window for most of the applicants.<sup>12</sup>

Column 4 of Table 1 describes the 73% of applicants who submitted applications at times and in markets where the pop-up was available. We label this group “pop-up eligible” because members received a warning if their application was deemed risky. Pop-up eligible applicants resemble the full population in their demographic characteristics and application behavior.

The second kind of intervention consisted of supplemental “reminders” to risky students. These reminders were delivered via text message or the messaging service WhatsApp, and contained information similar to the pop-up.

Our analysis of reminder interventions focuses on the 2020 application cycle, when Mineduc

---

<sup>12</sup>Demand predictions for early applicants in 2020 relied on data from the previous year. We did not have previous year demand data for students applying to non-entry grades in the Metropolitan Region, hence the pop-up was activated later for them (9% of total 2020 applicants). See below and Online Appendix F.

sent a sequence of up to three messages to applicants who submitted risky applications. Figure 2 outlines the time path of interactions with risky applicants in 2020, and presents images of each intervention. As in previous years, these interactions began with the pop-up intervention on students’ initial application submission. All applicants who had submitted risky applications as of day 20 of the application cycle received a text message from Mineduc. Mineduc sent another text message to risky applicants on day 27 (the day before application close) repeating this information and providing a link to the student’s choice application.

On day 25 of the application cycle, between the two text messages to all risky applicants, Mineduc and the NGO conducted an RCT evaluation of a WhatsApp intervention. We call this the *WhatsApp RCT*. The NGO chose a random subset of ten thousand risky applicants and sent them a WhatsApp message that listed their current choice application and displayed information about the number of seats available and the number of students applying to each school on that application.<sup>13</sup> It stated that the risk of non-placement was high, and suggested that students add schools to their applications to address this risk. Two factors motivated the RCT evaluation. The first was the idea that visual interventions sent through the popular messaging service might be an effective supplement to the other intervention approaches. The second was a constraint placed by the contractor that managed WhatsApp messaging, which capped the number of messages that could be sent.

The set of reminders implemented in 2020 built on a more limited reminder policy implemented in 2018. In 2018, Mineduc sent a single SMS message to all risky students four days before the application deadline. Mineduc did not send any reminder messages in 2019.

We evaluate the platform pop-up using a regression discontinuity design around the 30% risk cutoff. In 2019, the RD estimates capture the effect of the pop-up for applicants near the cutoff. In 2018 and 2020, the RD estimates capture the effect the pop-up and its interaction with the subsequent reminder interventions. Our goal in the RD analysis is to provide proof of concept that smart platform information interventions affect search behavior and placement outcomes, not to unpack the differential effects of interventions by medium and timing. In what follows, we present RD estimates separately by year. Readers who are interested in understanding the effects of pop-up absent their interactions with subsequent reminders can focus on the 2019 implementation year.<sup>14</sup>

We evaluate the WhatsApp reminder in a standard RCT framework. Because treatment and control in the WhatsApp RCT are drawn from the set of students who still have risky applications after receiving previous reminders, the RCT evaluation tells us about intensive margin treatment effects in a group that is negatively selected on its response to previous

---

<sup>13</sup>In addition to high application risk, the NGO imposed other restrictions on the sample universe for RCT randomization. To be RCT-eligible, applicants needed to be a) early-grade applicants in b) urban zones who c) did not have access to sibling priority. In addition, they d) had to have declined engagement with previous Mineduc outreach attempts (unrelated to application risk) sent via email. See Online Appendix C.3.

<sup>14</sup>See Online Appendix E for a detailed discussion of interactions between treatment types.



similar treatments. It also provides information on the distribution effects both close to the risk cutoff and higher in the distribution of application risk. Putting the RCT together with the RD yields a rich picture of how information on admissions chances shapes outcomes for students at different points in the risk distribution and at different points in the choice process.

In addition to our main analyses of the 2018-2020 platform pop-up and the WhatsApp RCT, we present supplemental findings from a 2017 pilot of the platform pop-up. The pilot differed from the main pop-up implementation in two ways. First, it was limited to markets that adopted centralized choice early in the national rollout, so the sample size was small. Second, cities used different risk cutoffs, ranging from 30% to 70%. The pilot provides another window into effects across the risk distribution.

### 3.3 Application risk and risk predictions

Predicted application risk is a critical input to the interventions we study. The NGO computed application risk in each market-year as follows. They first obtained the vector of reported school capacities for the current year, a projected number of applicants  $N$ , and a dataset of applications and student types (i.e. priorities). For the first few days of each market-year, these data consisted of the previous year’s joint distribution of applications and priorities in that market. For the remaining days, these data consisted of applications and priorities submitted so far in the current process.

The NGO’s algorithm resampled  $N$  (application list, student priority type) tuples from this dataset, drew lottery numbers, and simulated the matching process. Repeating this process 500 times, the NGO computed the probability of non-placement within each school-grade-priority group.

The NGO developed a web service that used the calculated probabilities to predict the risk of non-placement for any individual application. These are equal to the probability of not being assigned to any of the schools in the list, for the specific grade and priority of the applicant. For more details on simulation and demand prediction see Appendix F.

Risk predictions closely track applicants’ true non-placement risk. Panel A of Figure 3 describes the distribution of predicted placement probabilities across different values of true, ex-post placement probability. The ex post placement probability is constructed identically to the placement prediction, but using realized rather than predicted applications. Predicted values cluster around true placement probabilities across the distribution. The slope of the predicted value in the true value is 0.81, with deviations from one driven by slight but systematic underprediction of risk among the most risky applications. Our assessment is that the predicted risk measure provides a reasonable guide to true risk, particularly in comparison to applicants’ risk beliefs, which we discuss in detail below.

Many applicants submit risky initial applications. Panel G of Table 1 describes ex post

(or “true”) risk on the initial application attempt. Mean non-placement risk on the initial application is 24%. A majority— 55%— of applicants are almost sure to be placed. We classify individuals as facing zero risk if their nonassignment probability is less than 0.01. At the same time, many applicants submit very risky applications. 30% submit initial applications with risk non-placement risk higher than 30%. Median risk for students submitting applications with non-zero risk is 62%, and 25% of such applicants have non-placement risk of 92% or higher. Panel B of Figure 3 plots the histogram of the risk distribution for the first and final application attempts. In both cases, mass stacks on the edges of the distribution, at very high and low risk levels. Mass shifts slightly towards lower-risk applications between the initial and final submissions.

As reported in column 5 of Table 1, 20% of all applicants— 233,678 students over the three years— are classified as risky by the choice platform based on their initial application. Risky applicants are less likely to be economically vulnerable than other applicants and more likely to come from urban areas. They submit shorter initial applications than the sample population as a whole, but longer final applications, and are more likely to change their applications between their initial submission and the deadline. 45% end up being placed at one of their preferences in the first round, while 11% receive a second-round placement.

Columns 6 and 7 of Table 1 describe the sample of students critical to our empirical analysis of the effects of application warnings. Column 6 describes applicants near the cutoff for receiving a pop-up warning, defined here as having non-placement risk between 0.1 and 0.5. This group has slightly higher socioeconomic status, slightly longer initial and final applications, and similar rates of application changes to the full sample. Column 7 describes the sample of risky 2020 applicants in the text message RCT. Like the broader sample of risky applicants, this group is relatively high-income and characterized by longer choice applications and more frequent engagement with the choice process than the population as a whole.

### 3.4 Survey design

To learn more about how families engaged with the choice process, the NGO helped Mineduc conduct a survey of choice participants in 2020. The survey asked questions about several school choice topics. It included modules about preferences, beliefs, and search designed to provide context for the interventions we study here. The survey innovates over past surveys of choice participants (De Haan et al., 2015; Kapor et al., 2020; Wang and Zhou, 2020) by recruiting a larger sample and by asking about search in addition to preferences and beliefs. See Online Appendix G for survey text.

Mineduc contacted students using an email message sent from the official school choice email account. Mineduc sent the message following the application deadline, but before the release of placement outcomes. They chose this time to maximize applicants’ recall of their

school choice experience while ruling out the possibility that the survey might affect applicants’ portfolios. In total, Mineduc contacted 373,710 families. 48,929, or 13%, completed the survey. Column 7 of Table 1 describes survey respondents. They are slightly less likely to be economically vulnerable and rural than the population as a whole, but closely resemble the broader population in terms of application behavior.

## 4 Survey findings

### 4.1 Placement, enrollment, and student welfare

The main focus of our analysis is whether students receive any placement through the centralized mechanism. We made this choice *ex ante*, to reflect policymakers’ goals for improving the choice process. Evidence from our applicant survey supports the idea that placement vs. non-placement is a critical margin from a welfare perspective.

The survey asked respondents to report how satisfied they would feel if they were placed at the first-ranked school on their application, if they were placed at the last-ranked school, or if they were unplaced. At the time of the survey, applicants had submitted their applications, so responses reflect certainty over what the schools in question were. Applicants did not know the results of the placement process, so responses will not reflect *ex post* rationalization of known outcomes. Respondents rated these options on a seven-point scale, the same that is used for grading in the school/college context, with 1 being not at all satisfied and 7 being very satisfied. We interpret scores below 4 as dissatisfaction with the placement outcome, since 4 is the passing grade in the Chilean context.

Panel A of Figure 4 reports responses to this question. 99% of respondents report that they would be satisfied with a placement at their first-ranked school. 88% report that they would be very satisfied with this placement. 69% of students report that they would be satisfied if placed at their last-ranked school. In contrast, 88% of students report that they would be at least somewhat unsatisfied with not receiving any placement. Our interpretation is that while students prefer being placed at schools higher on the rank list, most are satisfied with a placement at any school on their list. In contrast, nearly all are unsatisfied with non-placement. Placement vs. non-placement appears to be a key welfare breakpoint.

Data on applicants’ choices to enroll in schools where they are placed validate our survey measures of satisfaction. They also suggest that enrollment choices are a useful measure of satisfaction with the placement outcome outside of the survey context.

Panels B and C of Figure 4 show that enrollment choices track stated measures of preference over schools. Panel B shows that the share of students enrolling in a placed school declines with the rank of the school, from 86% at the first-listed to school to 52% at schools ranked fifth or lower. Panel C shows that students are more likely to enroll in schools where stated

satisfaction is higher.<sup>15</sup> 93% of students placed in schools they give the highest satisfaction rating choose to enroll, compared to 40% at schools with the lowest rating.

## 4.2 Search costs and search strategies

We now turn to the question of how applicants search for schools. Our first result here is that getting to know a school well requires a lot of information, some of which may be costly to obtain. Our survey asked respondents what they needed to know about a school to feel that they knew it well. Respondents could select multiple options from a list of possibilities. As reported in Panel A of Figure 5, large majorities gave a long list of attributes. Some of these attributes are relatively easy to learn about from public sources. 83% said they would need to visit a school’s website, and 93% said they would need to learn about a school’s academic performance, which is also available online. Information on others, like extracurricular activities or school infrastructure, could likely be obtained upon a short visit to the school. However, some kinds of information that respondents value would likely be hard to find. For example, 66% of respondents said they needed to interview school staff. 79% said they required references from current families.

Our second result is that applicants do not feel that they know many schools well. We asked each respondent how well they knew a randomly-selected nearby school, a nearby school that was high-performing and expensive, and a nearby school that was low-performing and free.<sup>16</sup> We also asked respondents about a “fake” school— i.e., a school that did not exist. Panel B of Figure 5 reports the share of students that claim to know each school well. Only 17% of students report knowing the random nearby school and the high-performing, expensive school well. 14% report that they know the low-performing, free school well. Encouragingly, only 3% report knowing the fake school well. Search is costly enough that at the end of the choice process, most families do not feel well-informed about many nearby schools.

Consistent with the idea that applicants learn about schools before applying to them, respondents claim to know the schools on their applications better than they know randomly chosen nearby schools. Panel C of Figure 5 displays applicants’ responses to a question asking how much they knew about the schools on their submitted application. 73% of students claim to know their first-listed school well and 41% claimed to know the second-listed school well. Knowledge declines with application rank, but 30% of students who submit applications

---

<sup>15</sup>Panel C includes only on first- and last-ranked schools, since these were the schools for which the survey asked about student satisfaction.

<sup>16</sup>Schools in this question were selected from the alternatives within 2km from the residential location of the student that were not included in her application. We used the performance classification of the “Agencia de la Calidad de la Educación”, an institution that classifies schools in 4 tiers using standardized test scores, after taking into account socioeconomic status of the student body. We classify a school as “high-performing” if it is in one of the best 2 tiers, while “low-performing” if it is in the worst tier. “Expensive schools” are those that charge a monthly copayment of at least 35 USD on top of the voucher.

including at least five schools claimed to know the fifth school well. This is nearly twice the share claiming to know a randomly-chosen school well.

We now turn to the role of beliefs about admissions chances in search. Proposition 1 in our model provides conditions under which applicants who think they will be admitted to a school in their existing portfolio will be less likely to engage in additional search. Two survey findings suggest that this kind of behavior is widespread.

First, we asked applicants directly why they stopped adding schools to their application. Respondents could choose from four options: (1) there were no more schools to around to add, (2) there were schools around but they would rather not attend these schools, (3) it is hard to find more schools, and (4) they think they will be placed at one of the schools already on their application.

The most common reason applicants give for stopping search is that they think they will be placed in a school already on their list. As reported in Panel A of Figure 6, 35% of respondents chose this option. Another 17% said they stopped adding schools because additional schools were hard to find, a response that also invokes costly search. Together, these two search-related responses account for a majority (52%) of all responses. We interpret this as a likely lower bound on the share of respondents for whom costly search affected choice, since costly search might also have played a meaningful but not primary role for applicants giving other responses. The remaining 48% of respondents gave answers more in line with the traditional school choice problem, in which applicants list all available schools (“no more schools around”) or list schools preferable to an outside option (“I’d rather not be placed at remaining schools”).

Second, applicants who thought their chances of being placed were high were the most likely to say they stopped search because they thought they would be placed. Our survey asked respondents what they thought their chances were of being placed at any school on their submitted portfolio. Panel B of Figure 6 plots the share of students saying they stopped search because they thought they would be placed at one of their submitted options at each quintile of the distribution of subjective placement chances. Respondents become much more likely to give this reason for stopping search as their subjective placement beliefs increase. 51% of respondents in the top quintile of the subjective belief distribution said they stopped search because they were confident in their placement chances. In contrast, only 9% of respondents in the bottom quintile gave this reason for stopping search.

### 4.3 Optimism and search

Our first set of survey findings shows that search for schools is hard, and that beliefs about placement chances are a critical input to search strategy. Our second set of findings shows that these beliefs are wrong. To make this point, we compare respondents’ reported beliefs about placement chances to observed placement chances.

Panel A of Figure 7 shows the distribution of subjective and true placement chances for applicants with non-zero risk of non-placement. Applicants far overrate their placement chances. The mean subjective placement probability is 75%, 30 percentage points above the mean true placement probability of 45%. The graph shows a mass of subjective beliefs piling up around a placement probability of one. The densest part of the distribution of true placement chances for these students is near zero, with no corresponding mass in subjective beliefs. Panel B shows the distribution of optimism, defined as the difference between subjective and true placement chances. This distribution is shifted far to the right of zero. Many respondents overestimate their placement chances by fifty percentage points or more.

In a mechanical sense, the source of this optimism is that many applicants with low true placement chances think they are likely to receive a placement. Panel C of Figure 7 plots the distribution of subjective placement beliefs, binned into ten groups by true placement probability.<sup>17</sup> If beliefs were accurate on average, they would follow the 45 degree line. We instead observe a weak positive relationship with a large upward shift. The mean subjective belief for applicants with true admissions chances near zero is close to 70%.

For comparison, we also plot the distribution of the NGO’s predicted risk measure, as computed at the time of the application for the set of survey respondents. As in the full sample, risk predictions do not precisely track the final risk values. However, it is clear that predictions are much closer to true placement probabilities than are subjective beliefs.

Several pieces of evidence indicate that our belief measures are credible. We have already shown that beliefs are related to stated reasons for stopping search. Additional results presented in Figure B1 show that our findings on the distribution of beliefs are consistent whether we frame the question in terms of placement chances or in terms of non-placement risk, and also that respondents’ overall assessments of application risk are closely related to the level of application risk implied by their beliefs about school-specific placement chances.

## 5 Warnings, search, and placement

### 5.1 The platform pop-up

Our survey findings make a strong case that beliefs about admissions chances are important inputs to the choice process, and that many applicants strategize on the basis of overly-optimistic beliefs. Together with our theoretical analysis, this suggests that applicants should respond to warnings about non-placement risk by adding more schools to their portfolios. We test this proposition using experimental and quasi-experimental research designs implemented

---

<sup>17</sup>Because applications cluster at the tails of the risk distribution, these groups are not deciles. The bottom group includes applicants with placement probability less than 1%, and the top group includes applicants with placement probability of 99% or more. The middle eight groups split the remaining observations into equally-sized bins.

in the Chilean and New Haven choice systems.

We focus first on the platform pop-up administered to Chilean students inside the choice system. Because all students with at least a 30% chance of non-placement received this warning, we evaluate it using a regression discontinuity design. In our visual analysis of RD outcomes, we display binned means together with global polynomial fits, to provide a sense of broad patterns in the data and how they relate to observed discontinuities. When computing estimates of RD effects, we use local linear specifications with a triangular kernel and a bandwidth of 0.1. This bandwidth approximates that given by optimal bandwidth calculations (Calonico et al., 2014).<sup>18</sup>

We first show that applicants’ observable characteristics are unrelated to which side of the 30% cutoff they fall on. Figure 8 and Panel A of Table 2 show how the share of students from rural areas and the share of low-income students vary by position relative to the cutoff. We see no visual evidence of a discontinuity in either variable. Table 2 reports separate estimates for the full sample and for each choice year. Consistent with the graphical evidence, cross-threshold differences in these attributes are small in economic terms. Because our sample size is quite large—roughly 41,000 applicants within the local bandwidth—our estimates are very precise, and some economically small effects are marginally statistically significant. We also note that there is no evidence of a discontinuity in density of the running variable near the cutoff, as displayed in Panel C of Figure 8. These findings are consistent with the observation that the 30% cutoff had no significance for applicants prior to policy implementation.

## 5.2 Search and choice behavior

Panels A through D of Figure 9 and Panel B of Table 2 show how receiving the platform pop-up changed choice behavior. Receiving a warning caused 21.4% of applicants to alter their submissions. Essentially all of these changes are additions to the application. Receiving a warning caused 21.6% of applicants to add at least one school to their application.<sup>19</sup> Students add an average of 0.34 schools, and ex post risk of non-placement falls by 3.3 percentage points, 13% of the below-threshold mean. These effects are stable across years.

In the context of proposition 2, the 21.6% of students who add a school to their application in response to the intervention correspond to the share complying with the intervention by engaging in further search. The second column of Table 2 displays instrumental variables estimates in which adding at least one school to the application is the endogenous regressor.

<sup>18</sup>We report estimates obtained with Calonico et al. (2014) bandwidth selection in Online Appendix Table B1 and Figure B2. Alternate approaches to RD estimation do not change our findings.

<sup>19</sup>These calculations compare students across the RD threshold. Hence, although “adding a school” is a subset of “altering an application,” the respective treatment effects need not be ordered in this way. In particular, the estimated share induced to add a school is slightly larger than the share induced to make any change because, at baseline, a slightly larger fraction of “control” students are changing their applications without adding schools.



The resulting effect estimates can be interpreted as LATEs for the policy compliers. Students induced to search by the policy add an average of 1.6 schools to their application list, and reduce their ex post non-placement risk by 15.5 percentage points, equal to 58% of the below-threshold mean. The share of compliers with the intervention policy is large, and the risk reduction within this group is substantial.

The changes applicants make to their applications are consistent with the idea that the intervention leads to additional search. Panel B of Table 2 and Panels E through H of Figure 9 show that most but not all students who change their applications do so by adding schools to the end. Receiving the warning raises the chance a student will add a school to the end of their application by 20.5 percentage points, about 95% of the share of students adding any school to their application. The frequency with which students add schools to the end of their application indicates that Proposition 1’s focus on students adding schools to the bottom of their rank list is empirically relevant. However, receiving a warning also causes 2 percent of applicants to add schools to the middle of their list. This suggests that at least some applicants are learning about new schools, and not just adding known schools to the bottom of their rank lists. Very few students add schools to the top of their rank list, indicating that for the most part students identify their top-choice schools early in the search process.

The platform pop-up does not cause students to drop schools from their rank lists. This is consistent with our model, in which students who find additional schools add them to their portfolio and do not “un-search.” We find some evidence that a small share (1%) of students re-order the existing schools on their application in response to the intervention, although the visual evidence here is not as compelling as our other plots (Panel H of Figure 9). One way to interpret this finding is that the warning may prompt a small share of students to revise their applications in response to changing preferences over time, as in Narita (2018). However, any such effect is second-order compared to the share of students adding schools.

### 5.3 Enrollment and welfare

Changes in application behavior translate to changes in choice outcomes. Panel A of Figure 10 and Panel C of Table 2 report the effect of receiving the warning on placement outcomes. Students receiving a warning are 3.8 percentage points more likely to be placed in one of their listed schools. As expected, this closely tracks the reduction in non-placement risk (within one standard error).

Warnings do not produce lower quality placements. As reported in Panels B and C of Figure 10 and Panel C of Table 2, overall rates of enrollment rise in proportion to changes in placement across the cutoff, and the rate at which students enroll in school conditional on placement is continuous through the cutoff value.

Proposition 3 showed that the effect of the intervention on individual welfare was propor-



tional to the change in placement rates, except as offset by declines in enrollment conditional on placement. Our findings suggest that there are no offsetting enrollment effects. The implication is that receiving the warnings intervention raises welfare (excluding search costs) for compliers with the information intervention by 21% ( $= 0.15/0.74$ )— the per-complier change in placement rate as a percentage of the below-threshold mean.

## 5.4 Decongestion vs. reshuffling

The goal of our analysis is to understand how beliefs about admissions chances affect school search and placement outcomes from the perspective of individuals. However, it is also useful to think about how information on placement chances affects market-level congestion. If the warnings policy causes applicants to place in undersubscribed schools, the individual gains we observe may come “for free,” in the sense that other students are not displaced. Congestion effects are important to consider because beliefs interventions do not guide applicants towards specific schools. This contrasts with the provision of information on school attributes, which may point applicants towards oversubscribed schools (say, those with high test scores).

We assess the congestion effects of the platform pop-up by looking at how receiving a warning affects placement rates at uncongested schools, defined as schools with excess capacity. As reported in Panel D of Table 2, receiving a warning raises the chances that students add at least one uncongested school to their application by 7.3 percentage points. Put another way, roughly one third of applicants adding at least one school add an uncongested school.

Most of the decline in application risk from receiving a warning comes from increased chances of placement in an uncongested school. Receiving a warning raises applicants’ chances of placing at an uncongested school by 1.9 percentage points, corresponding to an 8.8 percentage point increase for compliers with the warnings policy. This is equal to 57% of the overall risk reduction of 15.5 percentage points reported in Panel B of Table 2. We conclude that the warnings policy helps reduce congestion, a core goal of centralized choice systems (Abdulkadiroğlu et al., 2017).

## 5.5 Warnings across the risk distribution

### 5.5.1 The WhatsApp reminder RCT

We use the random assignment of reminder message interventions to study the effects of warnings about risky applications vary away from the 30% risk cutoff and on the intensive margin. Recall from section 3.2 that, in the 2020 choice process, randomly selected risky applicants received a WhatsApp text warning three days before the application deadline. 44 hours after that, on the day before the deadline all risky applicants received the same warning through an SMS. In this context, what random assignment does is raise the number of warnings to which risky applicants are exposed between the time they first fill out their

application and the application deadline. For non-risky students (below the 30% risk cutoff) treatment and control status are randomly assigned, but the “treated” group does not receive a risk warning.

Figure 11 presents the effects of the RCT by plotting outcomes for the treatment and control groups by application risk at the time of randomization into the text message treatment. Panel A shows that the number of warnings students receive (summing over all SMS and WhatsApp interventions) rises across the cutoff for both treatment and control groups, but rises more for the treatment group, which receives the additional WhatsApp message. The 0.48 difference in messages viewed for treatment relative to control among risky applicants reflects the share of applicants who opened WhatsApp and viewed the image.

Panels B and C of Figure 11 describe application behavior in the 44 hour window between the randomized message to the WhatsApp treatment group and the message to all risky students. Risky students randomly assigned to the WhatsApp treatment are more likely to add schools to their application and reduce their application risk than untreated students. On average, assignment to the treatment group causes 3.3% of risky students to add at least one school to their application. This corresponds to a 6.8 percentage point effect for each student that views the treatment image. These changes cause application risk to fall by 1.0 percentage points, or 2.1 percentage points per message view. The implied risk reduction for applicants who comply with the WhatsApp intervention by engaging in additional search is 29.7 percentage points, equal to 49% of mean risk in the RCT sample.

Search and risk reduction outcomes in the treatment group outpace those in the control group over the full distribution of risk values above 30%. To facilitate comparison between RCT and RD estimates, panels B and C of Figure 11 also display RD estimates for the WhatsApp treatment calculated across the risk cutoff within the treatment group. RD estimates show smaller effects than those estimated in the RCT. We see little evidence that students near the 30% risk cutoff respond more to information interventions than applicants elsewhere in the risk distribution.

Panels D and E repeat the analysis from Panels B and C, but now look at *all* application changes between the randomized warning and the application deadline. These measures include the effects of the final text reminder sent to all risky students. Despite the text message followup, gaps between treatment and control expand over time. As for the 44-hour outcomes, treatment-control comparisons span the full distribution of risk above the risk cutoff, except perhaps the very top. Average treatment effects in the RCT are larger for endline outcomes than for the 44-hour outcomes. 4.4 percent of students add a school to their application, and the mean risk reduction for these students is 30 percentage points.

Table 3 summarizes findings from the RCT and RD analysis of the WhatsApp intervention. Treatment and control groups are balanced on observable characteristics. For choice outcomes, we present both ITT effects reported in Figure 11 and IV estimates that take adding at least

one school as the endogenous regressor of interest.

Overall, compared to the platform pop-up, the share of compliers with the WhatsApp RCT is smaller. This makes sense given that the RCT population is negatively selected on the response to previous interventions. However, the percent reduction in risk per complier is similar, and the percentage point reduction in risk is larger.

We draw two conclusions from this analysis. The first is that the effects of warnings persist as we move up the risk distribution. The second is that there may be benefits to providing the same person with information multiple times. This is consistent with literature in behavioral economics indicating that information provision has the largest effects on choice when it is provided roughly at the time of choice. Providing multiple reminders may raise the chances that one reminder is received around the time applicants need it.

### 5.5.2 RDs at multiple cutoffs from the 2017 pilot

The 2017 pilot of the platform pop-up intervention provides additional evidence on the effects of warnings across the risk distribution. The pilot was essentially identical to the 2018-2020 intervention, but a) was limited to markets that had implemented centralized choice by 2017, and b) varied the cutoff across cities, with some cities having cutoffs of 30%, others 50%, and others 70%. Table B2 reports results from the pilot for the pooled sample, and split by risk cutoff. The sample size is roughly 3% as large as in our main analysis, so estimates are noisy. However, the pooled sample effects on the key add any school and change in application risk outcomes are quite similar to what we see in the main intervention. Splitting across cutoff values, point estimates are largest at the 50% cutoff, and smaller for the 30% cutoff than for the 70% cutoff. The pilot was underpowered to detect heterogeneity across cutoffs, so we interpret these differences cautiously.

Evidence from the pilot informs the evaluation of our main analysis in two ways. First, it shows that the 30% cutoff used in the main intervention was not chosen because treatment effects were particularly large at that margin. Mineduc’s goal was to provide the treatment inclusively to a broad set of applicants facing non-placement risk. Second, to the extent we do observe pop-up effects at different margins, effects are larger at risk levels above 30%. Both pieces of evidence reinforce the point that warnings interventions have large effects across the distribution of application risk.

## 5.6 Heterogeneous effects within Chile

We next consider the distribution of treatment effects across places, times, and people. We begin by describing the distribution of estimated treatment across markets defined at the city-year level. Our central finding is that the platform pop-up intervention increases search across a wide range of markets.

Panels A and B of Figure 12 display (unweighted) histograms of the RD estimates of pop-up treatment effect estimates split by city and year. Panel A shows effects on the “add any school” outcome and Panel B on the “schools added” outcome. Modal values across these cells are close to the reported overall effects of 0.22 and 0.34, respectively. The IQR of the estimated effect of treatment on adding any school is (0.11,0.29). Table 4 shows city-by-year specific effects for the add any school outcome in 20 large educational markets, sorted by 2020 applicant count. In 2020, effects in the three largest markets are each close to the nationwide average. Our interpretation is that the platform pop-up intervention is broadly effective. Full-sample findings are not driven by specific city-year combinations.

The effects of information interventions on beliefs might diminish as market participants gain experience with choice. For example, it might be difficult to learn about admissions chances the first year a centralized system is in place because no information on past outcomes is available. To test this, Table 5 repeats Table 2, but splits the sample by the number of years a city-by-grade combination has used the centralized platform. Shifts in search and risk are similar for city-grade-years with one, two, or three or more years experience using the centralized platform. There is no evidence that the effects of the platform pop-up intervention decline as experience with centralized choice rises. This is consistent with results in Kapor et al. (2020) showing large belief errors in a setting with a long history of choice.

The platform pop-up affects search behavior for both high- and low-SES students. Table 6 again repeats the analysis of Table 2, this time splitting by economic vulnerability. Rates of application modification and risk reduction are slightly larger for economically vulnerable applicants. Policy changes that reduce the strategic complexity of school choice are often driven by equity concerns (Abdulkadiroglu et al., 2006; Pathak and Sönmez, 2008). Smart matching platforms, however, appear to be valuable to both high- and low-SES applicants.

The platform pop-up causes additional search in both large and small choice markets, but applicants in bigger markets add more schools to their applications. Panels C and D of Figure 12 split the sample by different measures of market size, and report estimated pop-up RD effects on the add any school outcome within each subgroup. Panel C splits by the count of available schools within a 3km radius of applicant’s address. We consider ten (roughly) evenly sized groups, with the largest category consisting of applicants with between 54 and 105 schools within 3km, and the smallest of applicants with zero to four schools. Panel D splits the RD estimates by the size of the market as a whole, regardless of the count of schools near the student. We observe sizeable effects across all groups, with no clear pattern by market size. This contrasts with results reported in panels E and F of Figure 12, which shows the *count* of schools that applicants add. Here, we see a clear increase with both measures of market size.

## 5.7 Smart vs. static matching platforms in New Haven

In addition to our work in Chile, we partnered with the NGO and the New Haven, Connecticut school district to implement a similar warnings intervention during the 2020 choice process. New Haven is a medium-sized school district that has used centralized choice to assign students to schools since the mid-1990s. Starting in 2019, New Haven adopted a truncated deferred acceptance assignment mechanism. See [Kapor et al. \(2020\)](#) and [Akbarpour et al. \(2020\)](#) for more institutional details.

The warnings policy in New Haven was similar in broad strokes to the policies implemented in Chile. Families were allowed to submit applications to the choice process starting at the end of January, with an application deadline of March 2. Seven days before the deadline, the district identified applications with a non-placement risk of greater than 50% as risky. Application risk was computed using data from the previous year; i.e., an application was designated as risky if the chance of non-placement for a student submitting the application in the previous year was at least 50%.<sup>20</sup> All applicants identified as risky received an email stating they were at risk of non-placement. The email included a link to a website where they could input hypothetical applications and view the chances of admission at each school, again based on the previous year’s data.<sup>21</sup>

The New Haven policy differed from the Chile policy in two important ways. The first is that, in addition to warning all risky applicants, the district selected a randomly chosen fifty percent of non-risky applicants to receive an email that provided a “static” recommendation to learn more about admissions chances by visiting the same application simulator website. This differed from the smart platform intervention in that it did not include any information about the applicant’s *own* application risk.

The static intervention allows us to test whether the “smart” part of the smart matching platform is important. Static interventions that do not rely on applicant-specific risk predictions will in general be easier for policymakers to implement. If static interventions can produce similar results to “smart” interventions, they may be the better policy option.<sup>22</sup>

The second major contrast between our work in New Haven and our work in Chile is sample size: in Chile, 233,768 students received a warning about a risky application. In New Haven, the number was 740. This reduces statistical precision substantially.

Figure 13 presents a visual summary of our findings from the New Haven intervention.

---

<sup>20</sup>The district focused on major choice grades (pre-Kindergarten, Kindergarten, and grade 9), where choice probabilities are relatively stable across years. Two new schools opened in 2020. Risk scores were not computed for students applying to these schools.

<sup>21</sup>See Online Appendix H for a detailed description of the intervention procedures in New Haven, the distribution of application risk, and the relationship between our risk simulations and realized application risk.

<sup>22</sup>We note that an ideal test of the smart vs. static intervention would have been an RCT within the risky population. The NGO and policymaker opted not to do this because they wanted risky applicants to receive the best intervention available. By 2020, we had several years of evidence from Chile that the smart approach was effective.

These graphs plot the rate at which students make different kinds of application changes in each ten percentage point bin of the predicted risk distribution, with additional bins for risk values of zero and one. We display these statistics for 2020 applicants, who received a warning email when predicted risk was 50% or higher, and for a comparison group of 2019 applicants, who did not receive a warning regardless of risk score.<sup>23</sup> For non-risky applicants in 2020, the graphs split out the set of applicants who received the static prompt from those who were not contacted at all.

Panel A of Figure 13 shows results for application modification. Rates of application modification for low-risk applicants were similar in 2019 and 2020. In 2020, we observe a large jump in rates of modification at the 50% cutoff for the warning treatment, with no similar increase for the 2019 comparison group. Just below the cutoff, roughly 1.5% of applicants modify their application prior to the deadline, compared to 15.8% just above. As shown in Panel B, almost all of these changes involve lengthening the application. As shown in Panel C, the effect of these additions is to reduce application risk by two to four percentage points around the cutoff.

An observation that holds across all panels is that the static intervention does not affect search at all. The static prompt and no contact series look essentially the same at all tested values of risk. This suggests that the “smart” component of the smart platform is important to the effects we observe. Online Appendix H presents a detailed analysis of the static intervention RCT.

Table 7 summarizes the effects of the warnings policy. We report two kinds of effect estimates. The first are RD estimates using only the 2020 data. The RD specifications allow for separate slope terms above and below the cutoff value, and include all data except for mass points at risk values of zero and one. The second are difference-in-difference estimates where the first difference is 2020 vs. 2019 and the second difference is above vs. below the warnings threshold. The difference-in-difference specifications control for risk-group fixed effects in ten percentage point bins. Both RD and DD estimates pool across the static intervention and no-contact control group among non-risky applicants. We do this because average outcomes for these groups are essentially the same.

Panel A of Table 7 shows that predetermined characteristics are balanced across the cutoff, although estimates of changes in female and Black share are imprecise. Panel B shows that while essentially all above-threshold students received a warnings email, relatively few logged into the online simulator or ran a simulation. This suggests that the behavioral effects we observe come mostly from the warning and not from the simulator availability. This is consistent with the large effects we observe in the Chilean setting, which did not include a simulator component. It is also consistent with null results from the static intervention

---

<sup>23</sup>We compute predicted risk for 2019 applicants using a snapshot of predicted risk status as of seven days prior to the admissions deadline. This procedure parallels our approach to identifying risky applicants in 2020.

RCT.<sup>24</sup>

Panel C shows estimates of effects on different choice outcomes. The RD estimates indicate that crossing the threshold and receiving the warning causes 11.4 percent of applicants to add at least one school to their application. These are the compliers with the information treatment. Ex post realized application risk falls by 2.6 percentage points across the cutoff. This means that compliers with the policy reduce their application risk by 22.8 percentage points ( $= 0.026/0.114$ ), or 45% of the below-threshold mean risk level. Compared to the Chilean setting, the complier population is somewhat smaller, while risk falls by more per complier in absolute terms, and the reduction as a share of baseline risk levels is similar.

Comparing the RD and DD estimates confirms the visual impression from Figure 13 that behavioral changes are larger for less risky students in the risky group, though estimates are imprecise. A possible explanation for this finding is that the highest-risk applicants are those applying to a small number of highly desirable schools. These applicants may have outside options beyond the public system and be uninterested in listing additional inside options (Akbarpour et al., 2020).

Findings from the New Haven intervention provide two important pieces of evidence. First, information on admissions chances is an important input to search and choice behavior in a wide variety of contexts, including in US districts with long histories of centralized choice. Second, the “smart” part of smart matching platforms appears to be important to their efficacy at expanding search.

## 5.8 Stated preferences across contexts

We conclude our discussion of generalizability across settings by stepping away from program evaluation and focusing on a simpler measure: how much people say they want information about admissions chances. The surveys of Chilean applicants asked about the kinds of information they thought would be most helpful in filling out their applications. Respondents could choose multiple options from a menu that included measures of academic performance, prices, and school suggestions. Choice participants in New Haven were asked a similar question.<sup>25</sup> We report survey results in Online Appendix Figure B4.

In both cases, the kinds of information respondents were most likely to say they needed was information about admissions. Roughly 90% of respondents in both the Chile and New Haven surveys said they needed information on admissions chances. Similar numbers indicated a need for information on vacancies. For comparison, 93% of Chilean respondents said they would have liked more information on accountability scores, and 84% of New Haven respondents

---

<sup>24</sup>To be clear, the implication here is not that risk simulators cannot form part of an effective intervention. Rather, they indicate that effective interventions do not require simulation approaches, and that the version of the application simulator implemented in New Haven in 2020 was not critical to the success of that intervention.

<sup>25</sup>See Online Appendix H for details.

said additional school suggestions would have been helpful.

## 6 Conclusion

This paper shows that beliefs about admissions chances are a key determinant of the way applicants search for schools in centralized choice systems, that optimism about school placement chances leads applicants to search too little, and that the smart matching platforms that build live feedback on application risk into the choice system increase search and reduce non-placement risk.

The main implication of our findings is that policymakers seeking to reduce the strategic burden school choice places on participants need *both* to choose a strategyproof assignment mechanism *and* to provide choice supports that aid with search for schools. The strategic challenges posed by school search are central to families’ experiences of school choice as well as to their placement outcomes. This holds even when the centralized assignment mechanism itself is strategyproof.

The smart matching platforms we propose and evaluate in this paper appear to be an effective and generalizable approach to reducing the strategic burden of school search. Critically, these smart platforms are not researcher-driven proofs-of-concept. Nor are they one-off policy solutions, implemented for a specific context. They are *products* implemented at scale. The “policy intervention as product” model makes it relatively easy to deploy smart matching platforms in new settings. At the time of this writing, policymakers in Brazil, Peru, and Ecuador are in the process of implementing the techniques we discuss in this paper. The close collaboration between researchers, policymakers, and implementation partners that made this work possible may be a useful approach for conducting scalable interventions in other domains.

## References

- Abdulkadiroğlu, Atila and Tayfun Sönmez, “School choice: A mechanism design approach,” *American Economic Review*, 2003, *93* (3), 729–747.
- , Nikhil Agarwal, and Parag A Pathak, “The welfare effects of coordinated assignment: Evidence from the New York City high school match,” *American Economic Review*, 2017, *107* (12), 3635–89.
- , Parag A Pathak, Alvin E Roth, and Tayfun Sönmez, “The Boston public school match,” *American Economic Review*, 2005, *95* (2), 368–371.
- , —, and —, “The New York City high school match,” *American Economic Review*, 2005, *95* (2), 364–367.



- Abdulkadiroglu, Atila, Parag Pathak, Alvin E Roth, and Tayfun Sonmez**, “Changing the Boston school choice mechanism,” Technical Report, National Bureau of Economic Research 2006.
- Abdulkadiroğlu, Atila, Yeon-Koo Che, and Yosuke Yasuda**, “Resolving conflicting preferences in school choice: The “Boston Mechanism” reconsidered,” *American Economic Review*, 2011, *101* (1), 399–410.
- Agarwal, Nikhil and Paulo Somaini**, “Demand analysis using strategic reports: An application to a school choice mechanism,” *Econometrica*, 2018, *86* (2), 391–444.
- Agrawal, Ajay, Joshua Gans, and Avi Goldfarb**, *Prediction machines: The simple economics of artificial intelligence*, Harvard Business Press, 2018.
- Ainsworth, Robert, Rajeev Dehejia, Cristian Pop-Eleches, and Miguel Urquiola**, “Information, Preferences, and Household Demand for School Value Added,” Technical Report, National Bureau of Economic Research 2020.
- Ajayi, Kehinde and Modibo Sidibe**, “School Choice Under Imperfect Information,” *Economic Research Initiatives at Duke (ERID) Working Paper*, 2020, (294).
- Akbarpour, Mohammad, Christopher Neilson, Adam Kapor, Winnie van Dijk, and Seth Zimmerman**, “Centralized School Choice with Unequal Outside Options,” *Industrial Relations Section Working Paper*, 2020.
- Allcott, Hunt and Dmitry Taubinsky**, “Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market,” *American Economic Review*, 2015, *105* (8), 2501–38.
- Angrist, Joshua D, Guido W Imbens, and Donald B Rubin**, “Identification of causal effects using instrumental variables,” *Journal of the American Statistical Association*, 1996, *91* (434), 444–455.
- Ashlagi, Itai and Afshin Nikzad**, “What matters in school choice tie-breaking? How competition guides design,” *Journal of Economic Theory*, 2020, *190*, 105–120.
- Barahona, Nano, Cristóbal Otero, Sebastián Otero, and Joshua Kim**, “Equilibrium Effects of Food Labeling Policies,” *Available at SSRN 3698473*, 2020.
- Bergman, Peter, Jeffrey T Denning, and Dayanand Manoli**, “Is information enough? The effect of information about education tax benefits on student outcomes,” *Journal of Policy Analysis and Management*, 2019, *38* (3), 706–731.

- Bettinger, Eric P, Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu**, “The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment,” *The Quarterly Journal of Economics*, 2012, 127 (3), 1205–1242.
- Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik**, “Robust nonparametric confidence intervals for regression-discontinuity designs,” *Econometrica*, 2014, 82 (6), 2295–2326.
- Calsamiglia, Caterina, Chao Fu, and Maia Güell**, “Structural estimation of a model of school choices: The Boston mechanism versus its alternatives,” *Journal of Political Economy*, 2020, 128 (2), 642–680.
- Corcoran, Sean P, Jennifer L Jennings, Sarah R Cohodes, and Carolyn Sattin-Bajaj**, “Leveling the playing field for high school choice: Results from a field experiment of informational interventions,” Technical Report, National Bureau of Economic Research 2018.
- Correa, Jose, Rafael Epstein, Juan Escobar, Ignacio Rios, Bastian Bahamondes, Carlos Bonet, Natalie Epstein, Nicolas Aramayo, Martin Castillo, Andres Cristi et al.**, “School choice in Chile,” in “Proceedings of the 2019 ACM Conference on Economics and Computation” 2019, pp. 325–343.
- de Chile Ministerio de Educación, Gobierno**, *El primer gran debate de la Reforma Educacional: Ley de Inclusión Escolar* September 2017.
- Doraszelski, Ulrich, Katja Seim, Michael Sinkinson, and Peichun Wang**, “Ownership concentration and strategic supply reduction,” Technical Report, National Bureau of Economic Research 2017.
- Dynarski, Susan, CJ Libassi, Katherine Micheltmore, and Stephanie Owen**, “Closing the Gap: The Effect of Reducing Complexity and Uncertainty in College Pricing on the Choices of Low-Income Students,” *American Economic Review*, 2019.
- Fernandes, Daniel, John G Lynch Jr, and Richard G Netemeyer**, “Financial literacy, financial education, and downstream financial behaviors,” *Management Science*, 2014, 60 (8), 1861–1883.
- Fischer, Mira and Valentin Wagner**, “Effects of timing and reference frame of feedback: Evidence from a field experiment,” 2018.

- Gruber, Jonathan, Benjamin R Handel, Samuel H Kina, and Jonathan T Kolstad**, “Managing intelligence: Skilled experts and AI in markets for complex products,” Technical Report, National Bureau of Economic Research 2020.
- Gurantz, Oded, Jessica Howell, Michael Hurwitz, Cassandra Larson, Matea Pender, and Brooke White**, “Realizing your college potential? Impacts of College Board’s RYCP campaign on postsecondary enrollment,” 2020.
- Haan, Monique De, Pieter A Gautier, Hessel Oosterbeek, and Bas Van der Klaauw**, “The performance of school assignment mechanisms in practice,” 2015.
- Haeringer, Guillaume and Flip Klijn**, “Constrained school choice,” *Journal of Economic Theory*, 2009, 144 (5), 1921–1947.
- Hakimov, Rustamdjan, Dorothea Kübler, Siqi Pan et al.**, “Costly Information Acquisition in Centralized Matching Markets,” Technical Report, CRC TRR 190 Rationality and Competition 2021.
- Hastings, Justine, Christopher A Neilson, and Seth D Zimmerman**, “The effects of earnings disclosure on college enrollment decisions,” Technical Report, National Bureau of Economic Research 2015.
- Hastings, Justine S and Jeffrey M Weinstein**, “Information, school choice, and academic achievement: Evidence from two experiments,” *The Quarterly Journal of Economics*, 2008, 123 (4), 1373–1414.
- Hoxby, Caroline and Sarah Turner**, “Expanding college opportunities for high-achieving, low income students,” *Stanford Institute for Economic Policy Research Discussion Paper*, 2013, 12, 014.
- Hsieh, Chang-Tai and Miguel Urquiola**, “The effects of generalized school choice on achievement and stratification: Evidence from Chile’s voucher program,” *Journal of public Economics*, 2006, 90 (8-9), 1477–1503.
- Hyman, Joshua**, “Can light-touch college-going interventions make a difference? Evidence from a statewide experiment in Michigan,” *Journal of Policy Analysis and Management*, 2020, 39 (1), 159–190.
- Immerlica, Nicole, Jacob Leshno, Irene Lo, and Brendan Lucier**, “Information Acquisition in Matching Markets: The Role of Price Discovery,” *Available at SSRN*, 2020.
- Jensen, Robert**, “The (perceived) returns to education and the demand for schooling,” *The Quarterly Journal of Economics*, 2010, 125 (2), 515–548.

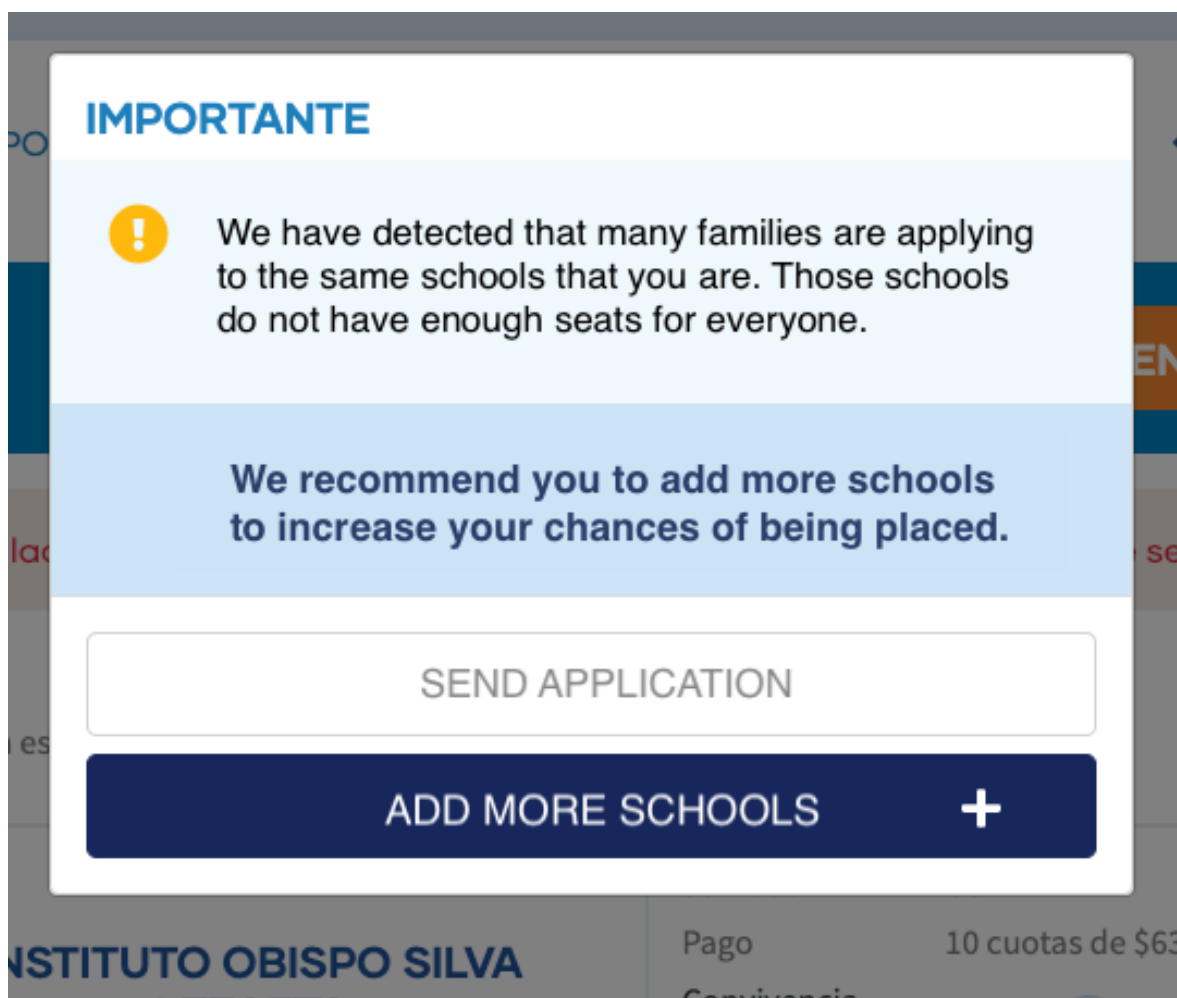
- Jin, Ginger Zhe and Phillip Leslie**, “The effect of information on product quality: Evidence from restaurant hygiene grade cards,” *The Quarterly Journal of Economics*, 2003, *118* (2), 409–451.
- Kapor, Adam J, Christopher A Neilson, and Seth D Zimmerman**, “Heterogeneous beliefs and school choice mechanisms,” *American Economic Review*, 2020, *110* (5), 1274–1315.
- Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao**, “Poverty impedes cognitive function,” *Science*, 2013, *341* (6149), 976–980.
- McCall, John Joseph**, “Economics of information and job search,” *The Quarterly Journal of Economics*, 1970, pp. 113–126.
- Milgrom, Paul and Ilya Segal**, “Clock auctions and radio spectrum reallocation,” *Journal of Political Economy*, 2020, *128* (1), 1–31.
- Narita, Yusuke**, “Match or mismatch? Learning and inertia in school choice,” *Learning and Inertia in School Choice (June 18, 2018)*, 2018.
- Neilson, Christopher**, “Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students,” Working Paper, Yale University 2020.
- , **Claudia Allende, and Francisco Gallego**, “Approximating the Equilibrium Effects of Informed School Choice,” 2019.
- Pathak, Parag A and Tayfun Sönmez**, “Leveling the playing field: Sincere and sophisticated players in the Boston mechanism,” *American Economic Review*, 2008, *98* (4), 1636–52.
- and —, “School admissions reform in Chicago and England: Comparing mechanisms by their vulnerability to manipulation,” *American Economic Review*, 2013, *103* (1), 80–106.
- Patterson, Richard, Nolan Pope, and Aaron Feudo**, “Timing is Everything: Evidence from College Major Decisions,” 2019.
- Rees-Jones, Alex, Ran Shorrer, and Chloe J Tergiman**, “Correlation neglect in student-to-school matching,” in “Proceedings of the 21st ACM Conference on Economics and Computation” 2020, pp. 467–468.
- Son, Suk Joon**, “Distributional Impacts of Centralized School Choice,” *Yale University Working Paper*, 2020.

**Thaler, Richard H and Wil Tucker**, “Smarter information, smarter consumers,” *Harvard Business Review*, 2013, *91* (1), 44–54.

**Wang, Tong and Congyi Zhou**, “High school admission reform in China: a welfare analysis,” *Review of Economic Design*, 2020, *24* (3), 215–269.

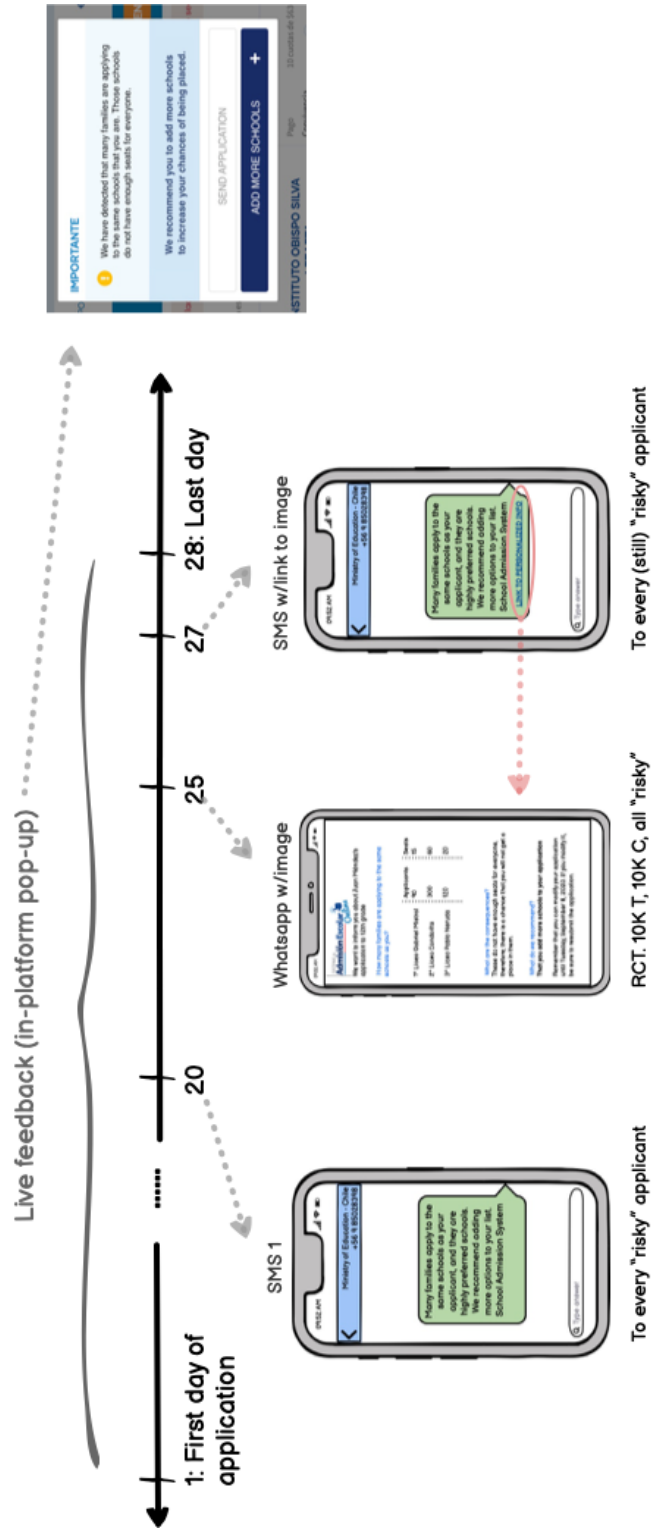
## Figures

Figure 1: Platform pop-up intervention– 2018 and 2019



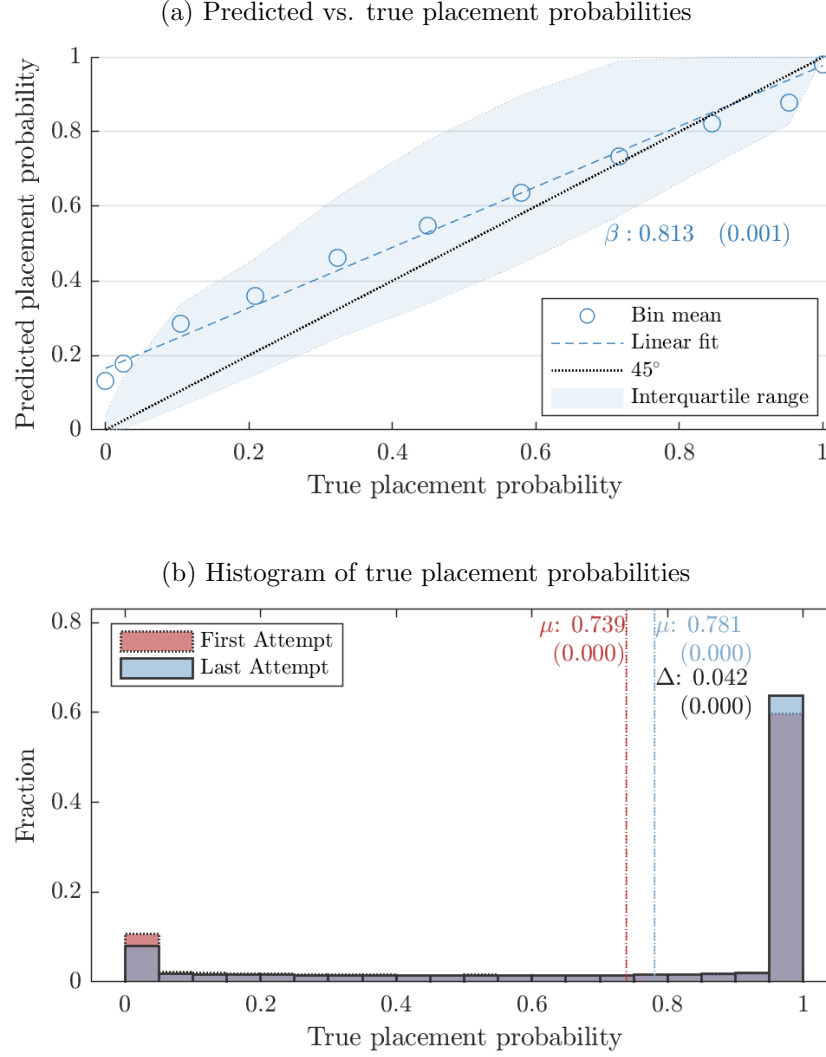
English translation of pop-up feedback shown to risky applicants on the application platform in 2018 and 2019. All applicants with predicted nonplacement risk of 30% or higher received this warning when they submitted their choice application. See section 3.2 for details.

Figure 2: Timeline of feedback interventions - 2020



Sequence of 2020 application feedback for risky applicants. All text translated to English. The platform pop-up on the right was shown to all risky applicants at the time they submitted their application. The SMS and WhatsApp messages shown at center were sent to (subgroups of) still-risky applicants based on contemporaneous risk predictions on the day of the application cycle listed on horizontal axis, where day 28 is the final deadline for application submission. The SMS messages on day 20 and 27 were sent to all risky applicants, while the WhatsApp image at center was sent to randomly selected applicants on day 25. The schools displayed in the WhatsApp image are those the student listed on her choice application. See section 3.2 for details.

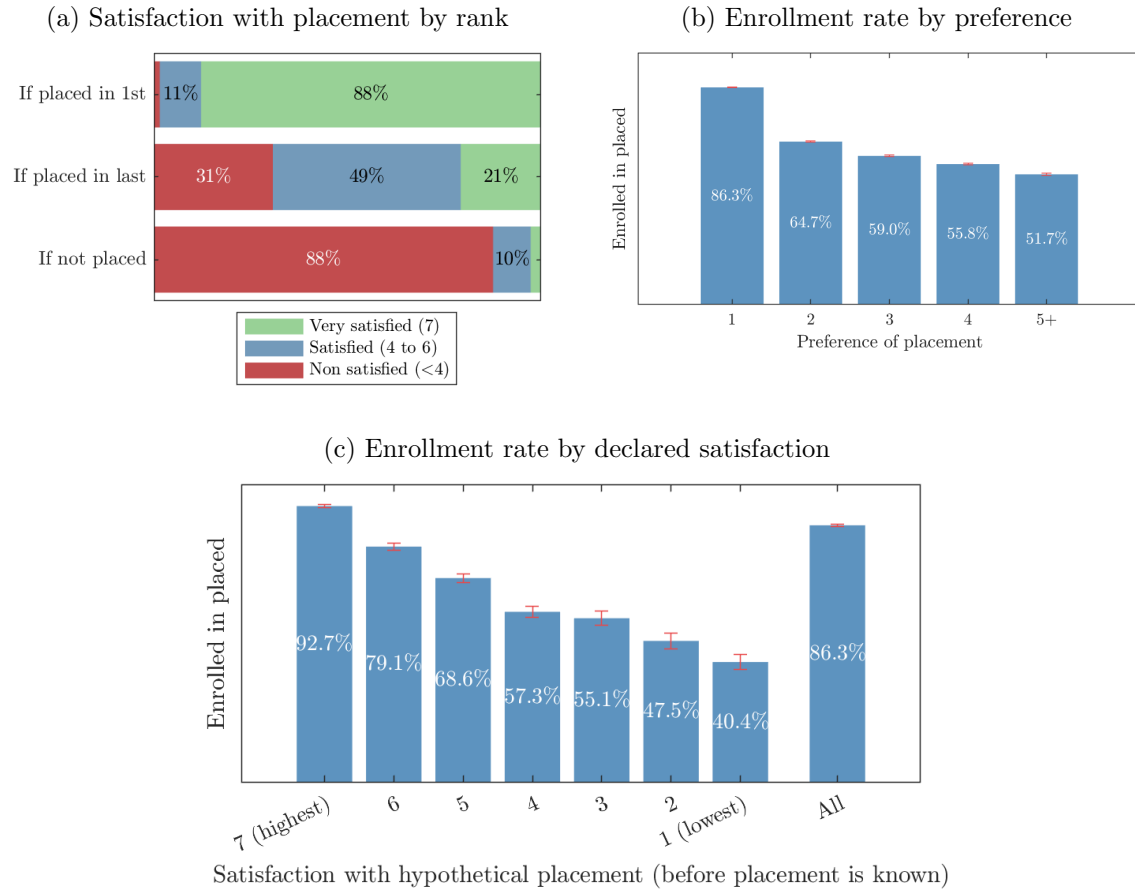
Figure 3: Distribution of placement probabilities and probability predictions



Panel A: binned means, linear fit and interquartile range of predicted placement probability by true placement probability. Points are centered means of 10 quantile-spaced bins of the support of the true placement probability  $\in [0.00; 0.99]$ . The last point at the right represents the mean of predicted placement probability for observations with true probability greater than 0.99. Placement predictions in Panel A combine observed applications at the time an individual submits her application with historical projections. See section 3.3 for details. Panel B: histogram of true placement probability for initial application attempt and final application submission. Vertical lines display means.

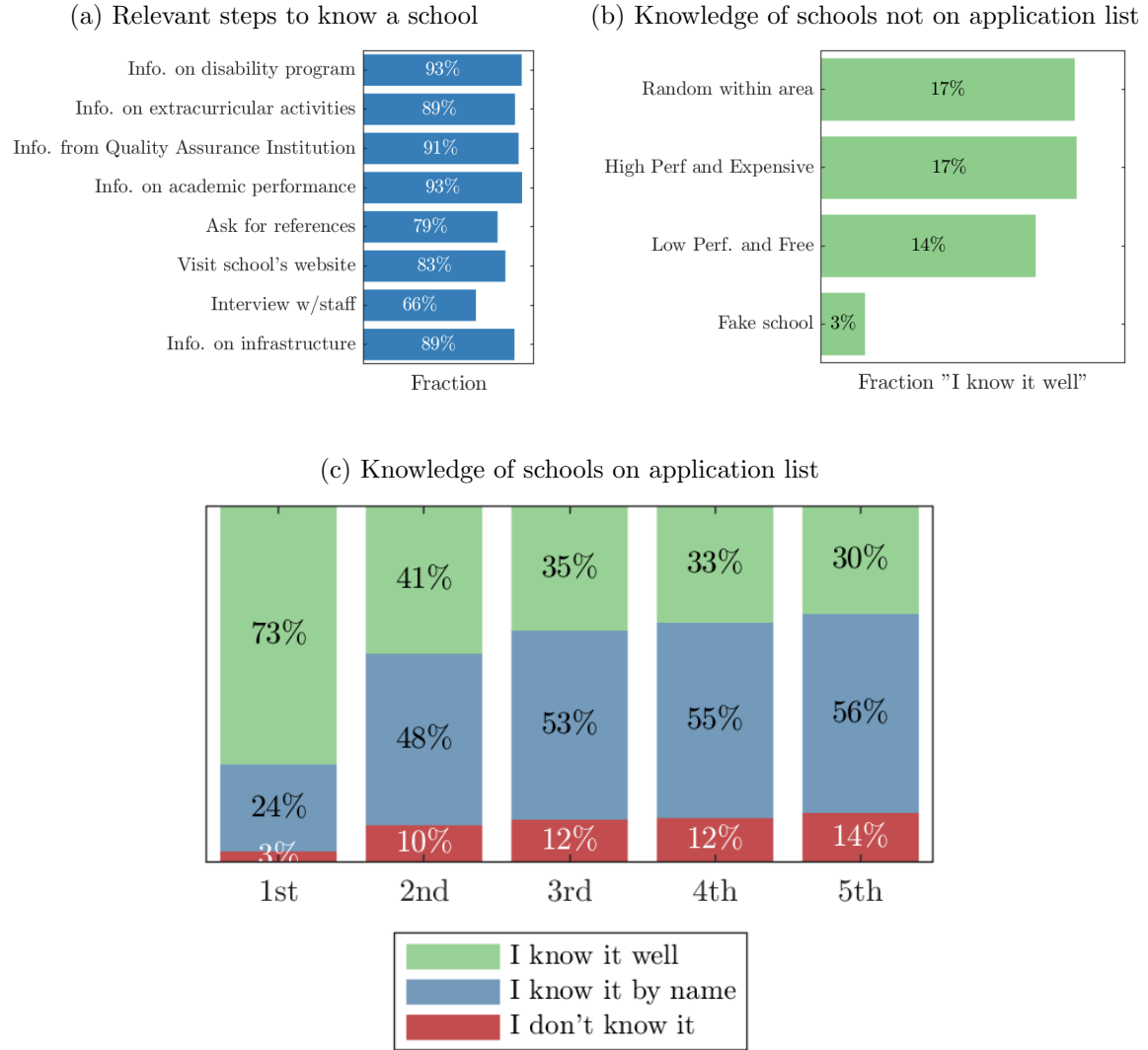


Figure 4: Satisfaction with placement outcomes



Panel A: stated satisfaction with hypothetical placement outcomes. Data are survey responses to questions about applicant satisfaction with being placed at their first-ranked school, last-ranked school, and nonplacement. Sample: survey completers. Results reported on a 1-7 scale, with 7 being very satisfied and 1 being not at all satisfied. Panel B: rates at which students enroll in the placed school, by rank of placed school. Unplaced students are not included. Sample: all placed students. Panel C: rate at which students enroll in the placed school, by survey reports of satisfaction with the placed school. Sample: survey completers who place in their first- or last-ranked school. See section 4 for details.

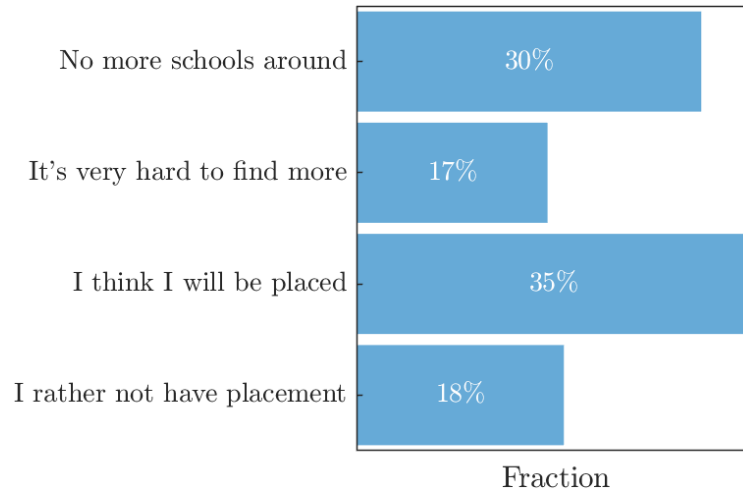
Figure 5: Knowledge of and search for schools



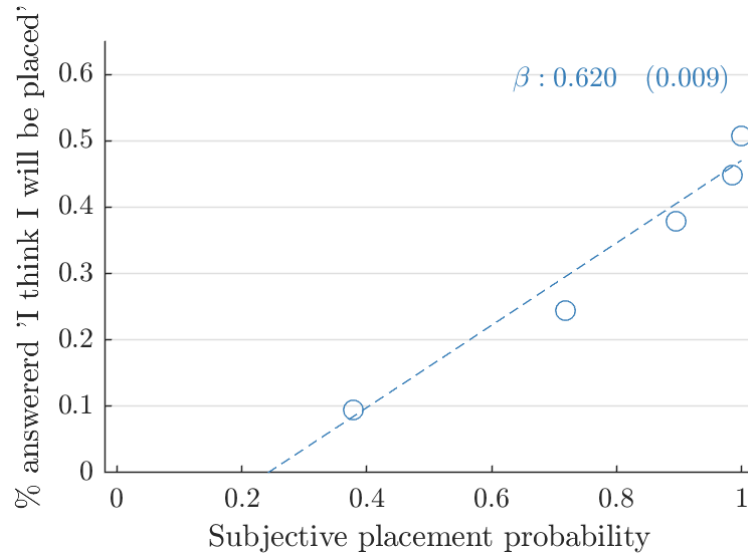
Panel A: share of survey respondents stating an understanding of listed attribute was relevant for “know[ing] a school well”. “Info from Quality Assurance Institution” is information on academic performance and other indicators not related to standardized tests from education regulators in charge of the evaluation of schools. Panel B: share of students stating that they “know well” schools not listed on their application, for schools of type listed on horizontal axis. All schools are within an applicant’s local area, defined as 2km from student’s location (home address reported on platform, replaced with centroid of application if geocoding was unreliable). “High performing and expensive schools” are those classified in 2 best tiers of performance (out of 4) by the Quality Assurance Institution, with a monthly copayment of \$35 USD or more. “Low performing and free” schools are defined as schools within the worst tier of performance, with no copayment. “Fake schools” are schools that do not exist in the student’s local area. Panel C: stated knowledge of schools on application list, by rank. See section 4.2 for details.

Figure 6: Reasons for stopping school search

(a) Stated reason for not adding more schools

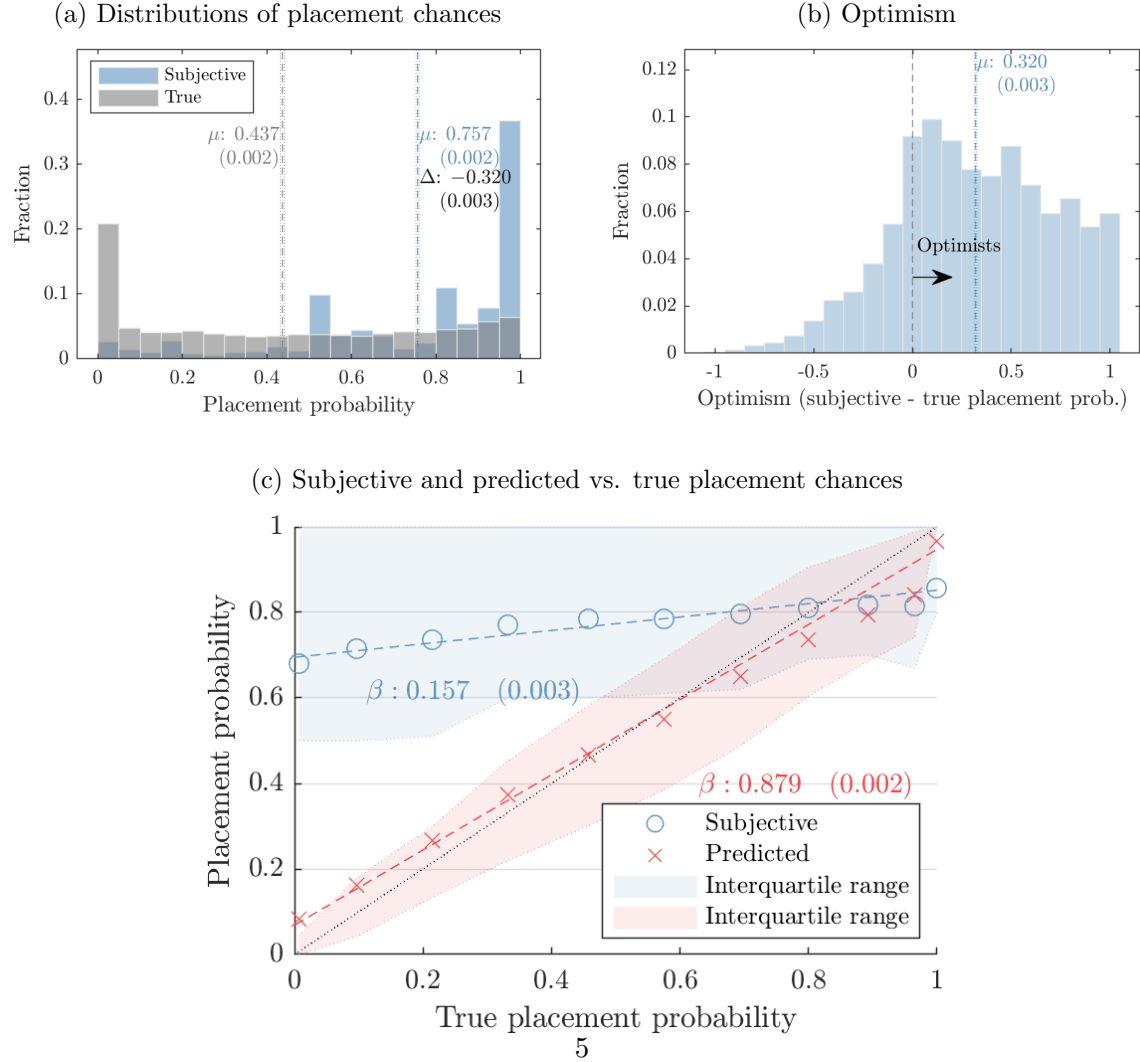


(b) Stated reason is "I will be placed" vs declared risk



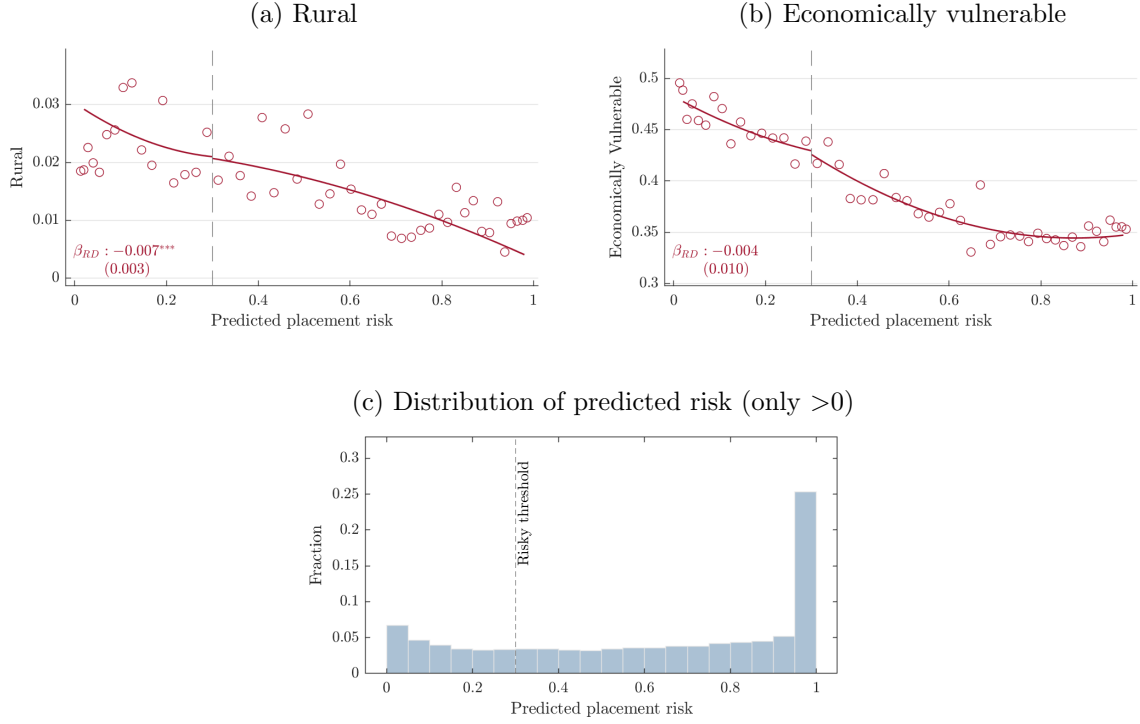
Panel A: survey reports of reason for not adding more schools to the choice application. Panel B: share of survey completers stating that they stopped search because they think they will be placed, by survey report of subjective placement probability. Sample in both panels: survey completers.

Figure 7: Subjective vs. observed application risk



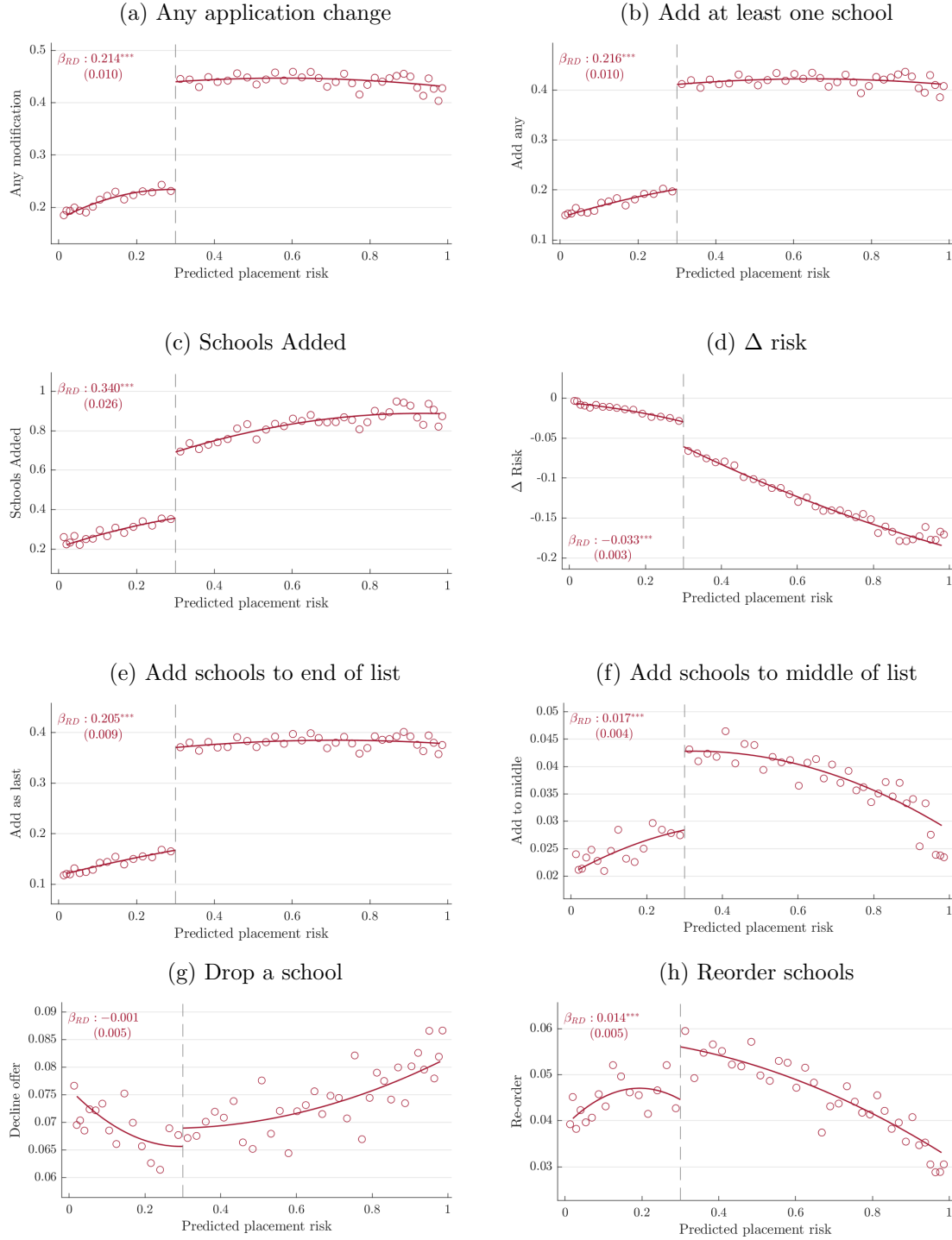
Panel A: distribution of true placement chances and survey-reported subjective placement chances. Vertical lines display means of each distribution. Panel B: distribution of optimism, defined as difference between subjective and true placement chances. Panel C: mean subjective placement belief within bins defined by true placement probability. The bottom bin includes applicants with placement probability less than 1%, and the top bin includes applicants with placement probability of 99% or more. The middle eight groups split the remaining observations into equally-sized bins. Dashed line is linear fit. Shaded areas are IQRs for subjective beliefs and risk predictions (within survey sample). 45-degree line displayed for reference. Sample: survey completers.

Figure 8: Balance in platform pop-up RD



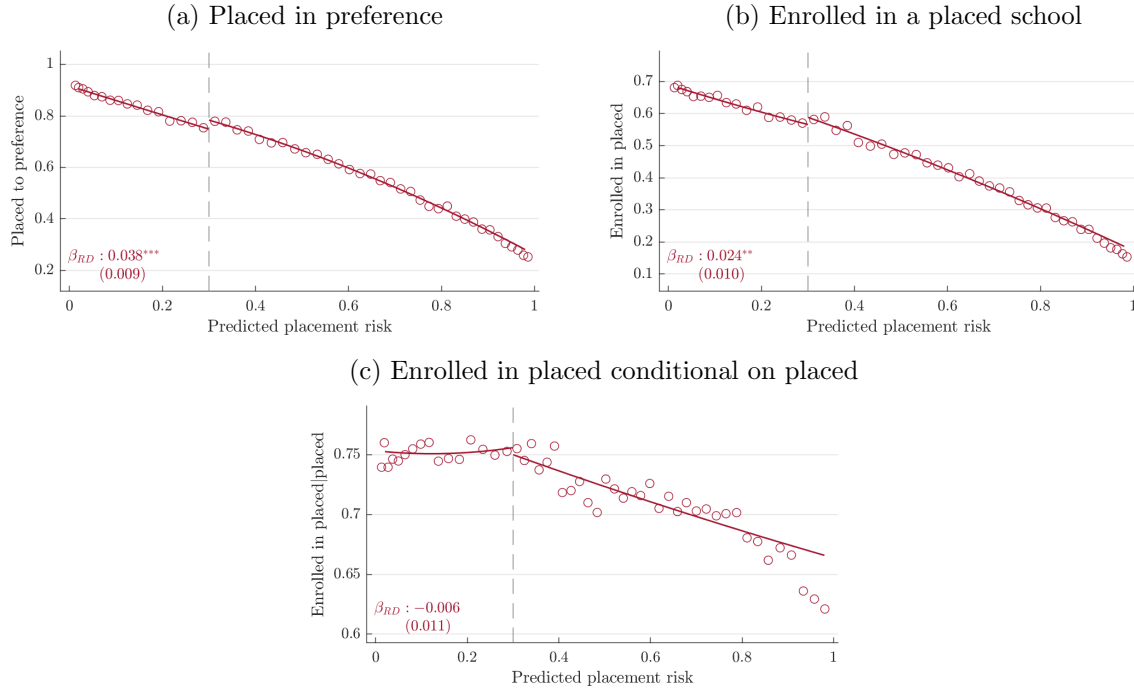
Binned means and global fits of predetermined characteristics by predicted risk for initial application. Points are centered means of 50 quantile-spaced bins of the support of the predicted placement risk  $\in [0.02; 0.98]$ . Solid line shows the quadratic fit. Reported coefficients and standard errors are from local linear specifications using  $\pm 0.1$  bandwidth. See section 5.1 for details. Because coefficients are local while polynomial fits are global, there may be minor differences between displayed fits and reported coefficients. Panel A: vertical axis is indicator for rural location. Panel B: vertical axis is indicator for economic vulnerability (a measure of socioeconomic status). Panel C: histogram of predicted placement risk for initial application attempt, conditional on being greater than 0.01. Vertical lines display the 30% risk cutoff.

Figure 9: Choice behavior and risk reduction in platform pop-up RD



Binned means and global fits of choice outcomes by predicted risk for initial application. Points are centered means of 50 quantile-spaced bins of the support of the predicted placement risk  $\in [0.02; 0.98]$ . Solid line shows the quadratic fit. Reported coefficients and standard errors are from local linear specifications using  $+ - 0.1$  bandwidth. See section 5.1 for details. Because coefficients are local while polynomial fits are global, there may be minor differences between displayed fits and reported coefficients. Outcomes by panel are as follows. Panel A: any application change. Panel B: add any school to application. Panel C: count of schools added. Panel D: change in risk from initial to final application. Panel E: add at least one school to end of choice application. Panel F: add at least one school to middle of choice application. Panel G: drop at least one school from application. Panel H: reorder existing schools.

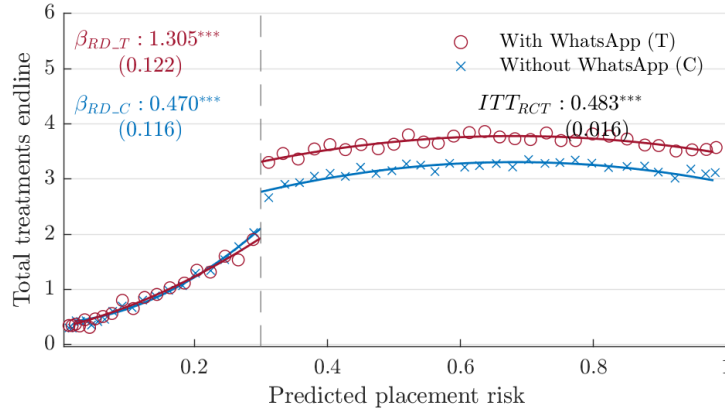
Figure 10: Placement and enrollment outcomes in platform pop-up RD



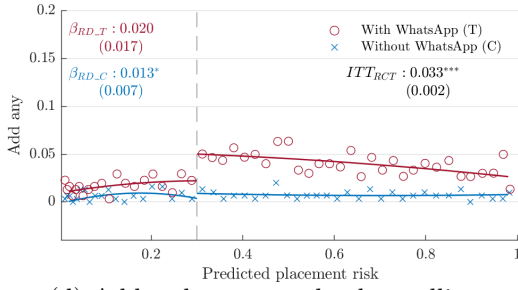
Binned means and global fits of placement and enrollment outcomes by predicted risk for initial application. Points are centered means of 50 quantile-spaced bins of the support of the predicted placement risk  $\in [0.02; 0.98]$ . Solid line shows the quadratic fit. Reported coefficients and standard errors are from local linear specifications using  $\pm 0.1$  bandwidth. See section 5.1 for details. Because coefficients are local while polynomial fits are global, there may be minor differences between displayed fits and reported coefficients. Outcomes by panel are as follows. Panel A: observed placement in any school. Panel B: receive placement and enroll in placed school. Panel C: enroll in placed school conditional on placement.

Figure 11: WhatsApp RCT outcomes

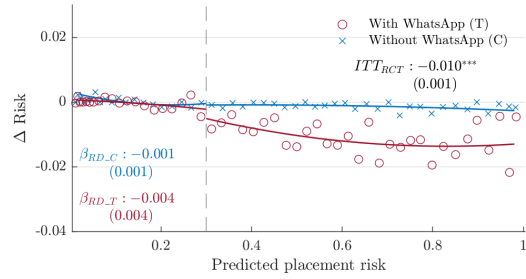
(a) Count of feedback messages received



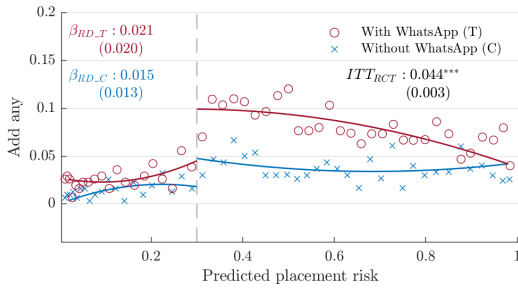
(b) Add at least one school – 44 hours



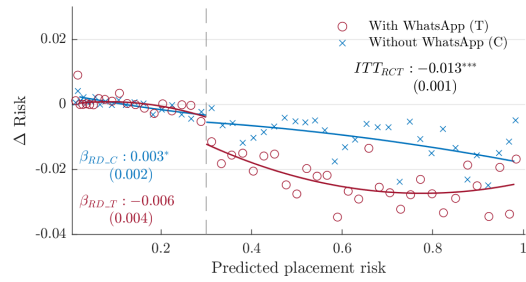
(c) Change in risk – 44 hours



(d) Add at least one school – endline



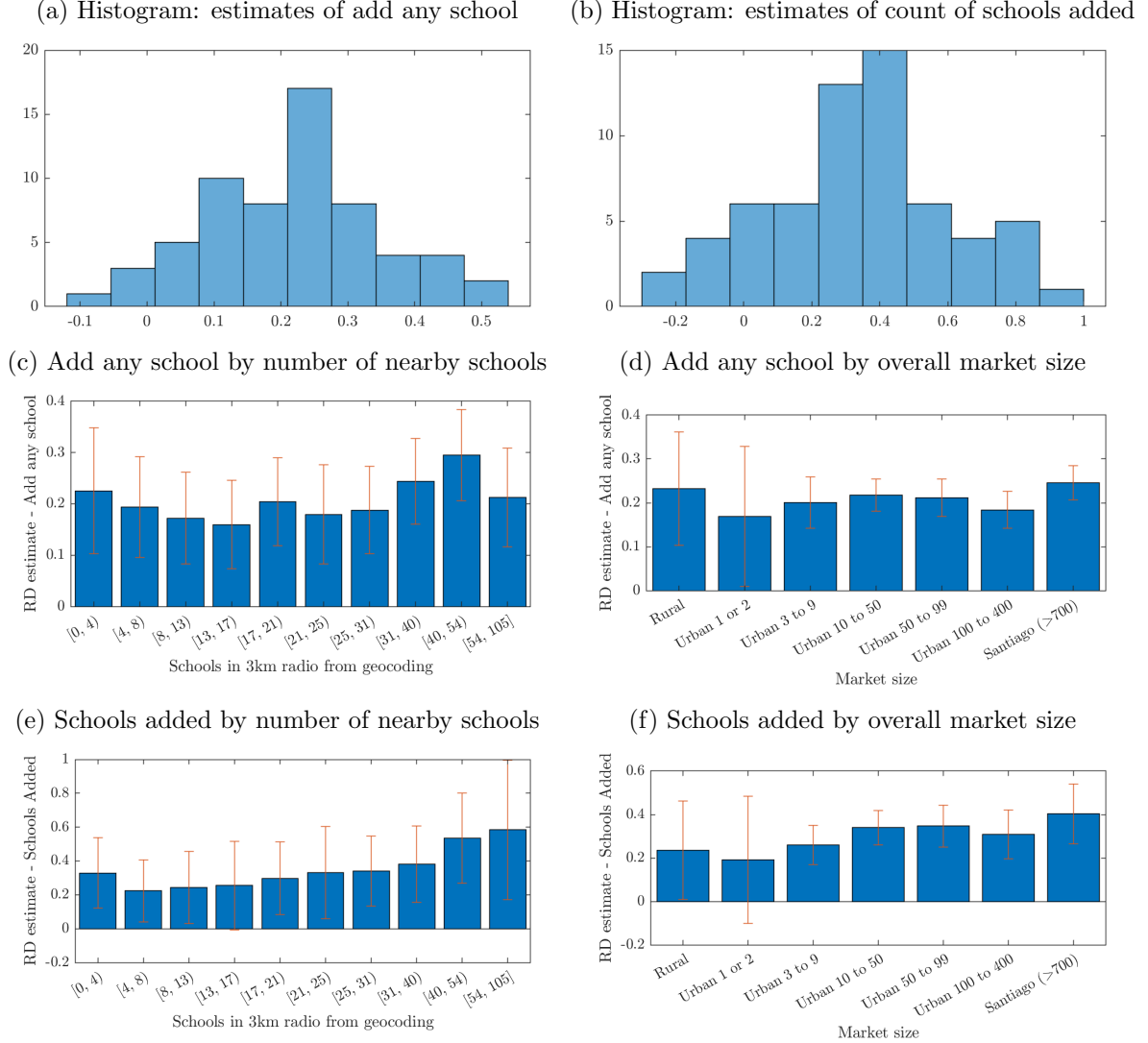
(e) Change in risk – endline



Binned means and global fits of message receipt, application behavior, and risk outcomes by predicted placement risk in RCT sample. Points are centered means of 50 quantile-spaced bins of the support of the predicted placement risk  $\in [0.02; 0.98]$ . Solid lines show the quadratic fit. Figures split by RCT treatment and control group, above and below treatment threshold. “With WhatsApp” group receives WhatsApp warning when above cutoff. “Without WhatsApp” group receives no warning. Below 0.30 predicted risk cutoff, the treatment group receives WhatsApp message with no risk-related information. Reported  $\beta_{RD}$  coefficients are RD estimates within treatment and control group, computed from local linear specifications using  $\pm 0.1$  bandwidth. See section 5.1 for details. Because coefficients are local while polynomial fits are global, there may be minor differences between displayed fits and reported coefficients.. Reported  $ITT_{RCT}$  estimate is the experimental RCT effect for all above-cutoff students on the listed outcome. Outcomes, listed in panel titles, are as follows. Panel A: count of warnings messages received over full application period. Panel B: add any school in 44-hour window between WhatsApp message and SMS followup. Panel C: change in risk within 44-hour window between WhatsApp message and application followup. Panel D: add any school between Whatsapp message and application close. Panel E: change in risk by application close. See section 5.5 for details.

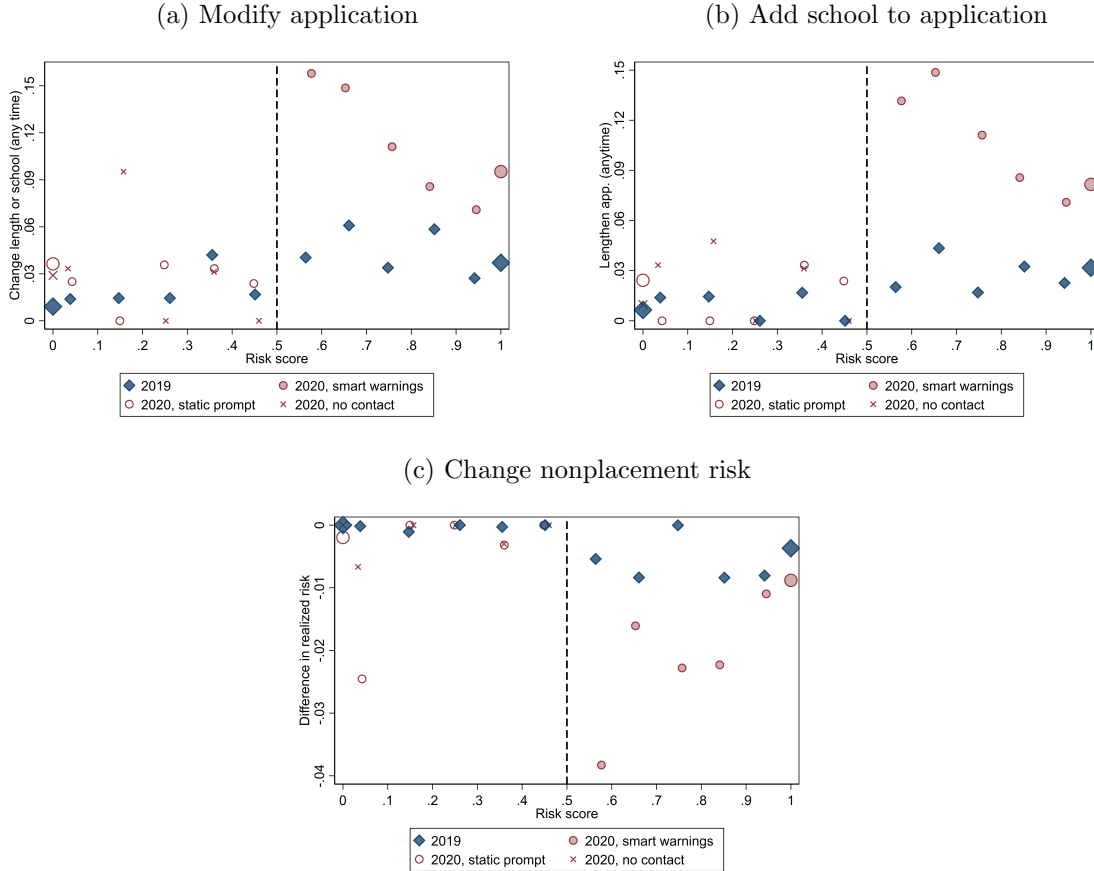


Figure 12: Platform pop-up effects over city-years and by market size



Panels A and B: distribution of estimated platform pop-up RD effects across city-year cells. Each city-year cell is one observation. Outcome in Panel A is add any school, outcome in panel B is count of schools added. Panel C: pop-up RD treatment effects on add any school split by count of nearby schools (within 3km of applicant address). Panel D: pop-up RD treatment effects on add any school split by overall market size, with size defined by the number of schools available to students in the city-year-grade cell and urban/rural status. Panel E: same as C, but with count of schools added as the outcome. Panel F: same as D, but with count of schools added as the outcome. See section 5.6 for details.

Figure 13: Smart and static warnings in New Haven



Outcomes of warnings intervention in New Haven centralized choice system. Figures show changes in application behavior by risk score as of 7 days prior to application deadline in 2019 and 2020. Points are centered binned means within intervals of width 0.1, except for top- and bottom-most points, which are for students with risk scores of 1 and 0, respectively. In 2020, all applicants with risk scores above 0.5 received the warnings intervention. Randomly chosen applicants with risk scores below 0.5 received a “static” prompt that provided a non-personalized message encouraging them to learn more about their assignment chances; the remaining non-risky applicants received no intervention. In 2019, no applicant received any intervention. Panel A: any change in application. Panel B: lengthen application. Panel C: change in risk from initial to final portfolio. See section 5.7 for details.

# Tables

Table 1: Descriptive statistics for Chilean choice applicants

	(1) All	(2) Economically Vulnerable	(3) Not Economically Vulnerable	(4) Pop-up eligible	(5) Risky (predicted risk>.3)	(6) Around Pop-up Cutoff	(7) RCT. sample (2020)	(8) Survey sample (2020)
N	1,168,706	575,521	593,185	848,795	233,678	84,517	19,213	48,929
%	1.00	0.49	0.51	0.73	0.20	0.07	0.02	0.04
<b>A. Demographics</b>								
Economically Vulnerable	0.49	1.00	0.00	0.51	0.37	0.42	0.25	0.42
Rural	0.05	0.07	0.03	0.06	0.02	0.02	0.00	0.04
<b>B. Rank-order list</b>								
Length initial attempt	2.77	2.61	2.93	2.70	2.36	3.04	2.79	2.74
Length final attempt	3.14	2.92	3.36	3.06	3.20	3.57	3.32	3.22
<b>C. Application behavior</b>								
Total attempts	1.41	1.35	1.46	1.38	1.74	1.51	1.53	1.45
Any modification	0.25	0.22	0.27	0.24	0.43	0.33	0.33	0.28
Add any	0.21	0.19	0.23	0.21	0.41	0.30	0.30	0.25
Add as last	0.18	0.16	0.20	0.18	0.38	0.26	0.27	0.22
Add to middle	0.03	0.02	0.03	0.02	0.03	0.03	0.03	0.03
Add as first	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.02
Change order	0.04	0.03	0.05	0.03	0.04	0.05	0.04	0.04
Change top 1	0.05	0.04	0.05	0.04	0.05	0.05	0.04	0.04
Delete any	0.05	0.04	0.05	0.04	0.03	0.04	0.04	0.04
<b>D. Placement</b>								
Placed in pref.	0.79	0.84	0.74	0.80	0.45	0.77	0.42	0.79
Placed 1st	0.54	0.61	0.47	0.56	0.18	0.39	0.17	0.53
Placed 2nd	0.13	0.13	0.14	0.13	0.12	0.19	0.12	0.14
Placed 3rd	0.06	0.06	0.07	0.06	0.08	0.10	0.07	0.06
<b>E. 2nd round</b>								
Particip. in 2nd round	0.09	0.08	0.10	0.08	0.15	0.12	0.20	0.09
Placed in 2nd round	0.07	0.06	0.07	0.06	0.11	0.09	0.16	0.07
<b>F. School capacity available after placement (at local market level defined for each student)</b>								
Share of total seats	0.42	0.41	0.42	0.42	0.50	0.39	0.44	0.50
Share of seats in free schools	0.46	0.45	0.47	0.47	0.55	0.44	0.52	0.55
<b>G. Classification by true risk of initial attempt</b>								
Mean risk	0.24	0.18	0.30	0.23	0.69	0.24	0.61	0.25
Zero risk	0.59	0.67	0.51	0.62	0.05	0.19	0.02	0.59
Risky (risk>.3)	0.30	0.23	0.37	0.29	0.86	0.37	0.84	0.31
.25 quantile  >0	0.28	0.23	0.32	0.29	0.52	0.15	0.41	0.30
.50 quantile  >0	0.62	0.56	0.66	0.64	0.79	0.28	0.63	0.65
.75 quantile  >0	0.92	0.91	0.93	0.94	0.99	0.42	0.87	0.94

*Notes:* N: 1,168,706 (20% from 2018, 41% from 2019 and 39% from 2020). All statistics are means in the population defined by the column header. "Pop-up eligible" (col. 4) are students who submitted applications that received a risk prediction. "Risky" (col. 5) is applicants whose first attempt had a predicted risks > 0.3. "Around pop-up cutoff" (col. 6) are applicants whose first attempt had a predicted risk in [0.1,0.5]. "RCT sample" (column 7) is applicants in treatment or control group of the 2020 RCT design. "Survey sample" (column 8) is applicants who completed the 2020 school choice survey. Selected row variable definitions are as follows. "Economically vulnerable" is an SES measure computed by Mineduc. "Rural" is an indicator if students live in rural areas. "Length of initial/final attempt" is the number of schools on an applicants first and final choice application. "Total attempts" is the number of times an applicant submitted an application to the centralized system. Application change and addition variables describe the share of applicants making different kinds of changes between their first and final submission. "Placed in pref/1st/2nd/3rd" are indicators for any placement or for placement in the listed rank. "2nd round" variables describe participation and placement outcomes in the second centralized placement round. "Share of total seats/seats in free schools" is the share of seats in all schools/in schools without fees unfilled after the first application round in a student's local market. True risk of initial attempt variables describe the nonplacement risk of an applicant's initial application, evaluated using ex post observed applications.

Table 2: RD estimates of platform pop-up effects

	(1)	(2)	(3)	(4)	(5)
	All		2018	2019	2020
IV					
A. Balance					
Economically Vulnerable	-0.004		-0.014	0.016	-0.012
	(0.010)		(0.029)	(0.018)	(0.013)
Rural	-0.007		-0.002	-0.009	-0.008
	(0.003)		(0.007)	(0.005)	(0.003)
B. Choice Behavior					
Any modification	0.214		0.164	0.217	0.224
	(0.010)		(0.025)	(0.018)	(0.013)
Add any	0.216		0.176	0.224	0.223
	(0.010)		(0.024)	(0.018)	(0.013)
Schools Added	0.340	1.573	0.379	0.317	0.344
	(0.026)	(0.090)	(0.068)	(0.050)	(0.033)
$\Delta$ Risk	-0.033	-0.155	-0.039	-0.040	-0.029
	(0.003)	(0.013)	(0.009)	(0.007)	(0.004)
Add as first	-0.003	-0.012	-0.007	-0.005	-0.000
	(0.003)	(0.013)	(0.008)	(0.005)	(0.003)
Add to middle	0.017	0.078	0.017	0.023	0.014
	(0.004)	(0.018)	(0.012)	(0.007)	(0.005)
Add as last	0.205	0.949	0.172	0.207	0.213
	(0.009)	(0.018)	(0.023)	(0.017)	(0.012)
Drop any	-0.001	-0.003	-0.009	0.018	-0.008
	(0.004)	(0.019)	(0.010)	(0.008)	(0.005)
Re-order	0.014	0.065	0.026	0.005	0.015
	(0.005)	(0.022)	(0.013)	(0.009)	(0.006)
C. Choice outcome					
Placed to preference	0.038	0.178	0.033	0.086	0.020
	(0.009)	(0.041)	(0.026)	(0.018)	(0.011)
Enrolled in placed	0.024	0.113	0.008	0.055	0.018
	(0.010)	(0.049)	(0.029)	(0.020)	(0.013)
Enrolled in placed placed	-0.006	-0.025	-0.021	-0.009	0.003
	(0.011)	(0.045)	(0.031)	(0.022)	(0.013)
D. Congestion-related outcomes					
Add any uncongested	0.073	0.339	0.052	0.081	0.075
	(0.007)	(0.026)	(0.016)	(0.012)	(0.009)
$\Delta$ prob. placed to uncongested	0.019	0.088	0.015	0.032	0.014
	(0.003)	(0.014)	(0.008)	(0.007)	(0.004)
NL	20,359	20,359	2,834	6,076	11,449
NR	21,145	21,145	2,776	6,015	12,354

*Notes:* Local linear RD estimates of pop-up effects from warning pop-up on application platform. Computed using triangular kernel with bandwidth 0.1. Heteroskedasticity-robust nearest neighbor variance estimator with minimum of 3 neighbors reported in parentheses; computed as in [Calonico et al. \(2014\)](#). We report estimates in the pooled sample and for each year. IV (column 2) shows the instrumental variable specifications, where the endogenous regressor is the add any school indicator. Panel A: predetermined covariates. Panel B: measures of choice behavior from initial to final application.  $\Delta$  risk is change in application risk from first to final attempt. “Add to X” are additions of schools in given place on list, relative to initial application submission. Panel C: outcomes of choice process. “Enrolled in placed” is equal to one for students who receive a placement and enroll in the placed school. “Enrolled in placed | placed” is the enrollment rate in the placed school, conditional on receiving a placement. Panel D: congestion attributes of behavior and placement outcomes. “Uncongested” schools are those with excess capacity. NL and NR are sample counts to the left and right of the cutoff, respectively.

Table 3: WhatsApp RD and RCT results

	(1)	(2)	(3)	(4)
	RCT		RD	
	ITT	IV	ITT	IV
<b>A. Balance</b>				
Economically Vulnerable	-0.019 (0.006)		-0.012 (0.039)	
<b>B. Message receipt</b>				
WhatsApp read	0.466 (0.005)		0.528 (0.030)	
SMS reminder received	-0.028 (0.004)		0.459 (0.034)	
Total treatments before final SMS	0.506 (0.016)		0.845 (0.116)	
Total treatments endline	0.483 (0.016)		1.305 (0.122)	
<b>C. Outcomes in clean 44 hours before SMS followup</b>				
Any modification	0.035 (0.002)		0.015 (0.017)	
Add any	0.033 (0.002)		0.020 (0.017)	
Schools Added	0.075 (0.007)	2.281 (0.136)	0.103 (0.042)	5.260 (3.194)
$\Delta$ Risk	-0.010 (0.001)	-0.297 (0.018)	-0.004 (0.004)	-0.209 (0.131)
<b>D. Endline outcomes</b>				
Any modification	0.046 (0.004)		0.012 (0.021)	
Add any	0.044 (0.003)		0.021 (0.020)	
Schools Added	0.112 (0.011)	2.550 (0.175)	0.138 (0.065)	6.681 (4.764)
$\Delta$ Risk	-0.013 (0.001)	-0.301 (0.022)	-0.006 (0.004)	-0.307 (0.206)
Placed to preference	0.022 (0.007)	0.495 (0.168)	0.059 (0.032)	2.851 (3.204)

*Notes:* ITT and IV effects of 2020 WhatsApp warnings intervention. RCT columns: effects of random assignment to treatment group vs. control group for students with predicted risk > 0.30. Robust SEs in parentheses. N=17,970. RD columns: regression discontinuity evaluation in treatment group around 0.30 cutoff. RD specifications computed using local linear fit with a bandwidth of 0.1. Standard errors are heteroskedasticity-robust nearest neighbor variance estimator with minimum of 3 neighbors, as in [Calonico et al. \(2014\)](#). ITT column shows effects of group assignment. IV columns show the instrumental variable specification, where the endogenous regressor is the add any school indicator, instrumented with group assignment for the RCT, and with a dummy of crossing the risky threshold for the RD. Panel A: balance tests on predetermined characteristics. Panel B: message receipt. “WhatsApp read” is an indicator equal to one if applicant views the WhatsApp treatment message. “SMS remainder received” is indicator for receiving SMS reminder 44 hours later. Panel C: outcomes within 44 hour window between WhatsApp intervention and followup SMS. Panel D: endline choice behavior and placement outcomes. See section 5.5 for details.

Table 4: RD estimates of platform pop-up effects on adding any school, by city and year

City	2020 applicants	2018	2019	2020
Santiago	158,057		0.24 (0.04)	0.25 (0.02)
Viña - Valparaíso	26,215	0.01 (0.08)	0.28 (0.07)	0.22 (0.05)
Concepción - Talcahuano	24,548	0.21 (0.08)	0.15 (0.06)	0.25 (0.05)
Coquimbo - La Serena	13,994	0.18 (0.10)	0.38 (0.10)	0.11 (0.07)
Rancagua	11,971	0.16 (0.10)	0.09 (0.09)	0.06 (0.07)
Antofagasta	12,722	0.24 (0.14)	0.36 (0.09)	0.23 (0.07)
Iquique - Alto Hospicio	10,251	0.25 (0.09)	0.23 (0.09)	0.25 (0.07)
Temuco	10,176	0.22 (0.10)	0.31 (0.08)	0.29 (0.06)
Puerto Montt - Puerto Varas	8,864	0.31 (0.15)	0.28 (0.08)	-0.02 (0.09)
Talca - San Clemente	8,913	-0.03 (0.13)	0.11 (0.09)	0.17 (0.07)
Arica	5,905	0.10 (0.16)	0.48 (0.12)	0.14 (0.13)
Curicó	6,827	0.11 (0.15)	0.26 (0.14)	0.18 (0.10)
Chillán	5,536	0.39 (0.26)	0.21 (0.10)	0.09 (0.09)
Los Andes - San Felipe	5,006	0.11 (0.32)	0.03 (0.24)	0.42 (0.13)
Los Ángeles	5,477	0.45 (0.13)	0.02 (0.16)	0.34 (0.11)
Calama	5,565	0.00 (0.21)	0.32 (0.17)	0.08 (0.10)
Copiapó	6,181	0.23 (0.13)	0.53 (0.11)	0.33 (0.08)
Osorno	4,542	0.04 (0.12)	0.25 (0.16)	0.23 (0.16)
Valdivia	4,599	0.10 (0.23)	0.37 (0.12)	0.13 (0.18)
Algarrobo a San Antonio	4,705	0.43 (0.15)	-0.10 (0.16)	0.45 (0.11)
Chile	454,226	0.18 (0.02)	0.22 (0.02)	0.22 (0.01)

*Notes:* RD estimates of smart platform pop-up effects on adding at least one school to the choice application, split by city and year. Cities are sorted by count of 2020 applicants. Santiago is not displayed for 2018 because centralized admission had not yet been rolled out. Estimates from local linear specifications, computed using triangular kernel with bandwidth 0.1. Heteroskedasticity-robust nearest neighbor variance estimator with minimum of 3 neighbors reported in parentheses; computed as in [Calonico et al. \(2014\)](#). See section 5.6 for details.

Table 5: RD estimates of platform pop-up effects by market-level choice experience

	(1)	(2)	(3)	(4)	(5)	(6)
	1st year		2nd year		3rd+ year	
	IV		IV		IV	
<b>A. Balance</b>						
Economically Vulnerable	-0.011		0.023		-0.029	
	(0.024)		(0.015)		(0.015)	
Rural	-0.001		-0.008		-0.009	
	(0.006)		(0.004)		(0.004)	
<b>B. Choice Behavior</b>						
Any modification	0.181		0.235		0.206	
	(0.023)		(0.016)		(0.015)	
Add any	0.194		0.232		0.210	
	(0.022)		(0.015)		(0.014)	
Schools Added	0.319	1.648	0.370	1.596	0.318	1.517
	(0.075)	(0.311)	(0.043)	(0.142)	(0.031)	(0.104)
$\Delta$ Risk	-0.037	-0.190	-0.032	-0.138	-0.033	-0.159
	(0.009)	(0.039)	(0.005)	(0.018)	(0.004)	(0.018)
Add as first	-0.011	-0.056	-0.004	-0.019	0.003	0.015
	(0.007)	(0.039)	(0.005)	(0.020)	(0.004)	(0.017)
Add to middle	0.028	0.147	0.017	0.072	0.012	0.057
	(0.010)	(0.049)	(0.006)	(0.026)	(0.006)	(0.027)
Add as last	0.184	0.951	0.222	0.956	0.197	0.940
	(0.021)	(0.050)	(0.015)	(0.026)	(0.014)	(0.028)
Drop any	0.004	0.021	-0.008	-0.033	0.004	0.021
	(0.010)	(0.052)	(0.007)	(0.029)	(0.006)	(0.030)
Re-order	0.007	0.038	0.021	0.091	0.010	0.046
	(0.011)	(0.059)	(0.008)	(0.033)	(0.007)	(0.033)
<b>C. Choice outcome</b>						
Placed to preference	0.059	0.306	0.049	0.211	0.017	0.083
	(0.023)	(0.119)	(0.014)	(0.060)	(0.013)	(0.063)
Enrolled in placed	0.009	0.046	0.041	0.177	0.014	0.067
	(0.025)	(0.130)	(0.016)	(0.071)	(0.016)	(0.077)
Enrolled in placed placed	-0.045	-0.214	0.006	0.025	0.000	0.001
	(0.028)	(0.131)	(0.017)	(0.069)	(0.015)	(0.066)
<b>D. Congestion-related outcomes</b>						
Add any uncongested	0.059	0.303	0.075	0.322	0.079	0.376
	(0.015)	(0.068)	(0.011)	(0.040)	(0.010)	(0.041)
$\Delta$ prob. placed to uncongested	0.025	0.127	0.015	0.065	0.021	0.101
	(0.009)	(0.042)	(0.005)	(0.021)	(0.005)	(0.021)
NL	3,819	3,819	8,573	8,573	7,967	7,967
NR	3,571	3,571	8,880	8,880	8,694	8,694

*Notes:* Local linear RD estimates of pop-up effects from warning pop-up on application platform, split by years elapsed since city-grade combination first began using the centralized choice process. Computed using triangular kernel with bandwidth 0.1. Heteroskedasticity-robust nearest neighbor variance estimator with minimum of 3 neighbors reported in parentheses; computed as in [Calonic et al. \(2014\)](#). IV estimates reported in second column of each set show the instrumental variable specifications (fuzzy RD), where the endogenous regressor is the add any school indicator. Panel A: predetermined covariates. Panel B: measures of choice behavior from initial to final application.  $\Delta$  risk is change in application risk from first to final attempt. “Add to X” are additions of schools in given place on list, relative to initial application submission. Panel C: outcomes of choice process. “Enrolled in placed” is equal to one for students who receive a placement and enroll in the placed school. “Enrolled in placed | placed” is the enrollment rate in the placed school, conditional on receiving a placement. Panel D: congestion attributes of behavior and placement outcomes. “Uncongested” schools are those with excess capacity. NL and NR are sample counts to the left and right of the cutoff, respectively.

Table 6: RD estimates of platform pop-up effects by applicant’s socioeconomic status

	(1)	(2)	(3)	(4)
	Vulnerable		Not vulnerable	
	IV		IV	
<b>A. Balance</b>				
Economically Vulnerable	0		0	
	(0.000)		(0.000)	
Rural	-0.012		-0.004	
	(0.004)		(0.003)	
<b>B. Choice Behavior</b>				
Any modification	0.225		0.206	
	(0.015)		(0.013)	
Add any	0.227		0.209	
	(0.015)		(0.012)	
Schools Added	0.327	1.445	0.350	1.673
	(0.035)	(0.115)	(0.036)	(0.133)
$\Delta$ Risk	-0.041	-0.179	-0.028	-0.136
	(0.005)	(0.020)	(0.004)	(0.016)
Add as first	-0.001	-0.005	-0.004	-0.017
	(0.004)	(0.020)	(0.003)	(0.016)
Add to middle	0.013	0.057	0.020	0.095
	(0.005)	(0.023)	(0.006)	(0.025)
Add as last	0.217	0.957	0.197	0.943
	(0.014)	(0.025)	(0.012)	(0.025)
Drop any	-0.003	-0.014	0.001	0.006
	(0.006)	(0.028)	(0.006)	(0.027)
Re-order	0.020	0.090	0.009	0.045
	(0.007)	(0.031)	(0.006)	(0.031)
<b>C. Choice outcome</b>				
Placed to preference	0.019	0.085	0.052	0.251
	(0.014)	(0.062)	(0.012)	(0.056)
Enrolled in placed	0.008	0.036	0.036	0.173
	(0.016)	(0.071)	(0.014)	(0.067)
Enrolled in placed placed	-0.009	-0.035	-0.004	-0.017
	(0.016)	(0.065)	(0.014)	(0.062)
<b>D. Congestion-related outcomes</b>				
Add any uncongested	0.071	0.315	0.075	0.359
	(0.011)	(0.040)	(0.009)	(0.035)
$\Delta$ prob. placed to uncongested	0.020	0.090	0.018	0.088
	(0.005)	(0.022)	(0.004)	(0.018)
NL	8,878	8,878	11,481	11,481
NR	8,721	8,721	12,424	12,424

*Notes:* Local linear RD estimates of pop-up effects from warning pop-up on application platform, split by applicant socioeconomic status. “Economically vulnerable” individuals are the lower-SES group. See section 3 for details. Computed using triangular kernel with bandwidth 0.1. Heteroskedasticity-robust nearest neighbor variance estimator with minimum of 3 neighbors reported in parentheses; computed as in Calonico et al. (2014). IV estimates reported in second column of each set show the instrumental variable specifications (fuzzy RD), where the endogenous regressor is the add any school indicator. Panel A: predetermined covariates. Panel B: measures of choice behavior from initial to final application.  $\Delta$  risk is change in application risk from first to final attempt. “Add to X” are additions of schools in given place on list, relative to initial application submission. Panel C: outcomes of choice process. “Enrolled in placed” is equal to one for students who receive a placement and enroll in the placed school. “Enrolled in placed | placed” is the enrollment rate in the placed school, conditional on receiving a placement. Panel D: congestion attributes of behavior and placement outcomes. “Uncongested” schools are those with excess capacity. NL and NR are sample counts to the left and right of the cutoff, respectively.



Table 7: RD and DD estimates of warnings effects in New Haven

Outcome	RD		Diff. in Diff.	
	$\beta$	SE	$\beta$	SE
<i>A. Demographics</i>				
Female	0.136	(0.089)	0.001	(0.033)
African American	0.073	(0.078)	0.040	(0.031)
Hispanic	0.001	(0.084)	-0.003	(0.032)
White	0.011	(0.073)	0.006	(0.025)
<i>N</i>		740		3918
<i>B. Interaction with Simulator</i>				
Warnings email	0.998	(0.008)		
Pr(Any login)	0.130	(0.077)		
Number of Logins	0.117	(0.093)		
Pr(Any sim. run)	0.066	(0.068)		
<i>N</i>		740		
<i>C. Choice Outcomes</i>				
Change length or school	0.116	(0.045)	0.042	(0.015)
Lengthen app.	0.114	(0.043)	0.053	(0.013)
Insert new school	0.040	(0.024)	0.012	(0.007)
Append new school	0.076	(0.037)	0.039	(0.011)
Change school	0.047	(0.026)	0.002	(0.009)
Shorten app.	-0.014	(0.009)	-0.008	(0.005)
Diff. in realized risk	-0.026	(0.013)	-0.007	(0.004)
Diff. in simulated risk	-0.035	(0.013)	-0.014	(0.004)
Any realized risk reduction	0.111	(0.041)	0.036	(0.010)
Any simulated risk reduction	0.116	(0.042)	0.046	(0.011)
<i>N</i>		740		3918

RD and difference-in-differences estimates of the effects of the New Haven, CT warnings intervention. The samples for these regressions consist of the universe of applicants to grades PreK, and K in the NHPS simulator study i.e. that have been randomized into either control or one of the two treatment groups or the equivalent comparison group in the 2019 application process. RD specifications are based on local linear fit, dropping observations with predicted portfolio risk of of less than 1% or more than 99%. For the difference-in-differences panel, no observations are dropped based on their risk score. Robust SEs in parentheses. See section 5.7 for details.