Smart Matching Platforms and Heterogeneous Beliefs in Centralized School Choice†

Felipe Arteaga *UC Berkeley* Adam Kapor *Princeton University*

Christopher Neilson *Princeton University* Seth Zimmerman *Yale SOM*

April 19, 2021

Abstract

A central point in the market design case for the use of strategyproof assignment mechanisms in school choice is that they relieve applicants of the need to strategize on the basis of beliefs about 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" providing 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 nonplacement risk. We then use RCT and RD research designs to evaluate live feedback policies at scale in the Chilean and New Haven choice systems. We find that 22% of applicants submitting applications where risks of nonplacement are high respond to warnings by adding schools to their list, reducing their nonplacement risk by 58%. These results replicate across settings and over time. We conclude that 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.

†We thank seminar participants at NYU Economics/Stern and the 2021 AEA meetings for comments and suggestions. We thank Manuel Martinez and Jan Knuf for research assistance and the implementation team of data scientists and developers at ConsiliumBots for their help throughout the project. The authors wish to thank the government partners that made this research possible, in particular the leadership at the *Sistema de Admission 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. Kapor: akapor@princeton.edu. Neilson: cneilson@princeton.edu.
Zimmerman: seth.zimmerman@yale.edu.

1 Introduction

The design of centralized school choice systems is among the leading applications of economic theory in the 21st century so far (Roth, 2015; Roth and Wilson, 2019). One of economists' main aims in their work with policymakers in cities including New York, Boston, New Orleans, and New Haven has been to create assignment mechanisms that are *strategyproof*; i.e., where students' best strategy on the application form is to list the schools they like, in the order they like them.[1] A central point in both the public and academic cases for strategyproof assignment mechanisms is that applicants do not need to form accurate beliefs about their own admissions chances to come up with this strategy. Eliminating the need to strategize on the basis of admissions chances, which are high-dimensional and may be hard for applicants to observe or compute, has the potential to reduce costly application mistakes for all students and lead to more equitable outcomes if some families are better able to acquire and analyze admissions information than others. These conclusions follow 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.

This paper makes the point that many students in centralized choice systems *do not* know which schools they like, and that costly search for schools places beliefs about admissions chances back in a central strategic role, even when the assignment mechanism is strategyproof. We first develop a simple model of costly search for schools, and show how over-optimism about admissions chances can reduce search and increase the risk of non-placement. We then turn to empirics, drawing on administrative records from centralized school choice systems in Chile and New Haven. Surveys of choice participants show that search for schools is in fact costly, that beliefs about admissions chances are a critical input to search, and that students submitting risky applications systematically underestimate nonplacement risk. Working with policymakers to evaluate a scaled response to this problem, we use experimental and quasi-experimental approaches to show that providing live feedback about application risk through a smart matching platform changes search behavior and raises placement rates. Our findings hold across years, cities, and countries. We conclude that fully leveraging the benefits of strategyproofness *inside* centralized choice systems requires supporting strategic behavior *outside* these systems.

We begin with a 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 for schools to add to their choice application. Applicants are endowed with a portfolio of schools

[1] Abdulkadiroğlu et al. (2005a,b); Abdulkadiroglu et al. (2006); Pathak (2017); Kapor et al. (2020); Akbarpour et al. (2020).

they know about, and may choose to pay a cost to learn about new schools, one school at a time. Once applicants decide to stop searching, they submit the application to a strategyproof assignment mechanism and receive at most one placement. Beliefs about admissions chances enter the model through the search process: the payoff to search depends on the probability the applicant is placed in the school she finds.

We use this model to show three results. First, for students who are sufficiently overoptimistic about their placement chances, reducing optimism tends to raise the return to search. Second, information interventions implemented after individuals have made their application choices 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 effects of information interventions rise in proportion to placement rates, except to the extent they are offset by declines in enrollment conditional on placement.

We bring our model to the data using administrative records 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. 1.2 million students applied through the centralized system over this period. Our second setting is the public school choice system in New Haven, Connecticut. We use data from 2020, when about 7,000 applicants applied to the centralized system. New Haven uses a "truncated" DA assignment mechanism in which applicants can list only a limited number of schools (in our case, six). Truncated DA mechanisms are not strategyproof, but they are less manipulable at any application length than the common alternative of Immediate Acceptance (Haeringer and Klijn, 2009; Pathak and Sönmez, 2013). They are also common in practice. For example, New York uses a truncated DA mechanism, while Boston uses a strategyproof DA. Studying the Chilean and New Haven settings together allows 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, we surveyed families submitting applications to the choice process about their search for schools, their preferences over schools, and their beliefs about their placement chances. We administered these surveys 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. Consistent with high search costs, choice participants have limited knowledge of relevant schools. Only 17% of respondents report that they know a random nearby school 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 directly 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. This latter response corresponds to the more traditional rationalization for short application lists, which is that applicants would prefer their options outside the choice system to being placed at the remaining available schools. We then ask participants about their beliefs about their placement chances, and relate elicited subjective beliefs to the reasons participants give for terminating school search. 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 the beliefs about admissions chances that applicants use to inform their search process are often wrong. On average, respondents submitting applications with non-zero risk of nonplacement 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%. Even applicants with true placement chances close to zero have average subjective beliefs of close to 70%.

Our fourth survey finding is that the welfare stakes are large, and reflected in participants' choices to enroll in their placed schools. 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. Looking within placed students, 99% of students who receive a placement at a school where they report in advance of placement decisions that they would be very satisfied choose to accept their placement at that school, compared to 61% of students placed at schools where they are unsatisfied.

We next ask whether smart platforms that bring live feedback on placement chances inside the placement process can expand search and reduce nonplacement risk. In both Chile and New Haven, we evaluate live-feedback systems that warned applicants submitting applications with high risk of nonplacement. 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 off-platform via email, and directed applicants to an application simulator, which they could use to assess their placement chances under different hypothetical applications. Policymakers assessed risk in advance of application deadlines by combining data from previous years and applications already submitted in the current application 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 to add schools to their applications, corresponding to the complier group in our model. These students reduced their nonplacement risk by an average of 15.5 percentage points, or 58% of average ex post application risk at the cutoff.

Results from the text-message RCT show that these findings are not local to the cutoff. 4.4% of students 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 nonplacement 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.

Digging deeper into the way students modify their applications provides further evidence that students are learning about new schools. For example, about 10% of students who add schools add to the middle or top of their application list, not the bottom. This suggests that students frequently find schools that they like more than at least some of the schools in their starting portfolio. Also consistent with our model, the intervention does not cause students to *remove* schools from their application lists. Essentially all of the application changes we observe are application additions.

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.

The focus of this paper is individual search behavior, not equilibrium outcomes. However, it is informative to consider crowdout effects. If participants receiving the intervention end up at schools where capacity constraints bind, individual welfare gains may be offset by losses for the students they displace. In our setting, 57% of the placements generated by the smart platform are at schools where capacity constraints are slack; i.e., the policy tends to reduce congestion, not just shift access to highly-desired schools.

Our findings replicate across markets and over time. Splitting the Chilean data by city and year shows that our findings are not driven by particular market-year combinations. In Chile, treatment effects are steady across years, 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, consistent with the idea that additional search is more rewarding in bigger markets.

Looking beyond Chile, 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, 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%. We conclude that costly search with limited information is important in many school choice settings, including in US districts with long histories of centralized choice.

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 larger school choice *process*, and that the divergence between the two places substantial information demands on participants. Many theoretical and empirical 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 contribution here is to provide extensive survey evidence that the search and information 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 behavior, 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 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.[2] Findings in this literature are mixed, with some interventions changing choices, and others finding precise zeros. Our intervention, which provides information on the matching-market equivalent of *prices* (Azevedo and Leshno, 2016) 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 something they like, they may not think it is worth it to conduct the due diligence necessary to add a new option to their portfolio, even if the information they receive suggests that option is appealing in some ways. Another important distinction between interventions on beliefs and interventions on attributes is that belief interventions do not guide people towards a small set of high-performing schools, and are thus more likely to 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), which distill complex datasets into the information people need to make the strategic decisions that inevitably remain. We bring AI-assisted "choice engines" (Thaler and Tucker, 2013) into a matching market setting; this contrasts with previous work focusing on attribute comparisons in product markets (Gruber et al., 2020). What makes this feasible is the close collaboration between policymakers, platform designers, and researchers to link the applicant-facing part of choice platforms to the back-end data and algorithm that generate the information the applicants need. Linking these elements of the choice process lets us construct tailored information interventions that reduce applicants' computational burden as much as possible, while also providing that information at the time of choice and from a trusted source. These are all features that past studies have shown are critical to effective information interventions (Mani et al., 2013; Fernandes et al., 2014; Fischer and Wagner, 2018; Patterson et al., 2019).

[2] 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).

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 admissions portfolios. 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 how changes in the rates of placement into schools and rates of enrollment conditional on placement provide information about the individual welfare effects of policies that reduce nonplacement risk.

Our analysis takes the perspective of an individual student searching for schools to (potentially) add to his or 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 must decide whether to take jobs as they arrive, and search terminates once the agent accepts an offer.

The strength of this approach is that it highlights the complicated strategic problem facing individuals even when the centralized assignment mechanism itself is strategyproof, and allows us to draw out the role that information about beliefs plays in this problem. The main limitation is that we do not consider the equilibrium properties of the matching of students to schools. Our choice to focus on the details of the individual strategic problem rather than the equilibrium outcome reflects the strengths of our data and the goals of our empirical analysis. For a consideration of limited admissions information and equilibrium in match markets see Immorlica et al. (2020) and Hakimov et al. (2021).

2.2 Model setup

Consider a strategyproof centralized assignment mechanism where applicants have limited information about what schools are available to them. Each person is endowed at time zero with consideration set $C_0 \equiv \{1, 2, 3, \ldots, K\} \subseteq \mathcal{J}$, where $\mathcal{J}$ is the set of all schools. Individuals receive utility u_j from placement at school j. Without loss of generality suppose $u_1 > u_2 > \ldots > u_K > 0$, and that utilities are measured relative to the outside option of nonplacement, 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.[3] Define $R_j = 1 - p_j$ as the risk of nonplacement at school j.

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

$$V_0 = p_1 u_1 + R_1 p_2 u_2 + \ldots + \prod_{j<K} R_j p_K u_k \tag{1}$$

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.[4] These search costs differ across individuals and are distributed according to $\Phi(C)$.

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 contrasts with much previous empirical work on information frictions in school choice, which documents errors in applicant beliefs about school attributes. Making this assumption allows us to focus on the novel aspect of our contribution, which is to understand the effects of erroneous beliefs about admissions chances. Our goal here is to show that belief errors about admissions chances create strategic challenges for students and can lead to welfare losses, even if other aspects of school search are well-functioning.

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. Then the value of the optimal application that includes s is given by

$$V = p_1 u_1 + \ldots + \prod_{j<r} R_j p_s u_s + R_s \sum_{k=r}^{K} \prod_{j<k} R_j p_k u_k,$$

[3]In the empirical settings that we consider, admissions outcomes are determined by lotteries which are independent across schools. In principle, additional uncertainty about 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).

[4]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 λ. In this case, C would denote the expected cost of searching until one additional school is found.

and hence

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

where

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

Γ_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. a is then a measure of optimism; as a gets bigger people think risk is smaller than it really is. Assume $a < 1$ so that people do not rule out all application risk. We can then rewrite $V - V_0$ as a function of true beliefs and belief errors, so that

$$V - V_0 = \Psi_r \times (1-a)^{r-1} \times (1 - R_s^*(1-a)) \times (u_s - \Gamma_r)$$

where $\Psi_r = \prod_{j<r} R_j^*$. Taking logs, and then taking the derivative with respect to a yields

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

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 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, and finally, increasing optimism reduces the value of adding school s by raising the expected value falling below s on the application. This is the third term of the sum, which 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.

How do these three channels combine to affect the value of search adding school s to the application? We show that for sufficiently high levels of baseline optimism, additional increases in optimism reduce the value of adding schools to the bottom of the application.

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

This proposition shows that if students believe that the outcome of search will be to find a school they can add to the bottom of their application existing multi-school application, then *reducing* optimism will *increase* the value of search for students who are sufficiently optimistic at baseline. As we show 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 accurate information on admissions chances can be important to the strategies students employ when participating in choice, even if they do not affect the applications students submit given their consideration set.

2.5 Information interventions and search behavior

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

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

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. In equilibrium, $E[Search|C_0, a] \leq C$; otherwise applicants would not have stopped searching.

Taking equilibrium behavior as a starting point, consider how a small change in optimism, Δa, 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. Assume that the distribution of search costs Φ is differentiable with pdf ϕ.

Proposition 2. *Once applicants have completed search, the effect of changing optimism by Δa 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. Note that

$$\Delta a \frac{dPr(\text{Search})}{da} = \phi(E[Search|C_0, a]) \times \max\{\Delta a \frac{dE[Search|C_0, a, \theta]}{da}, 0\} \geq 0.$$

Also note that search costs are immediately sunk once they are incurred. If an applicant searches and draws a school that has utility below the outside option value, he does not add it to the application, the value of future search is unchanged, and the applicant searches again. The implication is that

$$\Delta a \times \frac{dPr(\text{Add at least one school})}{da} = \Delta a \times \frac{dPr(Search)}{da}.$$

$\square$

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

The effect of adding schools to the application is to reduce nonplacement risk. We can bound the expected change in true nonplacement risk from below as the expected change from adding one school. Define not 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, we have

$$\frac{dRISK}{da} \geq \frac{dPr(Search)}{da} \times \frac{RISK_0}{P(Find)} \times \int\int 1[u>0]p dF(u,p)$$

where $P(Find) = \int\int 1[u>0]dF(u,p)$ is the probability an applicant finds a school to add to their application in a single search. In sum, we expect information interventions that change optimism relative to their equilibrium level to raise search rates, cause individuals to lengthen their applications, and to reduce nonplacement risk.

2.6 Enrollment choices and the welfare effects of information

We next consider how to use objects we observe in 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 μ_j, with outside the option normalized to zero. Following placement, they receive enrollment shocks ϵ_j, iid across schools. Students choose to enroll in the placed school j according 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 ϵ_j have distribution $G(\epsilon)$, which is differentiable with density function g and has an inverse that is also differentiable. Then we can write the utility terms as functions of enrollment probabilities. Let $q_j = P(\text{Enroll}|\text{place at } j)$ denote the probability of enrollment conditional on placement at j. We then have $q_j = 1 - G(-\mu_j)$ and $\mu_j = -G^{-1}(1 - q_j)$. $\kappa(q_j) = E[\max(-G^{-1}(1 - q_j) + \epsilon_j, 0)]$ is the value of expected placement utility u_j implied by the enrollment rate. Substituting in for μ_j with this function of enrollment probability, we see that the ex ante utility of placement at a school is then an increasing function of enrollment probability conditional on placement at that school:[5]

$$\frac{d\kappa}{dq_j} = \frac{dE[\max(\mu_j + \epsilon_j, 0)]}{dP(\text{Enroll}|\text{place at } j)} = \frac{P(\text{Enroll}|\text{place at } j)}{g(-\mu_j)} > 0.$$

We use this expression to 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 Δa 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 approximation a weighted sum of the resulting changes in enrollment probabilities, with all weights positive.*

Proof. First note that

$$\delta a \times \frac{dV}{da} = \Delta a \times \frac{d(1 - RISK)}{da} E[u_j|Place] + \Delta a \times (1 - RISK) \frac{dE[u_j|Place]}{da}$$

The first term in the sum is rises in proportion to the change in nonplacement risk. Unpacking

[5] Note that $\frac{dE[\max(\mu_j + \epsilon_j, 0)]}{d\mu_j} = P(\text{Enroll}|\text{place at j}) = q_j > 0$.

the second term, we have

$$\begin{aligned}\frac{dE[u_j|Place]}{da} &= \sum_{j\in\mathcal{J}} \frac{dPr(\text{Place at } j|\text{receive placement})}{da} \times u_j \\ &= \sum_{j\in\mathcal{J}} \frac{dPr(\text{Place at } j|\text{receive placement})}{da} \times \kappa(q_j) \\ &\approx \sum_{j\in\mathcal{J}} \frac{dPr(\text{Place at } j|\text{receive placement})}{da} \times \frac{d\kappa(q_j)}{dq_j} \times q_j \end{aligned} \tag{3}$$

The last line reflects a first-order approximation around an enrollment probability of $q_j = 0$, and uses the observation that the change in placement probabilites at all schools $j \in \mathcal{J}$ conditional on placement must sum to zero. We have $\frac{d\kappa(q_j)}{dq_j} > 0$ from above. $\square$

This last proposition indicates individuality 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 most of our analysis 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 5.6.

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 hoping to make the school choice process more transparent and more equitable adopted centralized choice for the first time (Ministerio de Educación, 2017). The centralized choice system was rolled out on a region-by-region basis, with adoption in all cities by 2019. The centralized process includes most of the public schools and private schools that accept school vouchers.[6] Since full scale-up in 2020 it covers the admission to all grades. 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 determine how

[6]7% of Pk-12 students are enrolled in expensive private schools that do not accept government vouchers; these schools do not participate in the centralized process. Also schools whose highest grade offered is K are not included.

students are allocated to oversubscribed schools, the mechanism combines coarse sibling, school employee, and alumni priorities with lottery-based tiebreakers.[7] Applicants may list as many schools as they want on their choice application.[8] This means that the mechanism is *strategyproof*. The approach Chile takes to centralized assignment is similar to that used in major US districts with assignment processes designed by economists, such as New York and Boston (Abdulkadiroğlu et al., 2005b,a). These districts both use DA assignment mechanisms with relatively long (New York) or unlimited (Boston) applications.

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 directly 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 B for additional discussion of school choice institutions in Chile.

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.[9] 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.[10]

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. The first portfolio an applicant submits contains an average of 2.8 schools. Following their initial submission, applicants are free to go back and change, add, or subtract schools up until the deadline. By the application deadline, 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.

[7] Some schools also use quotas for vulnerable students and, in a very small number of cases, for high-performing students.

[8] 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.

[9] See Online Appendix C for a discussion of our data sources.

[10] 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 most common change is to add a new school to the application: 21% of applicants have a school on their initial application that was not on their final 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. Comparisons of low-income to high-income students on these measures (columns 2 and 3) 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-listed 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 placements, compared to 74% of higher-income students. As reported in Panel E, 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.[11]

Nonplacement occurs despite slack capacity in most markets. Panel F of Table 1 displays an average over applicants of the share of seats in the applicant's 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, nonplacement risk was a major concern for education policymakers in Chile. Mineduc partnered with the NGO administering the choice platform to design a set of information interventions alerting applicants to nonplacement risk.[12] These interventions identified applicants whose submissions placed them at high risk of nonplacement, and notified them of the risks they faced prior to the close of the application deadline. The key design 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. This section summarizes the design of the interventions we study; see Online Appendix B for additional detail.

Mineduc conducted two kinds of information interventions over the period we study. The

[11]Applicants who do not participate in the first round are not included in our analysis.

[12]Our research team worked closely with Mineduc and the NGO to provide guidance on intervention design. See our Disclosure Statement for details.

first was an *interactive pop-up* embedded in the application platform. 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.[13] 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 2017 and has continued to use it since then. In 2017, 2018, and 2019, Mineduc activated the popup 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 popup effects in 2018 and 2019 excludes the students who submitted their first application attempt before the popup came online. These students made up 39% of all applicants in these years. In 2020, the popup was available over the full application window for most of the applicants.[14]

In 2020, Mineduc added a set of supplemental "reminder" interventions. Mineduc sent reminders to risky applicants via text and the messaging service Whatsapp over the 28-day application window. These messages contained information similar to the popup in several display formats.

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 popup 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 with text similar to the initial popup. 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 dates of the two text messages to all risky applicants, Mineduc and the NGO conducted an RCT evaluation of a Whatsapp intervention. The NGO administering the web service of the popup 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 the student's application.[15] It stated that the risk of nonplacement

[13] In 2017, the first year of the intervention, cutoffs were set at the city level and varied between 30% and 70% nonplacement risk. We do not use data from 2017 in this analysis.

[14] 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 pop-up was activated later for them (9% of total 2020 applicants). See below and Online Appendix E.

[15] In addition to high application risk, the NGO imposed other restrictions on the sample universe for RCT

was high, and recommended 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.

We evaluate the platform popup using a regression discontinuity design around the 30% risk cutoff. In 2018 and 2019, the RD estimates capture the effect of the popup for applicants near the cutoff. In 2020, the RD estimates capture the effect the popup and its interaction with the subsequent reminder interventions. Our goal in the RD analysis is to provide evidence on how changes in information about application chances affect search behavior and placement outcomes in general, not to unpack the differential effects of interventions by medium and timing. In what follows, we present estimates separately by year and for the pooled sample across years. Readers who are interested in understanding the effects of popup absent their interactions with the Whatsapp RCT can focus on the 2018 and 2019 implementation years.[16]

We evaluate the Whatsapp reminder in a standard RCT framework. Because applicants in the Whatsapp intervention receive multiple reminders, the RCT evaluation tells us about intensive margin treatment effects. 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.

3.3 Application risk and risk predictions

A critical input to both interventions is predicted application risk. The NGO computed application risk using the following procedure. They first constructed an estimate of the demand using data from the previous year or data from the current process. The former was used in the risk prediction process only for the 2 or 3 first days, while the later was used in all the following days. Combining the demand estimates with observed school capacities for the current application year, they recovered the probability of non-placement within each school-grade-priority group by simulating the assignment process 500 times.

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. The web service computed risk predictions for 73% of applicants over the 2018-2020

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 B.3 for details.

[16] See Online Appendix D for a detailed discussion of interactions between treatment types.

period. The remaining applicants were excluded because they applied in markets or at times in the application cycle when the web service was not operative. We label the applicants for whom the web service *did* compute application risk the "pop-up eligible" population, because this group received the pop-up warning if their application was deemed risky. Column 5 of Table 1 describes the pop-up eligible population, which closely resembles the full sample in pre-application characteristics and application behavior. For more details on the simulation and demand prediction see Appendix E.

Risk predictions closely track applicants' true nonplacement risk. Panels A and B of Figure 3 shows how ex post placement probabilities and observed placements vary with predicted placement probability. The ex post placement probability is constructed identically to the placement prediction, but using realized rather than predicted applications. Both outcomes rise slightly less than one-for-one with the predicted risk measure. The slope of true placement probability in predicted probability is 0.89, while the slope of observed placement is 0.85. For both outcomes, our predicted measure tends to slightly overstake risk amongst the riskiest applicants. This is reflected by the fact that observed values are slightly above the 45-degree line when placement chances are low. 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. Table 1 describes the distribution of application risk. Panel G of Table 1 describes ex post (or "true") risk on the initial application attempt. Mean nonplacement 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 of higher than 30%. Median risk for students submitting applications with non-zero risk is 62%, and 25% of such applicants have nonplacement risk of 92% or higher. Panel C 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. Applicants receiving warnings are less likely to be economically vulnerable than other applicants (37%, vs. 49% in the full sample) 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 nonplacement 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 students' engaged with the choice process, Mineduc and the implementing NGO conducted a survey of choice participants in 2020. The survey innovates over past surveys of school choice participants (De Haan et al., 2015; Kapor et al., 2020; Wang and Zhou, 2020) by recruiting a sample that is an order of magnitude larger, by asking about search in addition to preferences and beliefs.

The implementing NGO contacted students using an email 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. The survey covered a variety of topics, including students' preferences over schools, their beliefs about placement chances, and the process students used to search for schools. See Online Appendix F for survey text.

The NGO contacted 373,710 students using the official school choice email account. 48,929 students, or 13% of the total number of students contacted, completed the survey. Column 7 describes the sample of survey completers. 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 Search costs and search strategies

Results from the applicant survey provide strong evidence on three of our central claims. First, the cost of searching for schools represents an important constraint on the applications students submit. Second, beliefs about admissions chances are input to applicants' search decisions. Third, widespread and misplaced optimism about admissions chances suggests substantial scope for interventions targeting beliefs. Our survey findings also support the claim that the choice to enroll or not enroll in placed schools is a valid measure of satisfaction with the placed school.

The main focus of our analysis from this point forward is whether students receive any placement through the centralized mechanism. We made this choice ex ante, to reflect policy-makers' goals for improving the choice process. Evidence from our applicant survey supports the idea that placement vs. nonplacement is the critical margin from a welfare perspective. The survey asked respondents to report how satisfied they would feel if they were placed at the first-listed school on their application, if they were placed at the last-listed 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 factors like ex post rationalization of known outcomes. Respondents rated these options on a seven-point scale, 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-listed 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 a placement at all. 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 breakpoint from this perspective.

Panels B and C of Figure 4 show that enrollment choices map closely to measures of preference over schools, conditional on placement. Panel B shows that the share of students enrolling in a school declines with the rank of the placed schools, from an 86% rate at the first-listed to school to a 49% rate at schools ranked fifth or lower. Panel C shows that the rates at which students choose not to decline placements rise with stated satisfaction with the school.[17] 99% of students placed in schools they give the highest satisfaction rating choose not to decline, compared to 65% at schools with the lowest satisfaction rating.

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, much of which is likely costly to obtain. Our survey asked respondents what they needed to know about a school to feel that they knew it well. Students 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

[17] We use rates of non-decline here because we do not yet observe enrollment choices for the 2020 applicants. The choice not to contact the school to formally decline a placement is a leading indicator of final enrollment choices. Panel includes data only on first- and last-ranked schools, since these were the schools for which we asked about student satisfaction.

performance, which is also publicly available online. Other attributes might or might not be publicly available online, but could likely be obtained upon a short visit to the school. For example, 89% reported needing to know about the available extracurricular activities, and 89% reported needing to know about the school's infrastructure. Other things a majority of respondents said they needed to know would likely be hard to obtain. 66% 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 respondents how well they knew schools near them. For each respondent, we asked about a random nearby school, a nearby school that was high-performing and expensive, and a nearby school that was low-performing and free.[18] 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. That respondents do not feel they have strong knowledge of randomly chosen nearby schools is consistent with observation that families need to know a lot about a school to feel they know it well. Search is costly enough that at the end of the choice process, most families do not feel adequately informed about most nearby schools.

Consistent with the idea that applicants learn about schools before applying to them, respondents' claim to know the schools on their application list 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 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 school in their existing portfolio will be less likely to engage in additional search. Several 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: that there were no more schools to around to add, that there were schools around but they would rather not attend these schools, that it is hard to find more schools, and that

[18] 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.

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 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 indicates an important role for 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 conception of the 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").

The second survey finding suggesting that beliefs play a role in choice is the strong relationship between applicants' subjective placement beliefs and their stated reason for stopping search. 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. These respondents were nearly sure they would receive a placement. In contrast, only 9% of respondents in the bottom quintile gave this reason for stopping search. The average subjective placement chance these respondents reported was at 36%.

4.2 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 mean subjective placement belief binned into ten groups by true placement probability.[19] If beliefs were accurate on average, they would follow the 45 degree line, which we plot for reference. We instead observe a weak positive relationship with a large upward shift. Strikingly, the mean subjective belief for applicants with true admissions chances near zero is close to 70%. The third line on the graph is 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 true risk values, because application risk evolves over time. However, it is clear that predictions are on average much closer to true placement probabilities than are subjective beliefs.

Several pieces of evidence support the contention 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 A1 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.

We interpret our survey evidence as making a strong case that costly search is important for choice outcomes, that beliefs are an important input to search strategy, and that beliefs are often inaccurate. This evidence suggests that accurate information about admissions chances would be valuable to students participating in choice, and would likely cause many applicants to search more. We test this proposition in the next section.

5 Warnings, search, and placement

5.1 Interactive warnings and application risk

Our theoretical and descriptive analysis argue that beliefs about admissions chances are important inputs to the choice process, but that many applicants strategize on the basis of overly-optimistic beliefs. If true, applicants should respond to warnings about their risk of nonplacement by adding more schools to their portfolios. This section considers several experimental and quasi-experimental tests of this theory using data from scaled policies implemented

[19]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 Connecticut school choice systems.

We focus first on the interactive nonplacement risk warning administered to Chilean students inside the choice system. Because all students with at least a 30% chance of nonplacement received this warning, we can evaluate it using a regression discontinuity design. This is a sharp RD: all students above the threshold received the warning. 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 bandwidth of 0.1. This bandwidth approximates that given by optimal bandwidth calculations (Calonico et al., 2014). We report estimates obtained using the Calonico et al. (2014) approach in Online Appendix Table A1 and Figure A2. Alternate approaches to RD estimation do not change our findings.

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 effect estimates and robust standard errors calculated using local linear regressions and a bandwidth of 0.1. We show separate estimates for each choice year and in the pooled sample. 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. This is consistent with the fact that this 30% cutoff had no significance for applicants prior to policy implementation. We view these findings as consistent with quasi-random assignment around the cutoff.

Panels A through D of Figure 9 and Panel B of Table 2 show how receiving the warning 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.[20] Students add an average of 0.34 schools, and ex post risk of nonplacement 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 change 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

[20]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.

estimates in which adding at least one school to the application is the endogenous regressor. 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 nonplacement risk by 15.5 percentage points, or 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.

5.2 Search behavior

The kinds of changes applicants make to their applications are consistent with the idea that the intervention leads to additional search. Results reported in the 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 have identified their top-choice schools early in the search process.

The information intervention 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 that 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 share of students adding schools to the application.

5.3 Enrollment and welfare

Changes in application behavior translate to changes in choice outcomes. Panels A and B of Figure 10 and Panel C of Table 2 report the effect of receiving the warning on placement and enrollment. At present enrollment data are available only for the 2018 and 2019 implementations. 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 change in application risk (within one standard error).

Warnings do not produce lower quality placements. Overall rates of enrollment rise roughly one-for-one with rates of placement across the cutoff. The rates at which students enroll in school conditional on placement are continuous through the cutoff value.

Proposition 3 showed that the effect of the intervention on individual welfare was proportional to change in placement rates, except as offset by declines in enrollment conditional on placement. Our findings here 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 change in placement rate.

5.4 Warnings across the risk distribution

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.

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. For non-risky students (below the 30% risk cutoff) treatment and control status are randomly assigned, but the "treated" group does not receive any treatment. Panel A shows that the number of warnings treatments students receive (summing over all SMS and Whatsapp interventions) rises by 0.48 across the cutoff for both treatment and control groups, but that this value rises more for the treatment group, which receives the additional, randomly assigned Whatsapp message. The 0.48 increase in messages viewed 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 treatment group and the population message to risky students. We focus on the add any school and change in application risk outcomes. We find that 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, and that these differences span the full distribution of application risk above the 30% cutoff. 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.

These figures also display RD estimates calculated across the risk cutoff within the treatment group. RD estimates show modest but statistically imprecise increases in search. The imprecision is expected given the smaller size of the RCT sample. We see little evidence that students near the 30% risk cutoff respond more to the intervention than applicants elswhere 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 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 persist across the full distribution of risk above the risk cutoff, except perhaps at 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.

We draw two conclusions from this analysis. The first is that the effects of warnings persist as we move farther up the risk distribution. The second conclusion is that there may be benefits to providing the same person with information more multiple times. Though share of students induced to search by the followup interventions is smaller than the share induced to search by the initial platform popup, the effects per affected student are very large. 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 are ready to make choices.

5.5 Decongestion vs. reshuffling

We have thus far focused on outcomes from the individual perspective; i.e., holding other applicants' choices fixed. However, the overall effects of the warnings policy depend on crowd-out. If most schools applicants add to their portfolios are oversubscribed, the effect of the scaled warnings policy would be to reshuffle students across inside-option schools. Welfare gains would depend on relative preferences over inside-option schools, which we do not evaluate here. Conversely, if the policy results in more students being placed overall, thereby reducing market congestion, crowd-out effects may be limited.

We assess the congestion effects of the warnings policy by looking at how receiving a warning affects placement rates at uncongested schools. As reported in Panel D of Table 2, receiving a warning raises the chances students add an uncongested school to their application by 7.3 percentage points. Because admissions rates at these schools are equal to one, they have an outsize effect on placement outcomes. Receiving a warning raises the chances that students are placed at an added uncongested school by 2.2 percentage points, 57% of the overall placement effect. Effects on placement to an uncongested school overall are more noisily estimated.

Our goal in this paper is to make claims about the role of search in the individual choice problem, not to claim that informational interventions can reduce congestion overall, which depends on market conditions at baseline. However, the finding that many of the schools that students find and then place in are undersubscribed highlights the difference between approaches that provide information on beliefs and approaches that direct students to specific, often high-achieving schools.

5.6 Distribution of effects across years, cities, and countries

5.6.1 Across years and cities in Chile

We next consider how the effects of treatment assignment vary across contexts. A natural question for interventions of this type is whether the effects are specific to certain choice settings, where baseline risk levels, beliefs, sophistication, trust in choice institutions, and assignment mechanisms may differ. We explore this question in two ways. First, we report how effects vary across cities in Chile, where the choice platform is held constant but market conditions may differ. Then we consider an implementation of a similar warnings intervention in New Haven, Connecticut, where choice institutions and the choice platform are quite different.

Panels A and B of Figure 12 display histograms of the RD estimates of popup 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. Table 4 shows city-by-year specific effects for add any school in 20 large educational markets. Because centralized choice was not implemented in Santiago until 2019, the total number of market-year combinations is 59. Looking across markets and years, we find effects on proportion adding least one school ranging from -0.10 at the low end to 0.53 at the high end.

We interpret the consistently positive effects across place and time as strong evidence that these effects arise in a variety of choice settings. Beyond this, two specific features of the city-by-year distribution are important to note. First, the effects of the intervention are consistent over time. A natural hypothesis is that the effects of the intervention might decline with time, as applicants become more familiar with the choice options, the choice process, and perhaps the intervention itself. We see no evidence that this occurs. The finding that belief errors persist even in "mature" choice processes is consistent with the idea that acquiring information about admissions chances for a particular portfolio is very costly, and with the empirical results in Kapor et al. (2020).

Second, though students engage in additional search at similar rates across markets of different size, they tend to add more schools to their applications when markets are larger and they have more options to choose from. Panels C and D of Figure 12 split the sample by different measures of market size, and report estimated popup 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 in that radius. 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. Though the effects of the intervention on search engagement are similar across size, the benefits may be larger on average in bigger markets.

5.6.2 Alternate implementation in New Haven, CT

In addition to our main implementation 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 described 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 all applications with a nonplacement 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 nonplacement for a student submitting the application in the previous year was at least 50%.[21]

[21] The district focused on major choice grades (pre-Kindergarten, Kindergarten, and grade 9), where choice

All applicants identified as risky received an email stating they were at risk of nonplacement. 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. In addition to warning all risky applicants, the district selected a randomly chosen fifty percent of non-risky applicants to receive an email with a link to the simulator but no warning. See Online Appendix G 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. A 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.

Figure 13 presents a visual summary of our findings from the New Haven intervention. 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. 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.[22]

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 ponts around the cutoff.

Table 5 summarizes these and other 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.

Panel A of Table 5 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 behavior effects we observe

probabilities are relatively stable across years. Two new schools opened in 2020. Risk scores were not computed for students applying to these schools.

[22] 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.

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 results from the randomized evaluation of simulator access for low risk students. These results, which we report in Online Appendix G, suggest that simulator access without the warning email does not change behavior.

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 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).

Overall, we view findings from the New Haven warnings implementation as evidence that information on admissions chances is an important input to search and choice behavior in a wide variety of contexts. Our findings in New Haven supplement our main finding that the warnings intervention replicates across time and geography within Chile.

5.6.3 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. Our surveys of Chilean applicants asked about the kinds of information they thought would be most helpful in filling out their applications that they did not already have. Respondents could choose multiple options from a menu that included things like various measures of academic performance, prices, and school suggestions. We conducted a similar survey in New Haven.[23] In both cases, the kinds of information respondents were most likely to say they needed but did not have 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 students in Chile said they would have liked more information on account-

[23]See Online Appendix G for details.

ability scores, while 84% of New Haven students said additional school suggestions would have been helpful.

6 Conclusion

This paper shows that costly search for schools is a key determinant of school choice outcomes, that optimism about school placement chances leads many students to search too little, and that policies that help correct belief errors push many students towards further search and substantially reduce nonplacement risk. These findings have two broad implications. The first is that strategic decisions shape the endowments that participants bring with them into centralized assignment mechanisms. In our case, the endowment is the set of schools an applicant knows about when submitting his application. Sophisticated and informed participants may thus gain an advantage even when the centralized mechanism itself is strategyproof. However, it contrasts with traditional treatments of school choice problem, which treat endowments as exogenous. We show that the school choice game goes beyond the centralized system, and that the strategic challenges posed by the broader game are central to families' experiences of school choice as well as to their placement outcomes.

The second implication is on the policy side. Our findings suggest that school choice designers should take a two-pronged approach combining strategic simplicity inside the centralized mechanism with smart matching platforms that make the informed strategic play still required by the broader choice game easier to implement. This recommendation goes beyond economists' traditional mechanism design focus in school choice settings.

As a final point, we note that the policies we evaluate here 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 and now shown to work across many settings. The "policy intervention as product" model we use here makes it easy to deploy the tools we evaluate here. At the time of this writing, the methods we develop here are being deployed and extended in Brazil, Peru, and Ecuador.

References

Abdulkadiroğlu, Atila and Tayfun Sönmez, "School choice: A mechanism design approach," *American economic review*, 2003, *93* (3), 729–747.

_ , 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.

Azevedo, Eduardo M and Jacob D Leshno, "A supply and demand framework for two-sided matching markets," *Journal of Political Economy*, 2016, *124* (5), 1235–1268.

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 Michelmore, 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.

Immorlica, 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, "What really matters in designing school choice mechanisms," *Advances in Economics and Econometrics*, 2017, *1*, 176–214.

_ and Tayfun Sönmez, "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.

Roth, Alvin E, *Who gets what—and why: The new economics of matchmaking and market design*, Houghton Mifflin Harcourt, 2015.

_ **and Robert B Wilson**, "How market design emerged from game theory: A mutual interview," *Journal of Economic Perspectives*, 2019, *33* (3), 118–43.

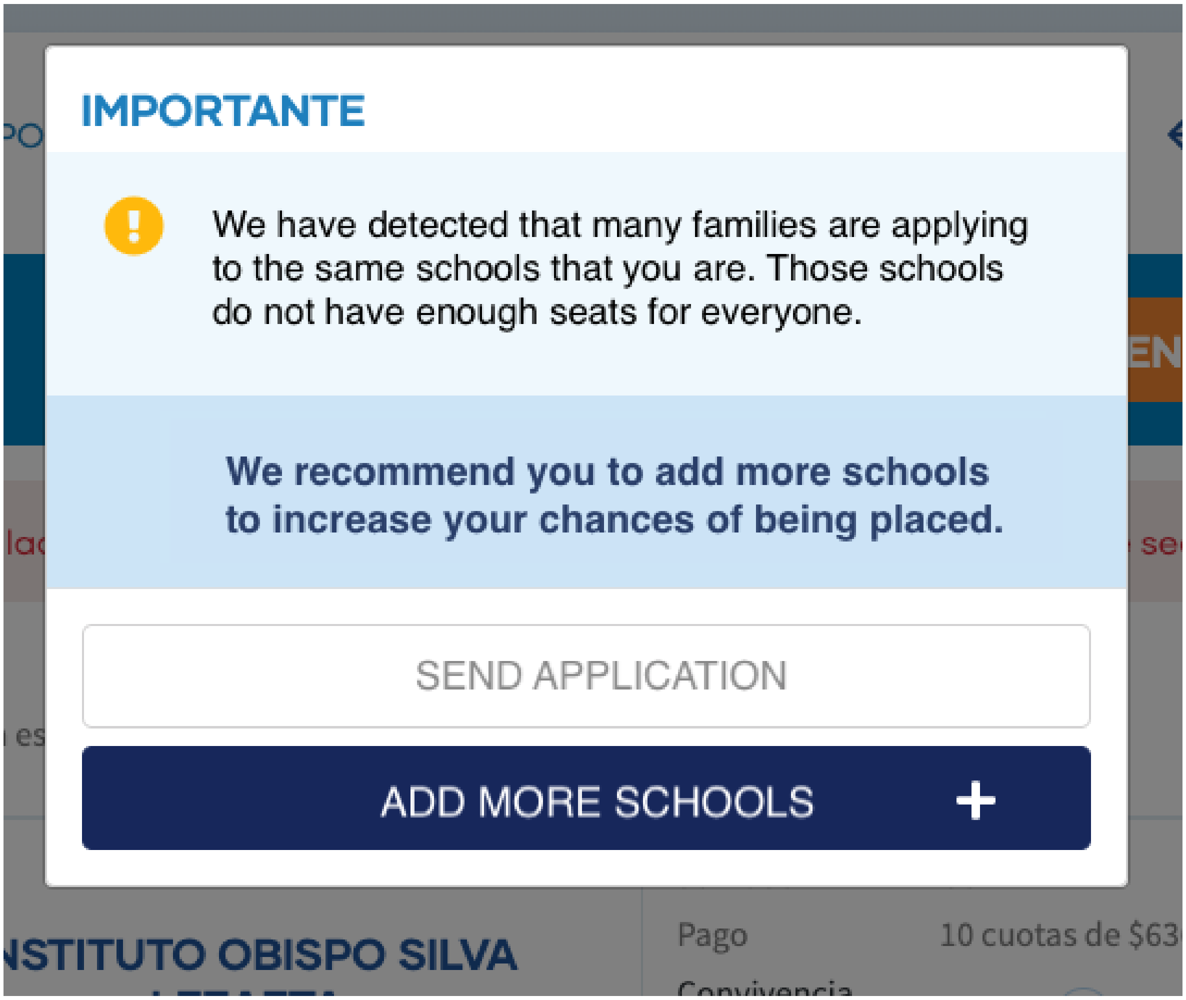
Son, Suk Joon, "Distributional Impacts of Centralized School Choice," *Yale University Working Paper*, 2020.

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.

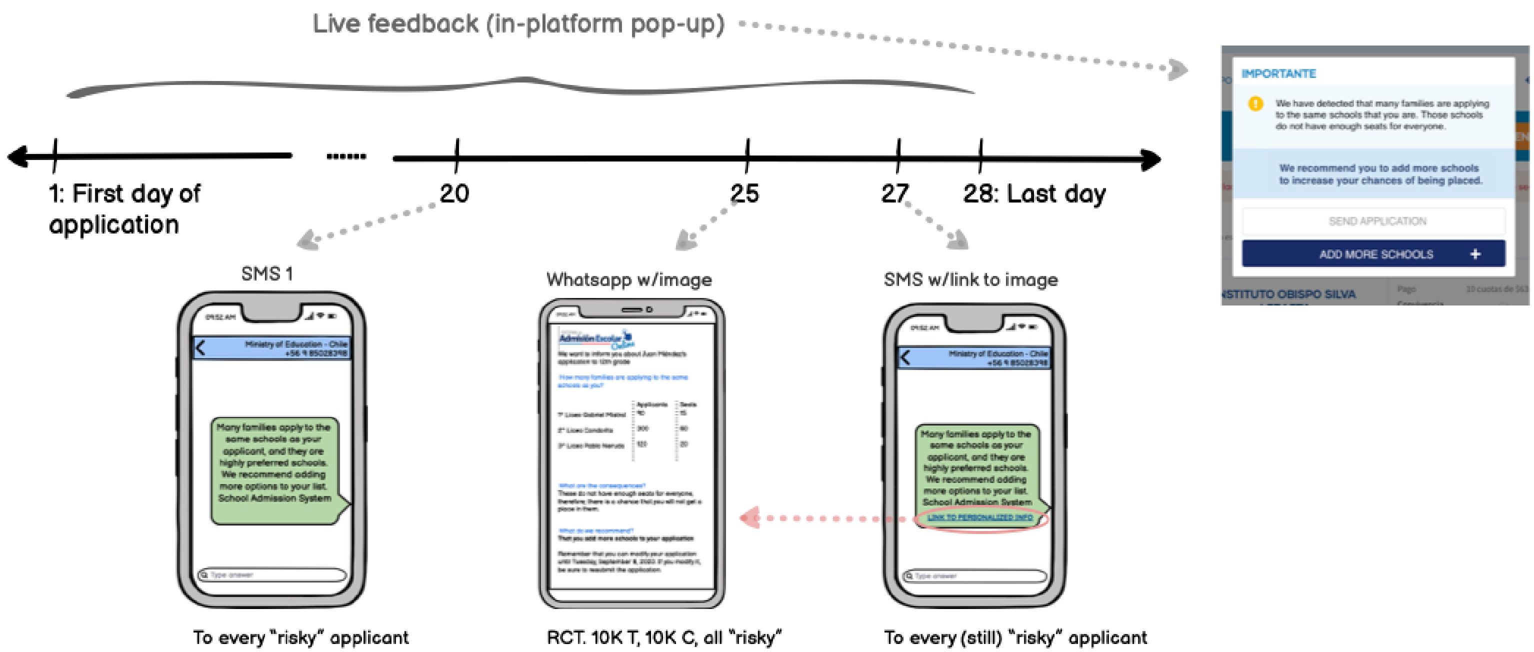
Figures

Figure 1: Platform feedback– 2018 and 2019



English translation of popup 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.

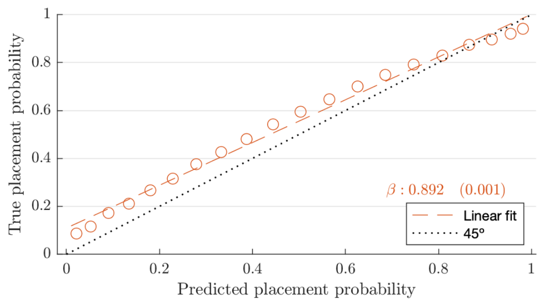
Figure 2: Timeline of feedback interventions - 2020



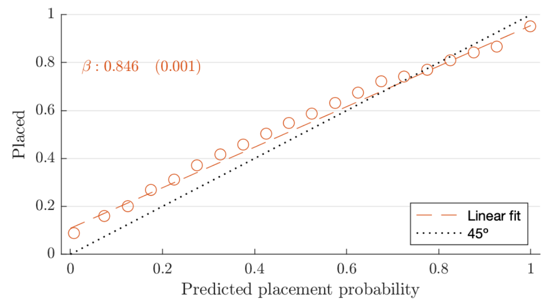
Sequence of 2020 application feedback for risky applicants. All text translated to English. The platform popup on the right was shown to all risky applicants at the time they submitted their application. The SMS and Whatsapp messages shown at center were sent to (subgroups of) still-risky applicants 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 his or her choice application. See section 3.2 for details.

Figure 3: Risk predictions and risk distribution

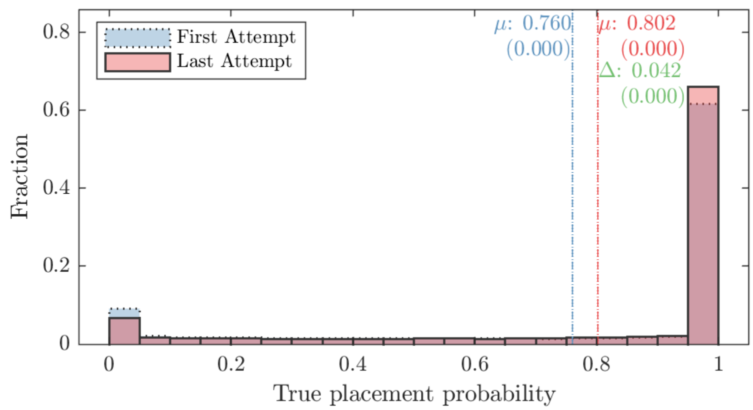
(a) Binscatter true vs predicted risk



(b) Binscatter placed vs predicted risk

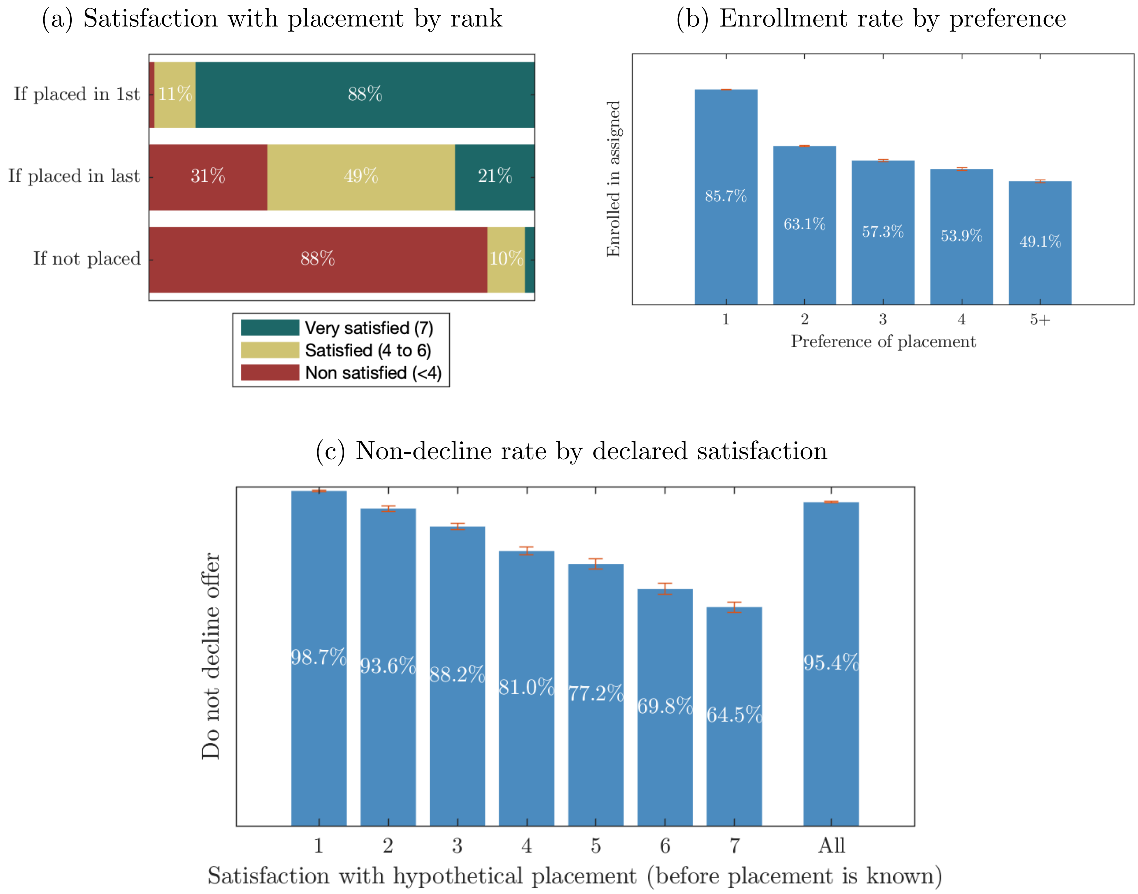


(c) Histogram of true risk



Panel A: binscatter of predicted placement probability (horizontal axis) vs. true placement probability (vertical axis). Panel B: binscatter of predicted placement (horizontal axis) vs. observed placement rate (vertical axis). Placement predictions in Panels A and B combine observed applications at the time an individual submits his orher application with historical projections. See section 3.3 for details. Panel C: histogram of true placement probability for initial application attempt and final application submission. Vertical lines display means.

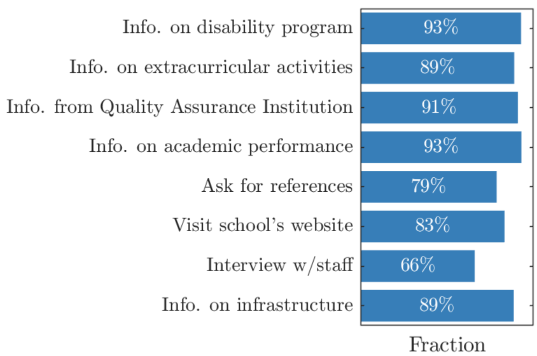
Figure 4: Satisfaction with placement outcomes



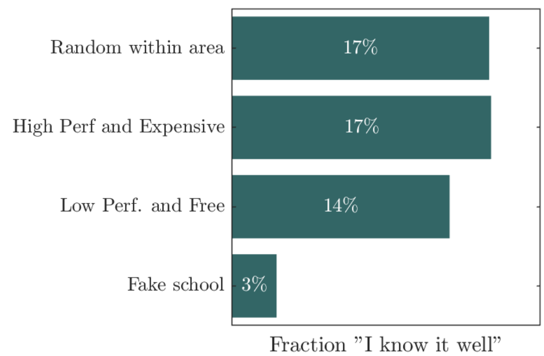
Panel A: Stated satisfaction with hypothetical placement outcomes. Data are survey responses to questions about satisfaction with being placed at first-ranked school, last-ranked school, and nonplacement. Sample: survey completers. Results reported on a 1-7 scale, with 7 being very satisfied and 1 being 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 *do not* decline placement offers, 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

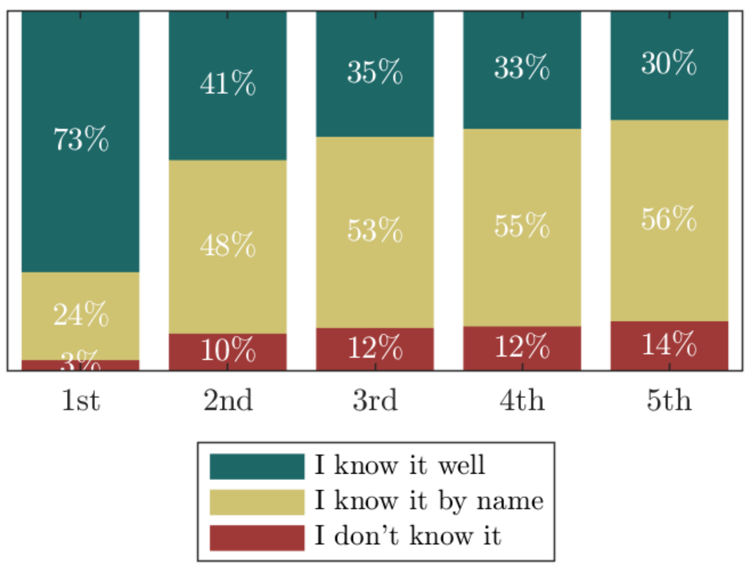
(a) Relevant steps to know a school



(b) Knowledge of schools not on application list

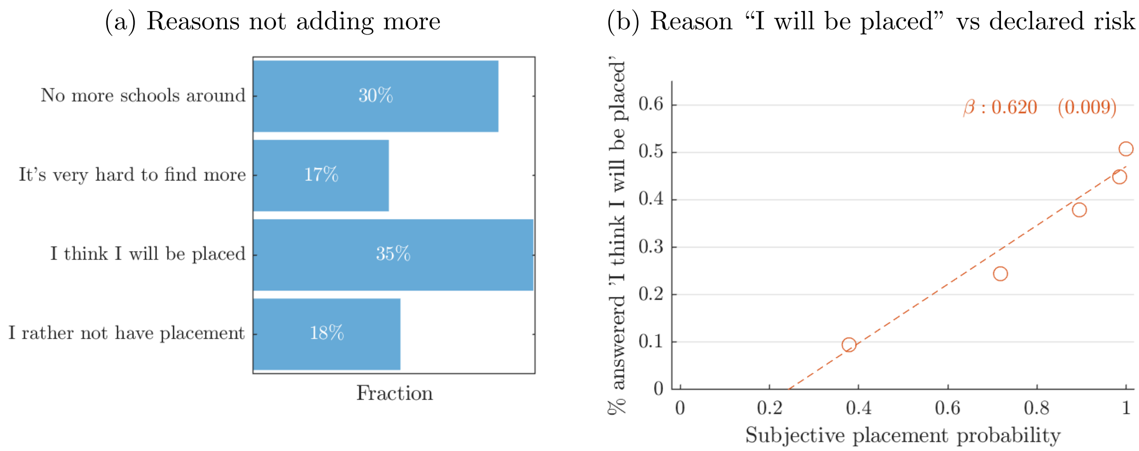


(c) Knowledge of schools on application list



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 local area, defined as 2km from student's location (home direction reported on platform, replaced with centroid of application if geocoding was unreliable). "High performing and expensive schools" are defined as classified within 2 best tiers of performance (out of 4) by Quality Assurance Institution. "Low performing and free" schools are defined as schools within the worst tier of performance (out of 4). "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.

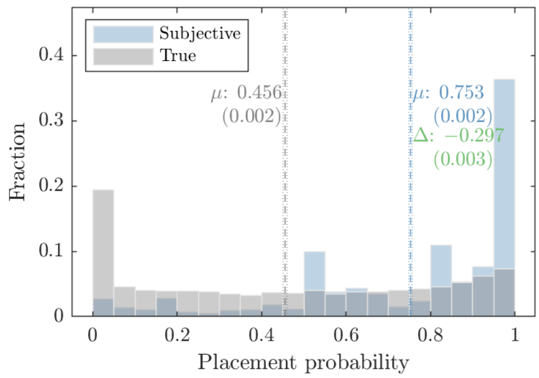
Figure 6: Reasons for stopping school search



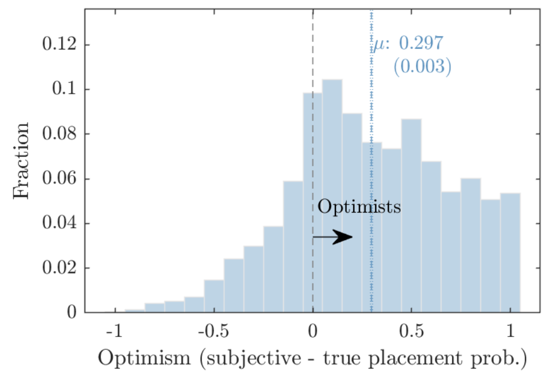
Panel A: survey reports of reason for not adding more schools to the choice application. Panel B: share of survey completers stating that they stopped search because they think they will be placed, by survey report of subjective placement probability. Sample in both panels: survey completers.

Figure 7: Subjective vs. observed application risk

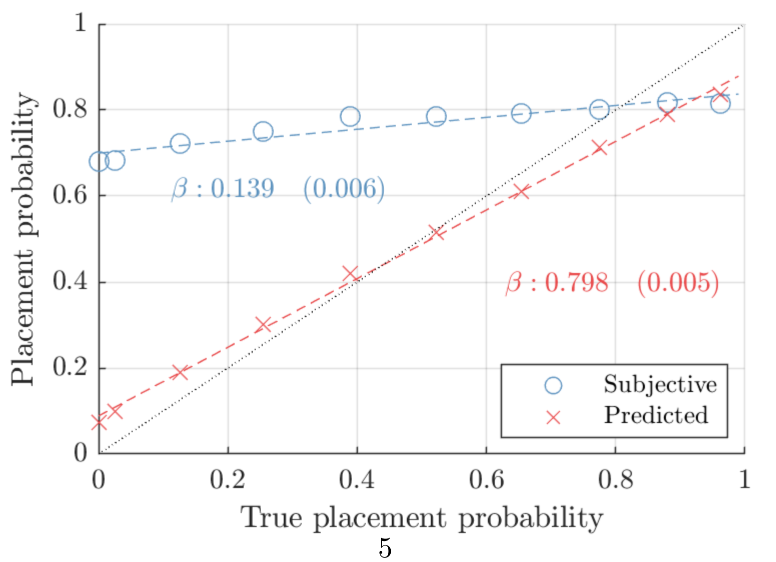
(a) Distribution of subjective and true placement chances



(b) Optimism



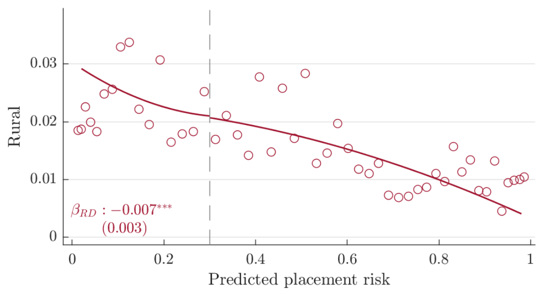
(c) Binscatter of subjective and predicted vs true placement chances



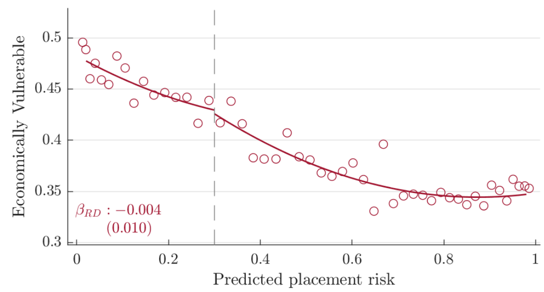
Panel A: distribution of true placement chances and survey-reported subjective placement chances. Vertical lilnes display means of each distribution. Panel B: distribution of optimism, defined as difference between subjective and true placement chances. Panel C: mean subjective placement belief by decile of true placement probability. Dashed line is linear fit. 45-degree line displayed for reference. Sample: survey completers.

Figure 8: Balance in smart platform RD

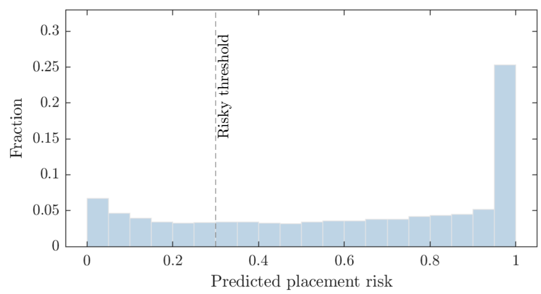
(a) Rural



(b) Economically vulnerable



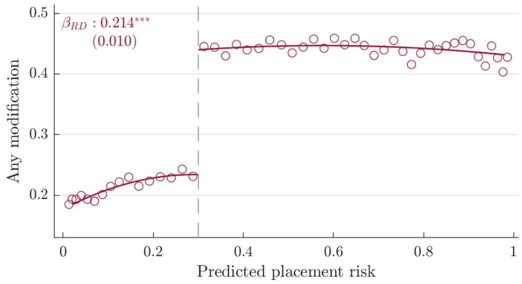
(c) Distribution of predicted risk (only >0)



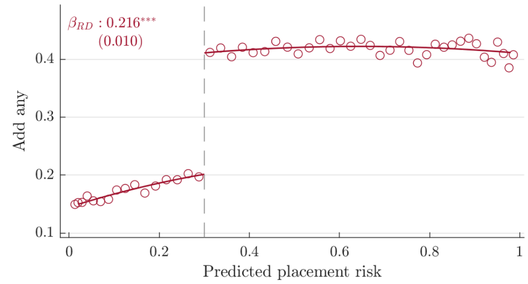
Binned means and global fits of predetermined characteristics by predicted risk for initial application. Points are centered means of 50 quantile-spaced bins of the support of the predicted placement risk $\in [0.02; 0.98]$. Solid line shows the quadratic fit. Reported coefficients and standard errors are from local linear specifications using $+-0.1$ 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 means

Figure 9: Choice behavior and risk reduction in smart platform RD

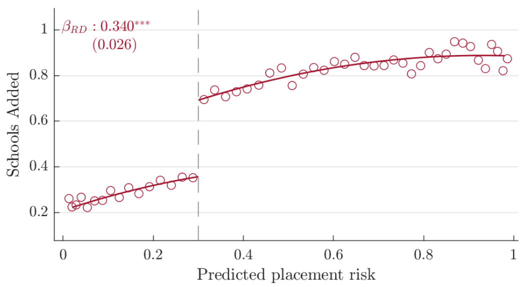
(a) Any application change



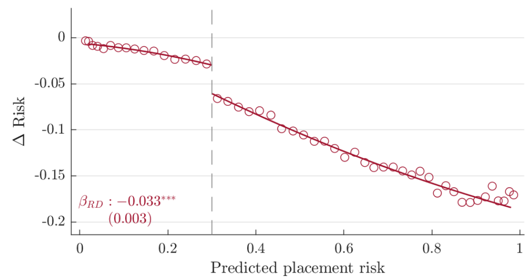
(b) Add at least one school



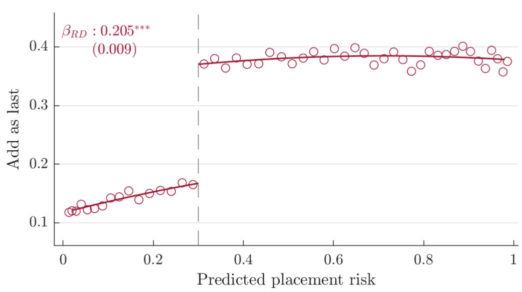
(c) Schools Added



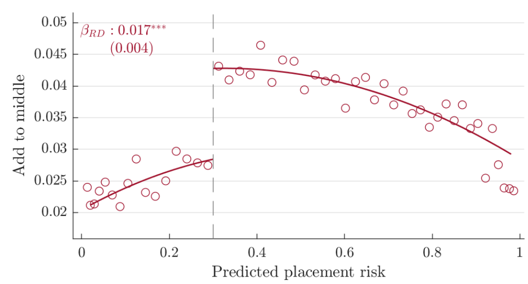
(d) Δ risk



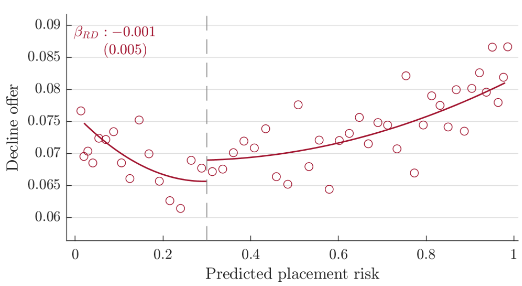
(e) Add schools to end of list



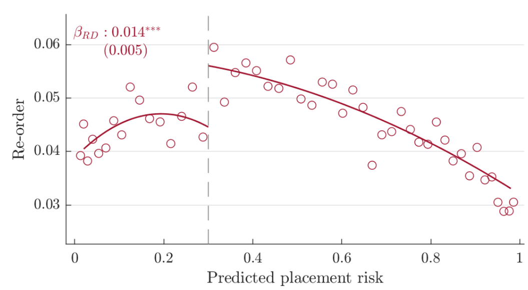
(f) Add schools to middle of list



(g) Drop a school



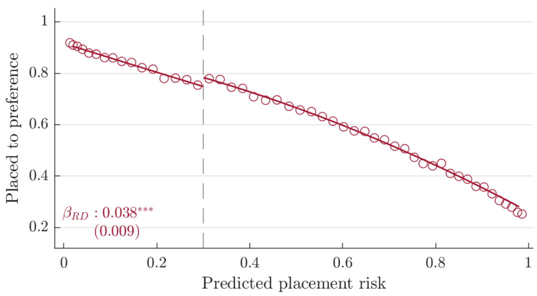
(h) Reorder schools



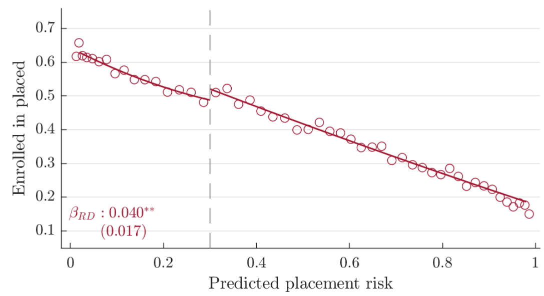
Binned means and global fits of predetermined characteristics by predicted risk for initial application. Points are centered means of 50 quantile-spaced bins of the support of the predicted placement risk $\in [0.02; 0.98]$. Solid line shows the quadratic fit. Reported coefficients and standard errors are from local linear specifications using $+-0.1$ 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 smart platform RD

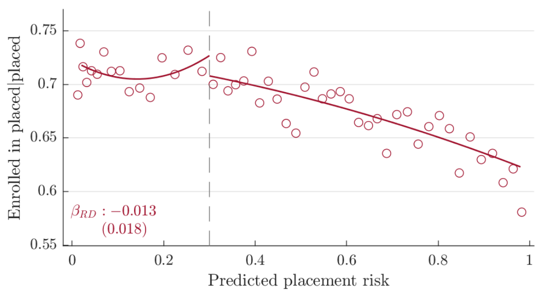
(a) Assigned in preference



(b) Enrolled in assigned



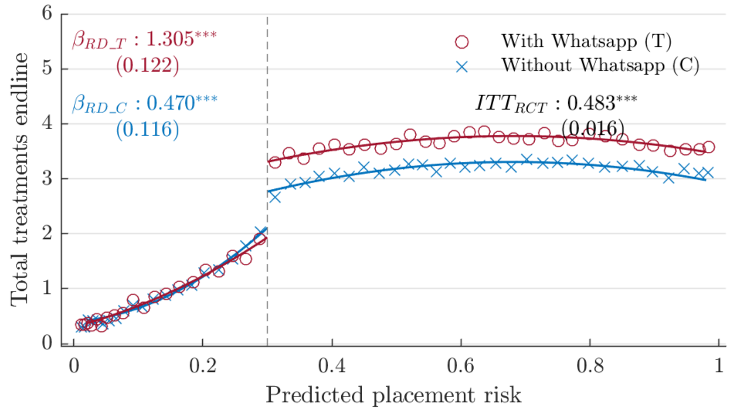
(c) Enrolled in placed conditional on placed



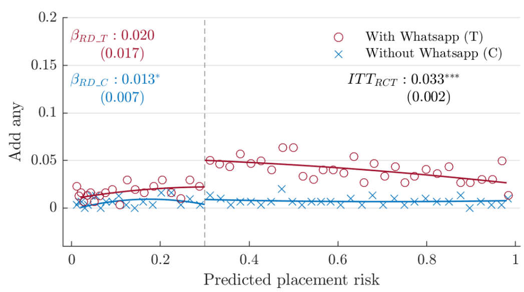
Binned means and global fits of predetermined characteristics by predicted risk for initial application. Points are centered means of 50 quantile-spaced bins of the support of the predicted placement risk $\in [0.02; 0.98]$. Solid line shows the quadratic fit. Reported coefficients and standard errors are from local linear specifications using $+-0.1$ 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: Randomized trial outcomes

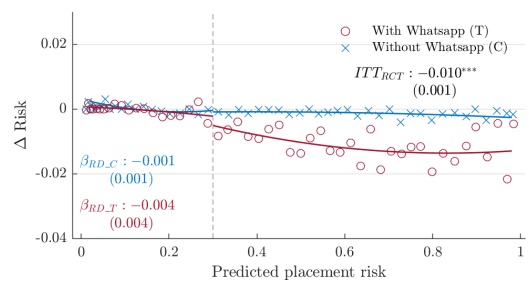
(a) Total treatments received



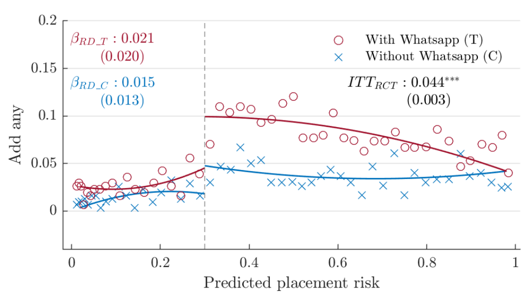
(b) Add at least one school – 44 hours



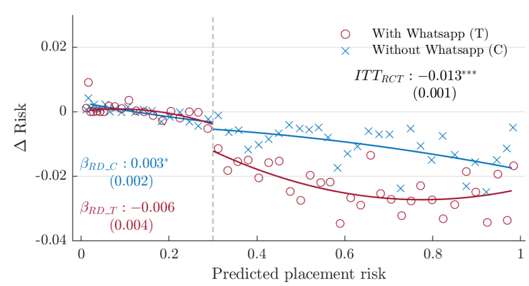
(c) Change in risk – 44 hours



(d) Add at least one school – endline



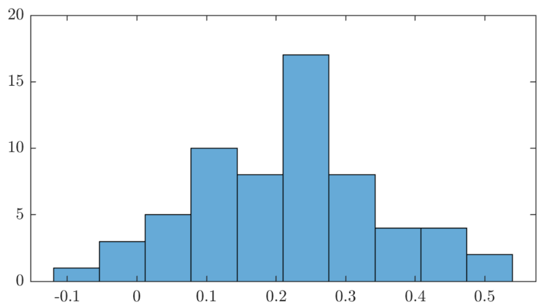
(e) Change in risk – endline



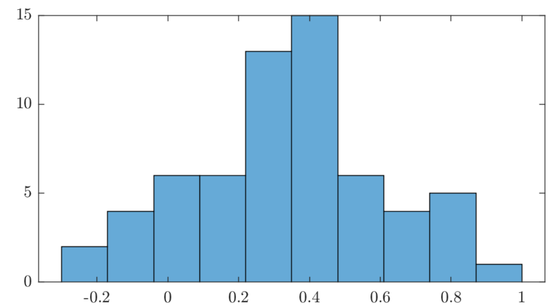
Binned means and global fits of application behavior and risk outcomes by predicted placement risk in RCT sample. Points are centered means of 50 quantile-spaced bins of the support of the predicted placement risk $\in [0.02; 0.98]$. Solid lines show the quadratic fit. Figures split by RCT treatment 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, neither treatment or control receives a warning, but they do receive a Whatsapp message with no information related to risk. Reported β_{RD} coefficients are RD estimates within treatment and control group, computed 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.. 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: add any school in 44-hour window between Whatsapp message and SMS followup. Panel B: change in risk within 44-hour window between Whatsapp message and application followup. Panel C: add any school between Whatsapp message and application close. Panel D: Change in risk by application close. See section 5.4 for details.

Figure 12: Effect distribution over city-years and by market size

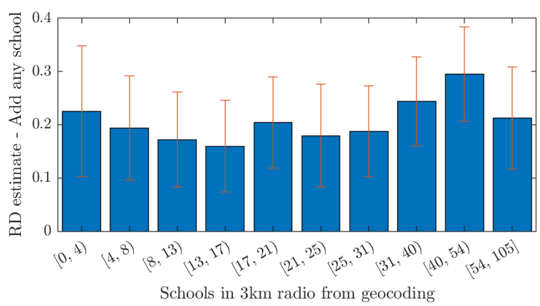
(a) Histogram: estimates of add any school



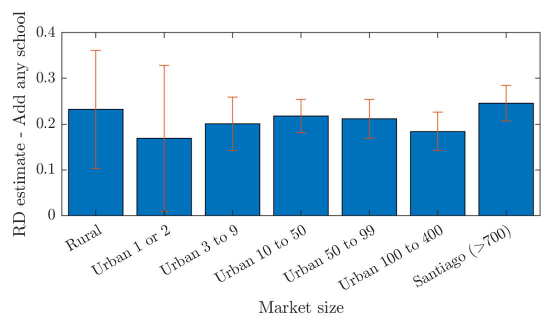
(b) Histogram: estimates of count of schools added



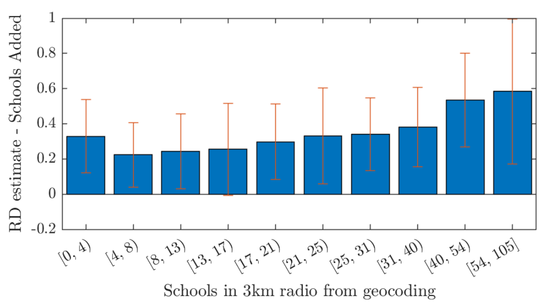
(c) Add any school by number of nearby schools



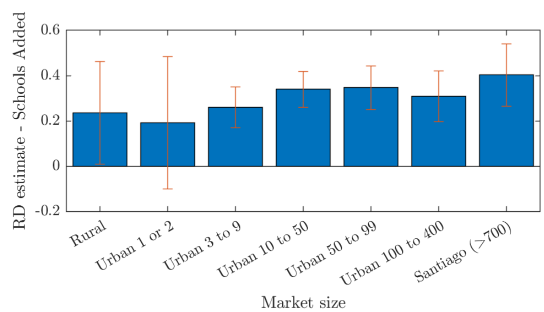
(d) Add any school by overall market size



(e) Schools added by number of nearby schools



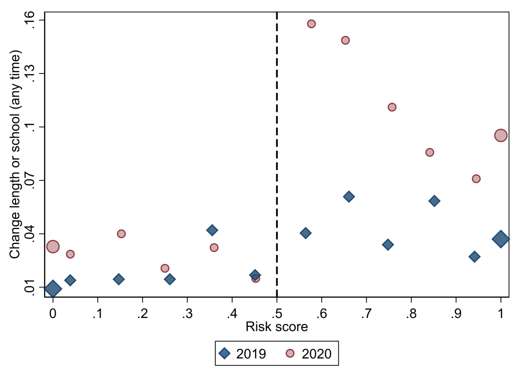
(f) Schools added by overall market size



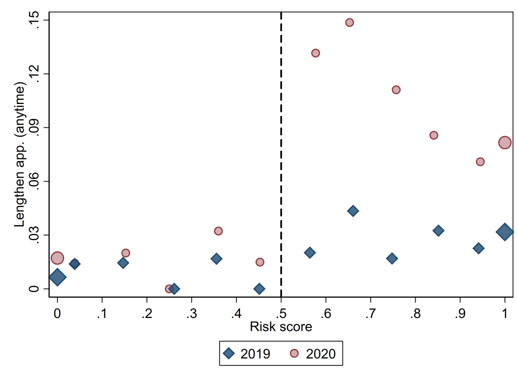
Panels A and B: Distribution of estimated popup RD effects across city-year cells. Each city-year cell is one observation. Outcome in Panel is add any school, outcome in panel B is count of schools added. Panel C: popup RD treatment effects on add any school split by count of nearby schools (within 3km of applicant address). Panel D: popup 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: Warnings intervention– New Haven, CT

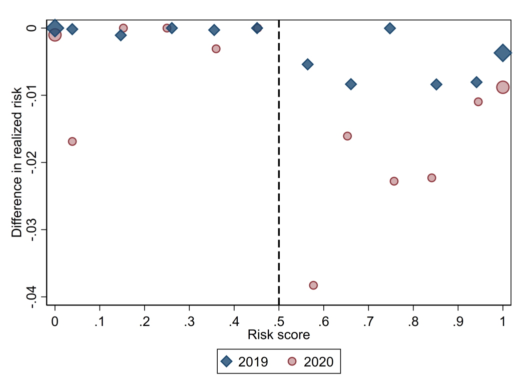
(a) Change application length or choices at any time



(b) Lengthen application at any date



(c) Difference in realized risk from initial to final portfolio



Outcomes of warnings intervention in New Haven, CT centralized choice. 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 above risk score of 0.5 received the warnings intervention; in 2019, no applicants received any intervention. Panel A: any change in application. Panel B: Lenghten application. Panel C: change in risk from initial to final portfolio. See section 5.6.2 for details.

Tables

Table 1: Descriptive statistics of 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
A. Demographics								
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. Rank-order list								
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
C. Application behavior								
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
D. Placement								
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
E. 2nd round								
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
F. School capacity available after placement (at local market level defined for each student)								
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
G. Classification by true risk of initial attempt								
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 represents means of the the population defined by the header of the column. "Risky" (column 5) is applicants whose first attempt had a predicted risks > 0.3. "Around pop-up cutoff" (column 6) are applicants whose first attempt had a predicted risk in [0.1,0.5]. True risk of initial attempt is the risk ex-post (calculated using the final set of applicants), not the predicted risk used to assign treatment. "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 applicants make 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 applicaiton, evaluated using ex post observed applications.

Table 2: Pooled Pop-Up effect

	(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)
Δ 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	0.040	0.193	0.008	0.055	
	(0.017)	(0.080)	(0.029)	(0.020)	
Enrolled in placed|placed	-0.013	-0.058	-0.021	-0.009	
	(0.018)	(0.078)	(0.031)	(0.022)	
Decline offer	-0.001	-0.006	0.007	-0.002	-0.003
	(0.005)	(0.025)	(0.016)	(0.011)	(0.007)
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)
Placed to uncong.	0.008	0.036	0.041	0.054	-0.018
	(0.007)	(0.033)	(0.018)	(0.012)	(0.010)
Placed to uncong. added pref.	0.022	0.100	0.019	0.036	0.015
	(0.004)	(0.017)	(0.010)	(0.008)	(0.005)
Enrolled in uncong. added	0.015	0.071	0.008	0.018	
	(0.005)	(0.022)	(0.008)	(0.006)	
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 popup effects from warning popup 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 estimate in the pooled sample and splitting by year. IV (column 2) shows 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. Δ 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: measures 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. "Decline offer" is an indicator for students who formally decline their placement offer, conditional on receiving an offer. Panel D: congestion attributes of behavior and placement outcomes. "Uncongested' schools are those with excess capacity.

Table 3: Whatsapp RD and RCT results

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

Table 4: RD estimates of popup effects on adding any school, split by city and year

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

Notes: RD estimates of smart platform popup effects on adding at least one school to the choice application, split by city and year. Cities sorted by count of 2020 applicants. Santiago 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: RDD and Diff-in-Diff estimates of Warnings in New Haven

	RD		Diff. in Diff.	
Outcome	β	SE	β	SE
A. Demographics				
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)
N		740		3918
B. Interaction with Simulator				
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)		
N		740		
C. Choice Outcomes				
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)
N		740		3918

RD and difference in difference 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. Applicants to eligible grades that have listed schools that did not exist in 2019 have not been randomized into treatment because their risk score cannot be predicted based on 2019 RatEx chances. Therefore we drop all those applicants that listed a new school in 2020. RD specifications based on local linear fit, dropping observations with predicted portfolio risk of of less than 1% or more than 99%. Robust SEs in parentheses. For the Diff. in Diff. panel, no observations are dropped based on their risk score. The standard errors reported in parantheses allow for heteroskedasticity. For the Diff. in Diff. panel, no observations are dropped based on their risk score. Standard errors reported in parantheses allow for heteroskedasticity.