A   Additional proofs

Note on footnote 5. 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 $\Psi_r = \prod_{j<r} R_j^*$. The derivative of the log of $V - V_0$ follows.

Proof of proposition 2. As $\Delta a \to 0$, we have

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

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

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

Finally, note that $U[\text{Search}|C_0, a]$ is decreasing in V_0. The value of search declines as one adds more schools to the consideration set. Applicants for whom the expected value of search is negative given their current consideration set also have negative values of search for all larger consideration sets. $U[\text{Search}|C_0, a] \leq C$ is therefore a sufficient condition for terminating search. $\square$

Proof of proposition 3. First, 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 $\mu_j = -G^{-1}(1 - q_j)$. The term

$$\kappa(q_j) = E[\max(-G^{-1}(1 - q_j) + \epsilon_j, 0)]$$

denotes the unique value of expected placement utility u_j that is consistent with enrollment rate q_j. 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:[1]

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

Next, note that

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

---

[1]Note that $\frac{dE[\max(\mu_j + \epsilon_j, 0)]}{d\mu_j} = P(\text{Enroll}|\text{place at j}) = q_j > 0.$
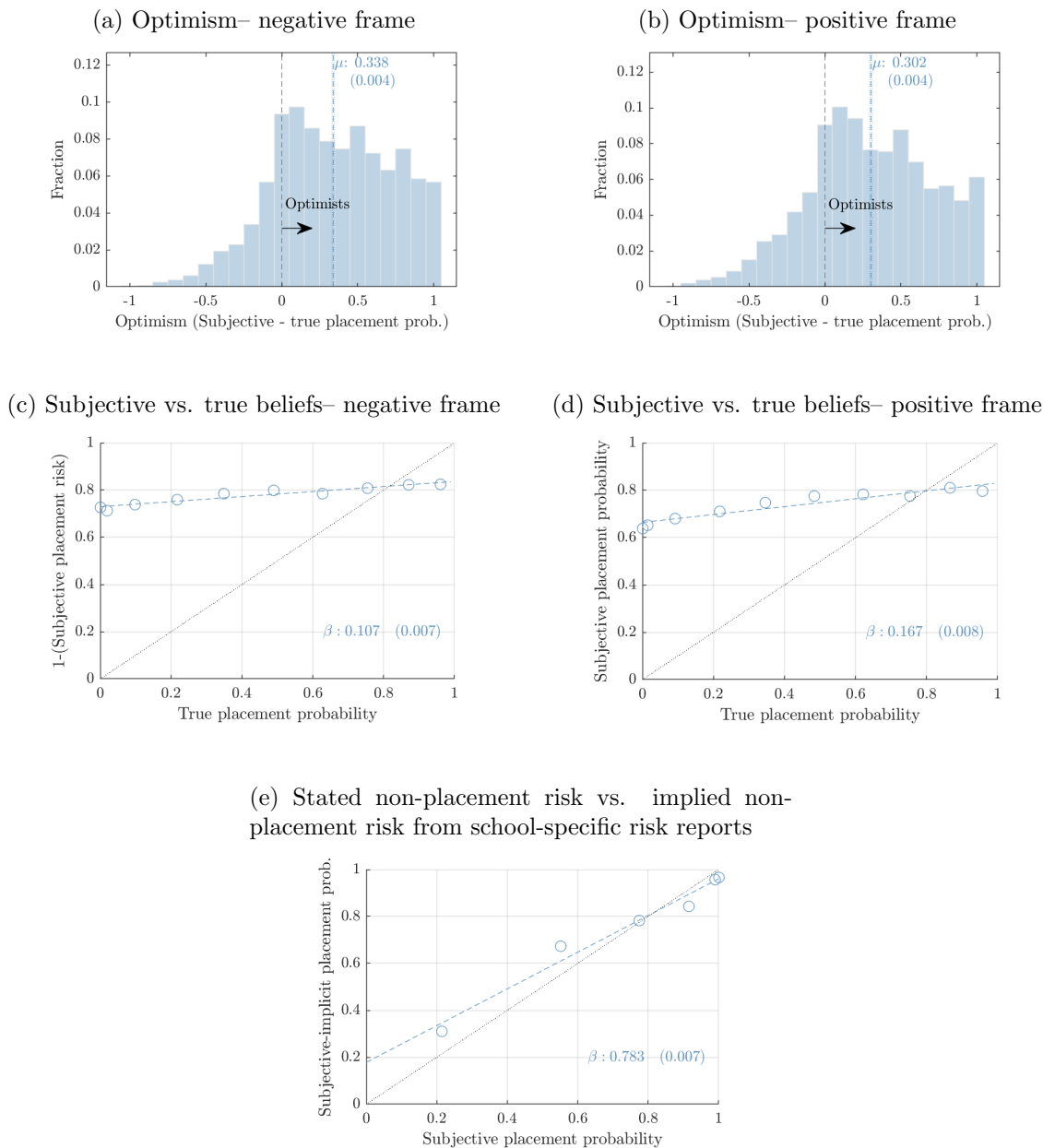
1

The first term in the sum is rises in proportion to the change in nonplacement risk. Unpacking 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. $\qquad\square$
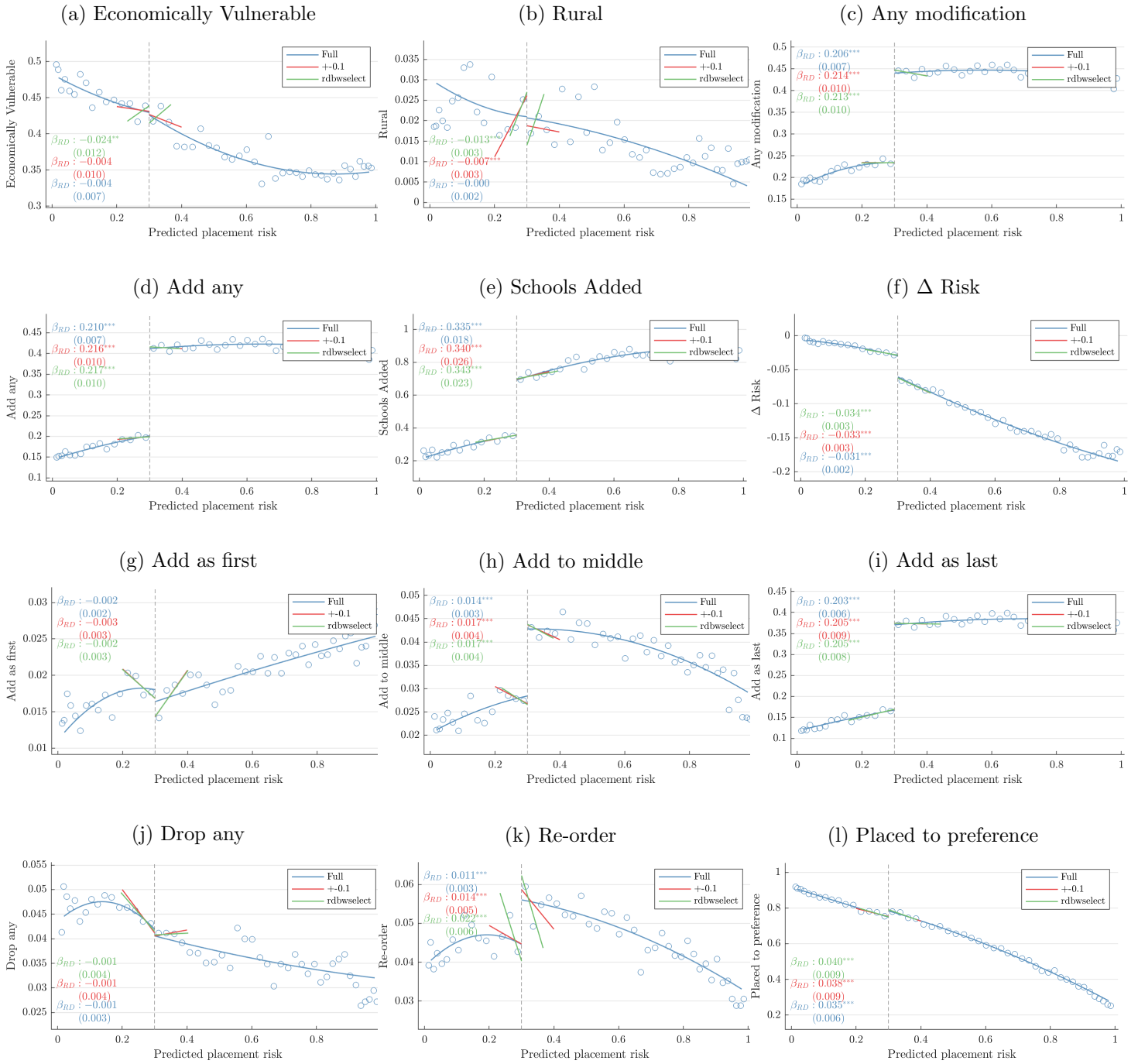
B Additional Figures and Tables

Figure B1: Alternate application risk framings

(a) Optimism– negative frame

(b) Optimism– positive frame

(c) Subjective vs. true beliefs– negative frame

(d) Subjective vs. true beliefs– positive frame

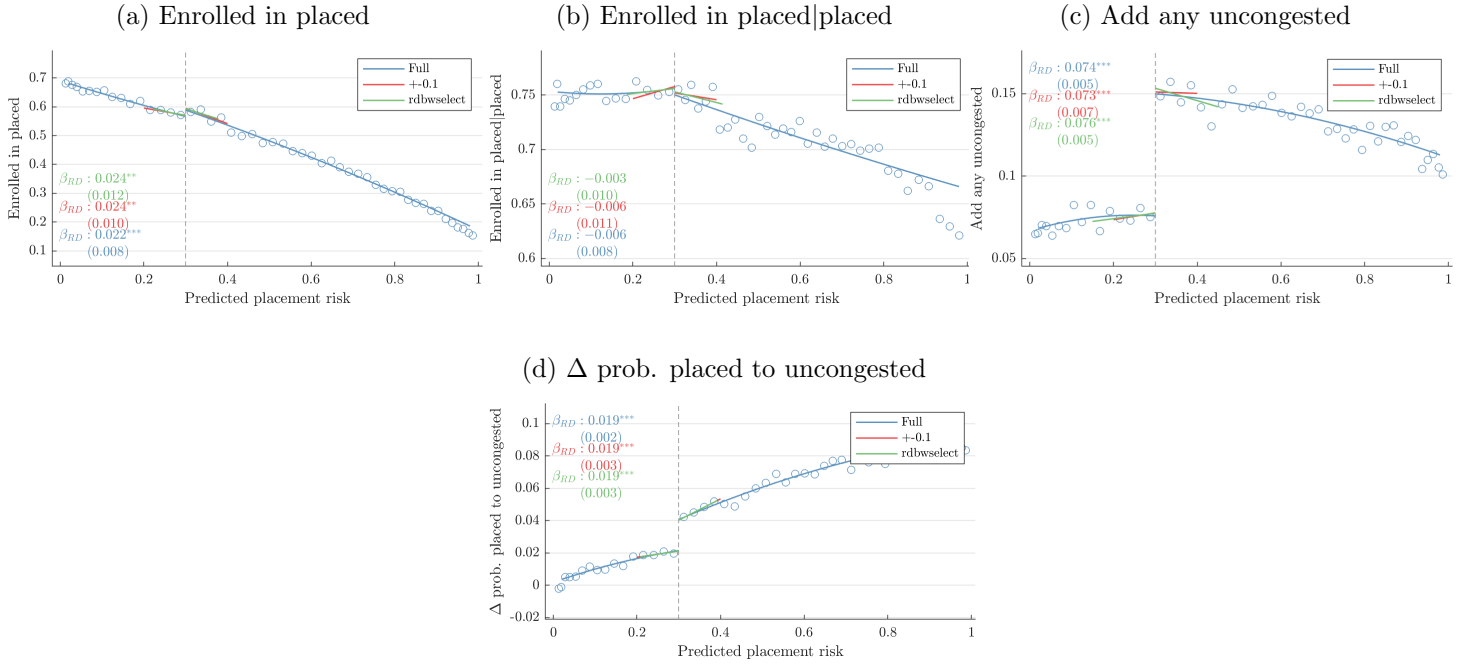(e) Stated non-placement risk vs. implied non-placement risk from school-specific risk reports

Panel A: distribution of optimism under "negative" framing for subjective risk question. Negative framing asks respondents about risk of non-placement under their submitted application. Panel B: distribution of optimism under "positive" framing for subjective risk question. Positive framing asks respondents about chance of placement under submitted application. Panel C: Binscatter of true placement probability vs. subjective placement probability under negative frame. Panel D: Binscatter of true placement probability vs. subjective placement probability under positive frame. Panel E: Comparison of subjective placement probabilty as calculated using response to question about overall placement chances (horizontal axis) vs. the placement chances implied by beliefs about placement chances at each school on their application ("subjective-implicit" placement risk; vertical axis). Subjective-implicit placement probability is calculated by multiplying out responses to school specific placement beliefs. Dashed line is linear fit. 45-degree line displayed for reference. Sample in all panels is survey completers. Panels A and B restrict to those with true non-placement risk > 0.01.

Figure B2: Multiple bandwidths RD plots of platform-based pop-up warning effects

(a) Economically Vulnerable

(b) Rural

(c) Any modification

(d) Add any

(e) Schools Added

(f) Δ Risk

(g) Add as first

(h) Add to middle

(i) Add as last

(j) Drop any

(k) Re-order

(l) Placed to preference

Notes: Pop-up warning RD effect fits and point estimates by bandwidth for outcomes listed in panel titles. "Full": global quadratic. "+/- 0.1": local linear within 0.1 bandwidth. "rdbwselect": optimal bandwidth selection using Calonico et al. (2014). See section 5.1 for details.

# Figure B3: Multiple bandwidths RD plots of platform-based pop-up warning effects

(a) Enrolled in placed

(b) Enrolled in placed|placed

(c) Add any uncongested

(d) Δ prob. placed to uncongested

Notes: Pop-up warning RD effect fits and point estimates by bandwidth for outcomes listed in panel titles. "Full": global quadratic. "+/- 0.1": local linear within 0.1 bandwidth. "rdbwselect": optimal bandwidth selection using Calonico et al. (2014). See section 5.1 for details.

Figure B4: Demand for information

(a) Information that you would have liked to have but you did not have (Chile)
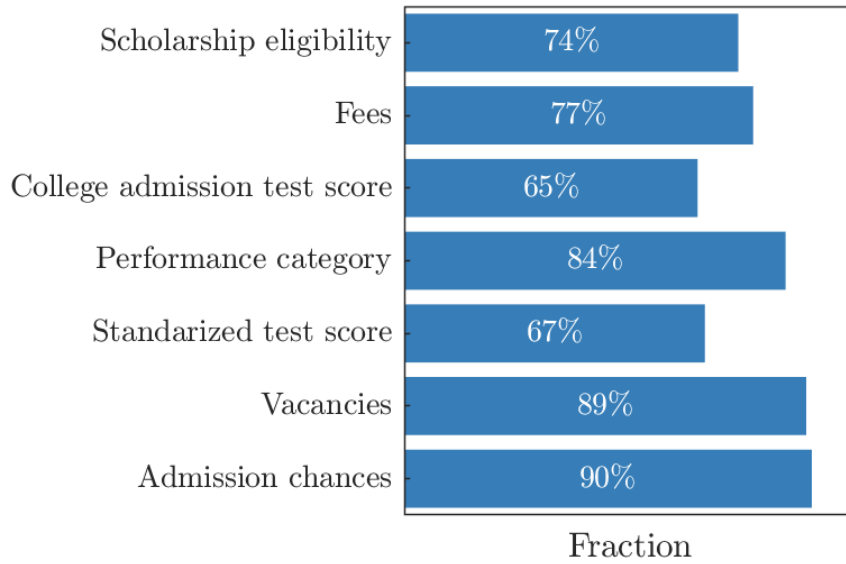
(b) Helpful additional information for future choice participants (New Haven)
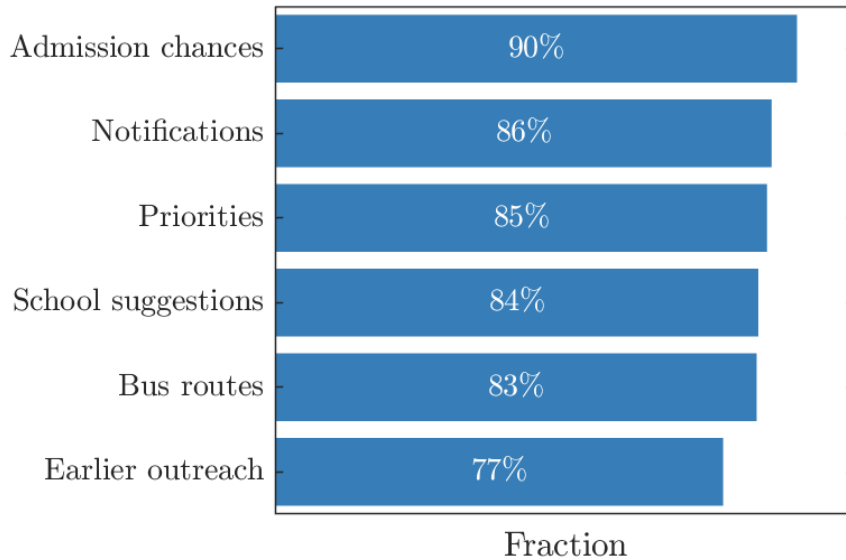
Share of survey respondents indicating desire for more information of the listed type, in response to the question listed in the panel title. Upper panel data source: Chilean school choice survey. See section 4 for details. Lower panel data source: email survey of 3,105 New Haven school choice participants in 2019 and 2020 (2,178 of those from applicants to simulator eligible grades). Bars refer to the following types of information (from top to bottom): personalized information about admission chances for school options, notifications about upcoming deadlines, detailed information about neighborhood and sibling priorities of school options, personalized suggestions for potential schools, information on bus routes to different school options, earlier outreach to allow more time to decide. See Online Appendix H for more details.

Table B1: Platform pop-up RD estimates with alternate bandwidths

Bandwidth	(1) Full Estimate	(2) +-0.1 Estimate	(3) Estimate	(4) BW left	(5) BW right	(6) N left	(7) N right
					rdbwselect		
A. Balance							
Economically Vulnerable	-0.004	-0.004	-0.024	0.07	0.07	13,853	14,095
	(0.007)	(0.010)	(0.012)				
Rural	-0.000	-0.007	-0.013	0.05	0.05	10,794	11,078
	(0.002)	(0.003)	(0.003)				
B. Choice Behavior							
Any modification	0.206	0.214	0.213	0.10	0.10	20,863	21,697
	(0.007)	(0.010)	(0.010)				
Add any	0.210	0.216	0.217	0.09	0.09	19,280	19,947
	(0.007)	(0.010)	(0.010)				
Schools Added	0.335	0.340	0.343	0.13	0.13	26,024	26,546
	(0.018)	(0.026)	(0.023)				
Δ Risk	-0.031	-0.033	-0.034	0.10	0.10	20,746	21,614
	(0.002)	(0.003)	(0.003)				
Add as first	-0.002	-0.003	-0.002	0.10	0.10	20,725	21,565
	(0.002)	(0.003)	(0.003)				
Add to middle	0.014	0.017	0.017	0.08	0.08	16,250	16,695
	(0.003)	(0.004)	(0.004)				
Add as last	0.203	0.205	0.205	0.14	0.14	29,060	29,434
	(0.006)	(0.009)	(0.008)				
Drop any	-0.001	-0.001	-0.001	0.10	0.10	21,241	22,045
	(0.003)	(0.004)	(0.004)				
Re-order	0.011	0.014	0.022	0.07	0.07	13,645	13,947
	(0.003)	(0.005)	(0.006)				
C. Choice outcome							
Placed to preference	0.035	0.038	0.040	0.09	0.09	18,068	18,676
	(0.006)	(0.009)	(0.009)				
Enrolled in placed	0.022	0.024	0.024	0.08	0.08	15,879	16,276
	(0.008)	(0.010)	(0.012)				
Enrolled in placed|placed	-0.006	-0.006	-0.003	0.11	0.11	18,177	18,105
	(0.008)	(0.011)	(0.010)				
D. Congestion-related outcomes							
Add any uncongested	0.074	0.073	0.076	0.15	0.15	31,104	31,361
	(0.005)	(0.007)	(0.005)				
Δ prob. placed to uncongested	0.019	0.019	0.019	0.09	0.09	18,826	19,513
	(0.002)	(0.003)	(0.003)				
N left	71,075	20,359					
N right	166,699	21,145					

Notes: Local linear and full sample quadratic polynomial RD estimates of pop-up effects from warning pop-up on application platform. Computed using triangular kernel with different bandwidths. "Full" bandwidth uses 2nd order polynomial fit, while "+-0.1" and rdbwselect uses 1st order (local) polynomial. Heteroskedasticity-robust nearest neighbor variance estimator with minimum of 3 neighbors reported in parentheses; computed as in Calonico et al. (2014). 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. Panel D: congestion attributes of behavior and placement outcomes. "Uncongested" schools are those with excess capacity.

Table B2: RD estimates of platform pop-up effects - 2017

Risky cutoff	(1) Pooled	(2)	(3) 0.3	(4) 0.5	(5) 0.7
		IV			
Any modification	0.237		0.107	0.352	0.137
	(0.059)		(0.112)	(0.086)	(0.140)
Add any	0.192		0.024	0.321	0.132
	(0.056)		(0.101)	(0.084)	(0.135)
Schools Added	0.428	2.225	0.109	0.651	0.394
	(0.174)	(0.739)	(0.350)	(0.214)	(0.322)
Δ Risk	-0.060	-0.313	-0.003	-0.103	-0.046
	(0.024)	(0.099)	(0.039)	(0.038)	(0.051)
Placed to preference	0.120	0.624	-0.014	0.210	-0.003
	(0.058)	(0.314)	(0.098)	(0.077)	(0.155)
NL	671	671	187	354	130
NR	647	647	194	334	119

Notes: Local linear RD estimates of pop-up effects from warning pop-up on application platform in 2017 implementation. Unlike 2018-2020 implementation, the 2017 cutoff varied by city, taking values of either 0.3 (as in the 2018-2020 implementation), 0.5, or 0.7. Computed using triangular kernel with bandwidth 0.1. Heteroskedasticity-robust nearest neighbor variance estimator with minimum of 3 neighbors reported in parentheses; computed as in Calonico et al. (2014). We report estimates in the pooled sample and for each different risky cutoff definition. IV (column 2) shows the instrumental variable specifications, where the endogenous regressor is the add any school indicator. Columns 3 through five split the sample by value of the risk cutoff. NL and NR are sample counts to the left and right of the cutoff, respectively.

C Chilean admission policy and interventions

C.1 Policy

In May 2015, Chile's congress approved the "School Inclusion Bill." The goals of the bill included addressing discrimination in school assignment (Ministerio de Educación, 2017). One major feature of the law was a change in the admissions process for public and private voucher schools in grades Pre-Kindergarten through 12. Schools of this type accounted for 92% of total primary and secondary enrollment. Between 2016 and 2020, the Ministry of Education implemented a nation-wide centralized school choice system. Rollout was staggered across regions and grades. Regions of Chile were divided into four sets. Each year a new set of regions was included in the system. For the first year following adoption in each region, only major "entry grades" were included in centralized choice. These grades were PK, K, 1st, 7th and 9th. In the second and following years of centralized choice in each region, all grades used the centralized system. The share of school-grade pairs included in the centralized system rose from 1% to 8% to 45% to 85% to 100% over the years 2016 through 2020. See Figure C1 for an illustration of the policy rollout.

Schools that only enroll pre-Kindergarten and Kindergarten students were excluded from the centralized system. Lower-grade applicants could enroll through the centralized system in schools that also offer higher grades, but not in standalone early-grade institutions.

Figure C1: Policy rollout

2016   2017   2018   2019   2020
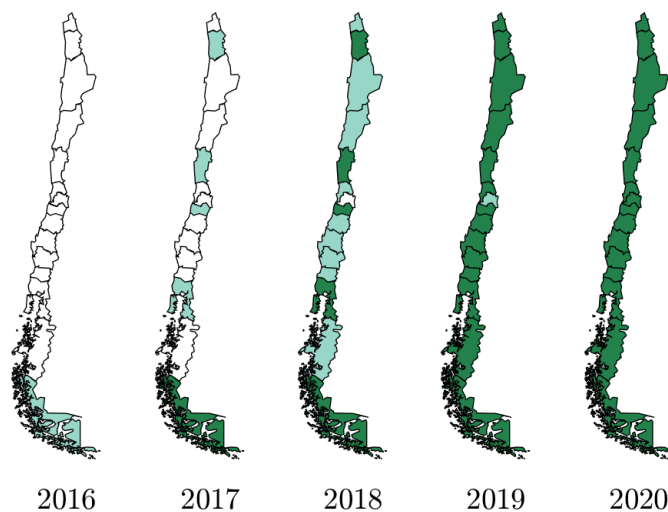
Notes: This figure shows maps of Chile representing the implementation progress of the centralized application system by year. White regions represent where the system has not been implemented yet, light green reflects implementation only in entry grades (PK, K, 1st, 7th and 9th grades), and dark green means implementation in all grades of the region.

The centralized assignment process used a Deferred Acceptance algorithm with multiple tiebreakers (DA-MTB) to assign students to schools.[2] The law dictated that assignment include priorities for the following groups, in order: siblings, applicants with parents working at schools, and former students (i.e. students who previously were enrolled at a school but

---

[2]For comparison over different lottery systems, see Ashlagi and Nikzad (2020).

left). Ties were broken with lotteries. The law also imposed quotas for economically vulnerable students (15% of seats) and special needs applicants (count decided by each school, with a cap of two students per classroom). Finally, a small set of high schools was allowed to use quotas for high-performing students (30% of seats). In 2020, 39% of the schools offered seats for applicants with special needs in at least one level, while only 0.3% (23 schools) had a quota for high-performing students.

Following the initial assignment round, there is a second application round for applicants who do not receive or do not accept their initial offer. This round offers seats at schools with remaining excess capacity after the initial allocation.[3] Applicants without an offer after the second round are administratively assigned to the closest school to their registered address that has excess capacity.[4]

Students submit their applications to each choice round through an online platform. In addition to collecting applications, this platform provides information about schooling options, including test scores, fees, infrastructure, enrollment, religion, and extracurricular activities. See Figure C2 for an illustration of the search engine and information. The Mineduc IT department developed the website, and a team from the Industrial Engineering Department of Universidad de Chile coded and ran the algorithm. The seed for the pseudorandom number generator is a mapping from the characteristics of the last six earthquakes recorded in Chile for a given date.

---

[3]There is also a waitlist process between rounds that fills declined offers in over-demanded schools. This process works as follows. After assignments are made for the first application round, placed students are given the option accept their placements, to accept their placement, but consider waitlist options, and to reject the placement. The default choice for non-responders is to accept the placement. Mineduc then reruns the DA process using the already-submitted rank-order lists, with the set of applicants restricted to students who were unplaced, rejected their placements, or chose the "accept, but consider waitlist" option. The set of seats consists of those opened up by applicants rejecting placed schools. These additional "interim" placements account for a very small share of placements overall; in 2020, for example, they made up 1.2% of all placements. We do not include these interim placements in our main-text analysis of placement counts.
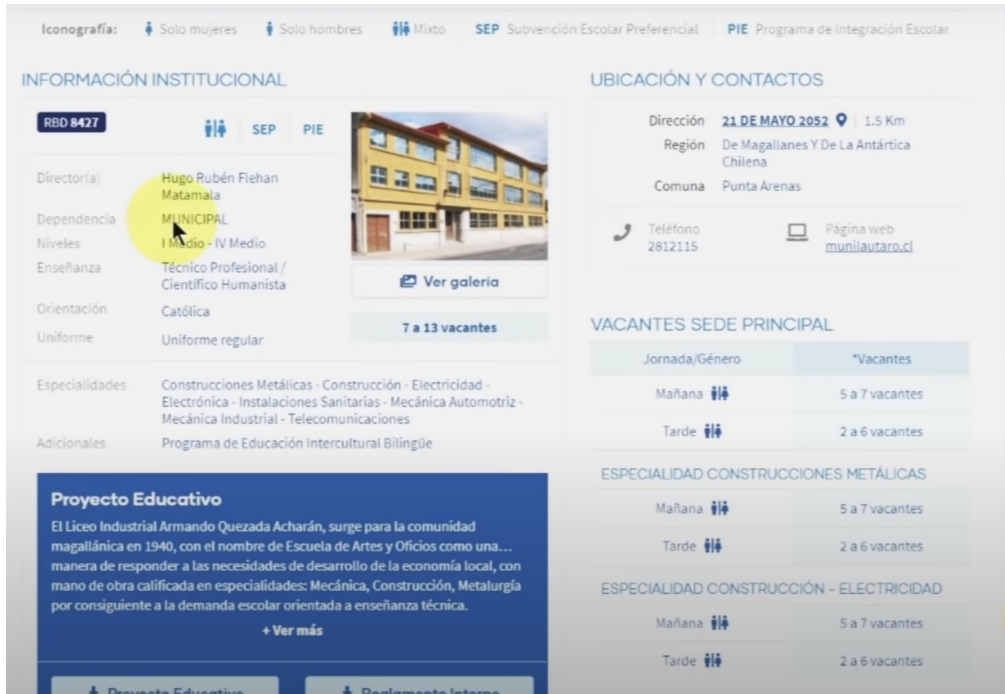
[4]In 2016, Mineduc also implemented default assignment to the closest school for the *first* application round. We do not use 2016 data in our analysis.

Figure C2: Application platform screenshots

(a) Gallery of schools
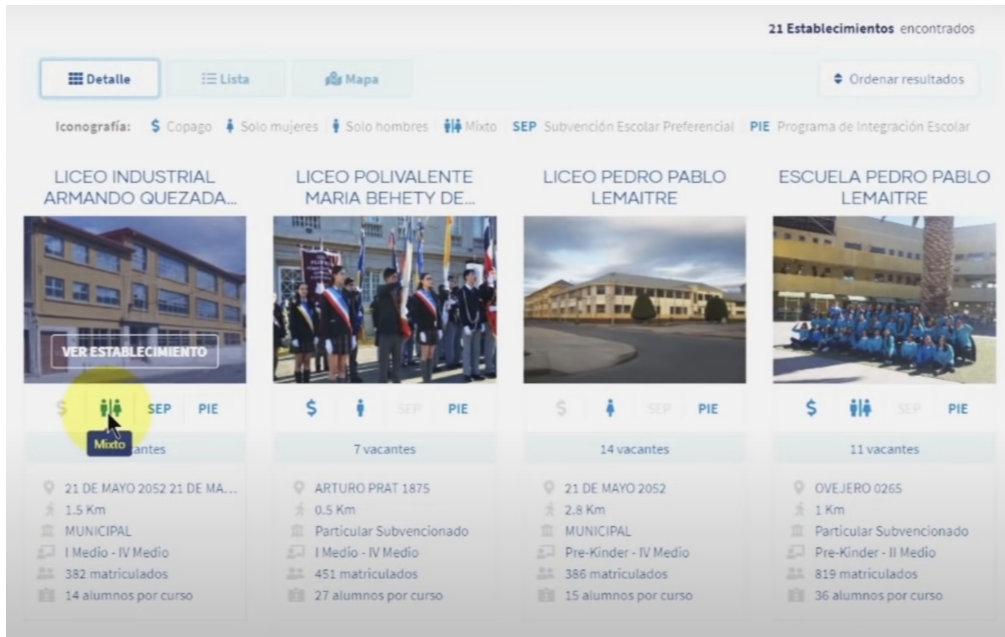
(b) Detailed information of a school

Notes: Panel A shows an example of an applicant's view of the gallery of schools, what includes the main photo and a few characteristics as distance to home, enrollment, or price. Alternatively, users can also choose to see a list of schools or a map with all the options available close to home. Panel B is a screenshot of the information for a specific school, including educational project, an estimate of the seats available, religion, etc.

C.2 Policy rollout, placement outcomes, and outside options

C.2.1 Policy implementation and policymaker concerns

Rollout of the centralized system proceeded more or less as planned. One concern for policymakers following the adoption of the centralized system was the proportion of applicants not assigned to any school they had listed. Reasons for this concern included the disutility of waiting for certainty about enrollment, the potential for unrealized matches between families and schools, the costs of aftermarket coordination, and the effect on the new system's reputation. To the latter point, policymakers were concerned that families might expect the centralized school choice system to assign all applicants.

C.2.2 Placement outcomes over time

Table C1 reports placement statistics for each year. It displays both aggregate placement statistics and statistics for the set of markets that first entered the centralized system in a given year, for each entry year between 2016 and 2020. Results in Panel A show that the share of applicants assigned to any preference has decreased. This is true both overall, and, to a lesser extent, within the set of markets entering the centralized system in each year between 2016 and 2020. Panels B and C show the share of applicants (in the full applicant sample) who are unplaced and are entering the schooling system or do not have the option to continue in their current school (Panel B, "Share not placed in pref. without continuation option") and the share who are unplaced but have the option to continue (Panel C, " Share not placed in pref. with continuation option"). Most of the rise in nonplacement comes through the latter channel.

Panel D of Table C1 shows that mean application length declines over time, both overall and within each year-of-entry group.

Table C1: Aggregate placement results for stable populations by year

	2016	2017	2018	2019	2020
A. Share placed in preference					
Aggregate	0.884	0.872	0.828	0.771	0.786
Since 2016	0.884	0.840	0.799	0.767	0.791
Since 2017		0.874	0.821	0.807	0.850
Since 2018			0.831	0.788	0.820
Since 2019				0.740	0.770
Since 2020					0.606
B. Share not placed in preference without continuation option					
Aggregate	0.088	0.087	0.089	0.104	0.094
Since 2016	0.088	0.122	0.121	0.152	0.117
Since 2017		0.085	0.110	0.116	0.096
Since 2018			0.081	0.102	0.094
Since 2019				0.100	0.096
Since 2020					0.079
C. Share not placed in preference with continuation option					
Aggregate	0.028	0.041	0.083	0.125	0.120
Since 2016	0.028	0.038	0.080	0.080	0.092
Since 2017		0.041	0.069	0.077	0.054
Since 2018			0.088	0.110	0.086
Since 2019				0.159	0.135
Since 2020					0.315
D. Mean of application list length					
Aggregate	3.4	3.5	3.2	3.2	3.0
Since 2016	3.4	3.2	3.2	3.2	2.9
Since 2017		3.5	3.3	3.1	2.9
Since 2018			3.1	3.1	2.9
Since 2019				3.4	3.3
Since 2020					2.9

Notes: In Panel B "Share not placed in preference without continuation option" also includes applicants that do not have a current school because they are entering the schooling system. "Since 201X" represents the zones and grades where the new centralized system was implemented in 201X. Therefore, each row represents a stable population. The category "Change of school not realized" reflects students that applied from a school that offers the next grade and were not assigned to any submitted preference. They kept the seat at their school of origin.

C.2.3 Options outside the centralized system

This subsection describes applicants' outside option behavior in detail. Figures C3a and C3b show the next-year enrollment of students that did not get a spot in the first round of the centralized process (i.e., those whom we classify as unplaced in our main text analysis). We split the sample by whether applicants have an option to continue in their current school or not. Figure C3a shows enrollment outcomes for applicants without the continuation option. Roughly 60% of these applicants go on to enroll in some SAE school. Roughly 20% enroll in a voucher school outside of the centralized system. These students are mostly pre-kindergarten or kindergarten applicants enrolling in standalone preschools that do not participate in the main system. Roughly 12% enroll in a private school that does not accept vouchers; private schools that decline vouchers tend to be quite expensive. For a small fraction of unplaced students (4-11% depending on the year), we do not observe any enrollment outcome.

Figure C3b shows that roughly three quarters of unplaced applicants who have the option to continue in their current school choose to do so. Almost all of the remaining applicants enroll in another SAE school.

Figure C3: Enrollment outcomes for unplaced students

(a) No continuation option

(b) Continuation option
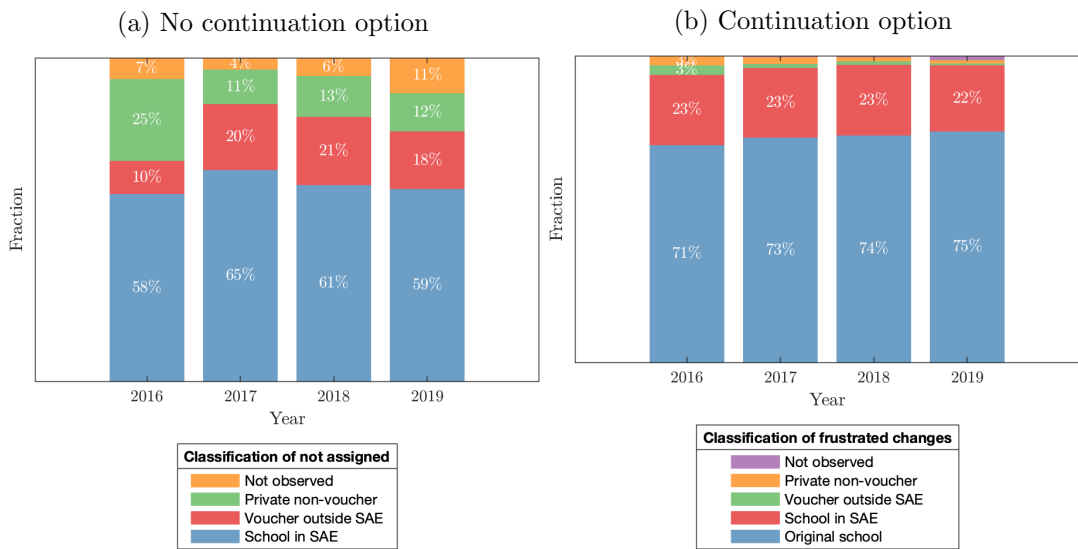
Notes: Panel A shows where we observe the applicants enrolled for the set of students that were not placed to any school in the assignment and did not have the option to continue at their current school. Panel B is the same for the set of applicants who are unplaced in the main assignment round but have the option to continue at their current school.

C.3 Feedback interventions

Mineduc worked with the Consiliumbots NGO to design and implement information interventions in the placement process each year between 2017 and 2020. Our analysis focuses on the 2018-2020 placement cycles, when the placement process was operating at or close to full scale. This section describes the interventions conducted in each year, including the 2017 intervention.

2017 There were two information interventions this year. The first was a live feedback plat-

form pop-up warning to risky applicants. The threshold that defined a "risky applicant" varied between 30, 50, and 70% placement probability, depending on the region. The pop-up was active starting with in the third day of the process. Before the end of the process, SMSs were sent to risky applicants who applied during the first two days that the pop-up was not active.

2018 The implementation team at Mineduc kept the "live feedback" (pop-up window) from 2017, fixing the definition of "risky" as applicants with a predicted probability of non-assignment higher than 30%. It was active from the 3rd day of application. Only applicants applying after this date are included in our platform pop-up analysis. Figure C4 displays the platform pop-up messages in 2018 and also in 2019-20.

Followup SMS messages were sent to risky applicants who applied during the first two days, when the pop-up was not active. A final SMS was sent four days before the end of the period to every risky applicant that did not receive the first SMS.

2019 As in 2018, but without the final SMS reminder to still-risky applicants. The platform pop-up was again active from the 3rd day of application. Only applicants applying after this date are included in our platform pop-up analysis.

2020 The platform pop-up was implemented again, this time from the first day of application. Mineduc sent an SMS to risky applicants eight days before the deadline.

Additionally, four days before the final application deadline, the NGO randomized the assignment of WhatsApp messages within risky applicants. The reason for the randomization was a cap on the number of messages that could be sent in a day. The message included one of three types of images, each with a slightly different type of information related to the risk and the application. See Figure C5. Each image conveys the same idea as the pop-up warning. We pool over image types in our analysis of the WhatsApp RCT.

The set of students eligible for randomization into RCT treatment and control groups was restricted in several ways. First, the RCT sample included applicants to grades Pre-Kindergarten, Kindergarten, and 1st grade, from urban zones, without sibling priority at any school on their list. Second, the warnings RCT sample was layered on top of a parallel information intervention on school attributes also being conducted by the NGO and Mineduc. This second intervention involved sending emails about school attributes to choice applicants. The sample for the warnings RCT was drawn from the set of applicants who received the attributes email but had not yet opened it. The reason for this is the WhatsApp campaign was viewed at Mineduc as a reminder to check this report card. This sample selection approach does not affect the *internal* validity of the WhatsApp experiment, but does affect how one interprets the findings. Our view is that the approach to sample selection will tend to draw relatively low-interest applicants (those who had not opened other correspondence from the authority), but in an environment with relatively low search costs (they had access to the attribute report cards if they wanted them). Recall that that search costs may be relatively low for many applicants in this setting, given the attribute search tools embedded in the application platform (see section C.1).
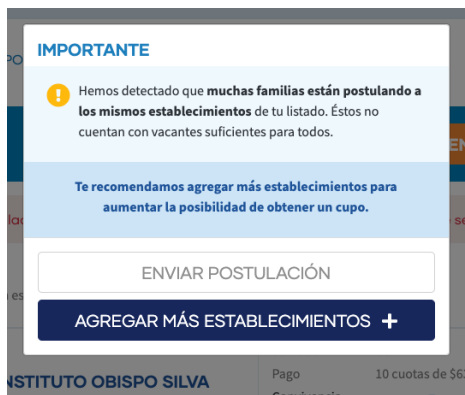
Furthermore, two days after the WhatsApp message went out, Mineduc sent a final reminder using SMS messages to remaining still risky applicants with a link to the same image attached in the WhatsApp message, encouraging applicants to add more schools.
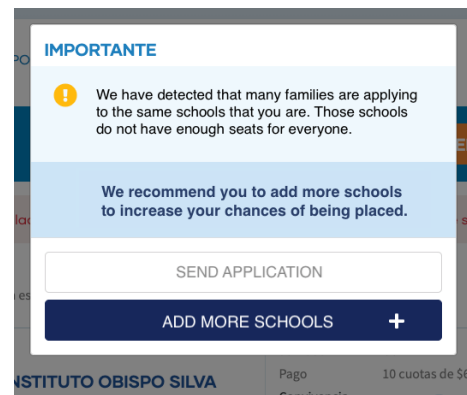
Figure C4: Platform pop-ups

(a) Platform pop-up 2018

(b) Platform pop-up 2019 - 20

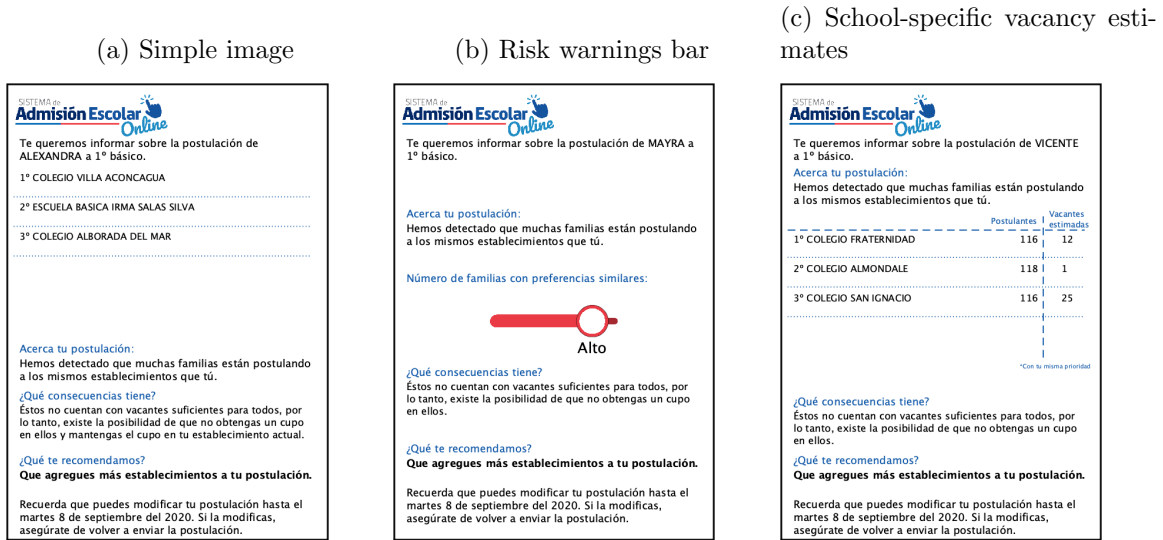(c) Translation pop-up 2019 - 2020

D Datasets

Datasets come from three sources. The first is publicly available administrative data. These data include all the inputs necessary to replicate assignment outcomes (including application lists, priorities, lottery numbers, and seat counts) as well as historical enrollment records.[5] The second source is confidential administrative records on application submission and edit histories (i.e. the time path of submitted applications for a given applicant) and priority groups for each applicant at each school. We use this data to construct risk predictions. The third source is survey data collected after the application process in 2020. This section describes each of these sources.

---

[5]These data are available on the Mineduc website, `https://centroestudios.mineduc.cl/`.

Figure C5: Whatsapp intervention warning images

(a) Simple image

(b) Risk warnings bar

(c) School-specific vacancy estimates

D.1 Public administrative data

Public data comprises all the necessary inputs to run the assignment if the algorithm is known. It includes rank order lists, priorities, vacancies, lottery numbers, and final assignments.

D.2 Confidential administrative data

In addition to the public data, we have access to applicants' priorities to every school, geocoding, and application edit histories. The details of data availability vary by year, particularly with regard to application edit histories and eligibility for the platform pop-up intervention. Looking across 2018, 2019, and 2020– the year range included in our main text analysis– we observe edit history data for 96.4% of applicants.

2016: (Not used for analysis.) We have daily copies of the application database. If an applicant files more than one application within a day, we observe only the last one.

2017: The NGO's risk classification web service stored the history of applications for students from the 20 most significant urban zones, covering 88% of the total. We have access to these data.

2018: The NGO's risk classification web service stored the history of applications for 84.2% of the applicants. We have access to these data. The initial intention here was to store application history for all applicants, but the remaining 15.8% of histories were not retained. The Mineduc IT team believes that the omission is because of a timeout they set to reach our web service. Applicants in this 15.8% did not receive the platform pop-up, and therefore are not included in our analysis of pop-up outcomes. Our platform pop-up analysis imposes two additional data restrictions, dictated by limits on the set of applicants eligible for the intervention. First, the NGO risk classification system considered only the 20 largest urban zones; students in other zones did not receive the

pop-up. Second, the risk classification system excluded applicants that applied during the first two days of the application process; the web service was not active until day three of the application cycle (see above). Overall, our pop-up analysis includes 44.3% of all 2018 choice applicants.

2019 The NGO's risk classification web service stored the history of applications for 99.9% of the applicants. Our analysis of the platform pop-up imposes the additional restriction that applicants must have applied after the first two days of the application cycle, because the web service providing the warnings was not active until day three (see above). Overall, our pop-up analysis includes 65.1% of all 2019 applicants.

2020 The risk classification web service stored the history of applications for all applicants. Our analysis of the platform pop-up excludes 9% of applicants applying to non-entry grades in the Metropolitan region, for whom there was no prior-year data on which to base initial risk predictions. These applicants did not receive the pop-up intervention. Overall, 91% of applicants 2020 are included in the pop-up analysis.

D.3   Survey Data

We sent an online survey to the parents of all applicants. The sample universe consisted of 373,710 households. This is slightly smaller than the number of applicants because we sent one email to each parent, and some households have multiple applicants. 66,282 (18%) started the survey, and 48,929 (13%) finished the survey. See section G for survey text.

E   Treatment first stages

This section describes how warnings treatments of different types interact with each other. As stated in section 3.2, our broad goal is to make the point that warnings about application risk affect application behavior, not to disentangle the effects of warning timing or media. Because some readers may nevertheless be interested in the precise combinations of treatments that applicants received in different years, we describe the details here.

Figure D1 shows how treatment status varies with the value of the initial application risk prediction used in assignment of the platform pop-up in 2018 and 2019. In 2018, there were two kinds of interventions: the initial pop-up, and a followup SMS sent to still-risk applicants. Panel A shows that crossing the risk cutoff raises the count of applications students receive from zero to 1.07. The latter number is slightly above one because applicants who respond to the warning by submitting a revised application can see the warning a second time if the revised application is also risky. Panel B shows that there is a positive association between initial application risk and eventual receipt of the SMS reminder, but that there is not much of a discontinuity in SMS receipt at pop-up cutoff. This is because risk predictions change over time as more applications come in, so designation as risky later in the process is (relatively) smoothly distributed with respect to the initial risk score. Our evaluation of the pop-up RD in 2018 thus reflects the effects of receiving the pop-up warning for initial students, in a setting where the warning may interact with follow-on SMS interventions.

In 2019, the only intervention was the platform pop-up. Panel C of Figure D1 shows how pop-up receipt varied across the 30 percent predicted risk cutoff. As in 2018, the count of
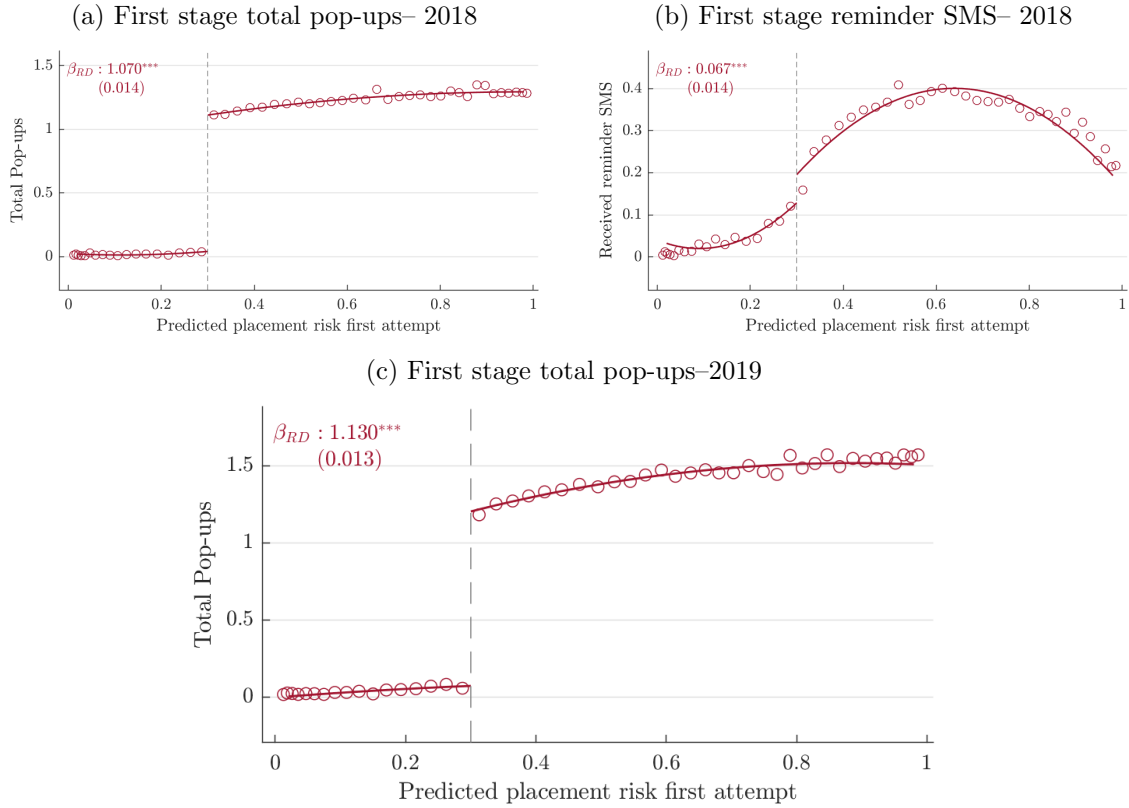
pop-up warnings applicants receive rises by slightly more than one across the cutoff, because applicants who revise their rank lists may receive more than one. The interpretation here is straightforward, since there were no follow-on interventions in this year.

Figure D2 shows how treatment outcomes for the platform pop-up, the two SMS warnings, and the RCT intervention vary with respect to the initial risk score cutoff for pop-up assignment. These graphs can help us understand how treatments change across the cutoff in our evaluation of the main pop-up RD in 2020. As shown in panel A, crossing the cutoff raises the count of platform pop-ups that applicants receive by a little more than one, just as in previous years. As shown in Panels B through D, receipt of follow-on SMS and Whatsapp interventions is postively correlated with initial risk, but discontinuities in follow-on treatments at the initial risk cutoff are not as sharp, because applicants respond and risk evaluations change. Relatively few students receive the Whatsapp RCT intervention, because this was conducted in a subsample. Our RD estimates of the effects of the 2020 platform pop-up intervention thus reflect the effect of the initial treatment, inclusive of interactions with SMS and Whatsapp follow-ons.

Figure D3 shows how treatment outcomes for the platform pop-up, the two SMS warnings, and the RCT intervention vary with respect to the risk score cutoff used to assign treatment in the Whatsapp RCT. This score was computed just before the Whatsapp RCT, later in the process than the risk scores we use for assignment platform pop-ups. The sample in these graphs is students assigned to the treatment or control group in the Whatsapp RCT. Most people in both the treatment and control groups for the Whatsapp RCT intervention have already received a platform pop-up and an initial SMS warning (Panels A and B). Reciept of the RCT intervention itself jumps by 42% at the cutoff for treated students; the difference from one reflects the rate at which applicants randomized into the treated group open the message and view the warnings image. Most risky individuals in both the treatment and the control group also receive the follow-on SMS.

In the main text analysis of the Whatsapp intervention, we present both RCT comparisons of outcomes for treatment and control groups, and RD evaluations across the risk cutoff for both treatment and control. The detailed analysis of treatment receipt presented in Figure D3 shows that the RCT estimates should be interpreted as intensive-margin effects of the additional Whatsapp intervention, in the context of the platform pop-up and SMS treatments. RD estimates of endline outcomes in the treatment group reflect the combined effect of the Whatsapp treatment, together with the effects of the discontinuous jump in platform pop-up and initial SMS warnings receipt at the cutoff. RD estimates of endline outcomes in the control group reflect similar discontinuous effects in the platform pop-up and initial SMS warning, but not the Whatsapp intervention itself.

Figure D1: RD first stages of 2018 and 2019 interventions

(a) First stage total pop-ups– 2018

$\beta_{RD} : 1.070^{***}$
(0.014)

Total Pop-ups

Predicted placement risk first attempt

(b) First stage reminder SMS– 2018

$\beta_{RD} : 0.067^{***}$
(0.014)

Received reminder SMS

Predicted placement risk first attempt

(c) First stage total pop-ups–2019

$\beta_{RD} : 1.130^{***}$
(0.013)

Total Pop-ups

Predicted placement risk first attempt

Warnings interventions in 2018 and 2019, by position relative to risk score prediction. Panel A: Count of platform pop-ups received in 2018 intervention. Panel B: count of reminder SMS messages received in 2018 intervention. Panel C: count of platform pop-ups received in 2019 intervention. 2019 intervention did not include SMS reminders. 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 3.2 for intervention details.

Figure D2: RD first stages in 2020 platform popup intervention
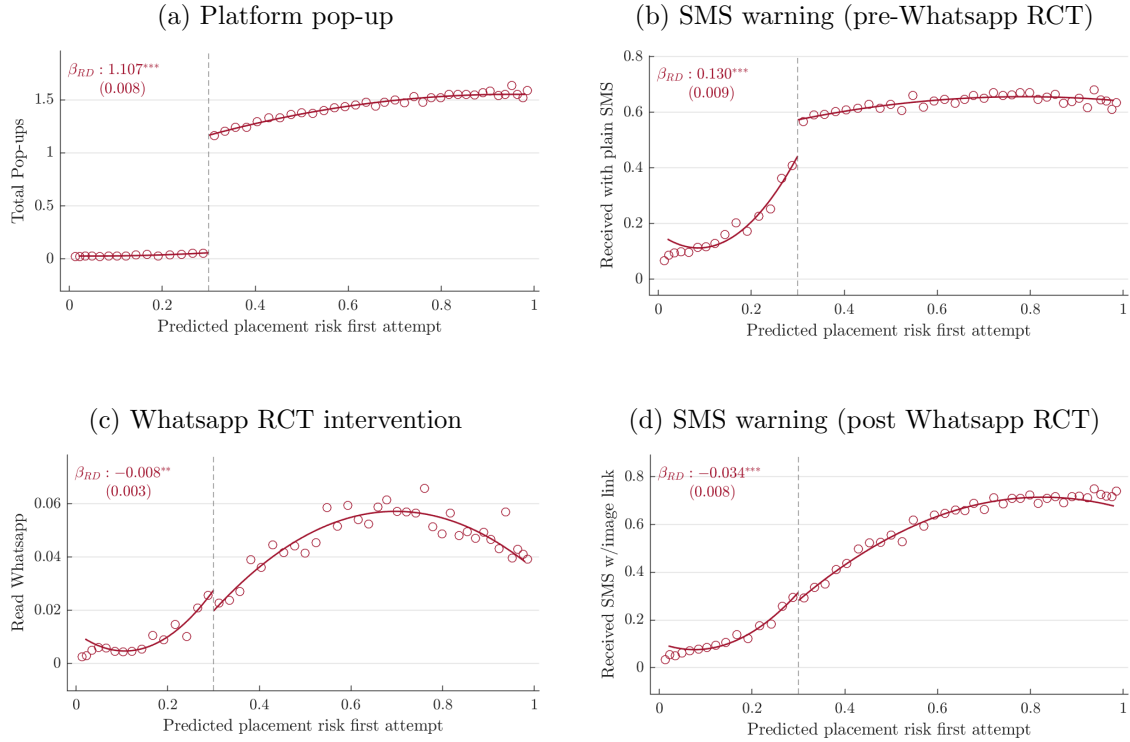
(a) Platform pop-up

(b) SMS warning (pre-Whatsapp RCT)

(c) Whatsapp RCT intervention

(d) SMS warning (post Whatsapp RCT)

Figure describes the receipt of various warnings interventions by risk prediction at the time of the on-platform pop-up intervention in 2020. For all graphs, risk predictions are as computed at the time of the initial platform pop-up. Panel A: count of platform pop-ups. Panel B: count of pre-Whatsapp reminder SMS interventions. Panel C: share of individuals reading the randomized Whatsapp intervention. Panel D: share of individuals receiving follow-up SMS, after Whatsapp. 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 3.2 for intervention details.

Figure D3: First stages of 2020 Whatsapp intervention
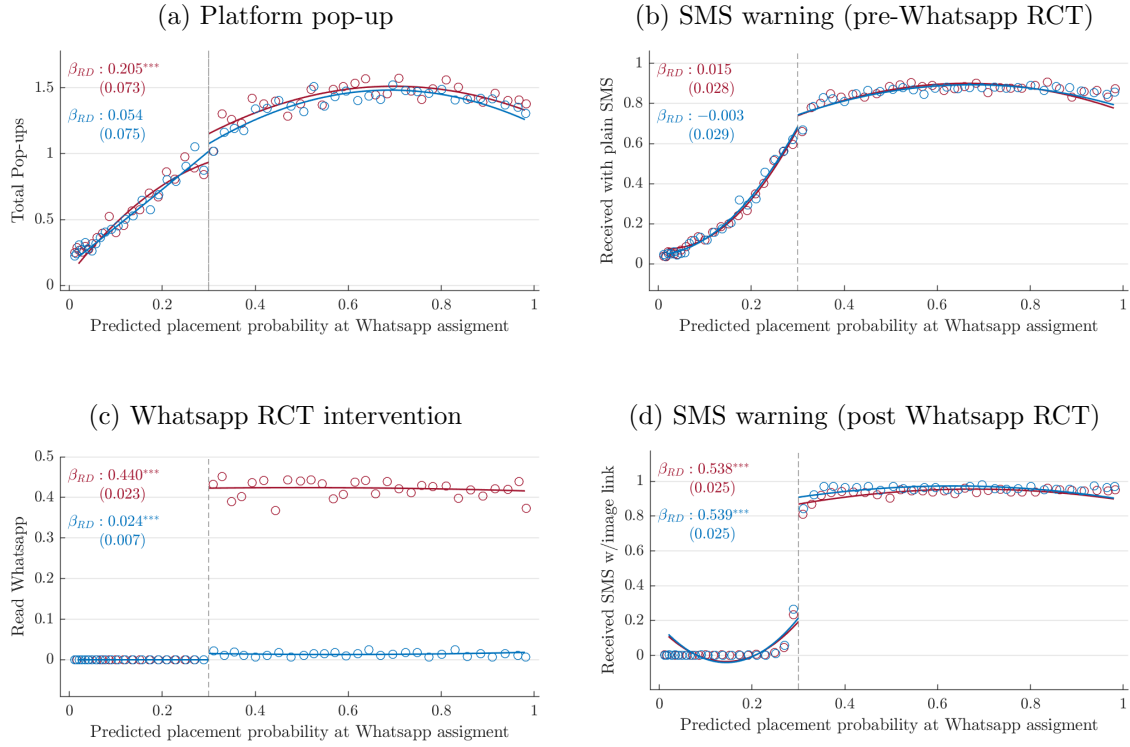
(a) Platform pop-up

(b) SMS warning (pre-Whatsapp RCT)

(c) Whatsapp RCT intervention

(d) SMS warning (post Whatsapp RCT)

Figure describes the receipt of various warnings interventions by risk prediction at the time of the Whatsapp RCT intervention in 2020. Sample is the treatment group in the RCT intervention only. For all graphs, risk predictions are as computed at the time of RCT intervention, not the initial platform pop-up. Panel A: count of platform pop-ups. Panel B: count of pre-Whatsapp reminder SMS interventions. Panel C: share of individuals reading the randomized Whatsapp intervention. Panel D: share of individuals receiving follow-up SMS, after Whatsapp. 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 3.2 for intervention details.

F Risk prediction

F.1 Obtaining admissions chances from application data

We define risk as nonplacement probability. In this section we describe how the NGO computed the simulated values of application risk that form the basis for the information interventions.

The NGO took a simulation-based approach. This required three inputs for each market-year: a seat count for each school and grade, a projected total number of applicants N, and a joint distribution of applications and student types (i.e., priorities). At each iteration of the simulation, we resample N applications from this joint distribution, and redraw the lottery tiebreakers. We then run the deferred-acceptance algorithm and determine the probability of admission for applicants in each school-grade-priority group combination conditional on not being admitted to a more-preferred option. For example, if 50 people in priority group p apply to school-grade j and did not get into an option they prefer, and 10 of these people are admitted to j, then the simulated placement probability for this school-grade-priority group cell is $10/50 = .2$. We repeat this process 500 times, and then compute the average admissions probability for each school-grade-priority group over all the simulations. This procedure is similar in spirit to Agarwal and Somaini (2018), but differs in that we average over probabilities within priority groups rather than cutoff scores. Given this set of (conditional) admissions probabilities for each school-grade-priority cell, applications are identified as risky if nonplacement risk is above 30%.

F.2 Simulated demand

We now describe how the NGO approximated the count and joint distribution of applications and priorities. The empirical distribution of applications and priorities is not observed until after the application deadline. In practice, the NGO had to project this distribution based on a combination of past data and early applications to the centralized system.

We begin the demand prediction procedure by identifying urban zones that correspond to education markets. We do this using the urban polygons of the census of 2017, similar to Neilson (2020). We merged two polygons if they were close enough or if we observed a substantial flow of students from one market to another. Each market is then defined by a set of polygons and a grade. A school belongs to a market if its location lies inside one of the polygons and offers the grade that defines the market. We consider schools located outside of every polygon as part of a "rural market" defined by region and grade.

Given these market definitions, the NGO took two different approaches to simulate the distribution of applications and priorities. The first approach was used for 2020 simulations. This approach uses the previous year's applications to estimate this year's congestion, given the current-year supply of seats in each academic program by grade combination. A challenge for this approach is that the menu of available programs is defined not just by school and grade, but also in some cases by gender, "shift" (morning, afternoon, or both), campus, and specialization (only for grades 11 and 12). In some cases, the menu of available options along these additional dimensions shifts over time. The approach the NGO took was to first match on all of these attributes, and keep available one-to-one matches. Then, for programs without a one-to-one match, they matched on school, grade, and gender only, and manually inspected

cases to determine the best of multiple options when more than one was available, allowing for one-to-one or one-to-many matches. Completely new academic programs were assigned values of zero risk; applicants listing new programs did not receive risk warnings. This approach followed from the idea that the goal was to identify known risks, rather than speculate on what options might prove to be risky, and with the knowledge that the there would be later followup based on observed same-year application counts. The advantage of this approach is that it can be applied from the beginning of the choice process, before any applications are submitted.

The second, which was used for the 2018 and 2019 processes, and beginning in the sixth day of the 2020 choice process, resamples from current applications to meet the expected count of applicants within each market. i.e., we estimate the count of applicants N in a given market, then resample N applications from the $M < N$ applicants who have already applied in the market. This method cannot be used for applicants at the very beginning of the application process; this is why students applying in the first few days of the 2018 and 2019 processes are omitted from the platform pop-up intervention (see above).

We compute the expected count of applicants, N, in each market as the number of students who participated in choice in the market in the previous year, or, in markets where there is no previous-year data, as the number of students who entered a new school at the start of the most recent school year (i.e. who switch schools between December and March).

F.3 Changes in risk predictions within the application cycle

Risk predictions change during the application process as more applications are filed. Applicants may have different values of predicted risk at the moment of application and when the SMS reminder is sent, even if they do not change their rank-order lists. Risk predictions presented on the platform pop-up and in the various SMS and Whatsapp interventions reflect live updates about the count and distribution of applications conducted every 3 days.

G 2020 Survey translation

Figure F1: 2020 Survey landing page

Maria, has sido invitado(a) a participar en la Encuesta de Satisfacción del Sistema de Admisión Escolar. Este es un esfuerzo conjunto entre el Mineduc e investigadores de la Universidad de Princeton. Tus respuestas servirán para mejorar el proceso de postulación y la información que se entregará a las familias en el futuro. Ten en cuenta que:

- Tus respuestas no afectarán en ningún sentido tus resultados en el Proceso de Admisión.
- La participación es completamente voluntaria, puedes detenerla en cualquier momento
- Todas tus respuestas son confidenciales.
- Solo el personal autorizado por el Mineduc tendrá acceso.

He leído la información sobre la Encuesta. Doy mi consentimiento para participar:

○ Sí

○ No

Siguiente →

Notes: This is the website displayed after applicants clicked the invitation link to participate in the 2020 survey. The link was sent by email. The translation to English is the following: *Maria, you have been invited to participate in the School Admission System Satisfaction Survey, a joint effort between Mineduc and Princeton University researchers. Your answers will help to improve the application process and the information that we will provide new applicants. Note that: (1) Your answers will not affect in any way your results in the Admission Process. (2) Participation is entirely voluntary; you can stop it at any time. (3) All your answers are confidential. (4) Only personnel authorized by Mineduc will have access. I have read the information about the Survey. I give my consent to participate. [Options: Yes or No]*

1. *(List of schools, a reminder of the filed application)*

2. First, we want to know how you evaluate the process of the School Admission System. Choose a grade from 1 to 7 for the following aspects
   [*Slider 1 to 20*]

   (a) Information on schools available (academic performance, collections, educational project, after school activities)

(b) Availability of information on the application process (relevant dates, website, etc).

(c) In general, what rating would you put to the application process?

3. How did you get information about of the application process? Select all that apply
   [*Select multiple*]

   (a) Through the Municipality

   (b) Through the current school/pre-school

   (c) Through the newspaper or radio

   (d) Through social networks (Facebook, Instagram, Twitter, Youtube)

   (e) Through friends or relatives

   (f) Through the website of the Ministry of Education (www.sistemadeadmisionescolar.cl)

   (g) Through the platform of the Ministry of Education Your Information

   (h) I did not inform myself

4. Select the social networks you used to get information about SAE?
   [*Select multiple*]

   (a) Facebook

   (b) Twitter

   (c) Instagram

   (d) Youtube

5. Select the traditional media outlets you used to get information about SAE?
   [*Select multiple*]

   (a) Newspaper

   (b) Radio

   (c) TV

6. When you add a school to your application, what do you consider a necessary steps to know well a school before applying?(Check all that apply).
   [*Select multiple*]

   (a) Knowing the infrastructure

   (b) Interview with the principal or a teacher

   (c) Visit the website of the school

   (d) Get referrals from someone you know

   (e) Academic Performance information

   (f) Knowing indicators from the Agency for Quality Education

   (g) Knowing the extracurricular activities offered

   (h) Know your project Educational Institutional (PIE)

7. Any other relevant step that we have not included here?
   [*Open text*]

8. How well do you know the schools in your application ?
   [*Knowledge scale: (I didn't know it, Only by name, I know it well)*]

   (a) [Name preference 1]
   (b) [Name preference 2]
   (c) [Name preference 3]
   (d) [Name preference 4]
   (e) [Name preference 5]

9. Because COVID-19, much of classroom activities have been suspended.Do you think this affected your application process in any of these dimensions?
   [*Select one*]

   (a) COVID-19 did not affect my application process
   (b) Without COVID-19, I would have known better the schools that I already know, but I would not have applied to more schools
   (c) Without COVID-19, I would have known more schools and perhaps I would have added them to my application

10. We note that during the application process you added schools to your initial list.¿Did you know these schools before the start of the application process?
    [*Knowledge scale (I didn't know it before applying, I knew it by name before applying, I knew it well before applying)*]

    (a) [Name preference added 1]
    (b) [Name preference added 2]
    (c) [Name preference added 3]

11. In order to convince yourself to add these schools:
    [*Select one*]

    (a) It was necessary to find out more about them
    (b) It was not necessary to search for more information

12. You applied to [Name preference 1] in first preference:From 0 to 100, how likely or how sure are you that you will get a seat on that option?
    [*Slider 0 to 100*]

13. Imagine if you would had put your second choice [Name preference 2] as your first choice:From 0 to 100, how likely or how sure are you that you would get a seat on that option?
    [*Slider 0 to 100*]

14. Imagine if you had put your third choice [Name preference 3] as your first choice:From 0 to 100, how likely or how sure are you that you would get a seat on that option?
    [*Slider 0 to 100*]

15. Some families are not placed in any option beacuse there is no sufficient seats.Using the same range of 0 to 100,How likely or how sure are you that [Applicant name] will be placed in one of the [Length application] schools in the application?
    [*Slider 0 to 100*]

16. Why you did not add more schools to your application?
    [*Select one*]

    (a) I know the other options well and I prefer to have no placement than to add those alternatives
    (b) I think I will definitely be placed in one of the schools I applied for
    (c) It is very difficult to find more schools
    (d) There are no more schools close enough (good or bad)

17. If you would had added more schools to your application. Do you think you would have higher changes to be placed to one school?
    [*Select one*]

    (a) No
    (b) Yes

18. Here are five schools. How well do you think you know these schools?
    [*Knowledge scale: (I didn't know it, Only by name, I know it well)*]

    (a) [College not considered in application 1]
    (b) [College not considered in application 2]
    (c) [College not considered in application 3]
    (d) [College not considered in application 4]
    (e) [College not considered in application 5]

19. From 1 to 10,How easy it is to find information on the academic performance of schools?
    [*Slider 1 to 10*]

20. Imagine that you spend time researching all schools that you do not know well.After you know them well, do you think you would add at least one of these schools to your application?
    [*Select one*]

    (a) No
    (b) Yes

21. From 0 to 100, how likely would you add it as your first preference?
    [*Slider 0 to 100*]

22. From 0 to 100, how likely would you add it below your last choice?
    [*Slider 0 to 100*]

23. During the application process, did you get any recommendations about adding more schools to your list?
    [*Select one*]

    (a) No

    (b) Yes

24. By what method did you receive the recommendation to add more schools?(Select all that apply)
    [*Select multiple*]

    (a) SMS

    (b) WhatsApp

    (c) E-mail

    (d) Web page

    (e) Other

25. By what method did you receive the recommendation to add more schools?- Other
    [*Open text*]

26. If [applicant name] get a seat in the following schools, from 1 to 7, how satisfied would you be?
    [*Slider 1 to 7*]

    (a) First preference: [Name preference 1]

    (b) Last Preference: [Name Last preference]

    (c) If you are not in any school in the regular period

27. Would you like to have had the following information on schools that did not have at the time of application?
    [*Yes or No*]

    (a) Information about your chances of being accepted

    (b) Standarized test score

    (c) Performance category

    (d) Price

    (e) Priority for economically-vulnerable students

    (f) SAT scores

    (g) Seats available

28. What is your preferred method of contact during the application process?
    [*Select one*]

    (a) E-mail

    (b) Other

    (c) SMS

    (d) Telephone

    (e) Whatsapp

29. What is your preferred method of contact during the application process? - Other
[*Open text*]

30. For registration purposes only, what is the highest educational level of the Mother (or Stepmother) of [applicant name]?
[*Select one*]

    (a) Educación Básica Completa

    (b) Educación Básica Incompleta

    (c) Educación Media Completa

    (d) Educación Media Incompleta

    (e) Educación incompleta en una Universidad

    (f) Grado de magíster universitario

    (g) No estudió

    (h) Titulada de un Centro de Formación Técnica o Instituto Profesional

    (i) Titulada de una Universidad

31. Do you know if [Field-nomPostulante] is a priority student (SEP)?
[*Select one*]

    (a) He/she is not a beneficiary of the preferential subsidy

    (b) I do not know

    (c) He/she is a beneficiary of the preferential subsidy

32. Do you have any other comments, complaints or suggestions to make us?
[*Open text*]

H   New Haven warnings policies and data

H.1 School choice institutions

New Haven, Connecticut uses a centralized mechanism to assign students to schools in all grades. As reported in Akbarpour et al. (2020), New Haven used a Boston- or Boston-like mechanism to assign students to schools from the the 1990s through 2018. In 2019, New Haven switched to a truncated deferred acceptance mechanism, which allowed applicants to list a maximum of four schools. In 2020, the district raised the maximum application length to six. The school choice process in New Haven takes place in the winter and early spring of each year, for enrollment the following fall. Following a series of informational events in January, the choice application opens in early February, with final applications due in early March. Applicants receive placement outcomes in early April.

Students from outside New Haven may attend NHPS schools through the district's interdistrict choice program. Choice "markets" in the the NHPS school choice system are defined by grade and by residency. Counts of available seats are determined separately for each school-grade combination, and then are further split by whether the applicant is from New Haven or a nearby town.

The first column of Table H1 reports descriptive statistics for all students applying through the choice system in 2020. 58% of applicants apply to a major transition grade– either prekindergarten, kindergarten, or grade 9. We focus on these transition grades in our analysis risk warnings. This is because we rely on prior-year risk predictions, and these are more stable in larger markets.

H.2 Warnings intervention

NHPS policymakers conducted two information intervention policies as part of the 2020 choice process for PreK, Kindergarten, and high school grades. The first was a warnings intervention applied to all risky applicants. In this intervention, students submitting applications with an estimated nonplacement risk of 50% or higher received an email warning one week before the application deadline. The email suggested that the applicant might want to add more schools to their rank list. The email also provided a link to an online risk simulator tool, where applicants could input hypothetical choice applications and learn about the chances of placement for those applications. Panel A of Figure H1 displays an example of the email sent to risky applicants.

The second intervention consisted of an email sent to a randomly chosen group of non-risky applicants. This email was identical to the warnings treatment email, but did not contain the line about high nonplacement risk. In the main text, we refer to this email as the "static" intervention, because it does not provide personalized feedback.

Panel B of Figure H1 displays the email sent as part of this second intervention. Randomization was stratified by market (grade by residency status). All school choice applicants could view the application simulator using their NHPS username and login, once they arrived at the simulator page. As we show below, simulator use by untreated individuals was rare. This makes sense because control group applicants did not receive information about the simulator's web address.

The application simulator website used in both interventions worked as follows. Applicants where first asked to state their beliefs about the admission chance for each of their choices, which were pre-loaded (see Figure H2a). Afterwards, their predicted admission chances were

31

displayed to them as shown in Figure H2b. Users then had the opportunity to add, remove or change schools and received immediate feedback on their changed portfolio risk. The schools available to them were shown both on a map and an alphabetical list (see figure H2c and H2d).

In contrast to the risk predictions we constructed for the Chilean choice intervention, the predictions in the New Haven setting relied only on prior-year applications. Specifically, NHPS computed portfolio risk predictions based on the admission chances that the same application would have had in the previous application year.[6]

Figure H3 describes the distribution of predicted placement probabilities for different values of realized placement probabilities. As with our predictions in the Chilean setting, risk predictions do not perfectly match ex post values, but do closely track them. One point of contrast with the Chilean setting is that our New Haven placement probability predictions tend to somewhat overstate true placement chances— lottery odds became somewhat worse for applicants in 2020, relative to submitting the same application in 2019. In practice, this meant that the risk warnings went to a *riskier* set of applicants than would have been the case had placement chances remained steady.

Two new schools entered the NHPS system in 2020. NHPS did not compute risk predictions for applications including these schools, and excluded applicants from the information intervention. Stepping down from the full sample to the sample of intervention students reduced the sample size as follows. 58% of all choice applicants applying in intervention-eligible grades. 46% submit applications by seven days before the application deadline and are included in the intervention procedure. 36% (of the full sample) applied to simulator-eligible grades, did so in time to be included in the intervention, and applied only to schools included in the simulator. These 36% of applicants formed the universe of applicants potentially subject to the warnings and simulator interventions.

Column 2 of Table H1 describes the 36% of students in the intervention-eligible universe. These students are less likely to be African American and more likely to be Hispanic than the full sample.[7] Columns 3 and 4 of Table H1 split the sample by treatment assignment of either the warnings intervention or the simulation intervention, which corresponds to predicted risk levels above and below 50%. The mean risk score in the former group is 89%, in the latter group it is 5.4%. 65% of all applicants in the intervention-eligible sample received an email. The remaining 35% where either assigned to the control group (33%) which did not receive any emails or the email could not be delivered (3%). 98.2% of students in the warnings group and 96% of students in the simulator intervention group successfully received an email

---

[6]To make sure applicants understood this, the text of the intervention stated that the warning was based on past data. See Figure H1.

[7]A few additional features of the data are worth noting. Nine students who have been assigned treatment later change the grade they are applying to or delete their profile altogether. These would be counted in the column 2 sample but are not considered to have an eligible grade. Thus the share of applicants to an eligible grade is not exaclty one. In addition, 14 students apply on February 26 after the application portfolio snapshot was taken at 7.00pm but before the last wave of treatment assignments is made. We classify these individuals as not having applied in time. This is why the share of applicants to an eligible grade that apply in time is smaller than one. As discussed in Kapor et al. (2020), New Haven residents applying to ninth grade have undersubscribed neighborhood high schools as their default placements. When computing application risk, the district defined these outcomes as placements, meaning that no in-district ninth-grade applicants were classified as risky for the purposes of the warnings intervention. The ninth grade applicants who did receive the warning were those applying through interdistrict choice programs. This is why the share of ninth graders in the warnings sample is smaller than the share in the eligible sample.

corresponding to their treatment group.

We also construct a comparison sample of choice applicants in 2019, consisting of all choice applicants applying in the major choice grades in that year. We construct this comparison sample to resemble as closely as possible the set of students who would have been included in a 2019 warnings intervention, had one taken place. We do this by considering only students in eligible grades who had submitted their application at least seven days in advance of the application deadline. We compute risk predictions for this group using the 2019 application data based on the state of their application as of seven days before the deadline. The students in this sample form the basis for the 2019 comparison group plotted in Figure 13. Column 5 of Table H1 displays descriptive statistics for this comparison sample. This group closely resembles the 2020 eligible sample on demographic characteristics and choice outcomes.

H.3 Simulator intervention results for non-risky applicants

Figure 13 and Table 7 display our main results for the RD and DD analysis of the warnings intervention in 2020. We discuss these findings in section 5.7 of the main text. Figure H4 displays RD-DD graphs for balance and first stage outcomes, parallelling the main results in Figure 13.

Table H2 reports our findings from the static RCT delivered to randomly chosen non-risky students. Panel I shows that standard balance tests pass. Panel II shows that treatment and control groups are balanced on the initial risk prediction ("risk score") as well as on the ex post risk associated with their initial application ("initial realized risk"). It also shows that there is no difference in *final* realized risk; i.e. the ex post risk of the final submitted application. The implication is that simulator access did not affect application risk.

Panel III shows that many applicants who receive the treatment email granting simulator access do in fact click the link and interact with the simulator. Treatment rasies the likelihood of simulator login by 23 percentage points and the share of applicants conducting simulator runs by 11 percentage points. In practice, the requirement that applicants state their beliefs about admissions chances at different schools prior to each use of the simulator may have placed a substantial burden on prospective simulation users. NHPS eliminated this requirement from subsequent implementations. Panel IV shows how treatment changes choice behavior. As with the headline risk values reported in panel III, we see no evidence of effects here.

Figure H1: Email communication with parents
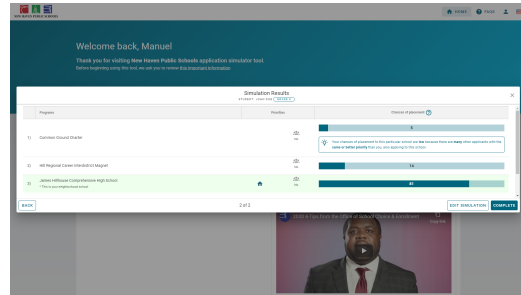

(a) Risky applicants

(b) Non-risky applicants

Figure H2: New Haven Simulator
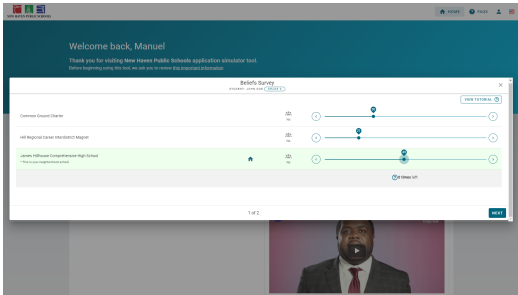
(a) Beliefs survey page

(b) Predicted admission chances
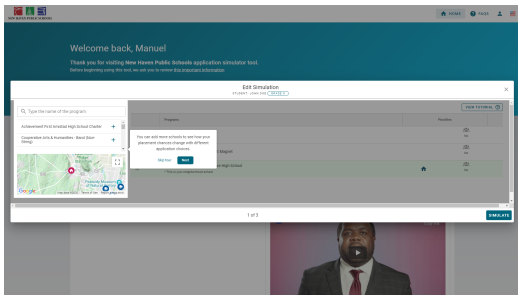
(c) Simulating portfolio changes
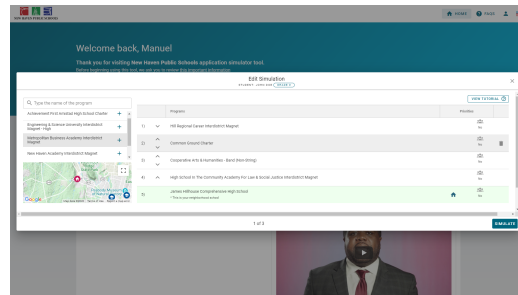
(d) New predicted admission chances

Figure H3: Observed vs. predicted placement probability in New Haven

Distribution of predicted placement probability by value of ex post observed placement probability. For each bin of observed placement, we display the mean and IQR of predicted values. 45-degree line displayed for reference. See section H for details.

Figure H4: Treatment balance and First Stage outcomes

(a) Female

(b) Black

(c) Hispanic

(d) Received Email

(e) Received Warnings Email
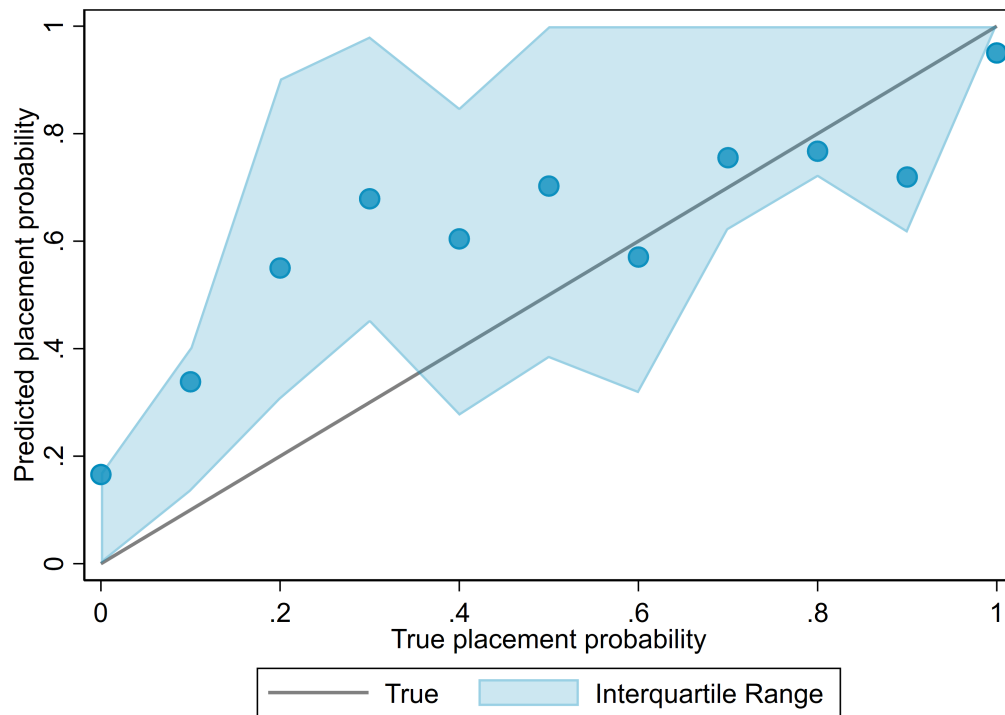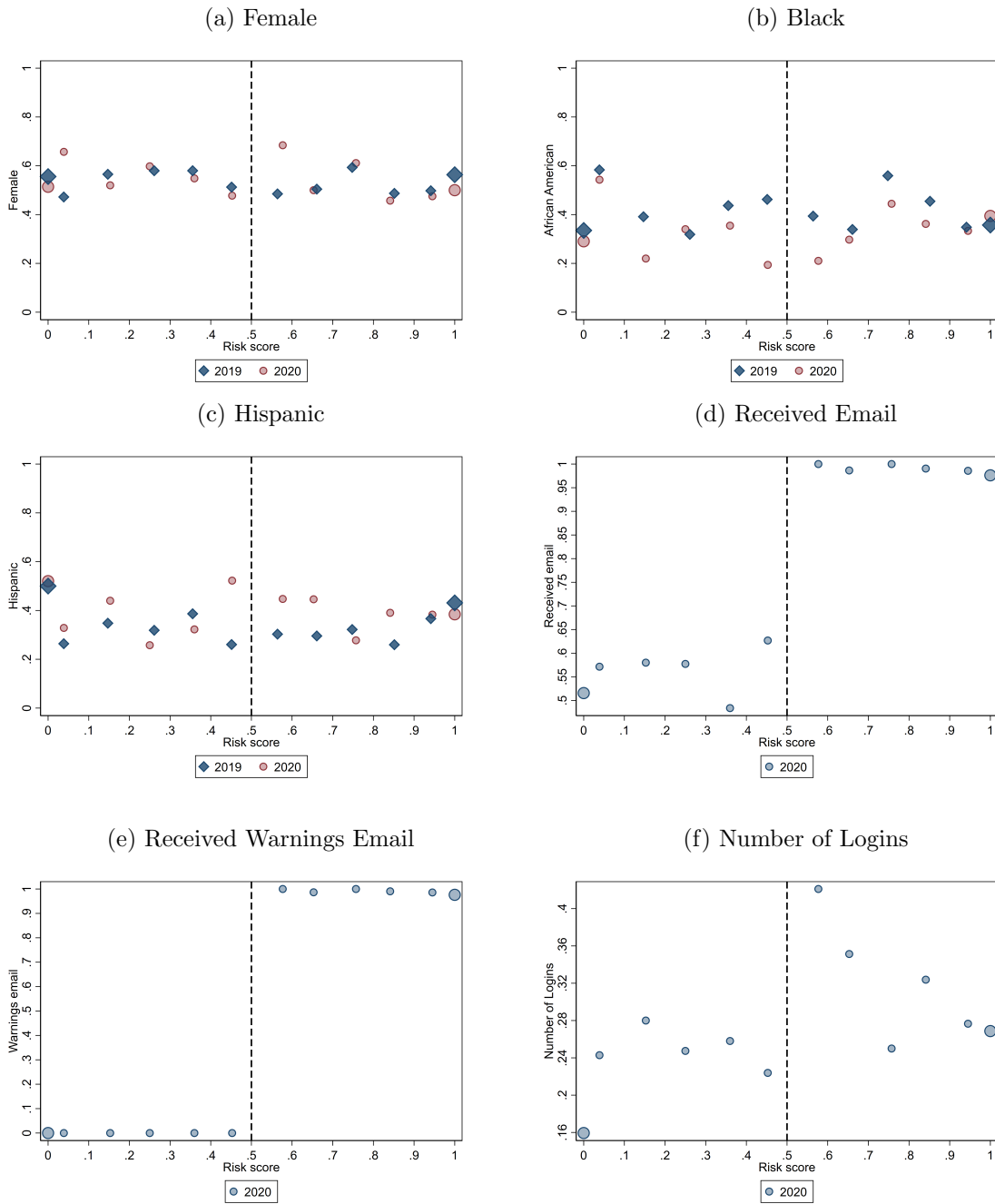

(f) Number of Logins

Balance and first stage effects for warnings intervention in New Haven, CT centralized choice. Figures show predetermined covariates and treatment receipt 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. Panels A-C display values for both 2020 application cohort and 2019 comparison group. Panels D-F display treatment receipt for 2020 cohort only; no warnings treatment or simulator intervention took place in 2019. See section H for details.

37

Table H1: Sample descriptives New Haven

	2020				2019
	All grades	Eligible	Warnings	Simulator	Comparison group
I. Demographics					
Female	0.513	0.539	0.510	0.530	0.547
African American	0.432	0.338	0.366	0.330	0.380
Hispanic	0.395	0.468	0.389	0.489	0.432
White	0.125	0.147	0.203	0.133	0.139
NH Resident	0.725	0.674	0.429	0.759	0.719
II. Simulator Eligiblity					
PreK3	0.075	0.105	0.268	0.046	0.129
PreK4	0.105	0.147	0.294	0.094	0.190
K	0.163	0.288	0.229	0.314	0.291
Grade 9	0.242	0.460	0.208	0.547	0.390
Apply to eligible grade	0.577	0.996	0.992	0.999	1.000
+ in time	0.458	0.991	0.988	0.994	1.000
+ only to simulator schools	0.363	1.000	1.000	1.000	1.000
III. Interactions with simulator					
Risk score	0.294	0.294	0.889	0.054	0.339
Warnings email	0.105	0.285	0.982	0.000	0.000
Received email	0.238	0.649	0.982	0.960	0.000
IV. Placements					
Placed 1st	0.259	0.320	0.132	0.399	0.304
Placed other	0.337	0.369	0.097	0.477	0.375
Unplaced	0.403	0.311	0.771	0.124	0.321
V. Choice outcomes					
Change length or school		0.056	0.100	0.033	0.030
Lengthen app.		0.042	0.092	0.020	0.019
Insert new school		0.018	0.025	0.014	0.011
Append new school		0.024	0.065	0.008	0.010
Change school		0.029	0.031	0.023	0.022
Shorten app.		0.011	0.007	0.013	0.006
Difference in realized risk		-0.006	-0.015	-0.002	-0.003
Difference in simulated risk		-0.006	-0.019	-0.000	-0.003
N	7027	2551	740	967	3150

Samples vary by column. The first column consists of all applicants in 2020. The second column consists of the sample that was eligible for treatment in 2020. Columns 3 and 4 represent the subsamples that have been assigned to either treatment group. The fifth column consists of those applicants in 2019 that would have been eligible for treatment had their been any. Statistics reported represent shares of applicants in the respective sample lr the mean difference in the last two rows of panel V.

Table H2: Treatment Balance RCT

	Control	Treatment	Difference	
	Mean	Mean	β	SE
I. Demographics				
Female	0.574	0.530	-0.043	(0.024)
African American	0.323	0.330	0.010	(0.022)
Hispanic	0.512	0.489	-0.026	(0.023)
White	0.114	0.133	0.018	(0.015)
Repeat grade	0.063	0.057	-0.003	(0.011)
II. Risk				
Risk score	0.046	0.054	0.002	(0.005)
Initial realized risk	0.124	0.136	-0.002	(0.009)
Final realized risk	0.123	0.134	-0.003	(0.009)
III. Interaction with simulator				
Received email	0.000	0.960	0.960	(0.006)
Warnings email	0.000	0.000	0.000	(0.000)
Pr(Any login)	0.018	0.224	0.198	(0.014)
Number of Logins	0.023	0.260	0.227	(0.017)
Pr(Any sim. run)	0.012	0.126	0.109	(0.011)
IV. Choice outcomes				
Change length or school	0.043	0.033	-0.005	(0.009)
Lengthen app.	0.022	0.020	0.002	(0.007)
Insert new school	0.016	0.014	0.002	(0.006)
Append new school	0.006	0.008	0.001	(0.004)
Change school	0.035	0.023	-0.007	(0.008)
Shorten app.	0.012	0.013	0.001	(0.005)
Difference in realized risk	-0.001	-0.002	-0.002	(0.001)
Difference in simulated risk	-0.000	-0.000	-0.001	(0.001)
Any placement	0.878	0.876	0.016	(0.012)
N	844	967	1811	1811

Statistics in this table are estimated from the sample of individuals in the control group and the Simulator-only (no warnings) treatment group. The column panels distinguish between these two subsamples. The reported coefficients β reflect regression estimates of the treatment indicator on outcomes, controlling for fixed effects of randomization time blocks and markets, i.e. resident status by grade.

Author disclosure statements

Felipe Arteaga: I am a research affiliate at ConsiliumBots, the NGO that implemented the choice platforms studied in the paper. I also worked in the implementation team of the centralized system (Ministry of Education of Chile) between 2015 and 2018.

Christopher Neilson: I am the founder and interim CEO of ConsiliumBots, the NGO that currently implements the feedback provision tools on the choice platforms studied in the paper. I have not received financial compensation from ConsiliumBots or any other source related this project.

Adam Kapor: I am an academic advisor to ConsiliumBots, the NGO that implemented the choice platforms studied in the paper. I have not received financial compensation from ConsiliumBots or any other source related this project.

Seth Zimmerman: I am academic advisor to ConsiliumBots, the NGO that implemented the choice platforms studied in the paper. I have not received financial compensation from ConsiliumBots or any other source related this project.