Online Appendix to “Smart Matching Platforms and Heterogeneous Beliefs in Centralized School Choice”

Felipe Arteaga

Adam J. Kapor

Christopher A. Neilson

Seth D. Zimmerman

January 13, 2022

Contents

A. Model Appendix

A.1. Additional Proofs

Note on Equation 3. Rewrite $V(\mathcal{C}) - V(\mathcal{C}_0)$ as a function of true beliefs and belief errors, so that

$$V(\mathcal{C}) - V(\mathcal{C}_0) = \Psi_r \times (1-a)^{r-1} \times (1 - R_s^*(1-a)) \times (u_s - \Gamma_r)$$

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

Proof of proposition 1. By assumption, s is the last-ranked school in $\mathcal{C} = \mathcal{C}_0 \cup \{s\}$. This implies that $\Gamma_r = 0$ and that $\frac{d\Gamma_r}{da} = 0$, so we may ignore the third term of the sum in equation 3. Rewriting equation 3 without this term and setting it to be negative (because we are looking for conditions under which the value of adding s is decreasing in a), we have

$$\frac{d \log(V(\mathcal{C}) - V(\mathcal{C}_0))}{da} = \frac{1-r}{1-a} + \frac{R_s^*}{1 - R_s^*(1-a)} < 0.$$

Rearranging then yields

$$a > 1 - \frac{r-1}{rR_s^*}$$

Note also a) that $r = N_0 + 1 \geq 2$ (by assumption) and b) that $rR_s^* > 0$ (since both r and R_s^* are positive numbers). $\frac{r-1}{rR_s^*} > 0$ then follows immediately.

Proof of proposition 2. Let $\mathcal{C}$ denote a consideration set that an applicant obtains following an optimal search strategy under optimism level a. Additional search takes place when $\kappa < U[\text{Search}|\mathcal{C}_0, a - \Delta_a]$. By optimality we have $\kappa > U[\text{Search}|\mathcal{C}_0, a]$. Hence the probability of additional search is equal to

$$Pr(\text{Search}; \mathcal{C}, a, \Delta_a) = 1\left(U[\text{Search}|\mathcal{C}_0, a - \Delta_a] > U[\text{Search}|\mathcal{C}_0, a]\right)\left[\Phi(U[\text{Search}|\mathcal{C}_0, a - \Delta_a]) - \Phi(U[\text{Search}|\mathcal{C}_0, a])\right] \geq 0.$$

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 the probability of adding at least one school is equal to the probability of search.

Finally, note that $U[\text{Search}|\mathcal{C}_0, a]$ is decreasing in $V(\mathcal{C}_0)$. 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. $U[\text{Search}|\mathcal{C}_0, a] \leq \kappa$ is therefore a sufficient condition for terminating search.

A.2. Enrollment Choices and the Welfare Effects of Information.

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

We model enrollment as a binary choice between the school where an individual is placed and the outside option. Timing is as follows. At the time of application, individuals observe school- (and person-) specific utilities μ_j, with the outside 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 to the rule

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

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

Let $q_j = Pr(\text{Enroll}|\text{place at } j)$ denote the probability of enrollment conditional on placement at j. An important observation is that, for each possible q_j, there is a unique utility level, $u^*(q_j)$, associated with enrollment probability q_j, and this utility level is increasing in q_j.[1]

Our approach is to consider a surprise change in optimism from a to a' after an applicant has engaged in optimal search. Our framework lets us break down the effect of a change in information on payoffs into two channels: a placement channel, and a utility conditional on placement

[1]We have $q_j = 1 - G(-\mu_j)$ and $\mu_j = -G^{-1}(1 - q_j)$. Define $u^*(q_j) = E[\max(-G^{-1}(1 - q_j) + \epsilon_j, 0)]$. By construction, when utility is equal to $u^*(q_j)$, the associated enrollment probability is q_j. Note that $\frac{dE[\max(\mu_j+\epsilon_j,0)]}{d\mu_j} = Pr(\text{Enroll}|\text{place at j}) = q_j > 0$. We have

$$\frac{du^*(q_j)}{dq_j} = \frac{dE[\max(\mu_j + \epsilon_j, 0)]}{dPr(\text{Enroll}|\text{place at } j)} = \frac{Pr(\text{Enroll}|\text{place at } j)}{g(-\mu_j)} > 0.$$

channel.

Before stating the result, we provide some definitions. Let $U^*(a', \mathcal{C}) = E(V(\mathcal{C}'); \mathcal{C}, a')$ denote the expected payoff an agent receives from his application after having followed an optimal search strategy, given an optimism level a', when endowed with consideration set $\mathcal{C}$, where the expectation is over the resulting consideration sets $\mathcal{C}'$ that may be obtained by further search.

Let $\mathcal{C}_0$ denote the initial consideration set that an applicant is endowed with. Let $\mathcal{C}_1$ be a consideration set that an applicant who is endowed with consideration set $\mathcal{C}_0$ reaches via an optimal search strategy under optimism a. By construction, an applicant with optimism a and consideration set $\mathcal{C}_1$ will not engage in further search. Hence we have $U^*(a, \mathcal{C}_1) = V(\mathcal{C}_1)$. However, when $a' \neq a$ the applicant may engage in additional search beyond the schools in $\mathcal{C}_1$.

Let $Pr(j|a', \mathcal{C})$ denote the probability that the person matches to school j when endowed with consideration set $\mathcal{C}$, optimism a', and the option to conduct further search if desired. Let $Pr(\text{place}|a', \mathcal{C}) = \sum_{j \in \mathcal{J}} Pr(j|a', \mathcal{C})$ denote the probability of any placement.

Proposition 3. Let $\mathcal{C}_1$ be a consideration set obtained by an optimal search strategy under optimism a. The individual utility gain from a change in optimism, $U^*(a', \mathcal{C}_1) - U^*(a, \mathcal{C}_1)$ where $a', a \in (0, 1)$, satisfies

$$U^*(a', \mathcal{C}_1) - U^*(a, \mathcal{C}_1) = \left[Pr(\text{place}|a', \mathcal{C}_1) - Pr(\text{place}|a, \mathcal{C}_1)\right] u^*(\overline{q}) + \sum_{j \in \mathcal{J}} \left[Pr(j|a', \mathcal{C}_1) - Pr(j|a, \mathcal{C}_1)\right] (u_j - u^*(\overline{q})),$$

where

$$\overline{q} = E(q_j|\text{place}, \mathcal{C}_1)$$

is the probability of enrolling in the inside option, conditional on receiving any placement, under consideration set $\mathcal{C}_1$. Moreover, for each $j \in \mathcal{J}$, the term $u_j - u^*(\overline{q})$ is nonnegative whenever $q_j \geq \overline{q}$.

Remark. This proposition shows that individual utility increases in proportion to the placement rate, except to the extent it is offset by declines in utility conditional on placement, where utility is an increasing function of enrollment probabilities.

This is particularly clear in our modal empirical case, which is when our intervention reduces optimism from a to $a' < a$ and the applicant adds a school $s \in \mathcal{J} \setminus \mathcal{C}_1$ to the bottom of his list. In this case, we would have $p_j(a, \mathcal{C}_1) = p_j(a', \mathcal{C}_1)$ for all $j \in \mathcal{C}_1$, and $p_s(a, \mathcal{C}_1) \leq p_s(a', \mathcal{C}_1)$ for $s \in \mathcal{J} \setminus \mathcal{C}_1$. When this happens, individual welfare increases at least proportionally to the placement rate as long as the probability of accepting a new placement satisfies $q_s \geq \overline{q}$.

We use this observation to guide our assessment of welfare effects.

Proof of proposition 3. We can write

$$\begin{aligned} U^*(a', \mathcal{C}_1) &= \sum_{j \in \mathcal{J}} Pr(j|a', \mathcal{C}_1) u^*(q_j) \\ &= \sum_{j \in \mathcal{J}} Pr(j|a', \mathcal{C}_1) u(\overline{q}) + \sum_{j \in \mathcal{J}} Pr(j|a', \mathcal{C}_1)(u_j - u^*(q)) \\ &= Pr(\text{place}|a', \mathcal{C}_1) u(\overline{q}) + \sum_{j \in \mathcal{J}} Pr(j|a', \mathcal{C}_1)(u_j - u^*(q)). \end{aligned}$$

The result follows. We have $u_j - u^*(\overline{q}) > 0$ if and only if $q_j > \overline{q}$ by monotonicity of $u^*(\cdot)$.

A.3. Additional Discussion

This subsection considers how violations of assumptions we impose in our theoretical model might mediate the effects of interventions that reduce optimism about application risk.

First, consider the possibility that applicants do not perfectly observe the utilities of schools they consider, but take schools' relative popularity as a signal of their utility. If so, informing an excessively optimistic applicant that the schools in his portfolio are more popular than he had thought may lead him to conclude that these schools are better, and schools outside his portfolio are worse, than he had previously believed, attenuating any increase in his incentives to search, and reducing the odds of placing a newly discovered school ahead of schools that the applicant already knows.

Second, our model makes the simplifying assumption that an applicant can surely discover a new school at a constant cost. Our assumption allows for uncertainty about the amount of effort needed to discover a school. For example, students may pay a flow cost k in order to discover an additional school with instantaneous probability λ. In this case, κ would denote the expected cost of searching until an additional acceptable school is found.

If search is uncertain, however, and in addition search costs are increasing or the chance of discovering a school is decreasing in the amount of effort that has already been exerted, then new information about placement chances may cause an applicant to engage in additional *unsuccessful* search before giving up. Thus this channel may also reduce the extent to which providing information to sufficiently optimistic applicants causes them to discover new schools.

B. ADDITIONAL FIGURES AND TABLES

Table B.I
More Descriptive Statistics for Chilean Choice Applicants

	(1) All	(2) Economically Vulnerable	(3) Not Economically Vulnerable	(4) Pop-up eligible	(5) Risky (predicted risk>.3)	(6) Around Pop-up Cutoff	(7) RCT. sample (2020)	(8) Survey sample (2020)
N	1,168,706	575,521	593,185	848,795	233,678	84,517	19,213	48,929
%	1.00	0.49	0.51	0.73	0.20	0.07	0.02	0.04
A. Application behavior								
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
Delete all	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
B. Placement								
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
C. School capacity available after placement								
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
D. Attributes of enrolled school								
Value added	0.08	0.03	0.12	0.08	0.12	0.15	0.17	0.11
Total enrollment per grade	88.79	89.52	88.06	87.50	83.29	101.42	70.10	92.62
E. Classification by true risk of initial attempt								
.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). Panels are as in Table I. All statistics are means in the population defined by the column header. "Pop-up eligible" (col. 4) are students who submitted applications that received a risk prediction. "Risky" (col. 5) is applicants whose first attempt had a predicted risks > 0.3. "Around pop-up cutoff" (col. 6) are applicants whose first attempt had a predicted risk in $[0.1,0.5]$. "RCT sample" (column 7) is applicants in treatment or control group of the 2020 RCT design. "Survey sample" (column 8) is applicants who completed the 2020 school choice survey. Selected row variable definitions are as follows. "Economically vulnerable" is an SES measure computed by Mineduc. "Rural" is an indicator if students live in rural areas. "Length of initial/final attempt" is the number of schools on an applicants first and final choice application. "Total attempts" is the number of times an applicant submitted an application to the centralized system. Application change and addition variables describe the share of applicants making different kinds of changes 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. Value added and school characteristic variables described in Online Appendix D. VA is calculated only for grades 8 and below. True risk of initial attempt variables describe the nonplacement risk of an applicant's initial application, evaluated using ex post observed applications. Panel F variables (School capacity available after placement) are calculated at a local market level defined for each student.

Table B.II
Descriptive Statistics for Chilean Choice Applicants- Alternate Samples

	(1) All	(2) Around Pop-up Cutoff	(3) RCT. sample (2020)	(4) Survey sample (2020)
N	1,168,706	84,517	19,213	48,929
%	1.00	0.07	0.02	0.04
A. Demographics				
Economically Vulnerable	0.49	0.42	0.25	0.42
Rural	0.05	0.02	0.00	0.04
B. Application behavior				
Length initial attempt	2.77	3.04	2.79	2.74
Length final attempt	3.14	3.57	3.32	3.22
Total attempts	1.41	1.51	1.53	1.45
Any modification	0.25	0.33	0.33	0.28
Add any	0.21	0.30	0.30	0.25
C. Placement				
Placed in pref.	0.79	0.77	0.42	0.79
Placed 1st	0.54	0.39	0.17	0.53
Particip. in 2nd round	0.09	0.12	0.20	0.09
Placed in 2nd round	0.07	0.09	0.16	0.07
D. School capacity available after placement				
Share of total seats	0.42	0.39	0.44	0.50
Share of seats in free schools	0.46	0.44	0.52	0.55
E. Attributes of enrolled school				
Enrolled at some school	0.97	0.96	0.91	0.97
Enrolled at placed	0.62	0.57	0.32	0.66
Have value added measure$\mid$grade$\leq$8	0.77	0.76	0.76	0.79
Value added$\mid$enrolled at placed	0.11	0.20	0.25	0.12
Value added$\mid$not enrolled at placed	0.04	0.06	0.11	0.09
School monthly fee (USD)	17.02	23.50	30.66	20.53
Share of vulnerable students	0.61	0.56	0.52	0.57
F. Classification by true risk of initial attempt				
Mean risk	0.24	0.24	0.61	0.25
Zero risk	0.59	0.19	0.02	0.59
Risky (risk>.3)	0.30	0.37	0.84	0.31

Notes. N: 1,168,706 (20% from 2018, 41% from 2019 and 39% from 2020). All statistics are means in the population defined by the column header. "Around pop-up cutoff" (col. 2) are applicants whose first attempt had a predicted risk in [0.1,0.5]. "RCT sample" (column 3) is applicants in treatment or control group of the 2020 RCT design. "Survey sample" (column 4) 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. Value added and school characteristic variables described in Online Appendix D. VA is calculated only for grades 8 and below. True risk of initial attempt variables describe the nonplacement risk of an applicant's initial application, evaluated using ex post observed applications. Panel F variables (School capacity available after placement) are calculated at a local market level defined for each student.

Table B.III
Platform Pop-Up RD Estimates of Main Outcomes (Table II) with Alternate Bandwidths

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Bandwidth	Full	+-0.1	rdbwselect				
	Estimate	Estimate	Estimate	BW left	BW right	N left	N right
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 undersubscribed	0.074	0.073	0.076	0.15	0.15	31,104	31,361
	(0.005)	(0.007)	(0.005)				
Δ prob. placed to undersubscribed	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, Cattaneo and Titiunik (2014).

Table B.IV
RD Estimates of Platform Pop-Up Effects on Adding Any School, by City and Year

City	2020 applicants	2018	2019	2020
Santiago	158,057		0.24	0.25
			(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 pop-up effects on adding at least one school to the choice application, split by city and year. Cities are sorted by count of 2020 applicants. Santiago is not displayed for 2018 because centralized admission had not yet been rolled out. Estimates from local linear specifications, computed using triangular kernel with bandwidth 0.1. Heteroskedasticity-robust nearest neighbor variance estimator with minimum of 3 neighbors reported in parentheses; computed as in Calonico, Cattaneo and Titiunik (2014). See section V.F for details.

Table B.V
RD Estimates of Platform Pop-Up Effects by Market-Level Choice Experience

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

Notes. Local linear RD estimates of pop-up effects from warning pop-up on application platform, split by years elapsed since city-grade combination first began using the centralized choice process. Computed using triangular kernel with bandwidth 0.1. Heteroskedasticity-robust nearest neighbor variance estimator with minimum of 3 neighbors reported in parentheses; computed as in Calonico, Cattaneo and Titiunik (2014). IV estimates reported in second column of each set show the instrumental variable specifications (fuzzy RD), where the endogenous regressor is the add any school indicator. Panel A: predetermined covariates. Panel B: measures of choice behavior from initial to final application. 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: 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. "Undersubscribed" schools are those with excess capacity.

Table B.VI
RD Estimates of Platform Pop-Up Effects by Applicant's Socioeconomic Status

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

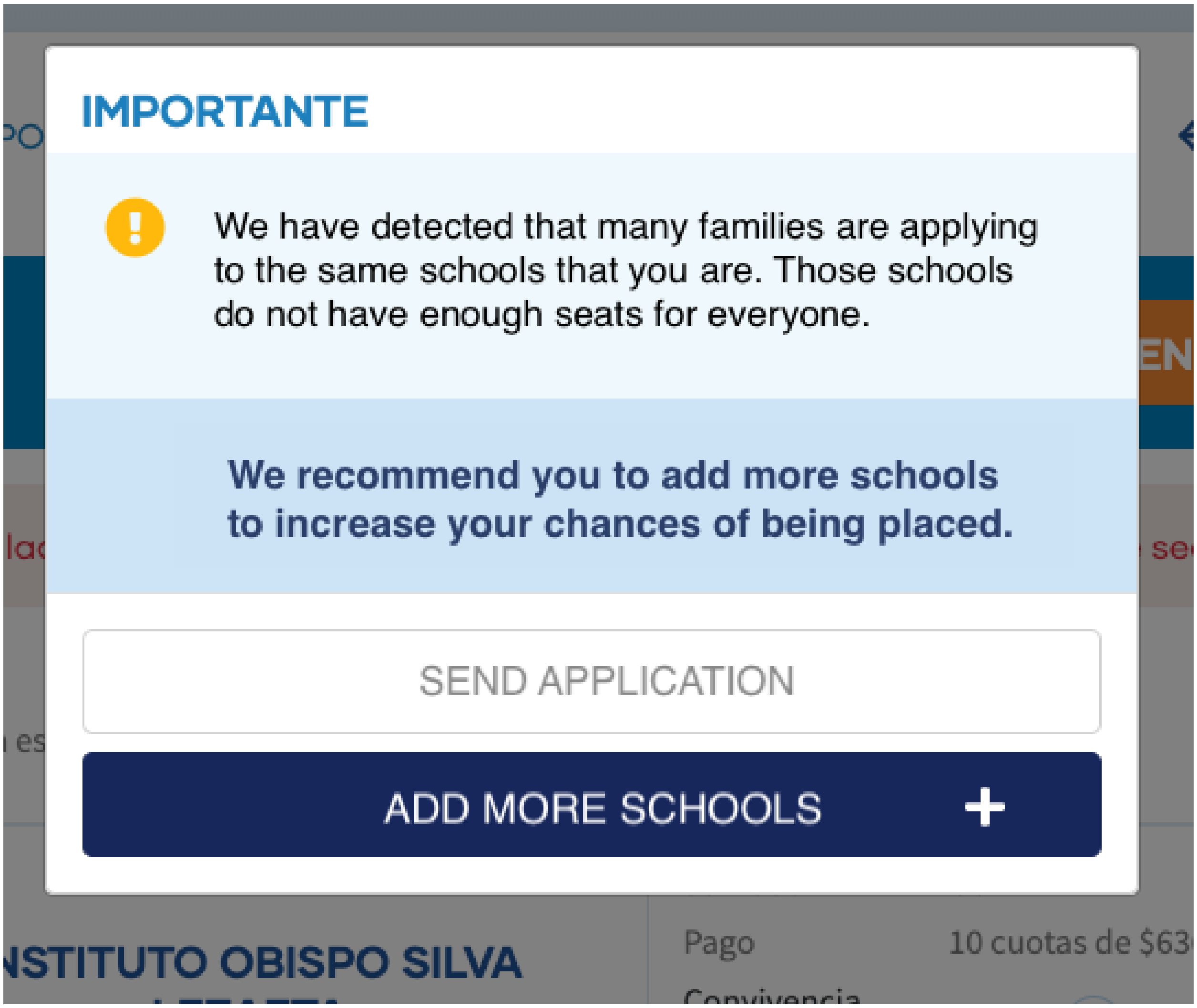
Notes. Local linear RD estimates of pop-up effects from warning pop-up on application platform, split by applicant's socioeconomic status. "Economically vulernable" individuals are the lower-SES group. See section III for details. Computed using triangular kernel with bandwidth 0.1. Heteroskedasticity-robust nearest neighbor variance estimator with minimum of 3 neighbors reported in parentheses; computed as in Calonico, Cattaneo and Titiunik (2014). IV estimates reported in second column of each set show the instrumental variable specifications (fuzzy RD), where the endogenous regressor is the add any school indicator. Panel A: predetermined covariates. Panel B: measures of choice behavior from initial to final application. 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: 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. "Undersubscribed" schools are those with excess capacity.

Table B.VII
RD Estimates of Platform Pop-Up Effects on Enrolled School Outcomes by Applicant's Socioeconomic Status

	(1)	(2)	(3)	(4)	(5)	(6)
	Economically Vulnerable			Not Economically Vulnerable		
		IV	$E[Y\|X=0.3^-]$		IV	$E[Y\|X=0.3^-]$
A. First stage and enrollment						
Add any	0.227		0.216	0.209		0.187
	(0.015)			(0.012)		
Enrolled	-0.009		0.979	-0.001		0.956
	(0.005)			(0.006)		
Have value added measure$\|$grade$\leq$8	0.061		0.730	-0.015		0.769
	(0.017)			(0.013)		
B. Attributes of enrolled school						
Distance (km)	-0.043	-0.214	3.294	0.112	0.516	2.841
	(0.466)	(2.345)		(0.281)	(1.292)	
Value added$\|$grade$\leq$8	0.001	0.006	0.103	0.034	0.151	0.160
	(0.018)	(0.091)		(0.014)	(0.061)	
Per teacher spending (1000USD)	1.202	5.992	30.224	0.519	2.361	30.926
	(0.358)	(1.855)		(0.289)	(1.327)	
Per student spending (1000USD)	0.001	0.005	2.252	0.002	0.008	2.240
	(0.022)	(0.112)		(0.020)	(0.091)	
With copayment fee	0.054	0.240	0.220	0.017	0.080	0.326
	(0.013)	(0.061)		(0.013)	(0.062)	
School monthly fee (USD)	3.925	17.390	15.534	0.507	2.378	26.851
	(1.043)	(4.794)		(1.202)	(5.639)	
Share of vulnerable students	-0.009	-0.040	0.604	-0.004	-0.018	0.539
	(0.004)	(0.018)		(0.004)	(0.018)	
Total enrollment per grade	14.382	63.825	104.184	4.702	21.964	94.904
	(2.758)	(12.771)		(2.193)	(10.347)	
NL	8,629	8,629		10,921	10,921	
NR	8,439	8,439		11,783	11,783	

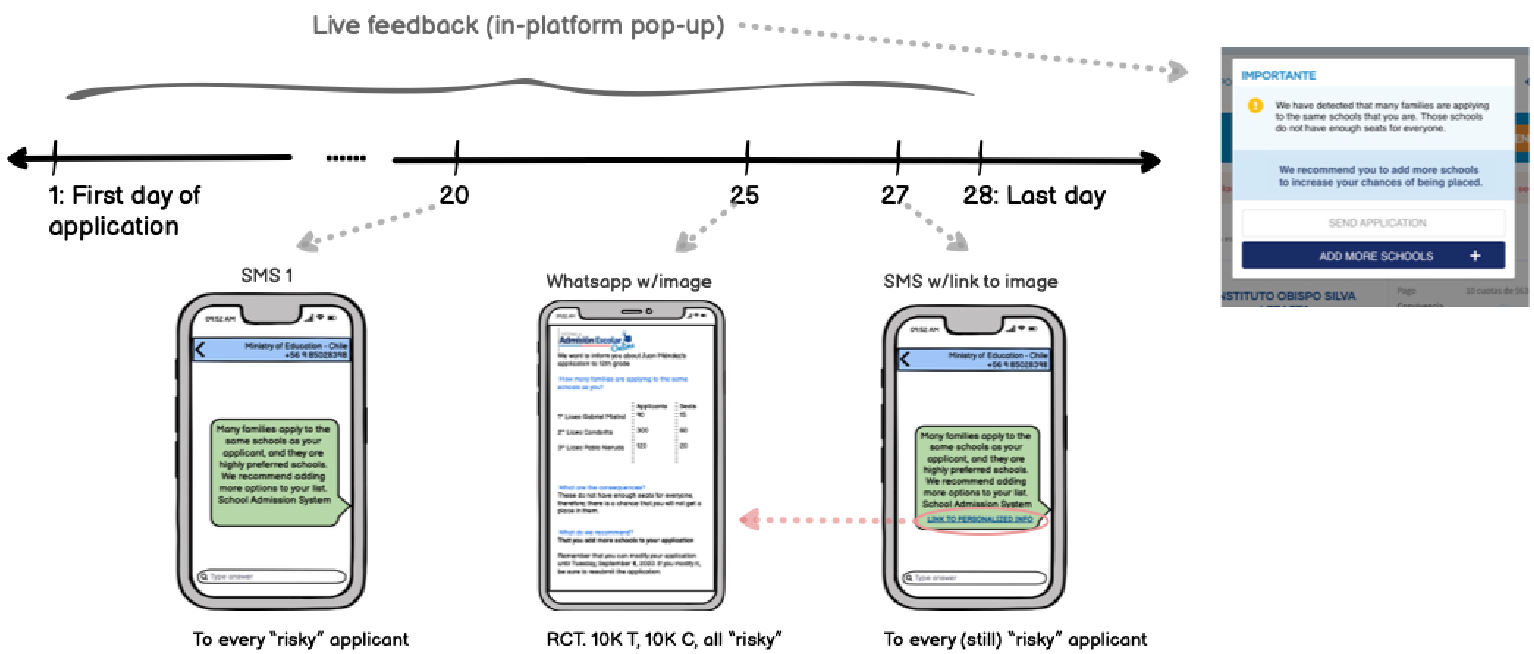
Notes. Local linear RD estimates of popup effects from warning popup on application platform. Computed using triangular kernel with bandwidth 0.1. Heteroskedasticity-robust nearest neighbor variance estimator with minimum of 3 neighbors reported in parentheses; computed as in Calonico, Cattaneo and Titiunik (2014). We report estimates split by whether students are economically vulnerable or not. IV estimates in columns 2 and 5 report instrumental variable specifications where the endogenous regressor is the "add any school" indicator. Columns 3 and 6 report below-cutoff means of the variable listed in the row in the analysis sample. Sample for value added outcomes is restricted to grades eight and below. Reported sample sizes are counts of enrolling students. See section V.E for discussion and Online Appendix D.4 for detailed variable definitions.

Figure B.I
Platform Pop-Up Intervention– 2018 and 2019



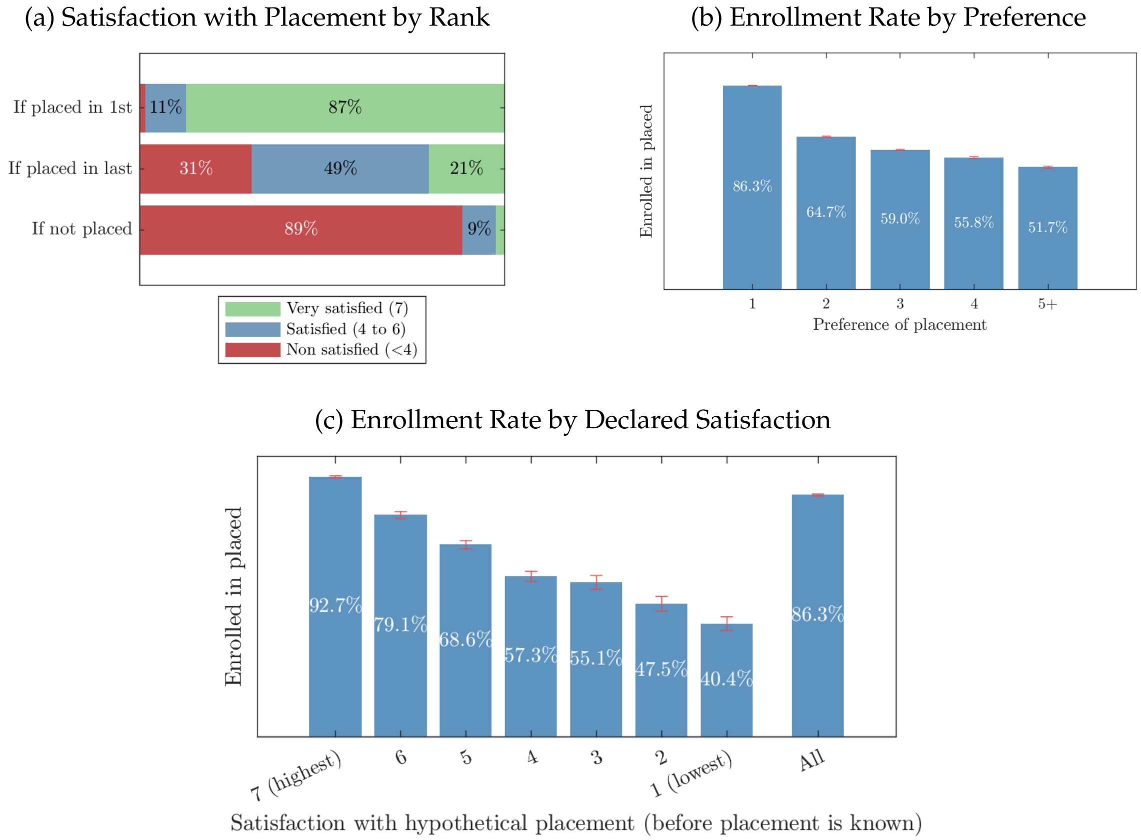
Notes. English translation of pop-up feedback shown to risky applicants on the application platform in 2018 and 2019. All applicants with predicted nonplacement risk of 30% or higher received this warning when they submitted their choice application. See section III.B for details.

Figure B.II
Timeline of Feedback Interventions– 2020



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

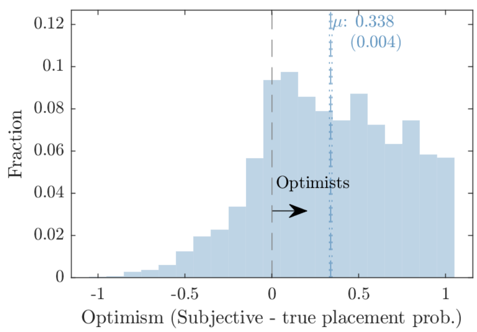
Figure B.III
Satisfaction with Placement Outcomes



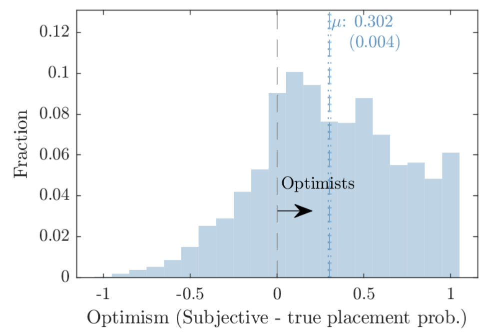
Notes. Panel A: stated satisfaction with hypothetical placement outcomes. Data are survey responses to questions about applicant satisfaction with being placed at their first-ranked school, last-ranked school, and nonplacement. Sample: survey completers. Results reported on a 1-7 scale, with 7 being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

Figure B.IV
Alternate Application Risk Framings

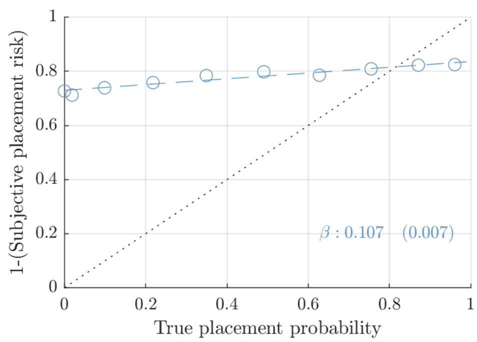
(a) Optimism– Negative Frame



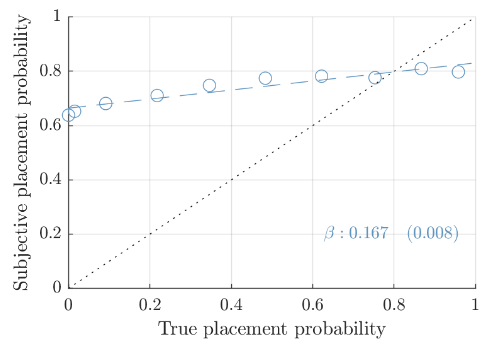
(b) Optimism– Positive Frame



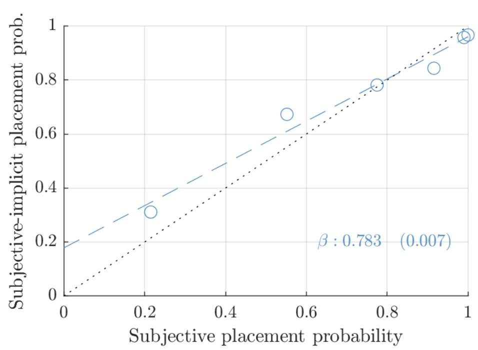
(c) Subjective vs. True Beliefs– Negative Frame



(d) Subjective vs. True Beliefs– Positive Frame



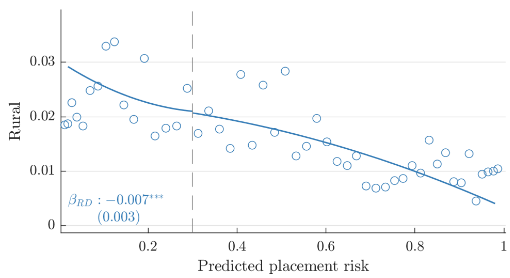
(e) Stated Non-placement Risk vs. Implied Non-placement Risk from School-Specific Risk Reports



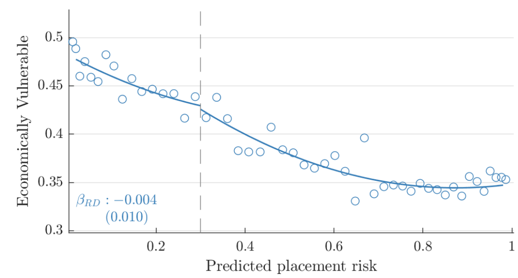
Notes. Panel A: distribution of optimism under "negative" framing for subjective risk question. Negative framing asks respondents about risk of non-placement under their submitted application. Panel B: distribution of optimism under "positive" framing for subjective risk question. Positive framing asks 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 probability 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 B.V
Balance in Platform Pop-Up RD

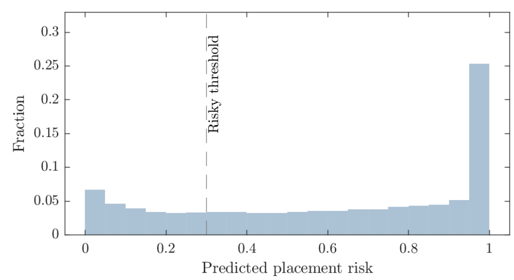
(a) Rural



(b) Economically Vulnerable

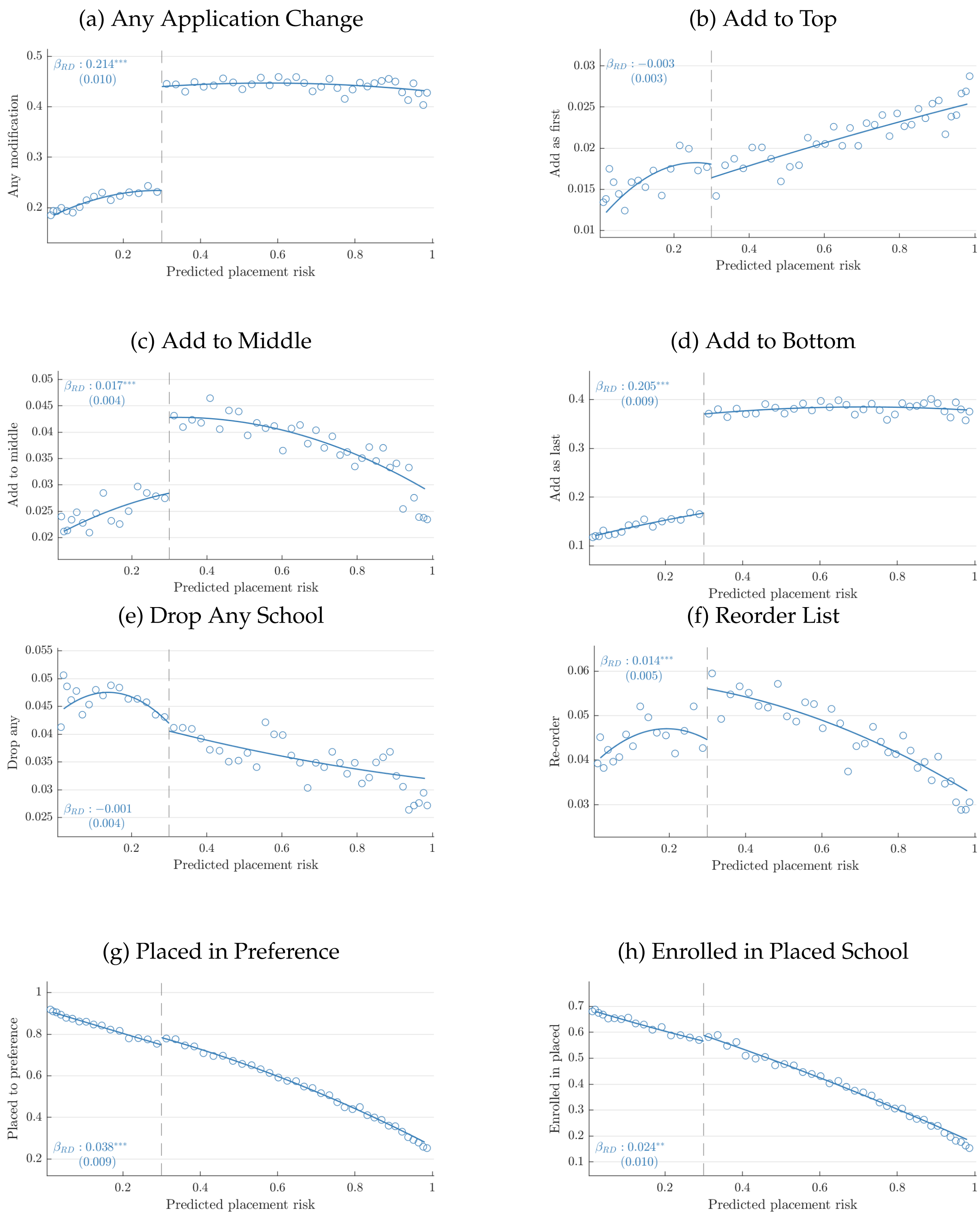


(c) Distribution of Predicted Risk (Only >0)



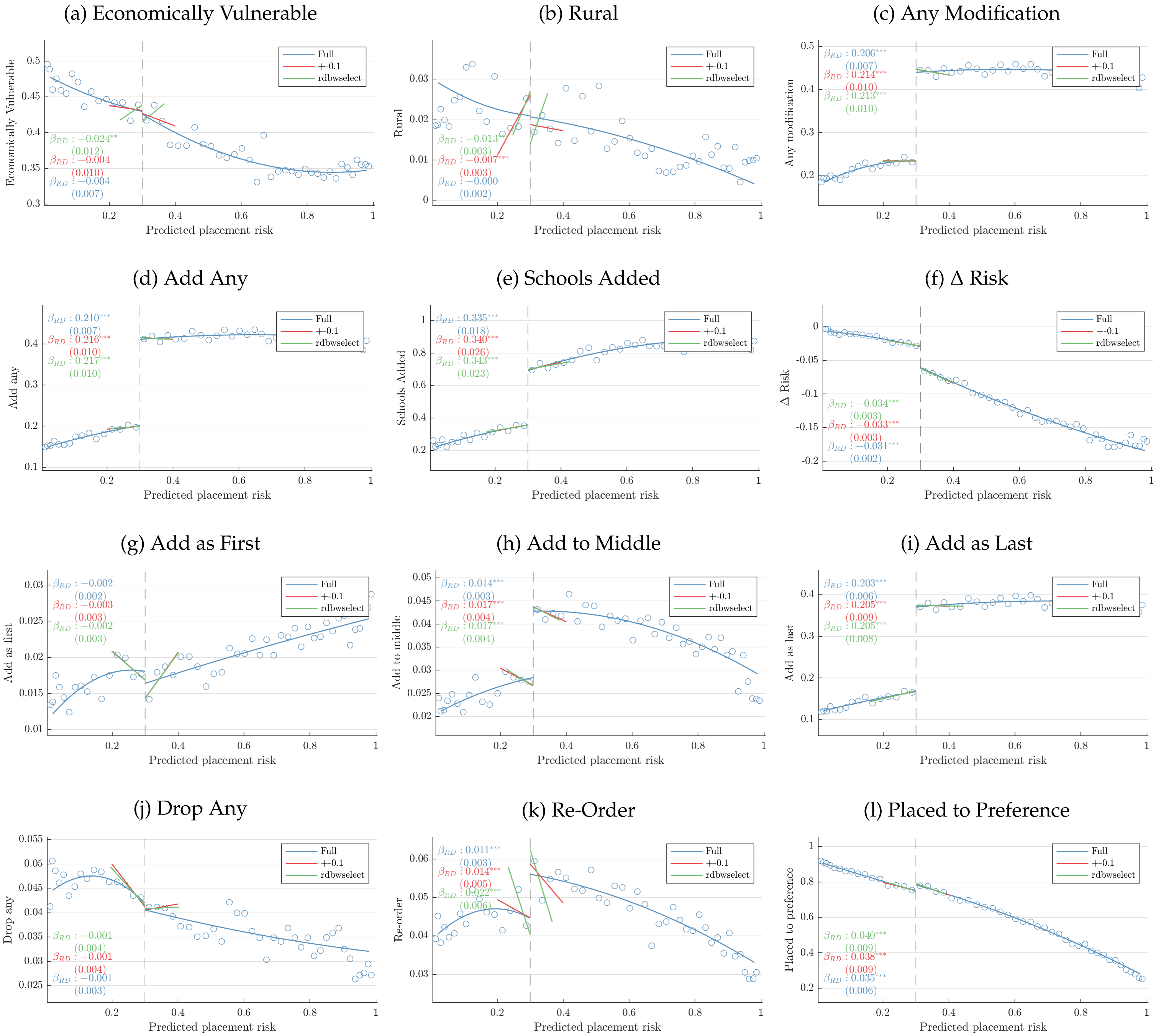
Notes. Binned means and global fits of predetermined characteristics by predicted risk for initial application. Points are centered means of 50 quantile-spaced bins of the support of the predicted placement risk $\in [0.02; 0.98]$. Solid line shows the quadratic fit. Reported coefficients and standard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

Figure B.VI
Additional Platform Pop-Up RD Figures



Notes. Binned means and global fits of choice outcomes by predicted risk for initial application. Points are centered means of 50 quantile-spaced bins of the support of the predicted placement risk $\in [0.02; 0.98]$. Solid line shows the quadratic fit. Reported coefficients and standard errors are from local linear specifications using $+-0.1$ bandwidth. See section V.A 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 school to top of list. Panel C: add school to middle of list. Panel D: add school to bottom of list. Panel D: drop any school from list. Panel F: reorder existing schools. Panel G: place to listed preference. Panel H: enroll in placed school.

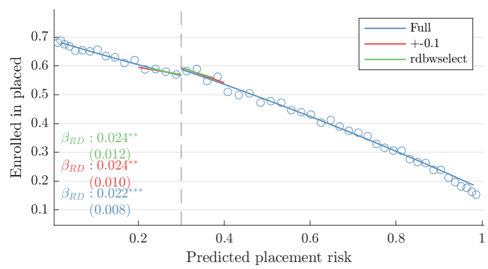
Figure B.VII
Multiple Bandwidths RD Plots of Platform-Based Pop-Up Warning Effects (Outcomes in Table II)



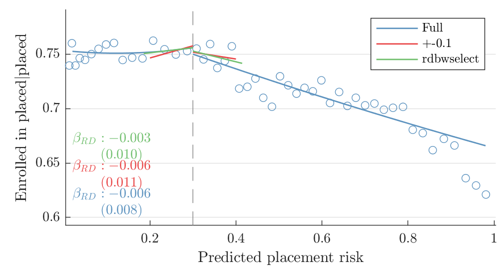
Notes. Pop-up warning RD effect fits and point estimates by bandwidth for outcomes listed in panel titles. "Full": global quadratic. "+/- 0.1": local linear within 0.1 bandwidth. "rdbwselect": optimal bandwidth selection using Calonico, Cattaneo and Titiunik (2014). See section V.A for details.

Figure B.VIII
Multiple Bandwidths RD plots of Platform-Based Pop-Up Warning Effects (Outcomes in Table II)

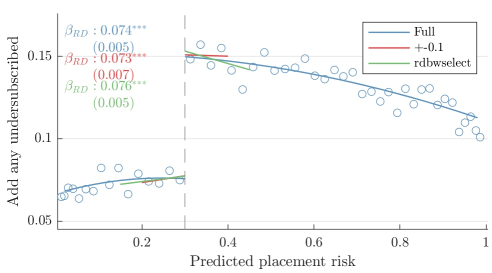
(a) Enrolled in Placed



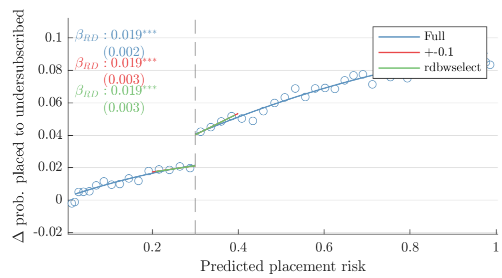
(b) Enrolled in Placed|Placed



(c) Add any Undersubscribed

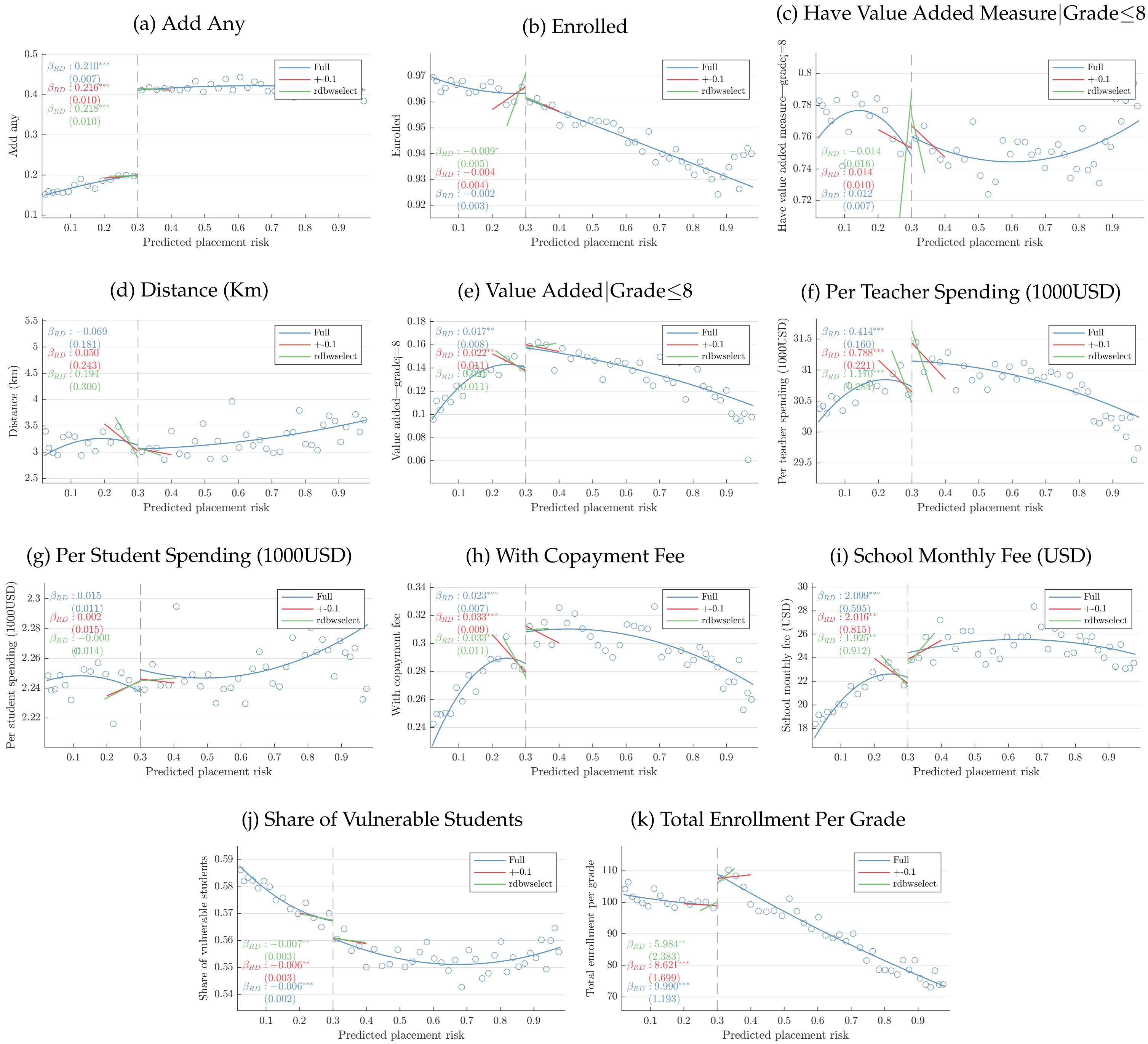


(d) Δ Prob. Placed to Undersubscribed



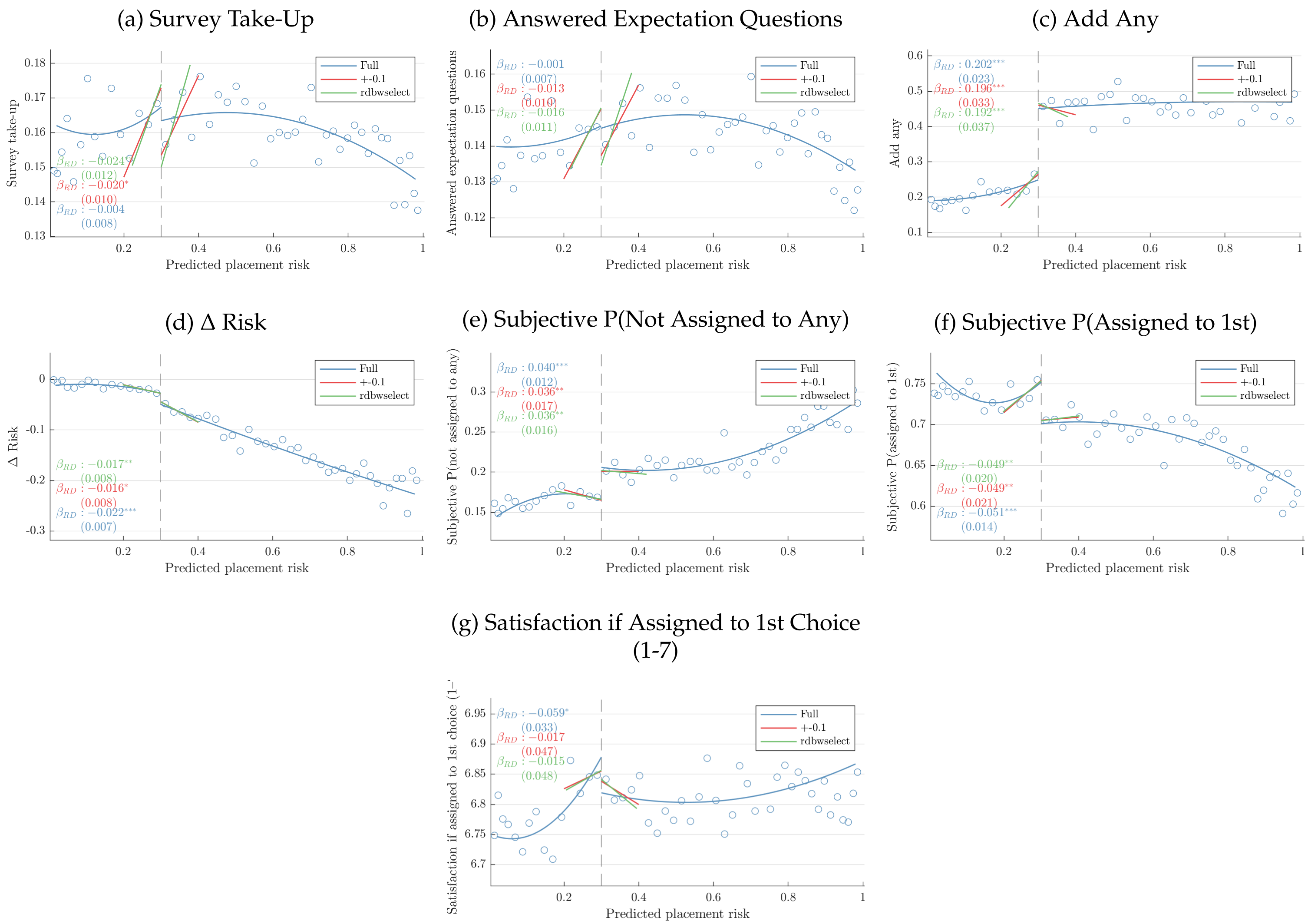
Notes. Pop-up warning RD effect fits and point estimates by bandwidth for outcomes listed in panel titles. "Full": global quadratic. "+/- 0.1": local linear within 0.1 bandwidth. "rdbwselect": optimal bandwidth selection using Calonico, Cattaneo and Titiunik (2014). See section V.A for details.

Figure B.IX
Multiple Bandwidths RD Plots of Platform-Based Pop-Up Warning Effects (Outcomes in Table III)



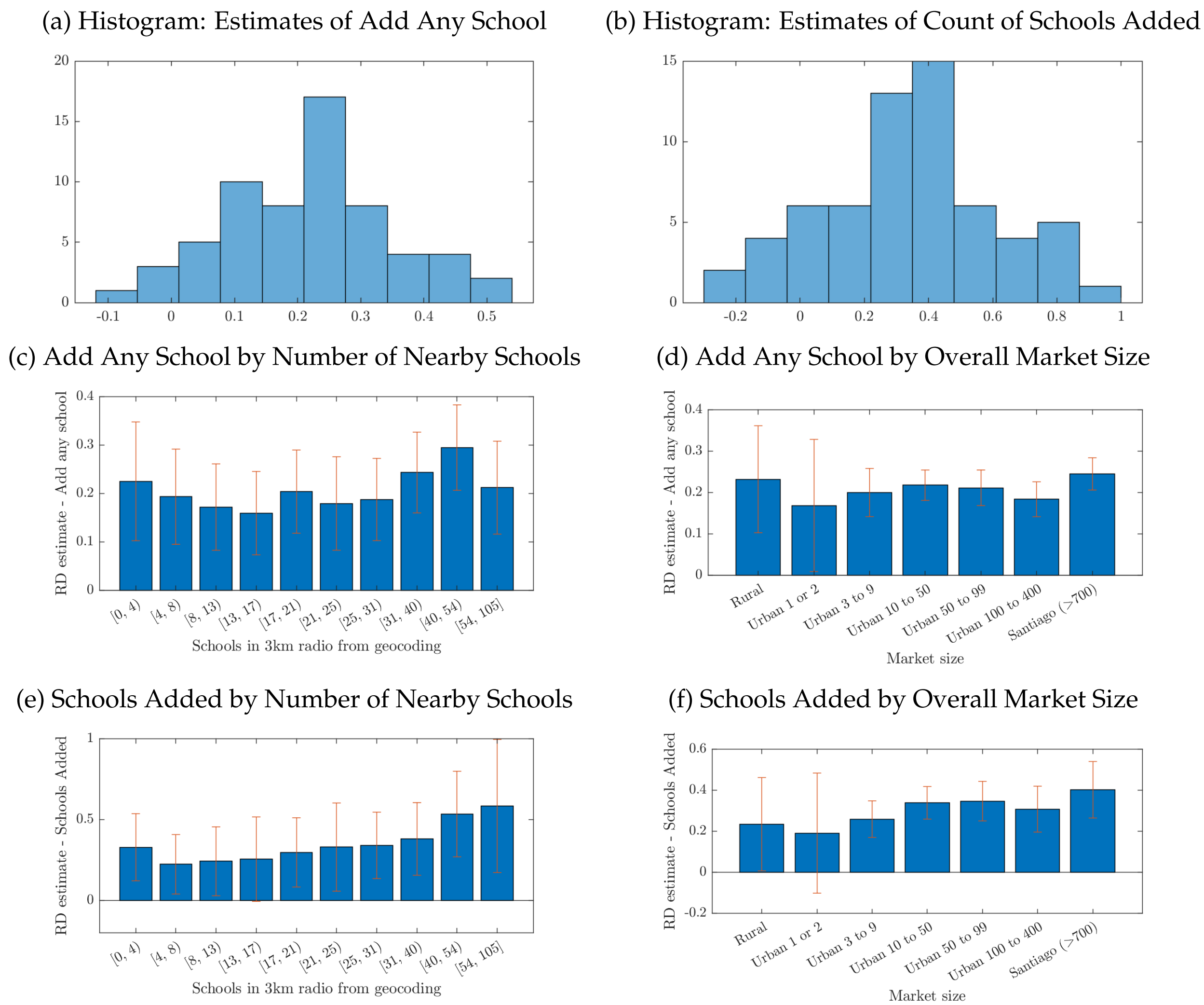
Notes. Pop-up warning RD effect fits and point estimates by bandwidth for outcomes listed in panel titles. “Full”: global quadratic. “+/- 0.1”: local linear within 0.1 bandwidth. “rdbwselect”: optimal bandwidth selection using Calonico, Cattaneo and Titiunik (2014). See section V.A for details.

Figure B.X
Multiple Bandwidths RD Plots of Platform-Based Pop-Up Warning Effects (Outcomes in Table V)



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, Cattaneo and Titiunik (2014). See section V.A for details.

Figure B.XI
Platform Pop-Up Effects over City-Years and by Market size



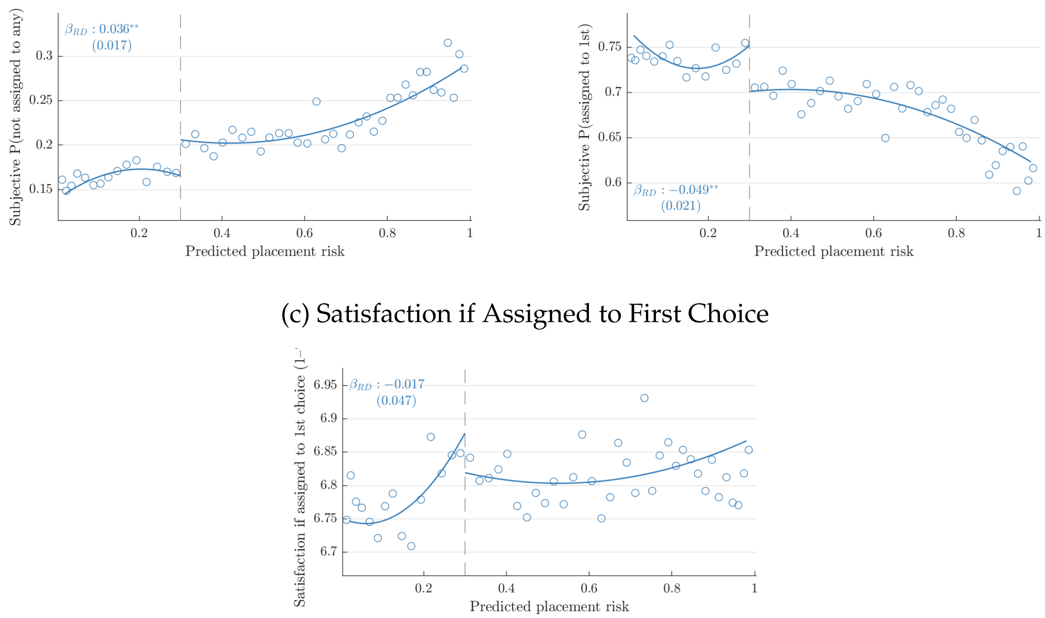
Notes. Panels A and B: distribution of estimated platform pop-up RD effects across city-year cells. Each city-year cell is one observation. Outcome in Panel A is add any school, outcome in panel B is count of schools added. Panel C: pop-up RD treatment effects on add any school split by count of nearby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

Figure B.XII
Intervention Effects on Beliefs and Preferences

(a) Not Assigned to Any School

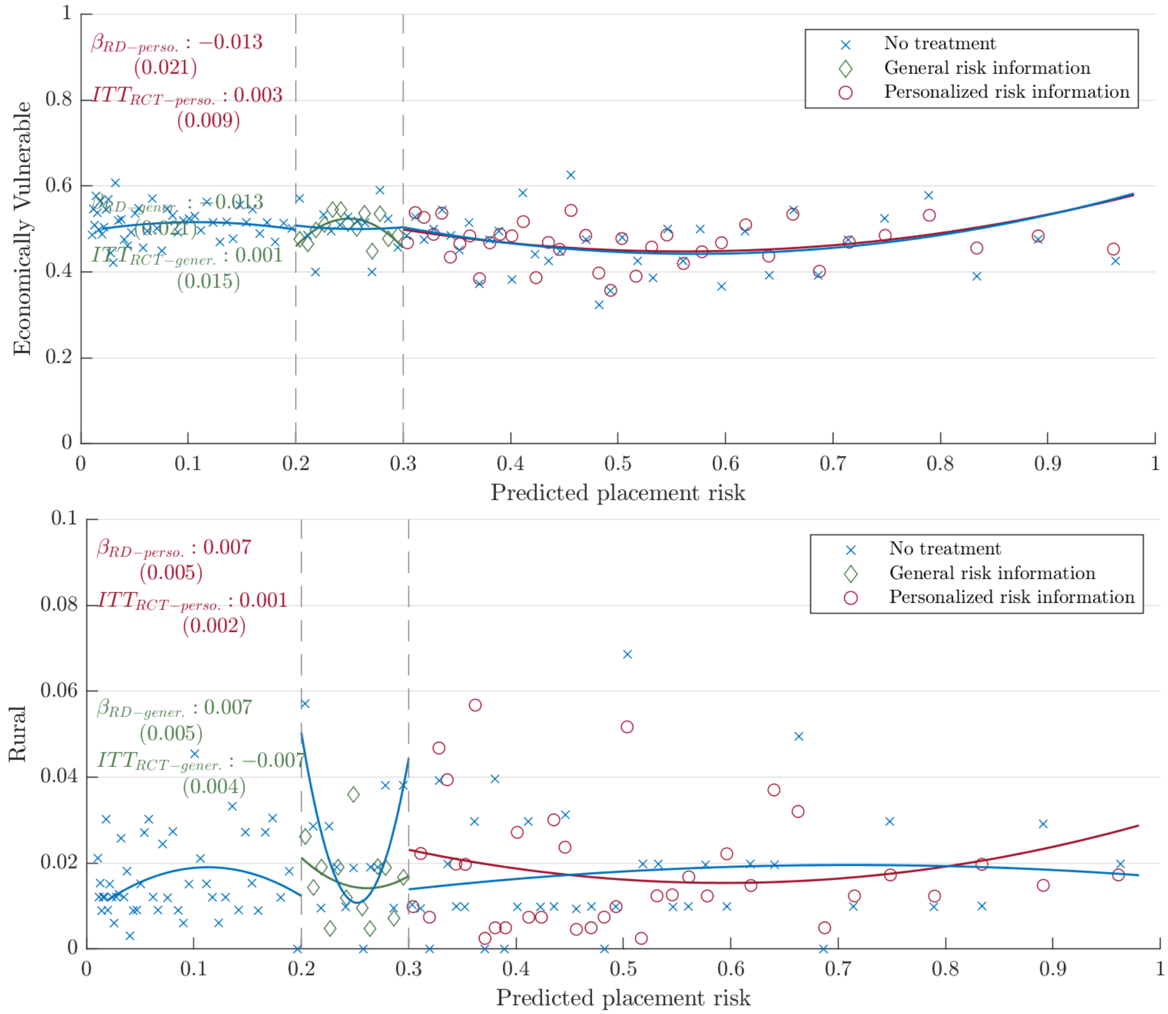
(b) Assigned to First Choice

(c) Satisfaction if Assigned to First Choice



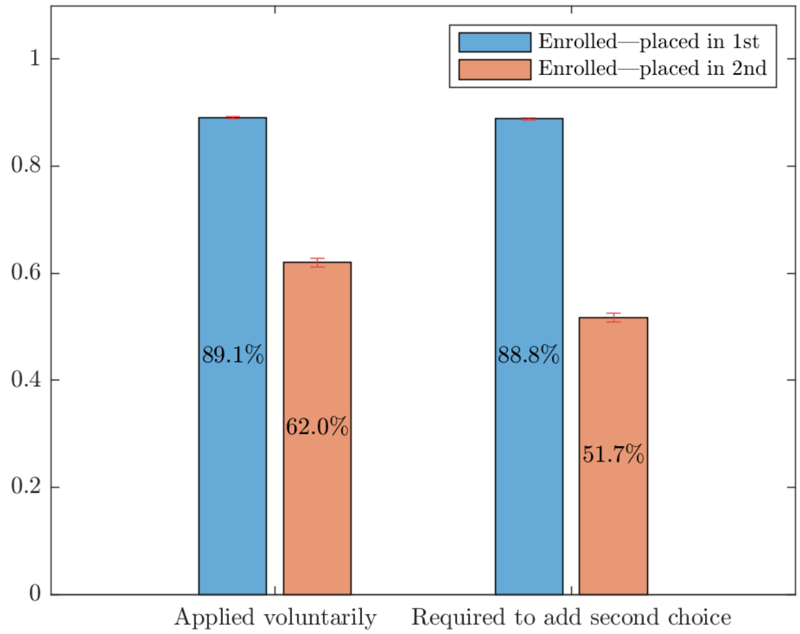
Notes. Binned means and global fits of choice outcomes by predicted risk for initial application. Points are centered means of 50 quantile-spaced bins of the support of the predicted placement risk $\in [0.02; 0.98]$. Solid line shows the quadratic fit. Reported coefficients and standard errors are from local linear specifications using $+-0.1$ bandwidth. See section V.A for details. Because coefficients are local while polynomial fits are global, there may be minor differences between displayed fits and reported coefficients. Sample is applicants who completed the belief modules of the endline survey. Panel A: outcome is survey-reported subjective belief about the chances of not being assigned to any school. Panel B: outcome is survey-reported subjective belief about placement chances at the first-listed school. Panel C: outcome is survey-reported satisfaction with placement at the first-choice school.

Figure B.XIII
Balance in 2021 WhatsApp RCT



Notes. Balance test results from 2021 WhatsApp RCT. Dependent variable is an indicator for (predetermined) economic vulnerability status (upper panel) and an indicator for rural location (lower panel). Treatments are as follows. "No treatment": control group that receives no WhatsApp message. "General risk information": treatment group that receives information about nonplacement risk in aggregate, not personalized to own application. "Personalized risk information": treatment group that receives information about own application risk, as in 2020 WhatsApp RCT. $\beta_{RD-general}$ is the RD estimate of general risk treament group against the control group at the 0.2 cutoff. $\beta_{RD-personal}$ is the RD estimate of the personalized risk information treatment group relative to the general risk treatment group. $ITT_{RCT-personal}$ and $ITT_{RCT-general}$ are RCT estimates of treatment effects for the personal and general info treatments (respectively) relative to the control group in the same risk range. See section 2 and Online Appendix C.3 for design details and additional results. Reported RD coefficients and standard errors are from local linear specifications using $+-0.1$ bandwidth. See section V.A for details.

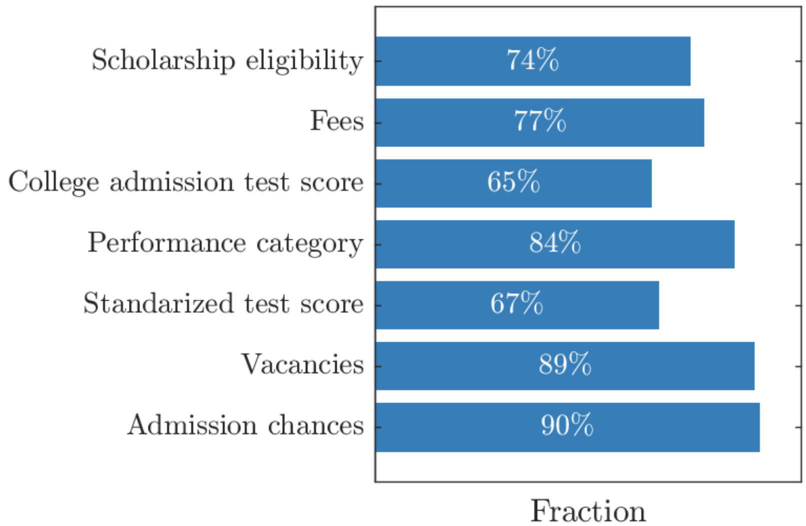
Figure B.XIV
Enrollment in Placed Rates Conditional on Placement Ranking



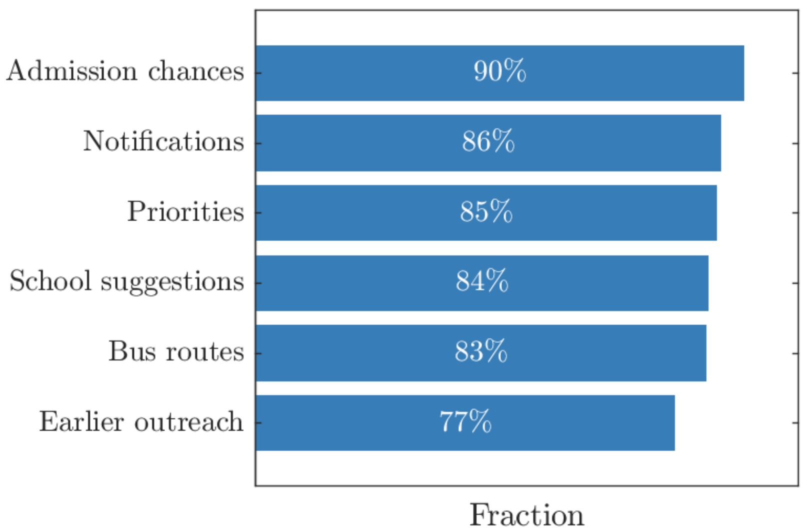
Notes. Share of students enrolling in the placed school, by rank of placed school and whether the applicant was forced to add the second school. The "applied voluntarily" group listed at least two schools on their initial application. The "Required to add second choice" group initially listed only one school and was required by the centralized system to add a second school.

Figure B.XV
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)



Notes. 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 IV 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 K for more details.

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 (Gobierno de Chile 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 C.I for an illustration of the policy rollout.

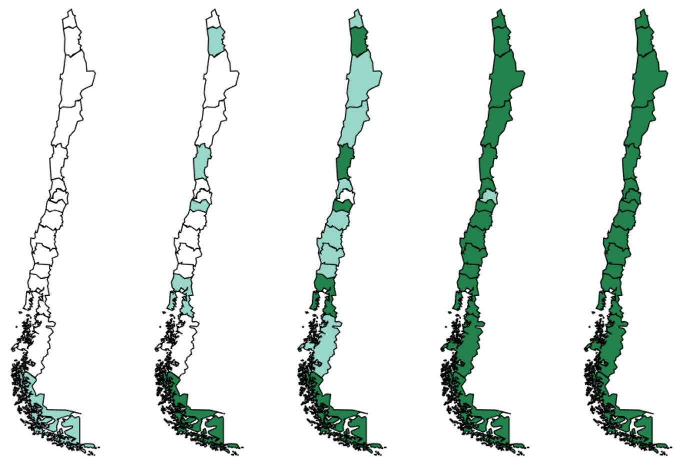
Schools that only enroll pre-Kindergarten and Kindergarten students were excluded from the centralized system. Lower-grade applicants could enroll through the centralized system in schools that also offer higher grades, but not in standalone early-grade institutions.

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 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

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

Figure C.I
Policy Rollout



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.

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]

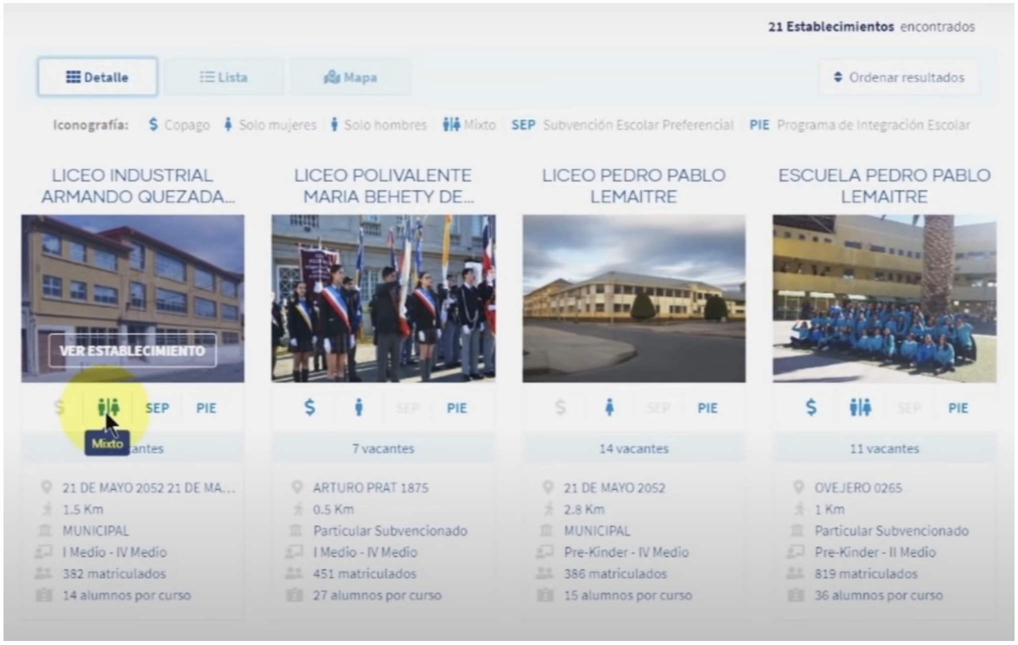
Students submit their applications to each choice round through an online platform. In addition to collecting applications, this platform provides information about schooling options, including test scores, fees, infrastructure, enrollment, religion, and extracurricular activities. See Figure C.II for 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.

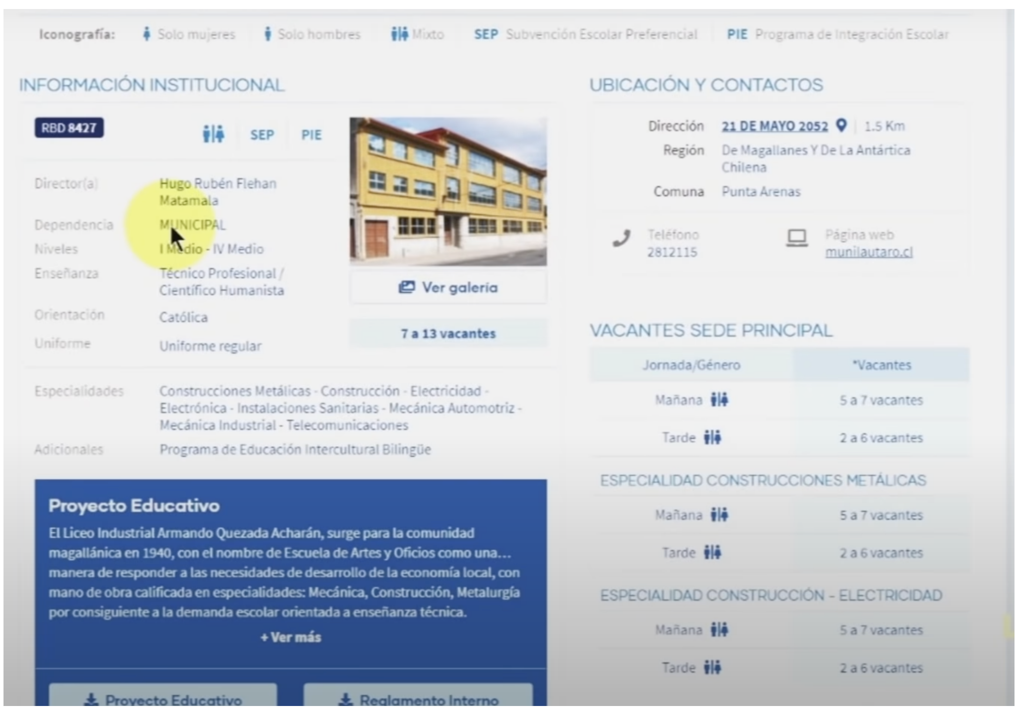
[4]In 2016, Mineduc also implemented default assignment to the closest school for the *first* application round. We do not use 2016 data in our analysis.

Figure C.II
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

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.

2. Placement Outcomes over Time. Table C.I 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 C.I shows that mean application length declines over time, both overall and within each year-of-entry group.

Table C.I
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.

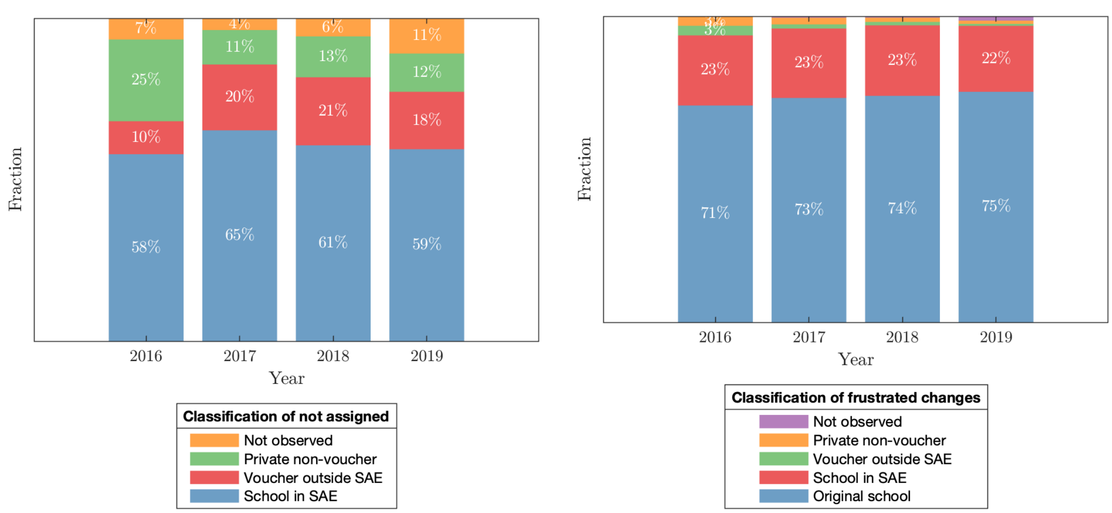
3. Options Outside the Centralized System. This subsection describes applicants' outside option behavior in detail. Figures C.IIIa and C.IIIb show the next-year enrollment of students that did not get a spot in the first round of the centralized process (i.e., those whom we classify as unplaced in our main text analysis). We split the sample by whether applicants have an option to continue in their current school or not. Figure C.IIIa 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 C.IIIb 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 C.III
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.

2016 SMSs were sent in three waves to applicants in the Puntarenas region (the only region with centralized choice at the time) in markets where there were overdemanded schools. The first wave was on day 25 of the application process. Three different general messages and no-message were randomly assigned and sent to all the applicants with fewer than five schools on their applications. For examples, see panels (a) to (c) of Figure C.IV. These are the interventions for which we report results in Table VI.

We also conducted an additional intervention in 2016 that provided a leading indicator on the promise of personalized risk warnings. On day 36, applicants with a predicted risk higher than 0.01 were randomly assigned to be the recipients of an additional personalized SMS. This message represented a waypoint on the path from impersonal, no-information nudges to the personalized smart platforms deployed in 2017 and later. It included a personalized reference to the number of schools that the applicant had already placed on his list, which is risk-relevant for many applicants, but did not refer directly to application risk. See panel (d) of Figure C.IV. On day 44, 5 days before the last application day, all applicants in the control group received the same SMS. The impact of the personalized SMS on the share of applicants who added schools prior to the message being sent to the control group is 0.051 (0.023), roughly double the size of the encouragement nudge effects reported in Table VI. The idea that personalized, risk-relevant information might be more effective than generic encouragement was an input to the choice to roll out smart platforms in subsequent years.

The personalized (but not "smart platform") SMS intervention was cross-randomized with the impersonal encouragement nudge interventions. When reporting the results of the encouragement nudge interventions in Table VI, we consider only changes to the application made in the "clean" 11 day window between the encouragement nudge and personalized non-smart intervention. This approach parallels our analysis of the 2020 WhatsApp RCT in section V.G, and lets us capture the effects of the encouragement nudge in the absence of interactions with the subsequent personalized intervention. Estimates of encouragement intervention effects that take final applications as the outcome show null results similar to those reported in Table VI, both in the full sample and in the subsample of applicants assigned to the control group for the personalized intervention. These results are available upon request.

2017 There were two information interventions this year. The first was an on-platform pop-up warning to risky applicants (Figure C.VI). The threshold that defined a "risky applicant" varied between 0.3, 0.5, and 0.7 values of predicted risk, depending on the region. The pop-up was active starting with the third day of the process. See Online Appendix I for a discussion of results from this intervention.

The second intervention consisted of SMSs sent to risky applicants who applied during the first two days, when the pop-up was not active. Four different personalized messages were randomized. The basic content was composed by three concepts (1) risk, (2) consequences of risk, (3) a recommended action, the message were different on the order of the concepts and the wording. For examples, see panels (a) to (d) of Figure C.V.

2018 The implementation team at Mineduc kept on-platform pop-up 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 C.VI 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 C.VII. Each image conveys the same idea as the pop-up warning. We pool over image types in our main analysis of the WhatsApp RCT. See Online Appendix J for a breakout of effects by image type.

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.

2021 The platform pop-up warning was implemented again, starting on the first day of the application window. As in the previous years, the risk cutoff for platform pop-up receipt was 0.3. Risk calculations were based on demand from 2020 and not updated with current demand until the 20th day. On day 20, the NGO re-calculated nonplacement risk for all applicants based on 2021 application data.

Starting on day 20, the NGO conducted a risk warning campaign with the goal of reaching out to applicants who had not received a platform pop-up warning based on the initial risk calculation, but who appeared to be at risk of nonplacement under the updated calculation. As in 2020, this campaign included a randomized WhatsApp component. Within the universe of applicants who had not received a platform pop-up warning, applicants with (updated) nonplacement risk above 0.30 were randomly assigned to either a control group that received no message or a treatment group that received a message with personalized application information similar to the pop-up warning. In addition, applicants with risk scores between 0.2 and 0.3 were randomly assigned to either a control group that received no message or a treatment group that received a *non-personalized* messsage about nonplacement risk in the aggregate. These two treatments form the basis for the analysis described in section 2. See Figure C.VIII for screenshots of the WhatsApp messages in the two treatment arms. The number and timing of warnings was the same across the two treatment arms, except for the difference in warning text.

Two points are important to make here. First, note that sample selection into the 2021 WhatsApp RCT is somewhat different than for the 2020 WhatsApp RCT, because the 2021 RCT sample universe consisted of applicants who had not yet seen a pop-up warning on the application platform, whereas in 2020 most participants in the WhatsApp RCT had already received a platform pop-up, as described in section V.G.

Second, risk calculations changed enough for enough applicants that the RCT sample universe of applicants who had not received platform pop-up had support across the distribution of updated risk scores.[5] Relatively low-risk applicants are over-represented among

[5]Risk updating was an important issue in 2021 because Covid-19 had depressed applications in 2020.

RCT-eligible applicants. However, as shown in Online Appendix Figure C.IX, we observe RCT-eligible applicants over the full range of updated risk values, and the share of RCT-eligible applicants is smooth through the 0.3 cutoff of the updated risk distribution. This latter point speaks to the validity of the RD analysis described in section 2.

Five days after the WhatsApp messages went out to treatment groups, 28% of RCT participants (and 30% of applicants overall) received a report card that included the risk measures for every school in the application. For risky applicants the report card also had a warning with a similar message to the platform pop-up. As in our analysis of the 2020 WhatsApp intervention, we focus our analysis of the effects of the 2021 WhatsApp RCT on changes made to the application in the "clean" five-day window between the randomized WhatsApp intervention and the followup message sent to all risky applicants. Findings reported in Figure VII and referenced in the main text are from this clean five-day window.

Table C.II summarizes findings from the 2021 WhatsApp RCT beyond those reported in Figure VII. As expected, both the RCT and RD designs are balanced on applicant observables (specifically, economic vulnerability) and on eventual receipt of the report card intervention. Roughly 80% of applicants in both the general and personalized information treatments view the WhatsApp image. There is no difference in viewership rates across treatment type. We observe changes in choice behavior both in the "clean" five-day window between the WhatsApp treatment and the report card and at the endline. As in the 2020 WhatsApp RCT, treatment effects grow over time. Effects on choice behavior (i.e., the share of "compliers" who add schools, the number of schools added) are much larger than in 2020. IV estimates of risk reductions for compliers are somewhat smaller than in 2020, likely because the population receiving the WhatsApp treatment is less risky at baseline.

We have also conducted analyses that exclude the 28% of applicants who received the report card intervention from the sample. These findings are available upon request and yield the same conclusions as those we report here.

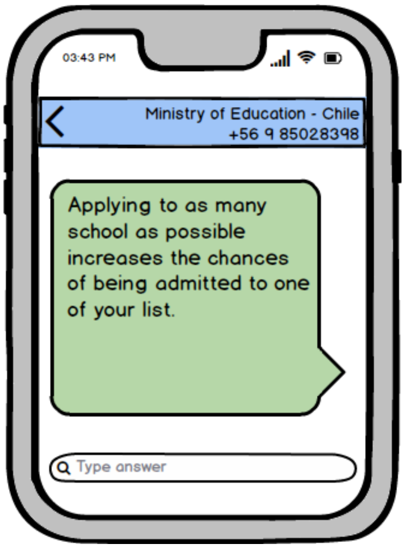
Table C.II
WhatsApp RD and RCT Results – 2021

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	General risk information				Personalized risk information			
	RCT		RD		RCT		RD relative to general risk treatm.	
	ITT	IV	ITT	IV	ITT	IV	ITT	IV
A. Balance								
Economically Vulnerable	0.001		-0.013		0.003		-0.008	
	(0.015)		(0.021)		(0.009)		(0.016)	
Rural	-0.007		0.007		0.001		0.006	
	(0.004)		(0.005)		(0.002)		(0.004)	
B. Message receipt								
WhatsApp read	0.793		0.798		0.792		0.005	
	(0.005)		(0.012)		(0.003)		(0.014)	
Report card intention to treat	0.004		0.013		0.008		0.026	
	(0.014)		(0.019)		(0.008)		(0.015)	
C. Outcomes in clean 5 days before report card								
Any modification	0.041		0.043		0.081		0.037	
	(0.003)		(0.007)		(0.003)		(0.008)	
Add any	0.039		0.038		0.080		0.038	
	(0.003)		(0.006)		(0.003)		(0.008)	
Schools Added	0.070	1.807	0.067	1.766	0.156	1.957	0.086	2.013
	(0.007)	(0.114)	(0.015)	(0.257)	(0.006)	(0.052)	(0.020)	(0.441)
Δ Risk	-0.002	-0.050	-0.001	-0.024	-0.013	-0.157	-0.000	-0.057
	(0.001)	(0.012)	(0.001)	(0.035)	(0.001)	(0.008)	(0.002)	(0.058)
D. Endline outcomes								
Any modification	0.054		0.053		0.099		0.055	
	(0.005)		(0.009)		(0.004)		(0.010)	
Add any	0.051		0.046		0.096		0.055	
	(0.005)		(0.009)		(0.004)		(0.009)	
Schools Added	0.098	1.921	0.085	1.856	0.203	2.108	0.148	2.978
	(0.015)	(0.236)	(0.022)	(0.313)	(0.010)	(0.067)	(0.025)	(0.611)
Δ Risk	-0.001	-0.029	-0.002	-0.034	-0.016	-0.163	-0.004	-0.149
	(0.001)	(0.017)	(0.002)	(0.049)	(0.001)	(0.009)	(0.003)	(0.061)

Notes. ITT and IV effects of 2021 WhatsApp warnings intervention. RCT columns (1) and (2): effects of random assignment to general risk information message vs. control group for students with predicted risk $\in (.2, .3]$. RCT columns (5) and (6): effects of random assignment to personalized risk information message vs. control group for students with predicted risk $> .3$. Robust SEs in parentheses. General information intervention N=6,819. Personalized information intervention N=18,763. RD columns (3) and (4): regression discontinuity evaluation of general risk information message vs no treatment around 0.20 cutoff. RD columns (7) and (8): regression discontinuity evaluation of personalized risk information message vs general risk information message around 0.30 cutoff. RD specifications computed using local linear fit with a bandwidth of 0.1. Standard errors are heteroskedasticity-robust nearest neighbor variance estimator with minimum of 3 neighbors, as in Calonico, Cattaneo and Titiunik (2014). ITT column shows effects of group assignment. IV columns show the instrumental variable specification, where the endogenous regressor is the add any school indicator, instrumented with group assignment for the RCT, and with a dummy of crossing the risky threshold for the RD. Panel A: balance tests on predetermined characteristics. Panel B: message receipt. "WhatsApp read" is an indicator equal to one if applicant views the WhatsApp treatment message. "Report card intention to treat" is indicator for receiving a report card 5 days later. Report card included additional information on nonplacement probability and school options. 28% of the RCT sample were assigned to received the report card. Panel C: outcomes within 5 days window between WhatsApp intervention and report card was sent. Panel D: endline choice behavior. See the description of the 2021 intervention in section C.3 for details.

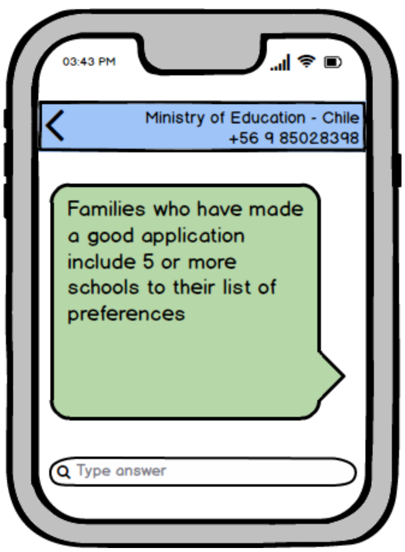
Figure C.IV
SMSs Intervention Warning Texts – 2016

(a) General SMS - "More Schools, Higher Chances"



(b) General SMS - "Range Suggestion"

(c) General SMS - "Role Model"



(d) Personalized SMS

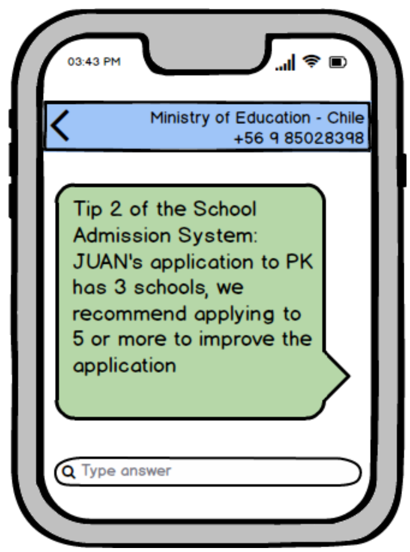
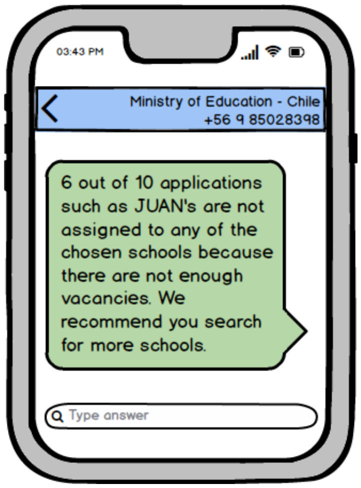
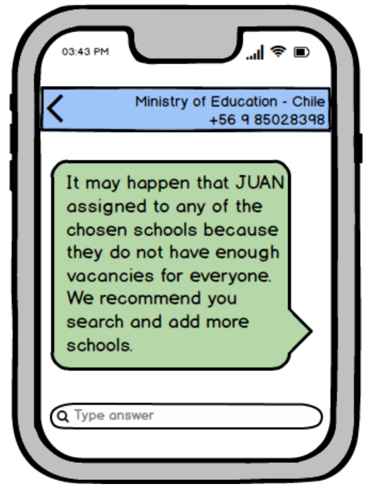


Figure C.V
SMSs Intervention Warning Texts – 2017

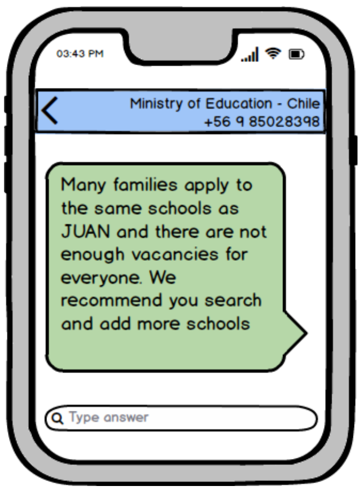
(a) Personalized Treatment - "Probability First"



(b) Personalized Treatment - "Consequences First"



(c) Personalized Treatment - "High Demand First"



(d) Personalized Treatment - "Recommendation First"

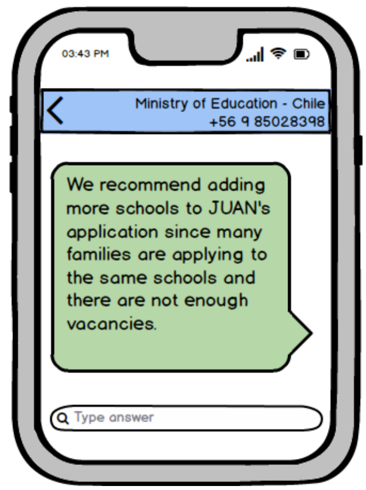


Figure C.VI
Platform Pop-Ups

(a) Platform Pop-Up 2018

(b) Platform Pop-Up 2019 - 20

(c) Translation Pop-Up 2019 - 2020

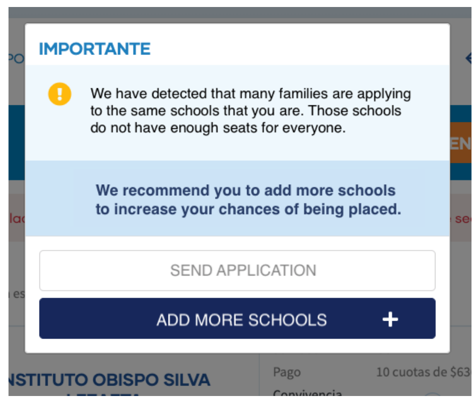
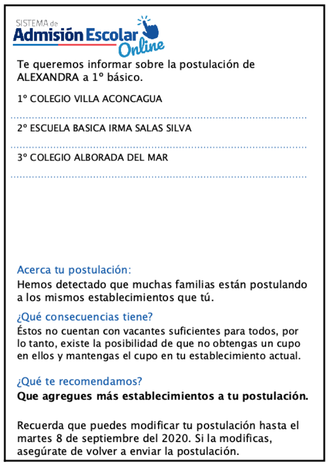
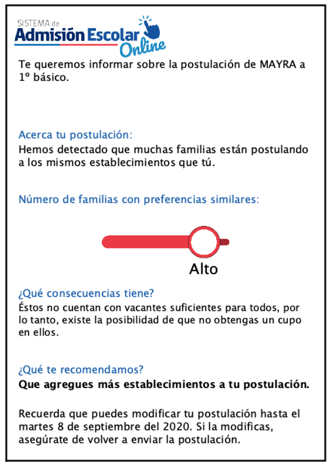


Figure C.VII
WhatsApp Intervention Warning Images – 2020

(a) Simple Image



(b) Risk Warnings Bar



(c) School-Specific Vacancy Estimates

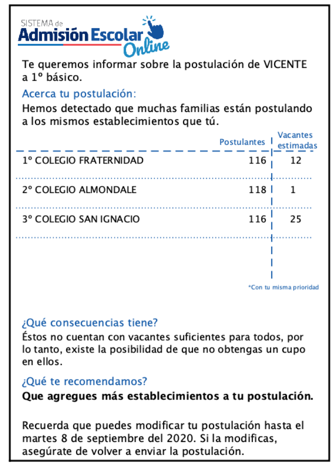
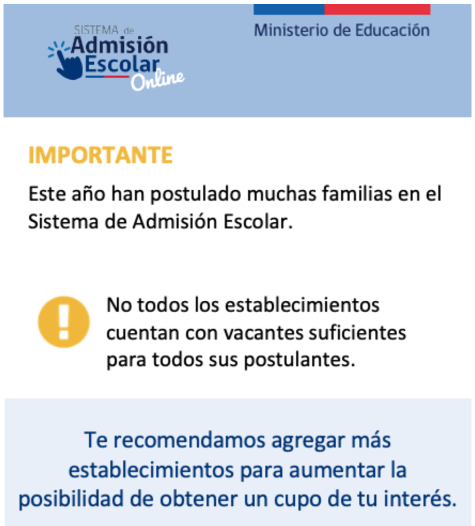
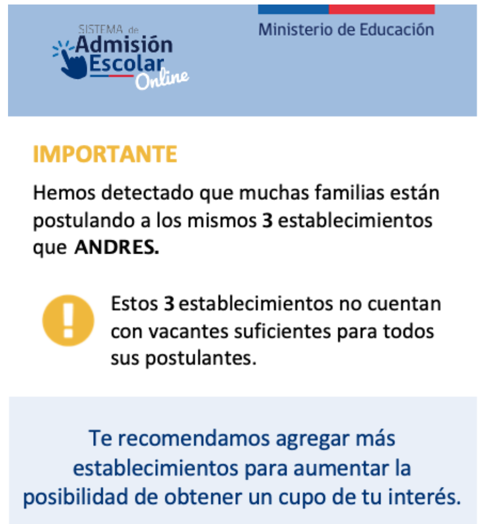


Figure C.VIII
WhatsApp Intervention Warning Images – 2021

(a) General Information Image



(b) Personalized Information Image



(c) Translation of General Information Image

(d) Translation of Personalized Information Image

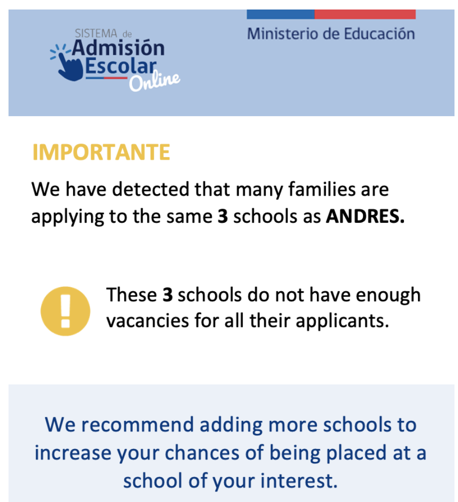
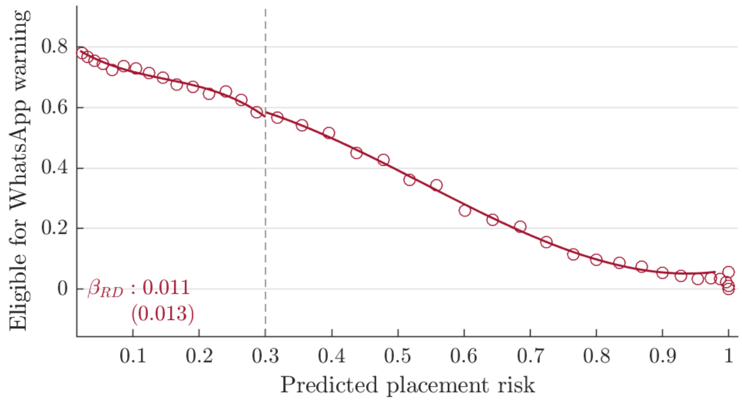


Figure C.IX
Eligibility for 2021 WhatsApp RCT



Notes. Share of 2021 applicants eligible for 2021 WhatsApp RCT by *updated* risk score. Only applicants who had not received the platform pop-up were eligible for the 2021 WhatsApp RCT. This graph plots of the share of such applicants in 50 quantile-spaced bins. Solid line shows a cubic fit. Reported coefficient and standard error are from local linear specifications using $+-0.1$ bandwidth. See section V.A for details. Some applicants below the 0.3 cutoff in the updated risk score received the platform pop-up, while some applicants above the 0.3 cutoff did not. This is because the platform pop-up assignment was based on initial risk scores, before the update. See the description of the 2021 intervention in section C.3 for details.

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.[6] 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.

D.1. Public Administrative Data

Public data comprises all the necessary inputs to compute the assignment. 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: 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

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

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.

2021 The risk classification web service stored the history of applications for all applicants.

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.

D.4. School Attribute Data

Tables I and III analyze the attributes of schools where students enroll. The variables we use are described in Neilson (2021). In particular, see Neilson (2021)'s school expenditure and value added estimation data supplements. A sketch of variables we use is as follows:

- Per teacher spending: The school's total wage bill for classroom teachers is divided by the number of classroom teachers.
- With copayment fee: Schools that charge an out-of-pocket fee beyond the government voucher are indicated with a 1 and schools with no additional fees are coded as a zero.
- School monthly fee (USD) is the average monthly out-of-pocket fee that families must pay at each school. This is the tuition net of the base voucher.
- Share of vulnerable students: The government of Chile identified students from disadvantaged families with a designation of "prioritario." This variable is share of students at a school that are "prioritario."
- Total enrollment per grade is the number of students enrolled by grade level. It is the cohort size.
- Per student spending: The school's total expenditures are divided by the number of students.
- Estimated value-added. This measure is estimated using student test scores in 4th grade and controls for measures of health at birth (such as birth weight, gestation), family demographics (parents' education, including mother's college entrance exams). Neilson (2021) shows that these measures are strongly correlated with alternate measures that control for baseline test scores, which can only be constructed for a few years of data. Because relatively few schools that enroll high school students (grades nine and up) also enroll the fourth graders whose scores form the basis for this measure, we used VA estimates only for students in grades eight and below.

In addition to the variables used in Neilson (2021), Table III also describes the effects of the warnings intervention on distance from home to school. We compute this as the Euclidean distance between the home and school, in kilometers.

E. Treatment First Stages

This section describes how warnings treatments of different types interact with each other. As stated in section III.B, 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 E.I 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 E.I 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.

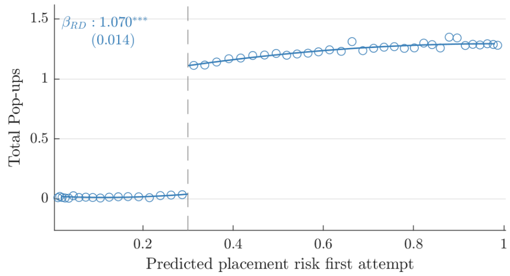
Figure E.II shows how treatment outcomes for the platform pop-up, the two SMS warnings, and the RCT intervention vary with respect to the initial risk score cutoff for pop-up assignment. These graphs can help us understand how treatments change across the cutoff in our evaluation of the main pop-up RD 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 E.III 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.

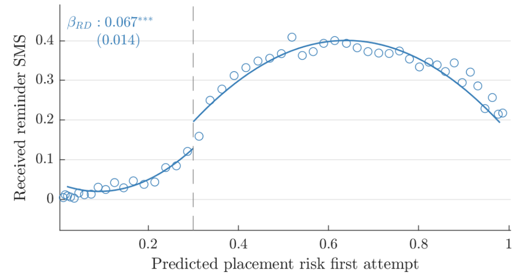
In the main text analysis of the WhatsApp intervention, we present both RCT comparisons of outcomes for treatment and control groups, and RD evaluations across the risk cutoff for both treatment and control. The detailed analysis of treatment receipt presented in Figure E.III shows that the RCT estimates 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 E.I
RD First Stages of 2018 and 2019 Interventions

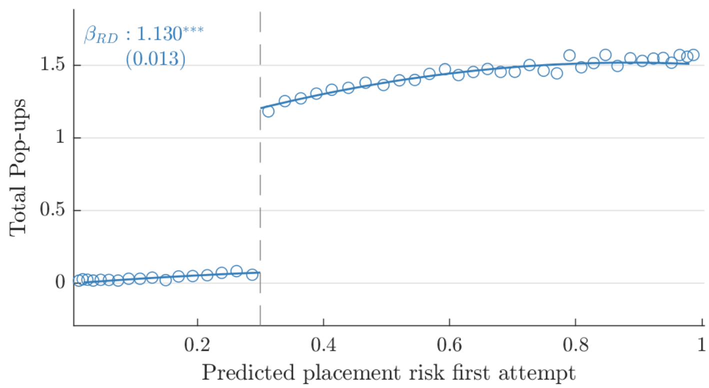
(a) First Stage Total Pop-Ups– 2018



(b) First Stage Reminder SMS– 2018

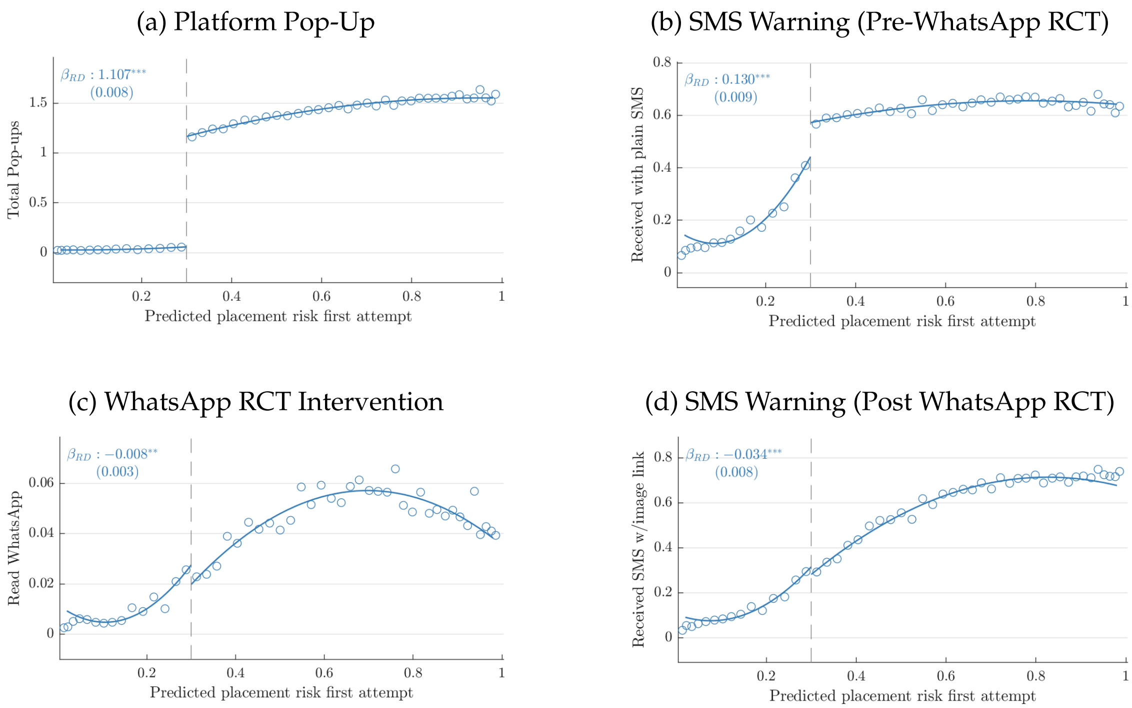


(c) First Stage Total Pop-Ups–2019



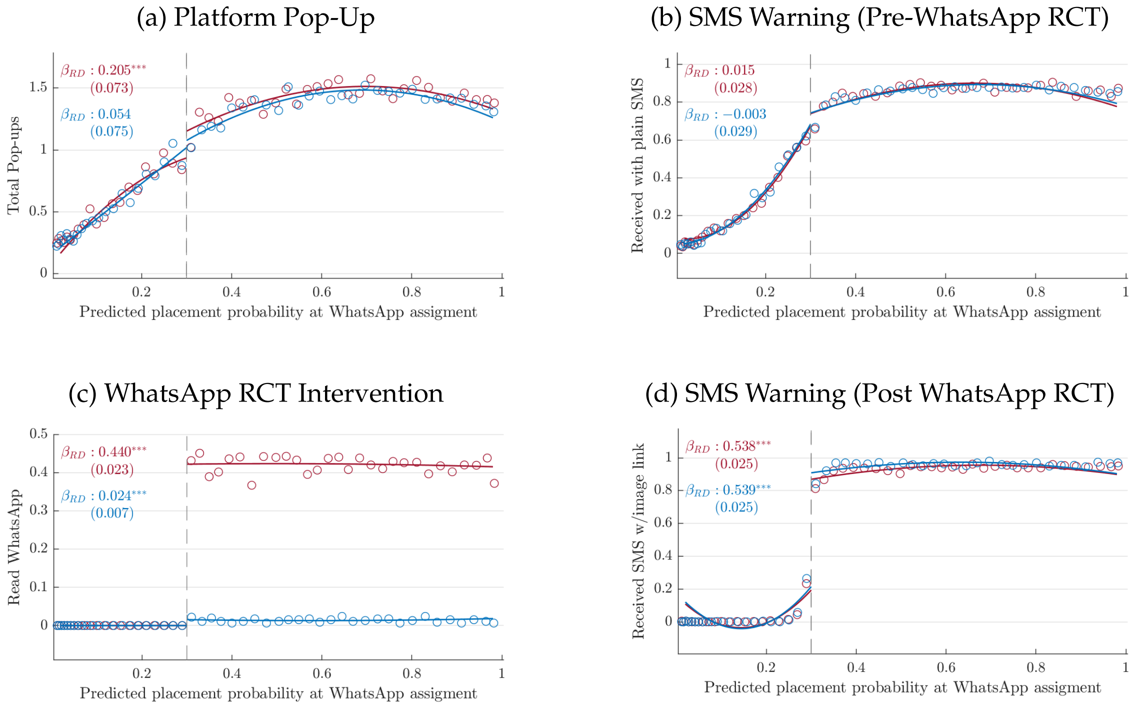
Notes. Warnings interventions in 2018 and 2019, by position relative to risk score prediction. Panel A: Count of platform pop-ups received in 2018 intervention. Panel B: count of reminder SMS messages received in 2018 intervention. Panel C: count of platform pop-ups received in 2019 intervention. 2019 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 III.B for intervention details.

Figure E.II
RD First Stages in 2020 Platform Popup Intervention



Notes. 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 III.B for intervention details.

Figure E.III
First Stages of 2020 WhatsApp Intervention



Notes. 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 III.B 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 (2021). 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 and 2021 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).

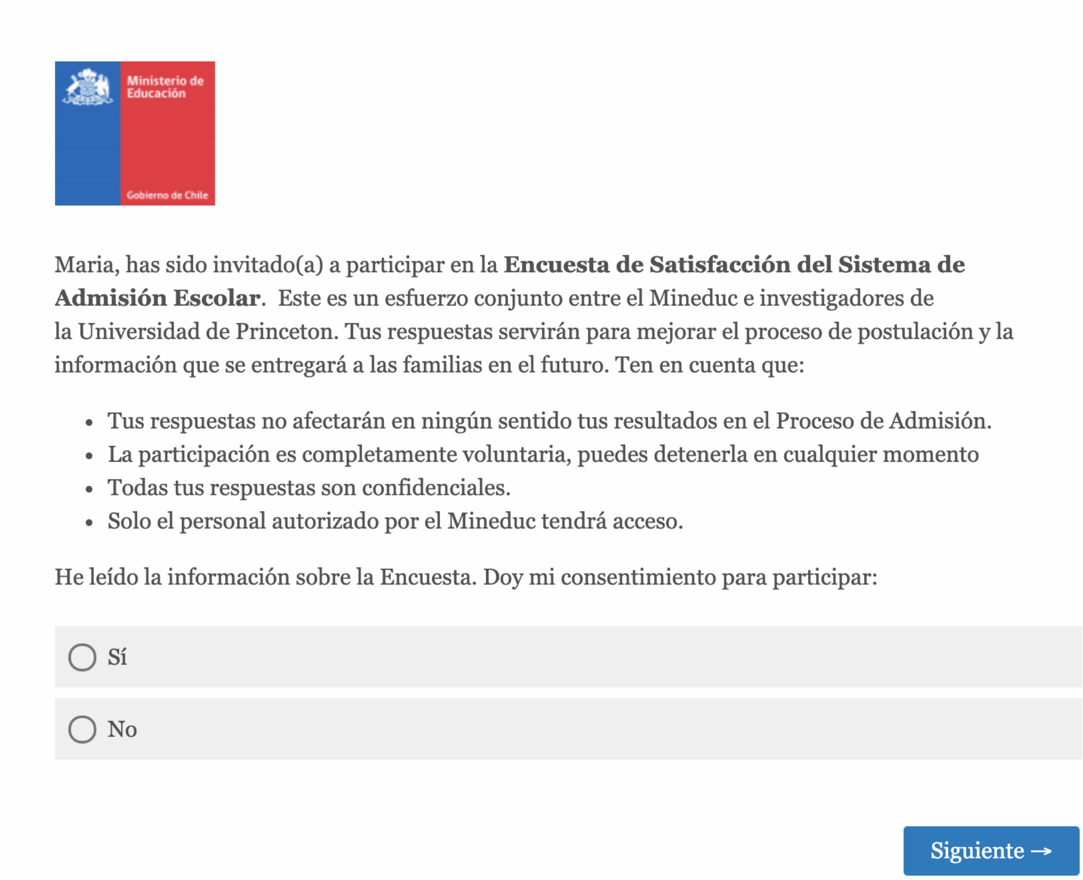
We compute the expected count of applicants, N, in each market as the number of students who participated in choice in the market in the previous year, or, in markets where there is no previous-year data, as the number of students who entered a new school at the start of the most recent school year (i.e. 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 G.I
2020 Survey Landing Page



Ministerio de Educación

Gobierno de Chile

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) [School not considered in application 1]
    (b) [School not considered in application 2]
    (c) [School not considered in application 3]
    (d) [School not considered in application 4]
    (e) [School 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. Additional School Quality Results

This appendix provides more detail on the effects of warnings on school quality reported in section V.E and Table III of the main text. The main finding is that the overall gain in school value added arises from both a shift towards oversubscribed schools, which are higher value added on average, and from shifts in value added within oversubscription status.

Online Appendix Figure H.I plots the distribution of school value added and per-teacher spending in schools that are oversubscribed and schools that are not. Because our goal is to understand how these measures vary by oversubscription status for a set of schools that a particular student might choose between, we restrict the sample schools offering pre-kindergarten in Santiago (the largest single "market" in our data).

On average, oversubscribed schools have higher value added and higher per-teacher spending than undersubscribed schools, but there is much dispersion within each category and the distributions overlap substantially. The implication is that the search treatment may in principle raise value added or teacher spending at the schools students attend by shifting students toward oversubscribed schools or by improving value added within oversubscription category.

Online Appendix Table H.I decomposes the overall gain in value added reported in Table III into within- and between- oversubscription status channels. Results are imprecise in some cases but nevertheless provide insight into the channels through which value added rises across the cutoff.

The first two rows repeat our main first stage (i.e., add any school) and value added results. The next two rows report that the warnings intervention pushed students to enroll in oversubscribed schools. This holds both overall and in the sample of students who enroll in schools where VA measures are available, though effects are larger in the full sample. Note that these effects on *enrollment* are not the same as results for *placement* in undersubscribed schools reported in Table V. This is as expected given the imperfect compliance with placed outcomes reported in Table I.

Rows five and six show the within vs. between decomposition results. $E[VA|type]$ is the mean value added for students given the type of school they attend, where type is either oversubscribed or undersubscribed. This rises across the cutoff because students are more likely to enroll in oversubscribed schools, where mean value added is higher. These between type gains account for about 20% (0.023/0.103) of the overall gains in value added we observe for compliers with the warnings treatment. $VA - E[VA|type]$ is the difference between the value added of the school where the student enrolls and the type specific mean. This rises across the cutoff because students attend higher value added schools within oversubscription type. Within-type shifts account for about 80% (0.080/0.103) of value added gains for compliers. The within and between effects mechanically add up to the full effect.

The final two rows report shifts in value added across the cutoff conditional on enrollment in either an over- or undersubscribed school. The goal of these specifications is to understand whether the within-type shifts come from over- or under-subscribed schools. We interpret the results with caution because, as documented in the upper rows of the table, the share of students enrolling in oversubscribed schools rises across the cutoff. The (imprecise) results from these specifications suggest that the within-type value added gains come from both over- and undersubscribed schools.

Gains from the shift towards oversubscribed schools and the shift within oversubscribed schools are consistent with the observations that a) most of the highest quality schools are obsersubscribed, b) all spots at these schools are by definition allocated in the centralized match, and c) changes to the application list induced by the warnings intervention raise the probability of being assigned to any school in the centralized match and also shift the distribution of schools to which students are assigned.

Gains from shifts within undersubscribed schools (where spots remain open after the match) suggest the possibility that, even absent capacity constraints, finding a good school may be easier within the centralized process than in the "scramble" that follows it. This pattern could arise if, for example, families searching for schools in the scramble feel pressure to accept the first offer they receive, because they are concerned the school will fill up. Understanding how search plays out in the scramble is a topic of possible interest for future work.

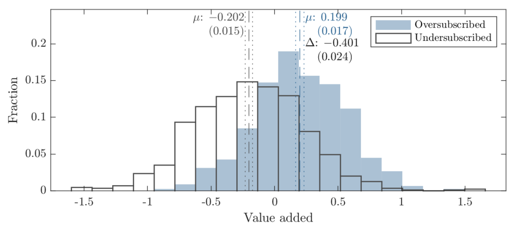
Table H.I
RD Estimates of Platform Pop-Up Effects

	(1)	(2)
	All	
		IV
Add any	0.216	
	(0.010)	
Value added (VA)	0.022	0.103
	(0.011)	(0.051)
Enrolled in oversubscribed ($type = over$)	0.039	0.178
	(0.010)	(0.048)
Enrolled in oversubscribed ($type = over$)$\vert$not missing VA	0.022	0.103
	(0.014)	(0.065)
$E[VA \vert type]$	0.005	0.023
	(0.003)	(0.014)
$VA - E[VA \vert type]$	0.017	0.080
	(0.011)	(0.050)
$(VA - E[VA \vert type]) \vert type = under$	0.029	0.131
	(0.020)	(0.091)
$(VA - E[VA \vert type]) \vert type = over$	0.013	0.063
	(0.012)	(0.059)
NL	10,782	10,782
NR	11,285	11,285

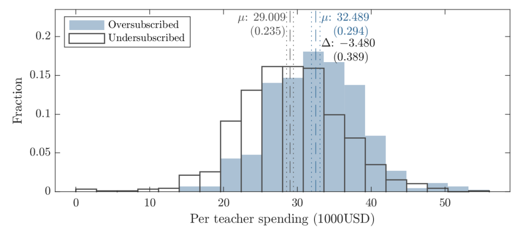
Notes. Local linear RD estimates of pop-up effects from warning pop-up on application platform. Computed using triangular kernel with bandwidth 0.1. Heteroskedasticity-robust nearest neighbor variance estimator with minimum of 3 neighbors reported in parentheses; computed as in Calonico, Cattaneo and Titiunik (2014). We report estimates in the pooled sample across years 2018-2020. IV (column 2) shows the instrumental variable specifications, where the endogenous regressor is the add any school indicator. “Add any” is the first stage indicator for adding at least one school to the choice application. “Value added (VA)” repeats main text results on value added at the enrolled school. “Enrolled in oversubscribed” is an indicator for enrolling in a school with a binding capacity constraint. “Enrolled in oversubscribed | not missing VA” restricts the sample to students enrolling in schools where value added measures are available. “$E[VA|type]$” is the mean value added for the enrolled school given over/underscription status, and “$VA - E[VA|type]$” is the deviation of value added at the enrolled school from the type-specific mean. “$VA - E[VA|type]|type = over/under$” additionally conditions on the sample on students enrolling in over- or undersubscribed schools, respectively.

Figure H.I
Distributions of School Value Added and Spending by Oversubscription Status

(a) Value Added



(b) Spending Per Teacher



Notes. Histograms of value added (panel A) and spending per teacher (panel B) for schools that are over- and undersubscribed. Means by over/undersubscription status and the difference between means reported in each panel and indicated by vertical lines. Sample is schools offering pre-Kindergarten in Santiago (the largest single market in our dataset).

I. RDs at Multiple Cutoffs from the 2017 Pilot

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

Table I.I
RD Estimates of Platform Pop-Up Effects - 2017

	(1)	(2)	(3)	(4)	(5)
Risky cutoff	Pooled		0.3	0.5	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. 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, Cattaneo and Titiunik (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.

J. 2020 WhatsApp Treatment Arms

The 2020 WhatsApp RCT tested the efficacy of different forms of the warnings intervention. We presented the personalized risk warning in three ways: 1) a text-only warning that nonplacement risk was high, 2) a visual risk display, with a red bar indicating high risk, and 3) a list of the schools the student had applied to displaying the count of applications and estimated places available at those schools. See Online Appendix Figure C.VII for images of each arm.

Our main analysis in Table IV pools across arms. Online Appendix Table J.I separates estimates by arm. All three treatments caused students to add schools to their application. Effects in the visual display and application list arms were 25-35% larger than for the text only arm. Making warnings more salient or informative improves performance, but most of the gains are from the 2.

Table J.I
RCT Estimates Split by WhatsApp Message

	(1)	(2)	(3)	(4)
	Personalized	Personalized - By image type		
	Pooled	Simple	Warning bar	Vacancy
Add any (clean)	0.033	0.028	0.035	0.036
	(0.002)	(0.003)	(0.004)	(0.004)
Add any (endline)	0.044	0.036	0.049	0.046
	(0.003)	(0.005)	(0.006)	(0.005)
N Treatment	8,995	3,009	2,971	3,015
N Control	8,975	8,975	8,975	8,975

Notes. Evaluation of 2020 WhatsApp RCT, splitting out by message type. Treatment arms are as described in Online Appendix J. Online Appendix Figure C.VII shows example images. Estimates are differences in the share of students adding any school to their baseline application between the treatment group and a control group that did not get any message. Column 1 contains pooled estimates of the treatments from columns 2-4. "Clean window" corresponds to outcomes measured 44 hrs after WhatsApp messages were sent. "Endline" outcomes were measured at the end of the application process (75 hrs after WhatsApp message). See section III.B for a description of the WhatsApp RCT.

K. New Haven Warnings Policies and Data

K.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 (DA-MTB), 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 K.I reports descriptive statistics for all students applying through the choice system in 2020. 58% of applicants apply to a major transition grade– either pre-kindergarten, 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.

K.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 K.I 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 "encouragement nudge" intervention.

Panel B of Figure K.I 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 K.IIa). Afterwards, their predicted admission chances were displayed to them as shown in Figure K.IIb. 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 K.IIc and K.IId).

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.[7]

Figure K.III 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.

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

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 K.I 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.[8] Columns 3 and 4 of Table K.I 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 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

[8] 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 exactly 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, Neilson and Zimmerman (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.

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 VIII. Column 5 of Table K.I displays descriptive statistics for this comparison sample. This group closely resembles the 2020 eligible sample on demographic characteristics and choice outcomes.

K.3. Intervention Results

Figure VIII displays our main results for the RD and DD analysis of the warnings intervention in 2020. We discuss these findings in section VII of the main text. Online Appendix Figure K.IV displays RD-DD graphs for balance and first stage outcomes, parallelling the main results in Figure VIII.

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

Panel A of Online Appendix Table K.II 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 nearly all above-threshold students received a warnings email, relatively few logged into the online simulator or ran a simulation. This suggests that the behavioral effects we observe come mostly from the warning and not from the simulator availability. This is consistent with the large effects in the Chilean setting, which did not include a simulator. The implication is not that risk simulators cannot form part of an effective intervention, but that effective interventions do not require simulation.

Panel C shows estimates of effects on different choice outcomes. The RD estimates indicate that crossing the threshold and receiving the warning causes 13.8 percent of applicants to add at least one school to their application. These are the compliers with the information treatment. Ex post realized application risk falls by 3.2 percentage points across the cutoff. This means that compliers with the policy reduce their application risk by 23.2 percentage points ($= 0.032/0.138$), or 42% of the below-threshold mean ex post risk of the initial application. 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 VIII that behavioral changes are larger for less risky students in the risky group, though estimates are imprecise. A possible explanation 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 additional inside options (Akbarpour et al., 2020).

Online Appendix Table K.III reports our findings from the encouragement nudge 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.

Table K.I
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 1^{st}	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

Notes. 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 or the mean difference in the last two rows of panel V.

Table K.II
RD and DD Estimates of Warnings Effects in New Haven

	RD		Diff. in Diff.	
Outcome	β	SE	β	SE
A. Demographics				
Female	0.113	(0.083)	0.001	(0.033)
African American	0.077	(0.072)	0.040	(0.031)
Hispanic	0.033	(0.082)	-0.003	(0.032)
White	-0.029	(0.073)	0.006	(0.025)
N		740		3918
B. Interaction with Simulator				
Warnings email	1.001	(0.010)		
Pr(Any login)	0.133	(0.074)		
Number of Logins	0.138	(0.090)		
Pr(Any sim. run)	0.068	(0.063)		
N		740		
C. Choice Outcomes				
Change length or school	0.146	(0.047)	0.042	(0.015)
Lengthen app.	0.138	(0.046)	0.053	(0.013)
Insert new school	0.053	(0.028)	0.012	(0.007)
Append new school	0.089	(0.038)	0.039	(0.011)
Change school	0.050	(0.028)	0.002	(0.009)
Shorten app.	0.004	(0.012)	-0.008	(0.005)
Diff. in realized risk	-0.032	(0.012)	-0.007	(0.004)
Diff. in simulated risk	-0.033	(0.016)	-0.014	(0.004)
Any realized risk reduction	0.156	(0.044)	0.036	(0.010)
Any simulated risk reduction	0.135	(0.045)	0.046	(0.011)
N		740		3918

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

Table K.III
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

Notes. 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.

Figure K.I
Email Communication with Parents

Office of Choice & Enrollment
54 Meadow Street, New Haven, CT 06519
NHPS School Choice Simulator Letter

NEW HAVEN
PUBLIC SCHOOLS OF CHOICE

Date: 2020/02/25

Dear Parent or Guardian of Jimmy Heckman,

Thank you for your participation in the New Haven Public Schools of Choice Program. You have successfully submitted an application for Jimmy!

Remember that you can make changes to your child's application until **March 2nd, 2020.**

We are contacting you because wanted to flag that the schools listed on the application you submitted for Jimmy have in the past have had many more applicants than slots. This means that the chances of being placed at those schools are often low. Our calculations suggests had you submitted this same application **last year**, there would have been **at least a 50% chance you would have not been placed at any of the choices listed on your application.**

This does not mean that you have to change your application or that you will not be placed anywhere this year. This message is meant only to be a timely alert. We advise you to apply to your favorite schools and rank them in the order that you like them. To increase your chances of being admitted to some school, what you may want to do is look into additional schools and add them to your application.

To learn more about your **chances of being placed in the schools on your application**, we strongly encourage you to use the **New Haven Public Schools Application Simulator**.

To use the Application Simulator:

1. Visit our simulator website.
(Please use a web browser different from Internet Explorer, such as Google Chrome or Microsoft Edge).
2. Log in using the **same** user name and password used in the NHPS School Choice website.

The **Application Simulator** can help you see what the chances of being placed in each school in your application would have been last year. You can also search, explore, and add new schools.

Note: The Application Simulator shows the chances of school placement based on data from last year, 2019. **The simulator does not show placement results for the 2020 School Choice Process**. The number of applications and available seats do not repeat exactly from year to year, so placement chances this year will differ somewhat from chances last year. Information from last year may still be a useful guide for families submitting choice applications.

For more information, please contact us:

- Email us at simulator@new-haven.k12.ct.us
- Visit the **Choice & Enrollment Office** at 54 Meadow Street, New Haven, CT 06519.
- Contact us by phone: **(475) 220-1438**.

Again, thank you for choosing New Haven Public Schools of Choice!

Sincerely,
Choice & Enrollment Office

(a) Risky Applicants

Office of Choice & Enrollment
54 Meadow Street, New Haven, CT 06519
NHPS School Choice Simulator Letter

NEW HAVEN
PUBLIC SCHOOLS OF CHOICE

Date: 2020/01/30

Dear Parent or Guardian of Jimmy Heckman,

Thank you for your participation in the New Haven Public Schools of Choice Program. You have successfully submitted an application for Jimmy!

To learn more about your **chances of being assigned to the schools on your application**, we strongly encourage you to use the **New Haven Public Schools Application Simulator**.

To use the Application Simulator:

1. Visit our simulator website.
2. Log in using the **same** user name and password used in the NHPS School Choice website.

The **Application Simulator** can help you discover the chances of being assigned to each school in your application and your chances of not being assigned to any school in your application. You can also search, explore, and add new schools.

Please note that the Application Simulator predicts the chances of school assignment **based on data from last year**. The results show the chance of school assignment if you had submitted an application last year. The simulator does not show placement results for the 2020 School Choice Process.

Remember that you can make changes to your child's application until **March 2nd, 2020.**

For more information, please contact us:

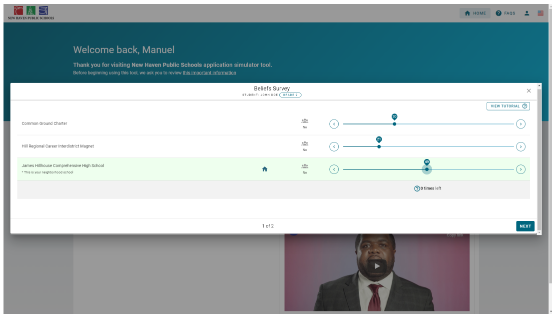
- Email us at simulator@new-haven.k12.ct.us
- Visit the **Choice & Enrollment Office** at 54 Meadow Street, New Haven, CT 06519.
- Contact us by phone: **(475) 220-1430** or **(475) 220-1431**. *Please note phone lines are extremely busy during this time of the year.*

Again, thank you for choosing New Haven Public Schools of Choice!

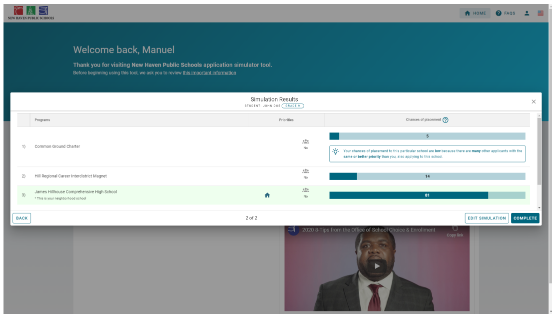
Sincerely,
Choice & Enrollment Office

(b) Non-risky Applicants

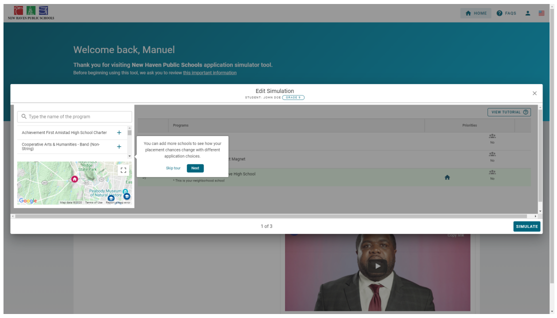
Figure K.II
New Haven Simulator



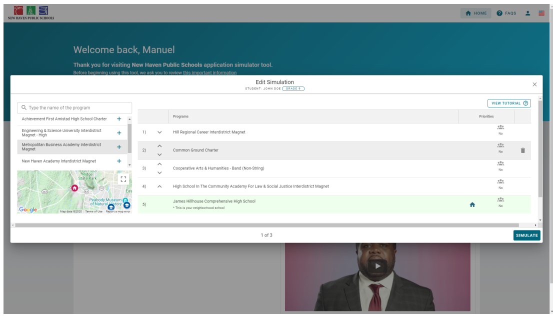
(a) Beliefs Survey Page



(b) Predicted Admission Chances

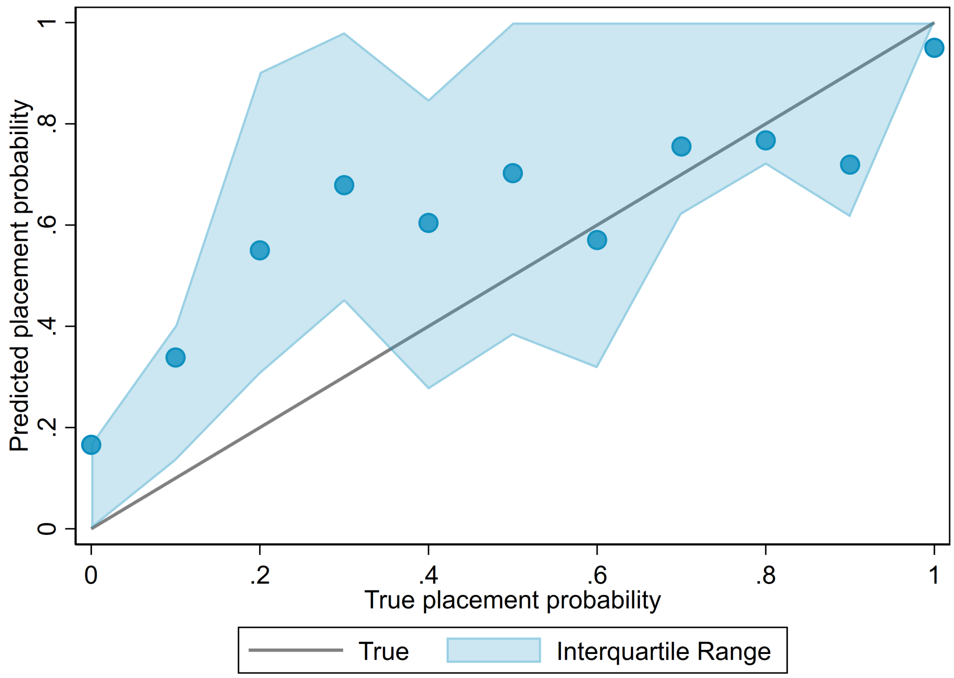


(c) Simulating Portfolio Changes



(d) New Predicted Admission Chances

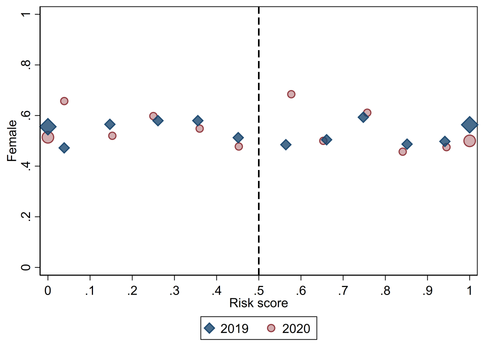
Figure K.III
Observed vs. Predicted Placement Probability in New Haven



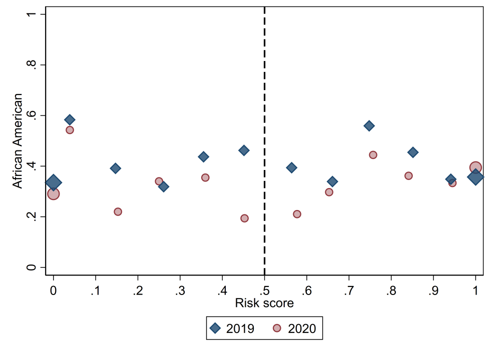
Notes. Distribution of predicted placement probability by value of ex post observed placement probability. For each bin of observed placement, we display the mean and IQR of predicted values. 45-degree line displayed for reference. See section K for details.

Figure K.IV
Treatment Balance and First Stage Outcomes

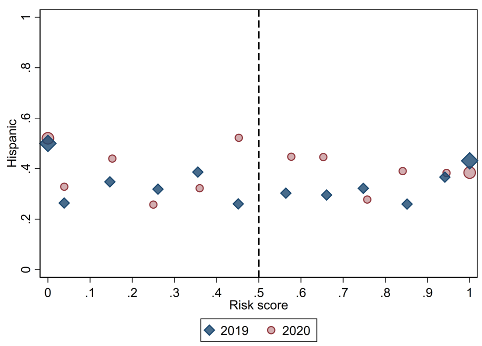
(a) Female



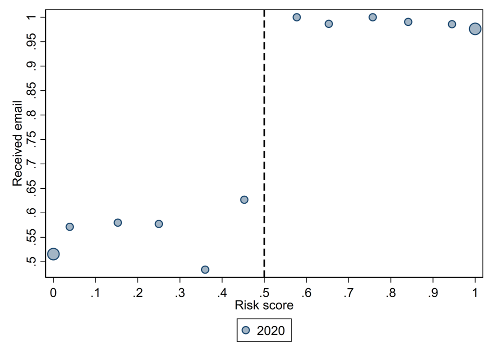
(b) Black



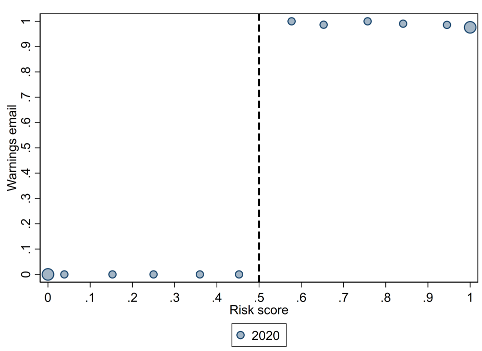
(c) Hispanic



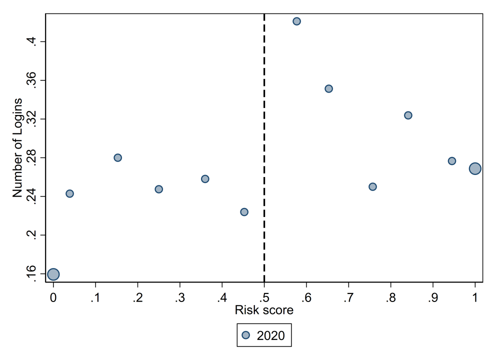
(d) Received Email



(e) Received Warnings Email



(f) Number of Logins



Notes. 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 K for details.

References

Agarwal, Nikhil and Paulo Somaini, "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism," *Econometrica*, 86 (2018), 391–444.

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).

Ashlagi, Itai and Afshin Nikzad, "What matters in school choice tie-breaking? How competition guides design," *Journal of Economic Theory*, 190 (2020), 105–120.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 82 (2014), 2295–2326.

Gobierno de Chile Ministerio de Educación, *El primer gran debate de la Reforma Educacional: Ley de Inclusión Escolar* September 2017.

Kapor, Adam J, Christopher A Neilson, and Seth D Zimmerman, "Heterogeneous Beliefs and School Choice Mechanisms," *American Economic Review*, 110 (2020), 1274–1315.

Neilson, Christopher, "Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students," (2021).

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.