NBER WORKING PAPER SERIES

SEARCH AND BIASED BELIEFS IN EDUCATION MARKETS

Patrick Agte
Claudia Allende
Adam Kapor
Christopher Neilson
Fernando Ochoa

Working Paper 32670
http://www.nber.org/papers/w32670

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2024

This project was funded in part by the National Science Foundation Grant Award #1919504, titled "Biased Beliefs and Search in Education Markets." The project was implemented in collaboration with the Government of Chile and the tech NGO ConsiliumBots. Christopher Neilson, the founder and CEO of ConsiliumBots. He is also the project lead for a publicly procured contract between ConsiliumBots and the Government of Chile to run the national student match. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Search and Biased Beliefs in Education Markets
Patrick Agte, Claudia Allende, Adam Kapor, Christopher Neilson, and Fernando Ochoa
NBER Working Paper No. 32670
July 2024
JEL No. I2,L0

ABSTRACT

This paper asks how search costs, limited awareness of schools, misperceptions of schools' attributes, and inaccurate beliefs over unknown schools affect families' search and application decisions in Chile's nation wide school choice process. We combine novel data on search activity with a panel of household surveys, administrative application data, randomized information experiments, and a model of demand and sequential search with subjective beliefs. Descriptively, households hold inaccurate beliefs and misperceptions along multiple dimensions which distort the perceived returns to search. Most importantly, they do not know all schools, and misperceive quality ratings of the schools they know and like. Improving the search technology would raise households' search effort and welfare. Correcting misperceptions about known schools' observables would cause students to match to schools with higher quality, equal to what can be achieved under a full-information benchmark. Models with out misperceptions would incorrectly predict quality reductions.

Patrick Agte
Yale University
Department of Economics
28 Hillhouse Avenue
New Haven, CT 06511
patrick.agte@gmail.com

Claudia Allende
Graduate School of Business
Stanford University
655 Knight Way
Stanford, CA 94305
and NBER
callende@stanford.edu

Adam Kapor
Department of Economics
Princeton University
280 Julis Romo Rabinowitz Building
Princeton, NJ 08544
and NBER
akapor@princeton.edu

Christopher Neilson
Yale University
87 Trumbull Street
New Haven, CT 06511
and NBER
christopher.neilson@yale.edu

Fernando Ochoa
19 West 4th Street
6th floor
New York, NY 10012
foa227@nyu.edu

A randomized controlled trials registry entry is available at AEARCTR-0007998

1. INTRODUCTION

Search and information frictions are a popular explanation for consumer demand for low-quality or expensive goods when cheaper, higher-quality alternatives exist, especially in markets with complex goods and/or many alternatives.[1] A researcher who observes consumers choosing a low-quality option may estimate search costs and conclude that they are high. These estimated costs, opaque to researchers, may motivate a market designer to pursue ad-hoc fixes or to reduce the scope of choice.[2] We study an alternative—or complementary—explanation: misperceptions and biases may distort the perceived returns to search. If people overestimate the quality of the alternatives they know, or underestimate the quality of those they have not yet looked into, then only a modest cost, reflecting further frictions, is needed to rationalize low levels of search. Moreover, a designer may help consumers by providing information.

We empirically investigate misperceptions, beliefs, and search effort in a complex, high-stakes decision: parents' choice of schools for their children in the context of Chile's nationwide school choice process. We ask: how do parents' (limited) awareness of schools, (inaccurate) perceptions of their characteristics, and (biased) beliefs over their distribution interact with their preferences and search costs to distort their information-acquisition efforts, application decisions, and school assignments?

To address this question, we construct a quantitative model of demand for schools with heterogeneous preferences and search costs, extending prior models by endogenizing search and incorporating relevant biases and misperceptions. We estimate heterogeneous misperceptions of schools' prices, government-provided quality scores, admissions chances and unobserved match qualities, as well as heterogeneous and potentially biased beliefs about the distribution of these four objects over unknown schools. Parents are initially endowed with limited infor-

[1]See e.g. Sorensen (2000), Handel and Kolstad (2015b), Agarwal et al. (2020), Bhattacharya et al. (2024), Ajayi and Sidibe (2020). Examples include markets for health plans, mortgages, and education.

[2]Proposed responses to high search and information-acquisition costs include simplifying choice sets (Brown and Jeon, 2023, Abaluck and Gruber, 2016), providing default options (Handel and Kolstad, 2015a), and encouraging delegation to intermediaries (Boehm, 2023).

mation, and may discover new schools, or obtain additional information about "known" schools, by paying a cost. Once this cost exceeds the subjective expected benefit, they stop searching and submit an application to an assignment mechanism. To estimate the model, we collect novel data on parents' search activity on a "school explorer" platform that we designed and implemented in collaboration with the Chilean Ministry of Education, and conduct field experiments to generate exogenous variation in perceptions and beliefs. We measure parents' beliefs, perceptions, preferences, and awareness of schools, before and after our interventions and their search decisions, via a novel panel of household surveys which we link to parents' explorer activity and to administrative application data.

We find that search costs interact with misperceptions and biased beliefs. Holding search costs and initial awareness of schools fixed, providing correct perceptions and rational expectations would raise aggregate welfare by the equivalent of a 0.25-kilometer decrease in distance, achieving 60% of the gains of a full-information zero-search-cost benchmark. Alternatively, holding misperceptions fixed, one would have to reduce estimated costs by 95% to achieve these gains. Decomposing biases and misperceptions, we find that the most important are misperceptions about observable characteristics–prices and quality scores—of known schools. Correcting these alone would achieve 45% of the welfare gains of full information. Moreover, doing so would achieve all of the gains from a full-information benchmark in the quality and value-added of assigned schools, and would entirely close the existing gap in these measures between students from low- and high-socioeconomic status (SES).[3] Had we estimated misspecified models without misperceptions and subjective beliefs, we would have obtained inaccurate estimates of search costs and reversed the sign of changes in school quality under counterfactuals.

We establish these findings via the following steps. First, we present descriptive analyses using a baseline survey, conducted several months before applications are

[3]In the absence of information interventions, low-SES students would be assigned to schools with an average quality score of 2.969, while high-SES students' assigned schools' average quality would be 3.093. Under this counterfactual, these numbers would increase to 3.193 and 3.181, representing 0.29 and 0.11 standard deviations respectively.

due, and subsequent "midline" and "endline" surveys, as well as administrative application data. We find that households know by name fewer than 50% of random nearby schools asked about at baseline, and know fewer than 20% of schools well. Thus, there is a role for search effort. Parents hold noisy but not systematically pessimistic beliefs about the distribution of academic quality scores of schools not yet investigated. However, they also perceive the characteristics of "known" schools with noise, leading them to systematically overestimate the quality of their first-choice school. Preference rankings over schools in surveys and administrative data are sensitive to perceived quality score, indicating that households value this characteristic. Moreover, conditional on elicited perceptions of quality, parents' preference rankings do not depend on the schools' "true" quality. Examining within-household updates across surveys, we also find that explorer search activity predicts knowledge of schools and accurate perceptions of their characteristics.

Second, we analyze our experiments. We embedded two randomized information interventions within the school explorer platform, which we shared with parents just after our baseline survey. The first treatment arm provided personalized information about the joint distribution of nearby schools' prices and quality scores, but did not provide information about specific schools. A second arm highlighted nearby low-cost high-quality schools, making them salient, in addition to this distributional information. These treatments caused changes in beliefs, search activity, knowledge of schools, and placements, with the effects driven by households with college-educated mothers, a proxy for high SES.

In addition, we used administrative application data to conduct a third randomized intervention that provided tailored feedback on the initial application parents submitted, targeting parents' misperceptions about known schools. This "feedback" intervention took place roughly a week before the final deadline, and provided personalized information about the price, quality, and admissions chance of schools to which parents had provisionally applied.[4] We find that this intervention corrected perceptions about school characteristics and had large impacts on the application decisions of low-SES parents.

[4]In addition, we use the platform to warn students with low chances of receiving a placement; see Arteaga et al. (2022).

Finally, we use the data, descriptive analyses, and experiments to inform a model of search and school choice. Chile uses a student-proposing deferred acceptance algorithm with lottery tiebreakers and no constraint on list length, following best practices in market design (Abdulkadiroglu et al., 2005, Correa et al., 2019). This mechanism produces rich preference data, and gives parents a simple dominant strategy at the time that applications are due: rank schools in order of their expected payoffs given the parents' information. However, the presence of search costs makes beliefs about admissions chances relevant to parents' search decisions (Arteaga et al., 2022). Our model incorporates admissions uncertainty and the need to pick a portfolio, which make the search decision difficult relative to standard settings (McCall, 1970, Weitzman, 1978).

Our empirical strategy exploits multiple measurements as well as shifters of information that are excluded from payoffs. We observe survey preferences at baseline and administrative rank-order lists at two times: just before the feedback intervention, and at the final deadline. We obtain measures of awareness, beliefs over unknown schools, and perceptions of known schools from up to three surveys per household. In the model, in the event parents know a school by name but not well, its subjective expected value is given by a "low-information" potential payoff, reflecting their best estimate of its expected utility given their beliefs. If they gather more information, they rank it according to a "high-information" potential payoff instead. Parents' search decisions, our randomly-assigned treatments, and passive learning over time provide variation in information across households, and within households between survey waves and measurements, that is excluded from these potential subjective payoffs.

We estimate the model in two steps. First, we estimate the parameters governing payoffs, awareness of schools, and perceptions of schools' characteristics via MCMC, extending Bayesian demand-estimation methods (McCulloch and Rossi, 1994), in school choice settings (Agarwal and Somaini, 2016, 2018, Kapor et al., 2020) with limited availability (Kapor et al., 2022) to our panel-data context.[5] Second, we estimate parents' beliefs over unknown schools, then impose optimality

[5]For choice with limited availability and/or awareness, see also Agarwal and Somaini (2022) and He et al. (2021). We differ by observing rank-order lists and measures of parents' awareness of schools.

of their search decisions to recover search costs and simulate counterfactuals. In counterfactuals we decompose the effects of all the misperceptions and biases that we model, and compare these to reductions in search costs.

Our results provide a unified treatment of the impacts of providing information about admissions chances (Arteaga et al., 2022, Ajayi et al., 2020, Gurantz et al., 2021, Hoxby et al., 2013, Hoxby and Turner, 2015, Luflade, 2017) and those of providing information about characteristics of schools (Hastings and Weinstein, 2008, Mizala and Urquiola, 2013, Corcoran et al., 2018, Cohodes et al., 2022, Andrabi et al., 2017, Allende et al., 2019, Bergman et al., 2020). While these literatures demonstrate that information can improve placement rates and outcomes, the estimated impacts depend on both the importance of the friction being targeted and the effectiveness or "takeup" of the intervention, and give little guidance about extrapolation. Our model and direct data on belief updating let us distinguish these factors and make comparisons across intervention types. While our results are consistent with large effects of targeted admissions-information interventions (Arteaga et al., 2022), we find that fully correcting parents' misperceptions of schools' price and quality would have much larger effects than fully correcting beliefs about admissions chances. Although high-SES parents responded more to our initial "search" interventions than did low-SES parents, full-takeup information interventions would have larger impacts on low-SES households.

Our paper is related to contemporaneous work estimating households' perceptions of school quality (Corradini, 2024) and eliciting perceptions of schools' characteristics from household surveys (Corradini and Idoux, 2023). We differ by modeling search and information acquisition. As our setting differs from theirs, we do not investigate the role of race.

We contribute to the empirical literature on consumer search (Sorensen, 2000, De los Santos et al., 2012, De Los Santos et al., 2017, Dinerstein et al., 2018, Hodgson and Lewis, 2023, Moraga-González et al., 2023, Agarwal et al., 2020). Our study is also closely related to an experimental literature studying the effects of information frictions on search behavior (Cortés et al., 2023, Bandiera et al., 2023, Belot et al., 2019, Carranza et al., 2022). We provide a novel model, dataset, design, and estimation strategy.

Our descriptive analysis of information frictions in a high-stakes setting parallels work on these topics in health economics (e.g. Handel and Kolstad (2015b)). As biases about the return to search are important in our setting, we do not pursue a rational-inattention approach (Brown and Jeon, 2023). Our model of information acquisition is closely tied to the institutional details of our setting and the features of the search technology we provided.

A limitation of this paper is that we estimate demand, and conduct "single-agent" counterfactuals, holding schools' prices, quality scores, match qualities, and admissions chances fixed. In our setting 41% of first-choice schools had excess capacity. In 2021, entry was heavily regulated and prices were fixed,[6] but other aspects of schools such as quality "markdowns" may adjust more easily. Understanding search and demand is a needed input for further research on equilibrium outcomes in this market.

The remainder of the paper proceeds as follows. Section 2 presents a motivating example. Section 3 describes the setting. Section 4 provides descriptive analysis. Section 5 describes and evaluates the experiments. Section 6 presents the model, section 7 describes estimation, and section 8 presents results. Section 9 concludes.

2. MOTIVATING EXAMPLE

The following stylized example illustrates the comparative statics that motivate our data collection and research design. Consider a household participating in a student-optimal stable matching (SOSM) mechanism with independent tie-breaking lotteries. The household knows a single school. If placed in it, the household will receive a payoff of $u_1 > 0$. However, the school will reject the household with probability $r_1 \in [0,1]$. There is also an outside option offering a sure payoff of 0. The household may choose to exert search effort to discover one additional school before it submits its rank-order list. If so, this school will have payoff and rejection chance $(u,r) \sim f(\cdot)$, for some distribution f over $\mathbb{R} \times [0,1]$.

To define the gain from search, let $\kappa = \{(u_1,r_1),\ldots,(u_N,r_N)\}$ denote the set of schools that the household knows and prefers to its outside option. The household will receive a payoff $u_j > 0$ in the event of a placement at j, but will be

[6]See our supplementary material S.1

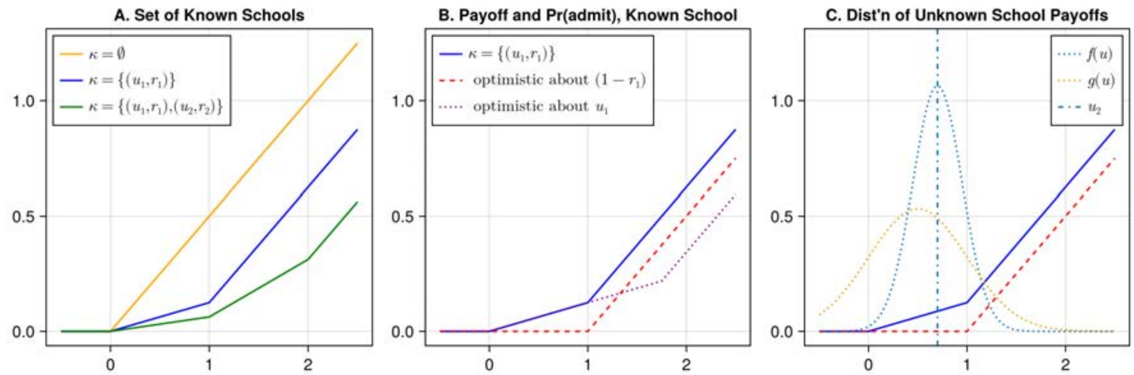


FIGURE 1.—Return to Search. Note: X-axis in all panels: utility u of new school. Panel (A): Gain from adding a new school, $U(\kappa \cup (u, 0)) - U(\kappa)$. We take $u_1 = 1$, $r_1 = 0.25$, $u_2 = 2$, $r_2 = 0.5$. Panel (B): Comparative statics of the gain from adding a new school, $U(\{(u_1, r_1), (u, 0)\}) - U(\{(u_1, r_1)\})$, with respect to u_1 and r_1. Panel (C): Comparative statics of the expected gain from adding a new school with respect to the distribution of u.

rejected by j with probability $r_j \in [0, 1]$, independently across schools. When the SOSM is used, it is well known that truthful reports are optimal. Supposing $u_{j_1} > u_{j_2} > \ldots > u_{j_N}$, the expected utility of the optimal portfolio is given by $U(\kappa) = \sum_{n=1}^{N} \left(\prod_{\ell=1}^{n-1} r_{j_\ell} \right) (1 - r_{j_n}) u_{j_n}$.

Panel A of Figure 1 plots the gain from search, $U(\kappa \cup \{(u, 0)\}) - U(\kappa)$, as a function of the utility u of the newly discovered school, for the case a single school is known (blue line), and for other sets κ. This gain from search is a piecewise-linear increasing convex function of u that is weakly decreasing in κ. To understand the shape, observe that when $\kappa = \{(u_1, r_1)\}$, the new school will be ranked first if $u > u_1$. In contrast, if $u_1 > u$, then the new school will be irrelevant unless the household is rejected by school 1, an event which occurs with probability r_1. Similarly, if $0 > u$ as well, the new school will never be relevant.

Panel B of Figure 1 illustrates the effects of misperceptions of school 1's characteristics. When the household overestimates the payoff u_1 (dashed line) or underestimates the rejection chance r_1 (dotted line), the perceived gain from search will be lower than the true value, which may cause reduced search effort.

Panel C of Figure 1 illustrates comparative statics with respect to distributions over the new-school payoff u. Distribution $f(u)$ has a higher mean but lower variance than $g(u)$. Whether the expected gain from search is higher under f or g depends on the set of known options. When the household perceives $r_1 = 0.25$ (blue line) the expected gain is slightly higher under f. However, if the (subjective) rejection chance is lower (dashed red line), then only the right tail is relevant, and the expected return will be higher under g. Thus the sign of the effect of shifting households' beliefs from g to f may depend on the set of "known" schools and their perceived payoffs and admission chances, motivating an empirical analysis.

Finally, Panel C also illustrates that access to a more informative signal and/or the ability to learn more about "known" schools will generally raise returns to search. To see this, compare a point mass at $u_2 = 0.7$ (vertical line) with a mean-preserving spread about u_2, coincidentally given by f. Because the gain from search is convex in u, the household would prefer to replace a point mass at u_2 with a draw from f. This may be interpreted as obtaining a more informative signal.[7]

3. EMPIRICAL SETTING

Our study takes place within the Chilean centralized School Admission System (SAE). Before presenting our interventions, we first discuss the setting and data.

3.1. *School Choice in Chile*

The SAE assigns applicants to schools based on a student-proposing deferred acceptance algorithm (Correa et al., 2019). The system accounts for 89% of primary school matriculation in the country, including almost all public schools and private schools that accept school vouchers. Seats at oversubscribed schools are rationed through quotas, coarse priorities, and lottery-based tiebreakers. Therefore, placement probabilities are student-school specific. Parents form school portfolios and submit a rank-ordered-list (ROLs) through a centralized platform. There are no restrictions on ROL length, which makes the mechanism strategy-proof.[8] The main

[7]One may interpret u_1 and u_2 as posterior means given the household's current information about schools 1 and 2, and f as a distribution over posterior means of payoffs that is induced by some signal structure.

[8]The mechanism is strategy-proof for single applicants. Parents submitting applications for multiple siblings in the same year may face strategic considerations. We abstract from this issue. Our main sample includes children with older siblings, but experimental results are similar in a subsample without such households. In cases of twins, we choose one twin.

application round starts in early August. Applications can be submitted and edited for roughly one month (Arteaga et al., 2022).[9]

We focus on parents applying to entry grades (pre-K, kindergarten, and first grade). In 2021, a total of $207,578$ students applied to $16,421$ programs in entry grades, representing 45% of all applicants and 35% of the total seats. Applicants to entry grades tend to have higher placement chances at their first-choice schools than other applicants. In the main application period of 2021, 65% of entry grade applicants were assigned to their first preference, and 92% were assigned to a school on their ROL, compared to 48% and 93% for non-entry grades. Furthermore, entry grade applicants apply to schools that are geographically closer. The median distance to the first-preference school is 0.89 km, whereas for non-entry grades, it is 1.01 km.

Efforts to provide accurate and easily accessible information about these schools predate our study. Since its beginning in 2016, the SAE application platform has provided a School Showcase (Vitrina SAE) website that allows parents to search for schools by name (Correa et al., 2019). This website provides information on searched schools' available seats, their prices, and their academic quality rating according to Chile's Education Quality Agency, a widely publicized measure that we take as our main measure of quality. [10]

Applicants list few schools on their ROL (Arteaga et al., 2022). In 2021, entry grade applicants listed an average of 3 different schools, despite having an average of 13 schools with available seats within 2km from their home. Moreover, omissions include many schools that may be desirable for parents. Of these 13, an

[9]We refer the reader to Correa et al. (2019) for a detailed description of the SAE mechanism, and to Arteaga et al. (2022) for a comprehensive description of the SAE stages and policy outcomes. A brief summary of the system can be found in Supplementary Material S.1.

[10]The Education Quality Agency (Agencia de Calidad de la Educación) is the main rating agency. It classifies schools into four categories (high, medium, medium-low, and insufficient performance) based on the distribution of students in learning levels, indicators of personal development, and results from the national SIMCE exam, adjusted for student characteristics at the school level. The SAE platform makes parents familiar with the quality measure as it is relevant for default assignment. If a student is not assigned to a school on their ROL, they are assigned to the closest school with available seats that is not in the "insufficient" category. Nationally, 15% of schools belong to the high category, 55% to the medium category, 24% to the medium-low category, and 6% to the insufficient category. Quality ratings are correlated with value-added measures, and survey data suggests that parents consider them as important factors when choosing schools. See supplementary material for details.

average of 6 are free for the student and have a high (4) or medium (3) academic quality rating. Nevertheless, applicants only apply to 33% of them on average. Arteaga et al. (2022) argue that short (suboptimal) ROLs are consistent with costly search, and that welfare stakes are large.

For these reasons, in 2021 the Ministry of Education launched a new school explorer platform (officially, *Más Información Mejor Educación* (MIME)) to help parents search for schools. The explorer was developed by an EdTech NGO and made available to the participants through the government website. We collaborated in its development and used it to host our interventions.

The school explorer platform aggregated all public information on all the schools in the country. Two features were key for our study. First, it allowed parents to search for schools on a map rather than by name. This allowed parents to discover and compare nearby schools that they had not been aware of, or whose names they had not known. Second, it used data on students' location and characteristics to provide individualized information.

The school explorer highlighted four key pieces of information for each school: (1) the distance to the student's home, (2) the out-of-pocket monthly fee which varied with the parents' socioeconomic status, (3) the quality category, and (4) a personalized predicted admission probability based on past-year data and the student's priority category.

The explorer allowed parents to perform two main activities. Clicking on a school location on the map ("school pin click") provided these four pieces of information and a link to a school profile. A second click ("school profile click") opened a detailed view of the school with photographs, information on the school's leadership and philosophy, and other materials.[11]

3.2. Study Design

Guided by the comparative statics described in section 2, we conducted multiple survey rounds and two field experiments in partnership with an ed-tech NGO and the Ministry of Education of Chile. Figure 2 provides an overview of the design and timing of the study. In the top section of the figure (in green), we summarize

[11]See supplementary material for screenshots and additional details on the explorer platform.

the data collection activities. In the middle of the figure (in blue), we present the evolution of three relevant model objects: the set of schools that the agent knows and may apply to at time t, the perception of the characteristics of those schools, and the beliefs about the characteristics of the schools that the agent does not know yet. In blue, we also describe potential actions the agent may take: searching for schools, and submitting an application while the platform is open. Finally, in the bottom section of the figure (in black), we describe the two interventions we implemented to generate exogenous variation in beliefs and knowledge levels.

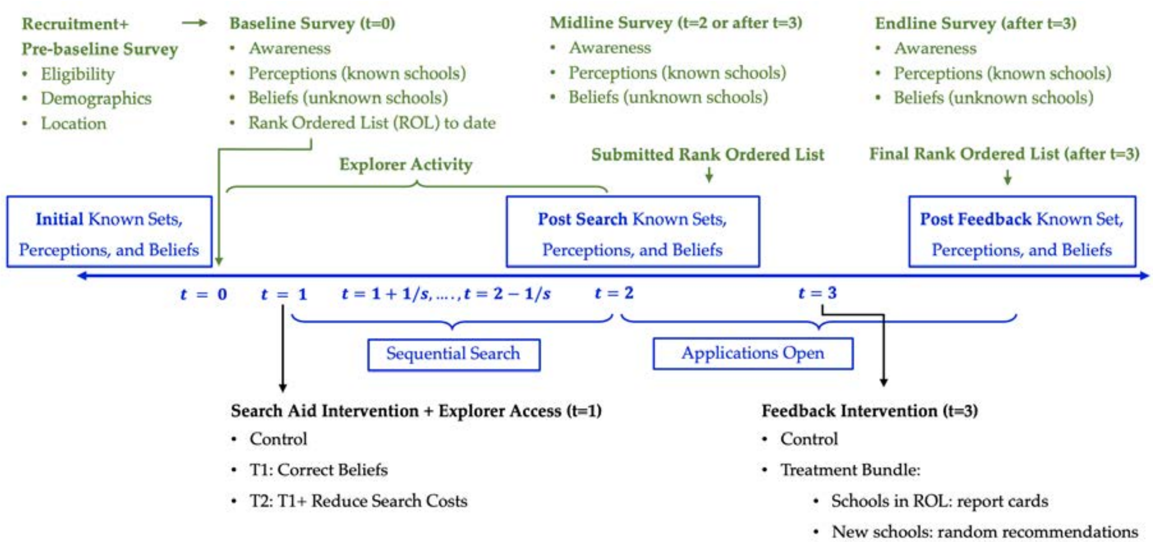


FIGURE 2.—Timeline for the Model, School Choice Process, Data Collection, and Interventions

Recruitment and Collaborations: The recruitment for the study was implemented through the Ministry of Education. Between May 25 and July 2, 2021, the government sent an email to potentially eligible parents through childcare-center principals across Chile that allowed them to sign up for the information program. As described in Figure 2, we implemented a registration form to identify eligible households and collect location and demographic information. To be eligible for the study, a parent would need to have a child applying to an entry grade through SAE for the first time.[12]

[12]In Chile pre-kindergarten and kindergarten are not mandatory to apply for first grade. Despite this fact, most schools in the country offer both pre-kindergarten and kindergarten.

A total of 3,948 parents who signed up for the study and completed the baseline survey were randomly assigned to one of three treatment arms. In our main analysis, we exclude 837 parents who never opened the school explorer platform, leaving us with a final sample of 3,111 parents.[13] Table A.VI shows the comparison between the universe of applicants applying for an entry grade in Chile (Column 1) and the experimental sample (Column 2). We find that parents who chose to participate in the study tend to be less likely to be eligible for the school voucher program, have longer rank-ordered lists, and are more likely to enroll in high quality schools.

3.3. *Data*

Figure 2 shows in green the primary sources of data that we use in this study.

Administrative data: We obtain information on school applications and school characteristics from administrative data provided by the Ministry of Education.

- **Application and Enrollment Data:** We obtained data on each application list submitted by parents during the application process in 2021. As described in Figure 2, families may submit an application once the application process starts and can change their submitted list of schools while the platform is open at no cost. The data provided contains the initial rank-ordered list of schools, the sequence and timing of any updates to this list, and the final assignment and enrollment outcomes. In addition, we use administrative data from the 2020 application process to construct a predicted (non-)assignment probability at each school for each applicant, which we provide to applicants as discussed below.
- **School Characteristics:** This data contains relevant information on schools, including their location, monthly school fees, and Education Quality Agency quality category.[14] We classified good inexpensive schools—those that cost less than 50k CLP per month and were of either high (4) or medium (3)

[13]The share of parents who did not open the school explorer platform is 22% in the control group, 21% in the treatment 1 group, and 20% in the treatment 2 group. Treatment group-specific information was only provided within the school exlorer platform.

[14]Around half of our sample households are eligible for school vouchers as part of the *Subvención Escolar Preferencial* (SEP) program and thus do not have to pay any fees for most schools.

performance—as "*highlight-worthy*". In our analysis, we further report results based on the value added measure used in Neilson (2021).

Explorer: We tracked all activities the parents performed in the explorer platform, including when and for which schools they opened the school pin and profile.

Surveys: We conducted four surveys to collect information on household characteristics, knowledge, misperceptions, beliefs, and preferences for our sample parents at key points in the search and application process.

- **Registration Form:** The initial registration form was used to recruit participants for the study and obtain information on demographics, family structure, and location.
- **Baseline Survey:** We implemented this online survey three months before families had to apply for schools. We consider it to measure objects at time $t = 0$ in our model. It was sent to eligible parents and collected a detailed list of schools that the parents knew,[15] their perceptions about the price and quality of those schools, measurements of subjective admissions chances, and a detailed elicitation of beliefs about the distribution of school characteristics in their neighborhood.[16] We also included questions about search behavior and the rank-ordered list of schools to which parents were planning to apply.
- **Midline Survey:** This survey was done over the phone in the final weeks of the application period. It collected a second measurement of parents' level of knowledge of schools, a second elicitation of beliefs about the distribution of school characteristics in the neighborhood, and measures of perceived price, quality score, and admissions chances at a set of schools partially overlapping with those asked about at baseline.
- **Endline Survey:** This online survey was sent to the universe of parents who submitted an application through the SAE system in 2021 and was part of a broader research study. We collected information on the application process

[15]We used a three-point scale: a parent could report not knowing a school, knowing it by name, or knowing it well.

[16]The survey elicited the perceived number of schools in 16 distinct price-quality categories within two kilometers of their home. Respondents were informed about the national quality distribution of schools to ensure a common understanding of the quality definition.

and knowledge of schools in the neighborhood. It also elicits perceptions of schools' price, quality, and admissions chances, and important factors in the search process.

53% of sample parents completed the midline survey and 15% completed the endline survey. Completion rates are not statistically different across treatment arms (Table A.VII).

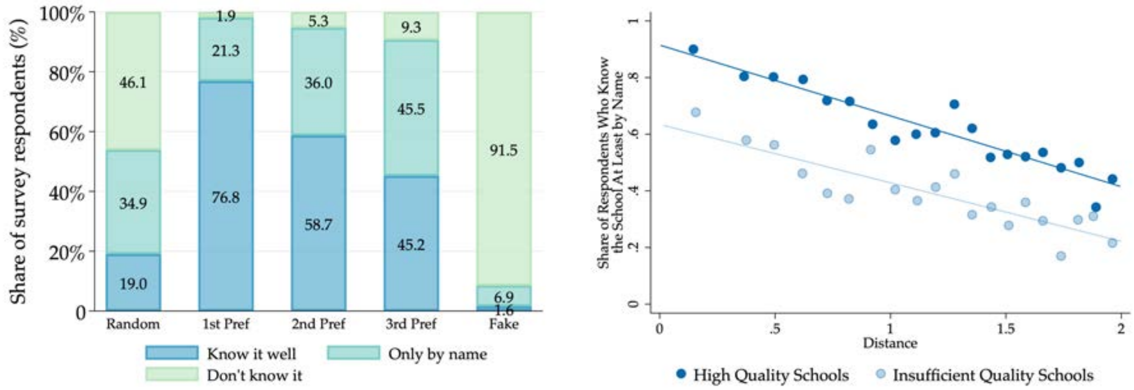
4. DESCRIPTIVE ANALYSIS

We combine our administrative and survey data to document a set of empirical patterns that are consistent with the ideas described in section 2.

4.1. Knowledge, Perceptions, and Beliefs

Limited knowledge of schools: We first show that parents have limited knowledge about the schools that are available in their neighborhood. We asked each respondent in the baseline survey to report how well they knew eight randomly selected schools that were located within 2km of the respondent's home. To measure the validity of the question, we further asked about two "fake" schools that did not exist. Additionally, we asked in the midline survey about the knowledge of schools on the application list. Figure 3a shows the responses to the different school types. When asked about a random school at baseline, 19% of parents say they know the school well and 35% of parents say they know the school by name. Consistent with the idea that applicants learn about schools before applying to them, respondents know schools on their applications better than they know the random schools in their neighborhood. Knowledge rates are also declining with the position of the school in the application. While 77% of families know their first preference well, only 59% of families know their second preference well. Reassuringly, we find that more than 91% of respondents say that they do not know the fake schools.

Figure 3b displays the probability of knowing a school as a function of the distance from the respondent's home to the school. The dark blue line shows this relationship for high quality schools and the light blue line shows the relationship for low quality schools. We document two patterns. First, knowledge decreases



(a) By Ranking

(b) By Distance and Quality

FIGURE 3.—Knowledge of Schools. Notes: Panel (A) plots the stated knowledge levels for five school categories: a random school within 2 km of the respondent's home, the top three schools the respondent ranked in the application, and a fake school. Responses to the random school and the fake school are based on the baseline survey and responses to the schools in the application list are based on the midline survey. Panel (B) uses baseline survey information to plot the share of parents who know a school at least by name by the distance of the school to the respondent's home (in km), separately for schools with high quality (dark blue) and insufficient quality (light blue). In both panels, we restrict the sample to the control group (N=1,318).

sharply with distance. Second, at every distance bin, high quality schools are considerably more likely to be known than low quality schools.

Incorrect beliefs about the distribution of unknown schools: Figure 4 documents that households further hold inaccurate beliefs about the distribution of school quality and price in their neighborhood. We elicited parents' beliefs about the number of schools within 2km of their homes and then asked them to allocate these schools across four quality bins and four price bins according to their beliefs. We find that the average parents underestimate the number of highlight-worthy schools in their neighborhood by five schools. Parents tend to overoptimistic about the distribution of school quality but underestimate the share of schools that are free (55% vs. 86%).

Misperceptions about the characteristics of known schools: Households inaccurately perceive the characteristics of schools that they say they know. Figure 5 plots the distribution of perception errors regarding the quality and price of three school types: a random school the respondent knows, a random school the respondent intended to list in the application, and the school the respondent intended to list

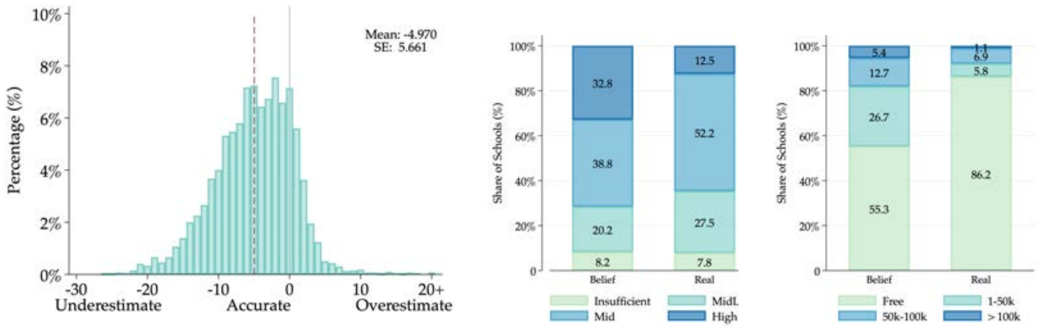
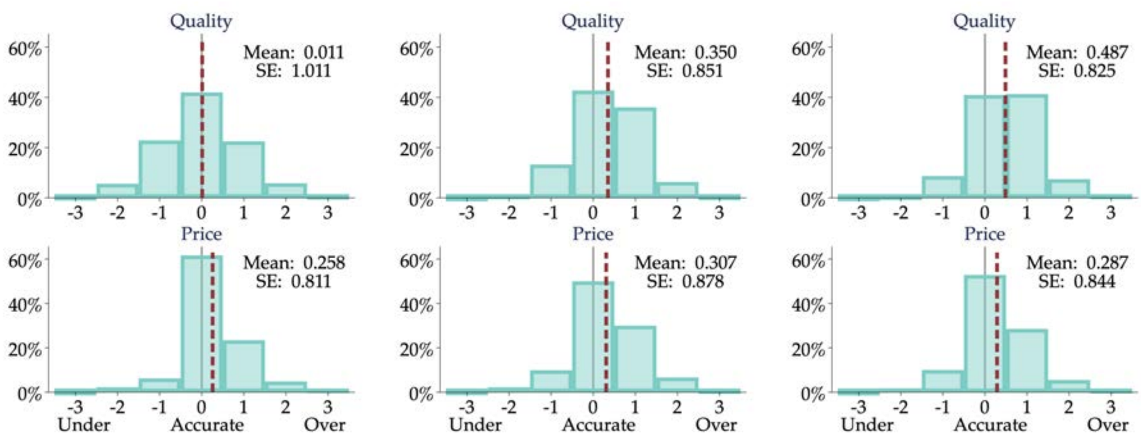


FIGURE 4.—Beliefs about the Distribution of School Attributes. Notes: The first figure shows the bias in the beliefs of the number of highlight-worthy schools within 2km of the parent's home. The second figure shows the perceived (left) and actual (right) share of schools in each of the four school quality categories. The third figure shows the perceived (left) and actual (right) share of schools in each of the four school price categories. Data on beliefs come from the baseline survey (N =3,948).



(a) Random School Known **(b)** Random School in App **(c)** First Preference

FIGURE 5.—Errors in Baseline Survey. Notes: Panel (A) shows the bias for perceived quality and price of a random school that the respondent knows in the baseline survey (N = 2,066). Panel (B) shows the bias for perceived quality and price of a random school in the baseline application list, excluding the first ranked school (N = 1,534). Panel (C) shows the bias on perceived quality and price of the first preference school at baseline (N= 2,523). All biases are measured as perceived quality minus real quality. Positive values indicate that the parent overestimates the quality of the school and negative values indicate that the parents underestimate the quality of the school. Quality is measured in four categories based on the classification of the Education Quality Agency. Survey answers on perceived school quality come from the baseline survey. Red dashed lines indicate the mean bias. The solid grey lines indicate the point of zero bias.

first in the application. We find that perception errors regarding quality are centered around zero for random schools, but that parents tend to overestimate the quality of schools to which they intend to apply, suggesting that these errors matter for application decisions. We further document that households also tend to overestimate the price they must pay for known schools. How these errors impact search decisions depends on the relative strength of preferences for price and quality.

TABLE I
EFFECTS OF PERCEIVED VS REAL CHARACTERISTICS ON RANKING

		Only Endline Survey Responses (1)		Combining All Survey Responses (2)	
Distance		-0.024*	(0.014)	-0.026***	(0.008)
Perceived Price Category	Free	0.311*	(0.182)	0.204**	(0.090)
	50k-100k CLP	-0.016	(0.205)	0.005	(0.111)
	100k+ CLP	0.283	(0.513)	0.097	(0.239)
True Price Category	Free	0.094	(0.214)	-0.111	(0.110)
	50k-100k CLP	0.226	(0.228)	0.080	(0.121)
	100k+ CLP	-0.375	(0.448)	-0.201	(0.223)
Perceived Quality	Low	-1.300	(1.133)	-0.810	(0.572)
	Medium	0.699***	(0.234)	0.584***	(0.136)
	High	1.762***	(0.261)	1.457***	(0.149)
True Quality	Low	-0.047	(0.429)	0.053	(0.234)
	Medium	0.144	(0.163)	0.176*	(0.098)
	High	-0.068	(0.200)	0.356***	(0.113)
Observations		1104		4168	

Note: This table shows the results of a rank-ordered logit choice model using perceived and actual school characteristics. The outcome is based on the submitted ranking in the SAE regular round. In Column 1, perceived price and quality come from responses in the endline survey. In Column 2, we extend the sample by also using information on perceived price and quality from the baseline and midline survey whenever the information is missing in the endline survey. Medium-low quality and 1-50k CLP are the omitted categories.

To show that these survey measures are capturing the perceptions that are relevant for households' decisions, we present results of a rank-ordered logit choice model with actual and perceived school attributes in Table I. The dependent variable is a school's rank in the application. Restricting the analysis to information from the endline survey (Column 1), we find that perceived price and quality are predictive of a school's rank, and that conditional on elicited perceptions the true price and quality are irrelevant. In Column 2, we repeat this exercise but increase our sample size by pooling responses from all survey rounds. True quality becomes predictive of a school's rank, but the coefficients are substantially larger for the perception measures.[17]

We also observe that households mispredict admission chances. We asked parents to report the probability of admission to their first preference school at baseline and compare respondents' reported beliefs to our calculations of objective placement chances. Figure A.1 shows that beliefs about admissions chances depart from objective chances in two ways: beliefs are biased upwards on average and exhibit compression, with households underestimating the share of schools with chances below 40% as well as the share of schools at which admission is nearly certain.[18]

4.2. Search Behavior

We next use data from the school explorer platform to provide descriptive evidence on search patterns.[19] Explorer usage is correlated with perceived returns to search. Table A.I shows that, conditional on the truth, the perceived number of high-quality and low-price schools is positively correlated with search effort, measured by the number of school pins clicked.[20] Table A.II shows that the decision to stop searching depends on history, consistent with sequential search.

We further find that platform behavior affects school knowledge in the midline survey. Table A.III shows that, controlling for baseline knowledge, clicking on a school in the explorer increases the likelihood that the parent knows the school well in the midline survey. These effects are especially pronounced for schools for

[17] True quality may correlate with rankings because of measurement error on survey beliefs and/or learning between survey rounds. Households' perceptions at the time that applications are due are more accurate than at baseline.

[18] This figure shows baseline perceived beliefs. We elicited perceived admissions chances again in the midline survey for a larger set of schools including this school, other known schools, and a random school. In the endline survey we elicited perceived admissions chances for the top three schools on the rank-order list as well.

[19] Most on-platform search occurs immediately after receiving access to the explorer (see Figure S.3).

[20] Moreover, parents who report a higher probability of searching for more information on schools at baseline are more likely to search for schools in the explorer (blue line in Figure A.2).

which the parents double clicked and opened the full profile, allowing them to view additional details about the school.

Consistent with an increase in knowledge, clicking on a school in the explorer also decreases the likelihood that the parent has misperceptions of school attributes. Table A.IV shows that, conditional on baseline perceptions and the true school attribute, clicking on a school in the explorer is associated with an increased likelihood that the respondent is correct about a school's quality and admission chances. These effects are again larger when the parent also opened the full profile of the school.[21]

Finally, Table A.V shows that platform behavior further affects beliefs about the distribution of schools' characteristics. To test this, we regress midline beliefs on the type of schools control group respondents clicked in the explorer. Column 1 indicates that, conditional on baseline beliefs and the actual number of schools within 2km of the household, clicking on more schools in the explorer is associated with an increase in the perceived number of schools in the midline survey. Similarly, we find that clicking on more highlight-worthy schools in the explorer is also associated with an increase in the perceived number of highlight-worthy schools in the midline survey (Column 2).

4.3. Heterogeneity by SES Status

Following the previous literature, we also analyze results separately for high and low SES households, measured by whether the mother completed college.[22] As shown in Table A.VI, high SES households submit longer applications and enroll in higher quality schools. Low SES households are more likely to be eligible for school vouchers and thus have access to cheaper schools. In the supplementary material, we show how knowledge, perceptions, and beliefs vary by SES status. We find little variation in school knowledge levels. However, high SES parents tend to have

[21] The coefficient for price is positive but insignificant. We also find similar associations for school quality when we use the absolute value of the difference between the perceived and actual value instead of a dummy for whether the respondent is correct.

[22] Existing work has studied the role of school availability, preferences, and beliefs to explain differences in schooling decisions by SES status (Burgess et al., 2015, Dizon-Ross, 2019, Attanasio et al., 2022). Previous research also documented how parental education affects child educational attainment (Black et al., 2005, Oreopoulos, 2006, Akresh et al., 2023).

more accurate beliefs about the distribution of schools in their neighborhood. For high SES parents, we also find lower mean errors in perceived school quality and prices for the first preference school and other schools in the application list. Low SES parents tend to have slightly more accurate beliefs about placement chances.

5. EXPERIMENTAL ANALYSIS

We embedded two treatment arms in the school explorer platform to generate exogenous variation in beliefs about the aggregate distribution of school attributes, and in the ease of finding low-price high-quality schools, early in the process. We also implemented another intervention, close to the application deadline, in which we mainly provided additional information about schools that parents knew. We describe each intervention in turn and then present the treatment effects.

5.1. Intervention Design

Search Aid Intervention: The search aid intervention was embedded in the school explorer platform, taking place several months before the final application deadline. It had two treatment arms and a control group. The first treatment provided information about the availability of schools and their price and quality distribution within 2km of a respondent's home.[23] The second treatment provided the same information but additionally showed where highlight-worthy schools are located on the map (Figure S.5). The control group also received access to the school explorer platform but did not receive any information about the distribution of schools or their characteristics.

The information was displayed when parents entered the explorer platform. After that, parents could navigate the map and click on each school to obtain additional information. For the control group and treatment 1 group, all schools were shown in the same color. In contrast, parents in treatment group 2 could directly observe the highlight-worthy schools on the map because these schools were highlighted in green, with an icon indicating their price and quality. In terms of our framework, the goal of the first treatment arm is to provide more accurate beliefs about the distribution of payoffs of unknown schools by providing accurate price

[23]Figure S.4-S.7 in the supplementary material show an example with screenshots of the platform.

and quality information. The second arm additionally changes the search technology, making desirable schools more likely to be found with fewer clicks.

Feedback Intervention: A second information intervention was implemented nationwide as part of a larger experiment. After applications were submitted, roughly a week before the final deadline, a random set of parents received a message that allowed them to receive tailored feedback on their applications (Figure S.8). Parents who opened the feedback intervention first received information on the schools that are currently included in their application (Panel A). If there was a high chance that the child would not receive any school based on the current application, a warning message was also shown to inform the parent that the application was risky (Panel B). The treatment further presented a full list of alternative schools within 2km of the respondent's home, sorted by quality, that were not yet included in the application.[24]

5.2. *Impact of Search Aid Interventions*

We use the randomization of the search aid interventions to study how search behavior and application outcomes change when households are provided with information about the distribution of schools in their neighborhood. Table A.VIII shows that the samples are well-balanced. For parent i, we estimate:

$$Y_i = \alpha + \beta_1 T1_i + \beta_2 T2_i + \theta_i + X_i + \epsilon_i. \tag{1}$$

Y_i is the outcome variable, θ_i are stratification dummies, and X_i are baseline controls selected via a double LASSO approach from Table A.VIII covariates.[25] We show results for the pooled sample and separately for high and low SES households, proxied by whether the mother completed college.

Our first result is that the search aid interventions affects beliefs in the midline survey. Both treatment arms increase the perceived number of total and highlight-worthy schools in their neighborhood (Table II, Panel A, Columns 1-2). Relative to their control group counterparts, households in treatment group 1 believe that

[24] See the supplementary material for an example and additional information.

[25] The double LASSO approach selects covariates that either predict the outcome variable or treatment assignments (Belloni et al., 2014).

TABLE II

TREATMENT EFFECTS OF SEARCH INTERVENTION

	Perceived Number of Schools		Number of Pin Clicks		Number of Schools Known	Enrolled School		
	All	Highlight-worthy	All	Highlight-worthy	At Least by Name	Highlight-worthy	Value Added	Distance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Pooled								
Treatment 1	0.801***	0.443***	0.616	0.227	0.167	-0.005	0.001	-0.183
	(0.291)	(0.118)	(0.514)	(0.216)	(0.202)	(0.021)	(0.019)	(0.253)
Treatment 2	0.750**	0.362***	-0.267	0.279	0.040	0.019	-0.013	-0.521**
	(0.293)	(0.121)	(0.474)	(0.209)	(0.190)	(0.020)	(0.019)	(0.229)
Control Group Mean	6.242	1.907	8.008	3.437	3.710	0.675	0.176	2.322
Observations	1671	1633	3111	3111	1076	2385	2319	2632
Panel B: Heterogeneity by Parental Education								
Treatment 1 × High SES	1.837***	0.544**	3.367***	1.216**	1.275***	0.020	0.081**	0.387
	(0.650)	(0.252)	(1.241)	(0.504)	(0.449)	(0.045)	(0.039)	(0.532)
Treatment 1 × Low SES	0.483	0.414***	-0.344	-0.119	-0.145	-0.012	-0.028	-0.374
	(0.324)	(0.135)	(0.550)	(0.235)	(0.226)	(0.023)	(0.022)	(0.285)
Treatment 2 × High SES	1.671**	0.573**	0.394	0.549	1.033**	0.106**	0.022	-0.244
	(0.671)	(0.282)	(1.171)	(0.492)	(0.401)	(0.044)	(0.041)	(0.485)
Treatment 2 × Low SES	0.525	0.307**	-0.458	0.195	-0.230	-0.009	-0.024	-0.613**
	(0.326)	(0.134)	(0.511)	(0.230)	(0.216)	(0.022)	(0.021)	(0.263)
p-value: Treat 1 x High SES = Treat 1 x Low SES	0.063	0.649	0.006	0.016	0.005	0.528	0.015	0.206
p-value: Treat 2 x High SES = Treat 2 x Low SES	0.125	0.396	0.506	0.516	0.006	0.020	0.331	0.507
Control Group Mean (High SES)	6.110	1.783	9.897	4.058	3.386	0.548	0.197	2.231
Control Group Mean (Low SES)	6.278	1.941	7.441	3.253	3.815	0.717	0.170	2.355
Observations 1 (High SES)	362	357	732	732	246	571	549	614
Observations 2 (Low SES)	1308	1275	2376	2376	829	1812	1768	2016

Note: This table presents the results of the search interventions on beliefs (Columns 1-2), search (Columns 3-4), knowledge (Column 5), and final school enrollment (Columns 6-8). In Panel A, we regress each outcome on indicator variables for both treatment arms, stratification dummies and baseline controls selected by LASSO. In Panel B, we further include the fully interacted effects of treatments and SES status. SES status is proxied by whether the mother completed college. Continuous outcomes are top-coded at the 99th percentile. The sample is restricted to parents who opened the school explorer platform.

there are 23% more highlight-worthy schools in their neighborhood. Panel B shows results by SES status. High SES households update their beliefs about the number of total and highlight-worthy schools in their neighborhood, but low SES households only update their beliefs about the number of highlight-worthy schools.

The updated beliefs affect the search behavior of parents (Columns 3-4). While we find limited effects for the pooled sample, we observe substantial increases in the number of school pin clicks among high SES households in the first treatment group. Consistent with increased search, we also observe knowledge gains in the midline survey (Column 5). High SES households in treatment group 1 report that they know 38% more schools at least by name. By contrast, we find null to negative effects for low SES households.

We next examine the effects of the search aid interventions on school enrollment (Columns 6-8). We again find limited effects on the pooled sample (Panel A), with the exception that treatment 2 households tend to enroll in closer schools. However, among high SES households, we find that the first treatment arm leads to a significant increase in the average value added of the enrolled school. Higher SES households in the second treatment arm are also 19% more likely to enroll in a highlight-worthy school.[26]

We also assess whether treatment effects on search vary by the perceived returns to search, measured as the reported likelihood of searching for more information on schools at baseline.[27] Figure A.2 plots the relationship with search effort and baseline beliefs separately for parents in the control group (blue) and treatment group 1 (red). As reported in section 4.2, we find that our baseline measure of the perceived returns to search has a strong positive correlation with observed search effort in the control group. We further find that the treatment effects are concentrated among parents who said that they were unlikely to search for additional information on schools. This is consistent with the idea that these parents underestimated the returns to search and that the treatment intervention corrected their beliefs. Table A.XI shows the effects in a regression specification, using a binary measure of the baseline variable. We find that treatment 1 increases the number of school pin clicks by 29% for parents who were unlikely to search for schools. By contrast, we find no effects for parents who already planned to search for more information on schools (Column 3). Treatment effects on enrollment outcomes are noisy but we find that treatment group 2 children whose parents were unlikely to search enroll in schools closer to their homes (Column 8).

[26] Table A.X shows treatment effects on additional application outcomes. We find that the second treatment arm increases the likelihood that the second-ranked school is highlight-worthy by 5 percentage points and decreases the likelihood that the parent already knew the school well at baseline. We also find some evidence that the second treatment arm increases the share of parent who submitted an application through the SAE platform. Treatment 2 also increases the likelihood that the child enrolls in the school to which the child was assigned, suggesting that treatment group 2 parents had better information at the point of the application.

[27] We asked respondents to separately report the likelihood of searching for more information on known and unknown schools at baseline. We take the maximum of both variables in this exercise.

5.3. *Impact of Feedback Treatment*

We next examine the impact of the feedback intervention. Table A.IX shows that the samples are also well-balanced for this intervention. For parent i, we estimate:

$$Y_i = \alpha + \beta F_i + \lambda_i + \gamma X_i + \varepsilon_i.$$

Y_i is the outcome variable, θ_i are stratification dummies, F_i is an indicator variable for whether the parent opened the feedback intervention and X_i are baseline covariates that control for the risk of the submitted application. We use the treatment assignment to instrument for using the feedback intervention. Standard errors are clustered at the market cluster level.

Column 1 in Table III shows how the first-stage results. 58% of parents who were assigned to the treatment group opened the feedback intervention (Panel A). These rates are similar for high and low SES households (Panel B). We next examine the effect on school-level knowledge using information from the endline survey. We find that the feedback intervention increases the likelihood that a parent has a correct perception about the price category of a school in their application by 17.2 percentage points (Column 2). We find similar effects for perceptions of school quality (Column 3). The treatment effects on the perceived knowledge of prices are driven by low SES parents who are more likely to have misperceptions than high SES parents. Columns 4-7 further show that the feedback intervention affects application behavior. We find that parents who received the feedback information are 9.7 percentage points more likely to change their application (Column 4). They are more likely to add any school and more likely to add a highlight-worthy school, showing that the information helped to increase search. We also find that treatment group parents are more likely to delete schools, suggesting that parents previously had incorrect perceptions of the attributes of some schools in their application. Consistent with the treatment effects on perceptions, we find that the effects on application behavior are concentrated among low SES households.[28]

[28]Table A.XII shows treatment effects on assignment and enrollment outcomes. We find some evidence that the feedback intervention increases the share of parents that are assigned to a school they added to the application after the initial deadline. We observe no effect on final enrollment outcomes, with

TABLE III

TREATMENT EFFECTS OF FEEDBACK INTERVENTION

	First Stage	Perceptions		Application			
	Open Feedback	Correct Price	Correct Quality	Changed Application	Added School	Added Highlight-worthy School	Delete School
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Pooled							
Feedback Treatment	0.578**						
	(0.018)						
Open Feedback		0.172*	0.227**	0.097**	0.061**	0.052**	0.020*
		(0.092)	(0.081)	(0.020)	(0.018)	(0.015)	(0.012)
F-Stat	1000.016						
Control Group Mean	0.000	0.362	0.553	0.028	0.026	0.018	0.009
Observations	2116	765	753	2116	2116	2116	2116
Panel B: Heterogeneity by SES Status							
Feedback Treatment × High SES	0.614**						
	(0.034)						
Feedback Treatment × Low SES	0.568**						
	(0.023)						
Open Feedback × High SES		-0.109	0.253*	0.021	0.017	0.026	-0.001
		(0.150)	(0.119)	(0.033)	(0.030)	(0.018)	(0.016)
Open Feedback × Low SES		0.300*	0.199*	0.122**	0.076**	0.062**	0.027*
		(0.116)	(0.105)	(0.025)	(0.022)	(0.018)	(0.015)
p-value: Open Feedback x High SES = Open Feedback x Low SES	0.290	0.034	0.731	0.017	0.124	0.133	0.186
F-Stat (High SES)	322.606						
F-Stat (Low SES)	616.414						
Control Group Mean (High SES)	0.000	0.441	0.667	0.037	0.029	0.008	0.008
Control Group Mean (Low SES)	0.000	0.343	0.528	0.026	0.025	0.021	0.009
Observations 1 (High SES)	460	165	160	460	460	460	460
Observations 2 (Low SES)	1656	600	593	1656	1656	1656	1656

Note: This table presents the results of the feedback intervention. In Column 1 in Panel A, we regress the outcome on an indicator variable for whether the parent was assigned to the feedback treatment group. In Columns 2-7 in Panel A, we regress each outcome on an indicator variable for whether the respondent opened the feedback information instrumented by whether the parent was assigned to the feedback treatment group. In Panel B, we further include the fully interacted effects of the treatment and SES status. All regressions control for market fixed effects and application risk groups. SES status is proxied by whether the mother completed college.

Taken together, we find that the search aid and feedback interventions affected search and application outcomes, but that the direction of the heterogeneity results differs. While the search aid interventions benefitted high SES households, low SES households benefitted from the feedback intervention. A potential explanation is that low SES households are more likely to have misperceptions about the schools to which they apply, as discussed above. A second explanation is that the timing of

the exception of a negative treatment effect on the value added of the enrolled school for high SES parents (but the estimate is only marginally significant).

the interventions matters. Low SES households may form applications later, and might thus only react to information when it is provided close to the application deadline, when the relevant decision is being made. The importance of timing has also been shown in several other settings (Richburg-Hayes et al., 2017). A third explanation is that the information being provided by the feedback intervention was easier to act on.

6. MODEL

We now present a model of households' preferences, beliefs, search efforts, and application decisions that is consistent with our setting and descriptive evidence.

Let I denote the set of households, and J the set of schools. Household i's maximal choice set, $J_i \subset J$, consists of all schools within five kilometers of i's house that offer a seat in i's grade level and for which i is eligible.[29] Time is discrete: $t = 0, 1, \ldots, T$. At time $t = 0$, our study begins. Applications are due at time T.

6.1. *Information, Preferences, Perceptions, and Application Portfolios*

Let u_{ij} be the "objective" expected payoff from that student i would receive from being placed in school j given all the information that could in principle be known at time T. This payoff is given by:

$$u_{ij} = z_{ij}\beta^z + x_{ij}^{rc}\beta_i^x + \delta_j + x_{ij}\gamma + \varepsilon_{ij}. \tag{2}$$

It depends on distance in kilometers (z_{ij}), observed price and quality category $x_{ij} \in \{1,2,3,4\}^2$, school effects δ_j, and idiosyncratic match value ε_{ij}. The vector x_{ij}^{rc} consists of price, quality category, distance, and a constant. We place normal random coefficients on these terms, with zero mean and arbitrary correlation matrix Σ^{rc}.

A challenge to tractability is that households may be misinformed about many objects. We address this challenge by modeling i's limited awareness of school j and misperceptions of its characteristics via a single index π_{ijt}, generalizing mod-

[29] Choice sets are larger than the 2km-radius neighborhoods used to present distributional information in our "search" interventions. Ineligibility for seats in the relevant grade level is rare. Empirically, single-sex schools are the main source of ineligibility (4.39% of total seats offered in 2021).

els of consideration sets (Goeree, 2008) to accommodate additional information frictions.

In particular, households have a family of potential subjective expectations of u_{ij}, indexed by the available amount of information. Let $\pi_{ijt} \in \{0,1,2\}$ indicate whether household i has zero information, low information, or high information about school j at time t, respectively.[30] At time t, student i's subjective expected utility from a placement in j is given by $\hat{u}_{ijt} = \hat{u}_{ij}^{\pi_{ijt}}$. The value of receiving no placement is normalized to zero. If i submits an application at time t, she will rank all schools $j \in J_i$ with $\hat{u}_{ijt} > 0$ truthfully, in descending order of $\hat{u}_{ijt}$.

If $\pi_{ijt} = 0$, then i does not know j well enough to apply to it.[31] We make the following parametric assumptions:

$$\hat{u}_{ij}^{(2)} = z_{ij}\beta^z + \hat{x}_{ij}^{2,rc}\beta_i^x + \delta_j + \hat{x}_{ij}^{(2)}\gamma + \varepsilon_{ij} \tag{3}$$

$$\hat{u}_{ij}^{(1)} = z_{ij}\beta^z + \hat{x}_{ij}^{1,rc}\beta_i^x + \delta_j + \hat{x}_{ij}^{(1)}\gamma + \hat{\varepsilon}_{ij}^{(1)} \tag{4}$$

$$\hat{x}_{ij}^{(1)} \sim \Gamma(\cdot|x_j), \;\; \hat{x}_{ij}^{(2)} = (x_j \text{ w.p. } p^h, \text{ otherwise } \hat{x}_{ij}^{(1)}) \tag{5}$$

$$\pi_{ijt}^* = z_{ij}\alpha^z + w_{ijt}\alpha^w + w_{ijt}^{rc}\alpha_i^{rc} + \eta_j + \nu_{ijt} \tag{6}$$

$$\pi_{ijt} = 1(\pi_{ijt}^* > 0) + 1(\pi_{ijt}^* > 1). \tag{7}$$

Subjective expected utilities $\hat{u}_{ij}^{\pi_{ijt}}$ from "known" schools—those with $\pi_{ijt} > 0$—differ from the true payoffs in two ways. First, households may hold inaccurate perceptions of observables x. Second, in the case of low information, households may misperceive econometrician-unobserved match quality ε.

In particular, the variables $\hat{x}_{ij}^{\pi_{ijt}} \in \{1,2,3,4\}^2$ denote i's perceptions at time t of the prices and quality scores of schools that they know at least by name. "Low-information" perceptions $\hat{x}^1$ are drawn from a multinomial distribution, $\Gamma(\cdot|x_j)$, which depends arbitrarily on the true value x_j. When the household gains further information, it learns the truth with probability p^h. Otherwise, it does not update

[30] There is nothing essential about three levels of knowledge. This restriction is motivated by our survey measures of knowledge of schools.

[31] To avoid notational special cases, we take $\hat{u}_{ij}^{(0)} < 0$ for all i, j.

about x. Random-coefficient terms $\hat{x}_{ij}^{\pi,rc}$ differ from x_{ij}^{rc} as well in that the true price and quality are replaced with the perceived values $\hat{x}_{ij}^{\pi}$.

"Match value" shocks ε_{ij} are normally distributed, iid across people and schools. The relevant shock is observed when $\pi_{ijt} = 2$. In the "low-information" event $\pi_{ijt} = 1$, the household observes a noisy measurement of this shock, $\tilde{\varepsilon}_{ij}$, and forms a subjective expectation $\hat{\varepsilon}_{ij}^1 = \hat{E}(\varepsilon_{ij}|\tilde{\varepsilon}_{ij})$ given this signal. We assume

$$\begin{pmatrix} \hat{E}(\varepsilon_{ij}|\tilde{\varepsilon}_{ij}) \\ \varepsilon_{ij} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu^l \\ 0 \end{pmatrix}, \Sigma^{\epsilon} \right).$$

This is a reduced form of a model of Bayesian updating with a possibly misspecified prior mean, prior variance, and signal precision, which we present in the following section. It allows parents to deviate from Bayesian updating by being pessimistic or optimistic when $\pi_{ijt} = 1$, and being systematically surprised as they learn more.

As a scale normalization, we fix the mean coefficient β^z on distance to -1.

Levels of knowledge π_{ijt} increase as a latent index of information π_{ijt}^* crosses thresholds. In this sense our model extends "consideration set" approaches (Goeree, 2008) to allow for additional information acquisition about "known" schools. Setting the thresholds at 0 and 1 normalizes the scale and location of π_{ijt}^*.

The terms w_{ijt} are potentially time-varying "knowledge shifters" that are excluded from preferences. In practice, w_{ijt} consists of three types of terms: (1) indicators for having received information treatments in current or prior periods; (2) indicators for having searched school j in the explorer model in current or prior periods; and (3) time indicators $1(t > s)$ for a set of periods s. The first set of terms captures variation induced by our experiments. The second set captures the effects of endogenous on-platform search effort. Time indicators proxy for the arrival of information over time outside of our explorer and interventions. We take this "off-platform" process as exogenous.

We place random coefficients on the time indicators in information: $w_{ijt}^{rc} \sim N(0, \Sigma^{\pi rc})$. The unrestricted covariance matrix allows variation in the timing of off-platform search. For instance, households may engage in "off-platform" search at a single, unpredictable time, leading to negative correlation between gains in knowledge across periods. The remaining terms are a school-level "discoverability" η_j and a shock ν_{ijt}. The shocks are independent across schools and households, but may be correlated over time: $\nu_{ijt} \sim N(0, \Sigma^{\nu})$.

We model schools' mean utility and discoverability as correlated random effects: $(\delta_j, \eta_j)' \sim N\left((x_j\overline{\beta}, x_j\overline{\alpha})', \Sigma^{\delta\eta}\right)$. This specification allows for selection in the set of known schools, as better schools may be systematically easier to find. If so, the next school to be discovered may in fact offer a lower payoff, in expectation, than the average "known" school.

Our specification is also flexible about the role of prices and quality ratings. Parents may value quality ratings per se and/or as a summary of underlying characteristics that they do not directly observe. They may also value characteristics that they observe but we do not which are merely correlated with quality ratings. By placing coefficients $\overline{\beta}$ on true price and quality, and γ on perceptions, we allow for both channels.[32]

6.2. *Timing, States, Beliefs, Search*

Our assumptions so far specify the applications that parents submit given their information and the distribution of that information in our data. These objects suffice for counterfactuals in which information is (re-)assigned exogenously. To predict search decisions under counterfactuals, however, further structure is needed. We begin with an overview of timing, states, and the search technologies available to parents. We then present the belief and perception objects needed to specify the subjective returns to search.

Timing: In practice, we allow four periods, $t = \{0, 1, 2, 3\}$. At time $t = 0$, our baseline survey takes place. At time $t = 1$, households are assigned to search treatments and have the opportunity to engage in on-platform search via our software. Within this period, households first receive treatments, causing their state to update. They then engage in endogenous search, potentially leading to further updates in π and other beliefs. When they choose to stop, the period ends. Consistent with de-

[32]The counterfactuals in this paper hold δ and η fixed. We are agnostic about the extent to which x "causes" δ or η.

scriptive evidence that almost all explorer use takes place at this time, on-platform search takes place only in this period. At time $t = 2$ ("just before feedback") household may submit initial applications to the official platform. They may also take the midline survey in this period.[33] At the beginning of period $t = T = 3$, "treated" households receive the feedback treatment. After this, final applications are submitted. Finally, the endline survey takes place.

States: An agent's state at the beginning of period t is given by

$$\Omega_{it} = (\{\hat{u}^1_{ij}, \hat{u}^2_{ij}, \hat{r}_{ij}, \pi^*_{ijt}, w_{ijt} : j \in J_i\}, \theta_i, \omega_{it}),$$

where $\hat{r}_{ij}$ are subjective perceptions of admissions chances, θ_i are fixed parameters relevant for preferences—in practice, the agent's random coefficients—and ω_{it} are the remaining parameters and latent objects determining the agent's beliefs. We discuss $\hat{r}$ and the components of ω in detail in the next subsection.

At time $t = 0$, agents are endowed with potential utilities and perceptions $(\hat{u}^1, \hat{u}^2, \hat{x}^1, \hat{x}^2, \hat{r})$, as well as an initial level of information π^*_{ij0} for each $j \in J_i$, and beliefs ω_{i0} relevant for the search decision. Agents observe $\hat{r}_{ij}$ and $\hat{u}_{ij}(\pi_{ij0})$ for each $j \in J_i$ with $\pi_{ij0} > 0$.

Search is sequential: Search decisions consist of a sequence of decisions to sample an additional school $j \in J_i$. The school explorer allows "pin clicks" as well as more-costly additional search actions. To exploit the observable variation in search intensity in the data, we model both decisions, closely following the structure of the explorer platform. A household may pay a cost to conduct a "pin click" on some j, which provides some information about it. Conditional on this information, the household may then pay an additional cost to conduct a "detail view" on this school. After this action, whether the household conducts a "detail view" or not, the household chooses whether to continue or stop searching.

Let s denote the number of searches the household has conducted so far. As a convention, we divide period 1 into sub-periods, with The sth search decision

[33]The relative timing of the midline survey and feedback treatment varied across households. For households who were given the midline survey before they received—or would have received—the feedback treatment, the midline survey takes place at $t = 2$; otherwise it takes place at $t = 3$.

taking place at time $t = 2 - 1/s$.[34] Once the household stops searching, we move to $t = 2$. the information shifters at any time t, w_{ijt}, contain an indicator for i having conducted a "pin click" at j in t or previous times, and an analogous indicator for having conducted a "detail view". As w_{ij} updates, so does π^*_{ij} and hence possibly π_{ij}.

Search technology: If the household samples a school, we take the decision of which school as exogenous from the parents' point of view. However, we allow the probability of finding certain schools to vary with our interventions. In particular, the "highlighting" of high-quality inexpensive schools in search treatment 2 makes those schools more likely to be found. If a household chooses to conduct a pin click, the next clicked school is $j \in J_i$ with probability

$$Pr_{ij} = Pr(\text{view } j|\text{continue}) = \exp(x_{ij}^{\text{click}}\gamma^{\text{click}}) / \sum_{j \in J_i} \exp(x_{ij}^{\text{click}}\gamma^{\text{click}}). \tag{8}$$

The variables x_{ij}^{click} consist of schools' price, quality, distance, and indicators for treatment 2 and being highlight-worthy in that treatment. Thus certain schools (nearby schools; high-quality schools) may be found more easily. Households may revisit already-clicked schools. However, by assumption only the first "pin click" and "profile click", if any, enter π_{ijt}.

Admissions Chances: Subjective admissions chances are given by

$$\hat{r}_{ij}^{\pi_{ijt}} = \min\{1, \max\{0, o_{i0} + o_{i1}(r_{ij} - o_{i0})\}\}, \;\; \pi_{ijt} \in 0, 1, 2, \tag{9}$$

where $(o_{i0}, o_{i1}) \sim N(\mu_o, \Sigma_o)$. This specification allows for optimism or pessimism (via the random intercept o_{i0}) as well as compression (via o_{i1}) which we find in survey data. Following the descriptive evidence, we also assume that admissions beliefs do not become more accurate with higher π_{ijt} or greater search activity, unlike perceptions of price and quality.

[34] All that matters is the order of timings; cardinal values are not meaningful.

The outside option, not receiving a placement, has $\hat{r}_{i0} = r_{i0} = 0$. The subjective expected payoff from the optimal portfolio at time t is given by

$$\hat{U}(\Omega_{it}) = \sum_{j=1}^{M_{it}} \left(\prod_{k=1}^{j-1} \hat{r}_{ijt} \right) (1 - \hat{r}_{ijt}) \hat{u}_{ijt}, \tag{10}$$

where $M_{it} = |j \in J_{it} : \hat{u}_{ijt} > 0|$ is the number of acceptable schools, and WLOG elements of J_{it} are sorted in descending order of $\hat{u}_{ijt}$.

Beliefs about unknown schools' observables: Beliefs about the price and quality of unknown ($\pi_{ijt} < 0$) schools at time t are given by a Dirichlet-Multinomial distribution. At time t, for $j \in J_i$ with $\pi_{ijt} = 0$, household i believes $x \sim \text{Multinomial}(\lambda_{it})$, with $\lambda_{it} \sim Dir(\Lambda_{it})$. The Dirichlet parameters Λ_{it} in turn are functions of the truth, treatment status, and the number of "known" ($\pi_{ijt} > 0$) schools with (perceived) chars $\hat{x}_{ijt}$. We allow heterogeneity in Λ_{it} via a finite mixture over latent types $g \in G$ as well.[35]

Enumerating all price and quality cells $x_1, \ldots, x_K$, The kth element of Λ_{it} in the event i is of type g is given by:

$$\Lambda_{igtk} = \Lambda_{g0k} + b_{1igt} * |\{j \in J_i : x_{ij} = x_k\}| + b_{2igt} * |\{j \in J_i : \pi_{ijt} > 0 \text{ and } \hat{x}_{ij}^{\pi_{ijt}} = x_k\}|. \tag{11}$$

The terms b_{1igt} and b_{2igt} vary with i's search treatment, provided that i has received the treatment by time t.[36] This specification allows information about the distribution of nearby schools' characteristics (treatment 1) and information about specific schools (treatment 2) to affect beliefs.

Our specification nests Bayesian updating. In the limit as J_i becomes large, this occurs when $b_{2igt} = 1$ and perceptions $\hat{x}$ are equal to x. It also accommodates full information about the remaining schools, which occurs when $b_{1igt} \to \infty$ and $b_{2igt} = -b_{1igt}$. By including multiple types, we allow heterogeneous partial updat-

[35] In practice, for each SES group, we have three latent belief types.

[36] We have $b_{1igt} = \tilde{b}_{1g0} + \tilde{b}_{1g1} 1(\text{treatment}_i = 1) 1(t \geq 1) + +\tilde{b}_{1g2} 1(\text{treatment}_i = 2) 1(t \geq 1)$, where treatment_i denotes i's search treatment assignment.

ing in response to our information treatments. For instance, type $g = 1$ may be more responsive to our interventions than $g = 2$.

Beliefs about match quality: If $\pi_{ijt} = 1$, household i observes ε_{ij} with classical measurement error $e_{ij} \sim N(0, \tilde{\sigma}^2)$, and forms a subjective expectation $\hat{E}(\varepsilon_{ij}|\tilde{\varepsilon}_{ij})$. Objectively,

$$\begin{pmatrix} \epsilon_{ij} + e_{ij} \\ \epsilon_{ij} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2 + \sigma_e^2 & \sigma_\varepsilon^2 \\ \sigma_\varepsilon^2 & \sigma_\varepsilon^2 \end{pmatrix}\right). \tag{12}$$

However, households believe that

$$\begin{pmatrix} \epsilon_{ij} + e_{ij} \\ \epsilon_{ij} \end{pmatrix} \sim N\left(\begin{pmatrix} \tilde{\mu} \\ \tilde{\mu} \end{pmatrix}, \begin{pmatrix} \tilde{\sigma}_\varepsilon^2 + \tilde{\sigma}_e^2 & \tilde{\sigma}_\varepsilon^2 \\ \tilde{\sigma}_\varepsilon^2 & \tilde{\sigma}_\varepsilon^2 \end{pmatrix}\right). \tag{13}$$

The parameter $\tilde{\mu}$ represents the subjective mean unobserved match-value shock at unknown schools. If $\tilde{\mu} < 0$, households are systematically pessimistic about their payoffs at unknown ($\pi_{ijt} = 0$) schools and at schools they do not know well ($\pi_{ijt} = 1$). Households may also hold inaccurate beliefs about the prior variance or the informativeness of the signal.

A Bayesian with the correct prior will shrink noisy signals toward their mean. If the mean is low relative to the (positively selected) values of the schools the household likes and knows well, this will tend to push schools that are not known well away from the top of rank-order lists and reduce their dispersion conditional on observables. Relative to this benchmark, our specification allows households to further penalize (or under-penalize) schools that are not known well, and to mispredict the mean and variance when considering the returns to search.

The primitives relate to the reduced form described in the previous subsection as follows. Let $\tilde{\rho} = \tilde{\sigma}_\varepsilon^2 / \left(\tilde{\sigma}_\varepsilon^2 + \tilde{\sigma}_e^2\right)$ denote the subjective informativeness of the signal received when $\pi_{ijt} = 1$. We have $\tilde{\rho} = \Sigma^\varepsilon_{1,2} / \Sigma^\varepsilon_{2,2}$ and $\tilde{\mu} = \mu^l / (1 - \tilde{\rho})$. See the Online Appendix for details.

Search Decisions: We model the sequence of pin clicks via a "one-period lookahead" heuristic. Households form beliefs over the distribution of payoffs, as specified in Equation 10, that they would obtain with one additional pin click and the optimal exercise of the option to conduct a detail view at the clicked school. As in the benchmark, they continue if and only if the expected value of this object outweighs the cost of an additional pin click.

Search costs: To conduct its sth pin click, a household must pay a cost

$$c_{i1s}^{\text{pin}} = x_i^c \gamma^{\text{cost}} + \bar{c}_i^{\text{pin}} - \sigma^{\text{pin}} \varepsilon_{i1s}^{\text{pin}},$$

where $\bar{c}_i \sim N(0, \sigma_{\bar{c}}^2)$. If instead it stops, it pays $c_{i0s}^{\text{pin}} = -\sigma^{\text{pin}} \varepsilon_{i0s}^{\text{pin}}$. The shocks $\varepsilon_{i1s}^{\text{pin}}$ and $\varepsilon_{i1s}^{\text{pin}}$ are independently standard Gumbel (T1EV) distributed. If the household conducts a "detail view" following this click, it pays $c_{is}^{\text{detail}} = \bar{c}^{\text{detail}} - \sigma^{\text{detail}} \varepsilon_{i1s}^{\text{detail}}$, otherwise it pays $-\sigma^{\text{detail}} \varepsilon_{i0s}^{\text{detail}}$, where the shocks are again drawn from independent Gumbel distributions.

The one-period-lookahead gain from conducting a pin click and detail view at school j is given by

$$\hat{V}_{ijt}^1 = \hat{E}(\hat{U}(\Omega_{it'}) | \Omega_{it}, \pi_{ijt'}^* = \pi_{ijt}^* + \alpha^{\text{pin}} + \alpha^{\text{detail}}) - \hat{U}(\Omega_{it}), \tag{14}$$

where t' is the subsequent period, and α^{pin} and α^{detail} are the elements of the information-shifter coefficients α^w corresponding to pin clicks and detail views, respectively.[37] The subjective expectation is over the change in $\hat{U}(\cdot)$ due to a possible increase in knowledge $\pi_{ijt'} > \pi_{ijt}$.[38] We provide details in the Online Appendix.

The one-period-lookahead gain from conducting a pin click only at school j is identical to Equation (14) except that α^{detail} does not occur in it. The gain from conducting a detail view, having already conducted a pin click, is analogous as well. The value of a pin click, integrating over the expected utility of the optimal

[37]An implication is that households are naive about information gains that will take place later, between periods 2 and 3, when making search decisions at $t = 1$. Consistent with this, our "feedback" intervention was not announced, and was a surprise to households.

[38]The gain is nonzero only if $\lfloor \pi_{ijt'}^* \rfloor > \lfloor \pi_{ijt}^* \rfloor$ and $\pi_{ijt}^* < 1$, i.e. if the clicks would cause an unknown school to become known, or a school that was known but not well to become known well. When evaluating unknown schools, households correctly anticipate the distribution of δ and distance, but believe x is distributed Dirichlet-Multinomial as in equation 11 independently of δ.

detail-view decision given the household's beliefs, is given by

$$\hat{V}_{it} = \sum_{j \in J_i} Pr_{ij} E \max\{\hat{V}^1_{ijt} - \bar{c}^{\text{detail}} + \sigma^{\text{detail}} \varepsilon^{\text{detail}}_{i1s}, \hat{V}^0_{ijt} + \sigma^{\text{detail}} \varepsilon^{\text{detail}}_{i0s}\}. \tag{15}$$

Households conduct an additional pin click if and only if this object is greater than the search cost.

Effects of information treatments: We model "search" treatments 1 and 2 as having direct impacts on belief parameters Λ. In addition, treatment 2 affects the search technology via Pr_{ij}, raising the probability of finding "highlight-worthy" schools. Further, all treatments including feedback enter information indices π^*_{ijt} directly. We allow overall effects of receiving feedback, as well as impacts on knowledge of the schools about which feedback was provided, via indicators in w.

7. ESTIMATION

All parameters are estimated separately for low-SES and high-SES households.

Identification comes from repeated measurements of rank-order lists, awareness, subjective beliefs and perceptions of schools' characteristics, together with variation over time within person and school in these observed objects that is induced by our treatment assignments and by households' search activity. In particular, our survey and administrative data generate multiple measurements, within household i and school j, of parents' awareness of schools $\pi^{\text{survey}}_{ijt} \in \{0, 1, 2\}$ reflecting knowing the school "not at all", "by name," or "well" respectively, parents' perceptions of schools' characteristics $\hat{x}^{\text{survey}}_{ijt} \in \{1, 2, 3, 4\}^2$, their perceived admissions chances, and the rank of j within i's rank-ordered application at time t, as well as measures of beliefs over the distribution of x.

The relationship between reported knowledge level $\pi^{\text{survey}}_{ijt}$ and perceptions $\hat{x}^{\text{survey}}_{ijt}$, and between received "treatments" and perceptions, pins down learning about observables. If parents' reported perceptions of x_j become more accurate when they know school j "well" than previously when they did not, we will conclude that the updating parameter p^h is large. Updates in reported knowledge, and changes in preference rankings and in perceptions $\hat{x}$ within person and school, tell us also the impact of information treatments and search "clicks" on knowledge π^*.

Because we observe the effects of treatments on intermediate outcomes (beliefs, perceptions, awareness) as well as on applications and assignments, we are able to observe the "takeup" of our information treatments, and therefore distinguish "takeup" of information from its impacts in the event households' beliefs were to fully update.

The extent to which households penalize schools known "by name" at time t, relative to those known well, and the these schools' dispersion in households' rank-order lists conditional on their observables, reveal the distribution of match shocks ε, their subjective mean μ^l, and the subjective signal-to noise-ratio $\frac{\tilde{\sigma}_\varepsilon^2}{\tilde{\sigma}_\varepsilon^2+\tilde{\sigma}_e^2}$. To pin down the remaining parameters, we use additional survey questions on the likely rankings of hypothetical schools.[39]

Within-school-j variation in $\hat{x}_{ijt}$ across households, and variation within households over time, let us distinguish mean-utility parameters $\overline{\beta}$ from the direct impact of perceptions γ, consistently with the descriptive patterns described in Table I. To address selection into search on the basis of baseline knowledge, our model makes use of pre- and post-search measurements of the relevant objects, exploiting within-person variation. In our second step, we model the search decision, allowing households' search costs to vary with baseline knowledge.

Repeated measurements of $\hat{x}$, $\hat{r}$, π, and rank-ordered preferences within person and school allow us to accommodate measurement error in perceived characteristics, in reported knowledge levels and in non-final rank-order lists. We allow measurement error on every survey variable, as well as on the "just-before-feedback" administrative rank-order list submitted at $t = 2$.[40]

Estimation: We estimate the model in two steps. First, we estimate the distribution of $(\hat{u}^\pi, \pi^*, \hat{x})$, and the parameters relevant for these objects, via a Gibbs sampler. We use the baseline survey rank-order list at time $t = 0$, submitted rank-order lists at $t = 2$ and $t = 3$, the sequence of explorer clicks and detail views, baseline

[39]We provide details on identification and estimation of match value belief parameters in the Estimation Appendix.

[40]Reassuringly we find that this latter measurement error is negligible.

and midline survey responses to questions on "how well do you know" school j, and perceived price and quality elicited at baseline, midline, and endline surveys. The index π^* underlying reported awareness π^{survey} has additive Gaussian measurement error whose variance we estimate, as do the payoffs that enter non-final rank-order lists. Perceived $\hat{x}$ are also misreported (drawn uniformly on $\{1,2,3,4\}$ independently of the beliefs that enter the household's decisions) with a probability that we estimate.

We estimate admissions-belief parameters μ_o, Σ_o and click probabilities γ^{click} offline via maximum likelihood, allowing for measurement error in surveyed admissions perceptions. In a second step, with draws of $\hat{x}$ in hand from step 1, we estimate beliefs over the distribution of unknown schools' prices and qualities, again allowing for measurement error.[41] We estimate beliefs over shocks ε as well at this stage.[42] Finally, we impose optimality of the search decision, and estimate search costs via MLE, taking the means from the first stage as point estimates of the relevant parameters. We provide details in the Online Appendix.

8. RESULTS AND COUNTERFACTUAL SIMULATIONS

8.1. Results

Table IV presents selected parameters that describe families' preferences, information, and school unobservables. A full set of estimates is available in Table A.XIII.

Panel A shows the main parameters that govern parents' preferences and information. The first section of the panel presents estimated subjective expected utility parameters. These are relative to the distance coefficient, which we normalized to -1. Preferences for perceived school attributes present the expected signs: parents

[41] A challenge is that our belief elicitation is over all schools in the area, while the beliefs relevant for search are those over unknown schools. If the household reports beliefs for N^u unknown schools and N^k "known" schools, unknown schools are distributed over cells according to a Dirichlet-Multinomial distribution with parameters as in Equation (11), while known schools are sampled with replacement, and the perceived $\hat{x}$'s are reported. This may be interpreted as the household sampling a set of schools from memory. We describe the estimation procedure in detail in the Estimation Appendix.

[42] To pin down $\tilde{\mu}$ and variance parameters, we use additional survey questions on where households would rank hypothetical schools with given observables that they have not yet found. See the Estimation Appendix.

value school quality and do not like prices. Low SES parents are relatively more responsive to prices, and relatively less responsive to quality. In addition, parents systematically underestimate the match value of the schools they do not know well ($\mu^l < 0$), implying that parents systematically update positively when learning a school well that had been known "by name".

Consistent with descriptive evidence, schools that are located closer to the household, and highlight-worthy (low price, high quality) schools are more likely to be known by parents, as are schools that parents have explored in the past. Parents who received information treatments are in general more likely to know schools as well. When a high SES household knows a school "well", it learns the true x's with probability 27%. For low SES households this probability is 5.4%.

At the bottom of panel A, we show the estimates for the parameters that govern subjective admission chances. High and low SES parents' beliefs exhibit similar optimism and compression on average. However, the standard deviation of the "compression" term, σ_{o1}, indicates substantial heterogeneity, with some households' beliefs more extreme than the truth.

Panel B describes the variance-covariance matrix for the estimated random coefficients. For all attributes, there is heterogeneity in preferences that is not explained by observable characteristics. Unobserved price and distance sensitivity are positively correlated, which implies that price-sensitive families are also less willing to travel. The covariance between price and quality and distance and quality is also positive, which implies that parents who value school quality tend to be less price and distance-sensitive.

Panel D shows the means and variance-covariance matrix for the school level random effects that enter the subjective utility and knowledge level equations. The intercept for the expected mean utilities is negative for both types of parents, and more expensive and higher quality schools have higher expected mean utilities. Schools with higher prices are more easily discovered by both types of parents. For quality, the sign differs: high quality schools have higher (lower) expected discoverability for high (low) SES parents. The positive covariance parameter implies that more desirable (high δ) schools are more likely to be known at baseline (high

η) and that the schools that families may find via search have lower mean utilities on average than those already known.[43]

Panel E shows search cost estimates for pin and profile (detailed) clicks. The first section of the panel presents the results for the deterministic component of the cost of clicking on a school. To allow selection on baseline knowledge, this is a linear function of a constant and two cost shifters (mean π_{i0} and the probability of non-placement given π_{i0}. The deterministic component of the cost of the pin clicks is negative and substantially larger in magnitude for the low SES parents, as shown in Figure A.5a, but variances are larger as well.[44]

Figure A.4 shows the estimated distortion functions. Perceived distributions of price and quality are compressed relative to the truth. In addition, both low- and high-SES parents systematically overestimate quality. While low-SES households tend to overestimate prices relative to true values, high SES households do not. This pattern may be due to some low SES families' being unaware of their targeted voucher eligibility, which makes most schools free for them.

Figure A.6 shows that our model fits the data well. The sub-figures compare the distribution of observed and predicted characteristics of the choices that parents make, as well as placement probabilities and search behavior (pin and detailed clicks), capturing key patterns in the choices that parents make. We also compare the distribution of latent variables (utility and components of the information index) conditional on the observed search actions and applications to the "ex-ante" distributions produced by simulating preference shocks, random coefficients, latent information terms, search, and application decisions conditional on exogenous characteristics only.

8.2. Counterfactuals

Table V provides a summary of the counterfactual simulations, pooling results across low SES and high SES households. In each counterfactual, we change some

[43] We plot the distribution of mean utility and discoverability in Figure A.3, showing a positive relationship that is steeper for high-SES households.

[44] We treat the first search differently. There is a separate mean and variance for the cost of the first click, see the Online Appendix for details. Our estimates indicate that, conditional on this click, high-SES households' decisions are more responsive to perceived values while low-SES households' decisions are more random.

TABLE IV

SELECTED MODEL ESTIMATES

	Param.	Low SES Coeff.	Low SES Std Err.	High SES Coeff.	High SES Std Err.
Panel A: Preferences and Information					
Subjective Expected Utility Parameters					
Perceived Price	γ_p	-0.534	(0.070)	-0.335	(0.168)
Perceived Quality	γ_q	0.652	(0.048)	0.814	(0.079)
Knowledge Shifters (Selected)					
Distance	α_z	-0.146	(0.029)	-0.116	(0.018)
Treatment 1	α_w	0.139	(0.121)	0.448	(0.128)
Treatment 2	α_w	0.195	(0.060)	0.198	(0.237)
Highlight-worthy	α_w	0.362	(0.102)	0.258	(0.048)
Single Click	α_w	0.933	(0.180)	0.433	(0.059)
Double Click	α_w	1.369	(0.228)	0.772	(0.072)
Pr(learn true x's | know well)	p^h	0.054	(0.012)	0.274	(0.032)
Subjective Admission Chances Parameters					
Optimism (mean)	μ_{o0}	0.685	(0.008)	0.683	(0.019)
Optimism (sd)	σ_{o0}	0.197	(0.011)	0.216	(0.033)
Compression (mean)	μ_{o1}	0.160	(0.012)	0.262	(0.049)
Compression (sd)	σ_{o1}	0.232	(0.039)	0.298	(0.049)

	Param.	Low SES Coeff.	Low SES Std Err.	High SES Coeff.	High SES Std Err.
Panel B: Variance Covariance of Random Coefficients (Utility)					
Variances					
Constant	σ_c	15.424	(1.890)	14.517	(1.507)
Distance	σ_d	0.535	(0.014)	0.698	(0.036)
Price	σ_p	1.790	(0.099)	1.238	(0.126)
Quality	σ_q	0.547	(0.120)	0.535	(0.097)
Covariances					
Constant-Distance	$\sigma_{c,d}$	-1.506	(0.069)	-1.885	(0.217)
Constant-Price	$\sigma_{c,p}$	-3.657	(0.281)	-2.657	(0.383)
Constant-Quality	$\sigma_{c,q}$	-2.222	(0.448)	-2.015	(0.335)
Distance-Price	$\sigma_{d,p}$	0.202	(0.030)	0.067	(0.048)
Distance-Quality	$\sigma_{d,q}$	0.115	(0.022)	0.169	(0.051)
Price-Quality	$\sigma_{p,q}$	0.182	(0.057)	0.104	(0.087)
Panel C: Match Value Shocks ε_{ij} Primitives					
Mean Subjective Expectation of ε_{ij} given $\tilde{\varepsilon_{ij}}$	μ_I	-0.413	(0.036)	-0.592	(0.087)
Variance Covariance of Errors (Σe)					
1	1	0.190	(0.026)	0.880	(0.153)
	2	0.045	(0.009)	0.189	(0.020)
2	2	0.618	(0.057)	1.414	(0.149)

	Param.	Low SES Coeff.	Low SES Std Err.	High SES Coeff.	High SES Std Err.
Panel D: Random Effects (δ and η)					
Coefficients on Mean Elements (Mean Utility)					
Constant	$\bar{\beta}$	-2.284	(0.094)	-4.193	(0.188)
Price	β_p	0.198	(0.033)	0.658	(0.130)
Quality	β_q	0.340	(0.023)	0.419	(0.051)
Coefficients on Mean Elements (Discoverability)					
Constant	$\bar{\alpha}$	-0.172	(0.056)	-0.599	(0.081)
Price	α_p	0.189	(0.043)	0.315	(0.016)
Quality	α_q	-0.015	(0.058)	0.059	(0.027)
Variances					
Mean Utility	σ_δ	0.522	(0.108)	0.938	(0.133)
Discoverability	σ_η	0.488	(0.131)	0.433	(0.035)
Covariances					
M. Utility - Discov.	$\sigma_{\delta,\eta}$	0.242	(0.052)	0.250	(0.041)

	Param.	Low SES Coeff.	Low SES Std Err.	High SES Coeff.	High SES Std Err.
Panel E: Search Costs					
Pin Click					
Std. dev. of $\bar{c}_i^{pin}$	$\sigma_{\bar{c}}$	6.723	(1.752)	1.223	(0.935)
Std. dev. of the shock	σ^{pin}	8.835	(2.187)	2.188	(0.950)
Intercept - first	$\bar{c}_0$	13.500	(54.153)	0.952	(0.705)
Std. dev. of the shock - first	σ_0^{pin}	0.771	(1.762)	1.452	(0.994)
Coefs on xc					
Constant	γ^{cost}	-10.493	(2.512)	-3.283	(0.565)
Mean π	γ^{cost}	1.778	(0.574)	0.768	(0.125)
Pr Place i	γ^{cost}	-1.711	(0.644)	-0.206	(0.176)
Profile Click (Detailed View)					
Mean	$\bar{c}^{det.}$	1.768	(0.540)	0.692	(0.134)
Standard Deviation	$\sigma^{det.}$	0.555	(0.471)	-0.529	(0.284)

Note: This table presents results from model estimation.

aspects of the information environment and simulate search behavior, applications, allocations (holding admissions cutoffs fixed), and final enrollment for the families in the sample. We report the expected utility of the final application under the true preferences ("welfare"), placement probabilities and expected rank on the submitted application, mean characteristics of placed schools conditional on receiving a placement, search activity, and perceived returns to the first search.

Gains from full information: We begin by comparing a baseline scenario (row 1) to a full-information benchmark (row 2). In the baseline, we simulate search decisions and applications using our model. This scenario differs from the data in that we remove the effects of our information treatments and re-simulate search

TABLE V

MAIN RESULTS

		Welfare	Placement		E(School Charact)		Search (N.Clicks)		
			Place	E(rank)	Quality	VA	Single	Double	V(1st)
Gains from Full Information									
(1)	Full model baseline	0.540	0.731	1.428	2.995	0.139	3.717	1.058	0.576
		(0.019)	(0.006)	(0.010)	(0.010)	(0.006)	(0.110)	(0.034)	(0.005)
(2)	Full information	1.307	0.844	1.609	3.194	0.215	-	-	-
		(0.016)	(0.005)	(0.012)	(0.010)	(0.006)	-	-	-
(3)	Gains (difference (2)-(1))	0.767	0.113	0.181	0.199	0.076	-	-	-
		(0.025)	(0.008)	(0.016)	(0.014)	(0.008)	-	-	-
	(% Change)	142.04%	15.46%	12.68%	6.64%	54.68%	-	-	-
Decomposition: sequential correction of beliefs and misperceptions									
(4)	Better Search (S^*)	0.753	0.770	1.456	3.051	0.157	5.490	-	0.759
		(0.019)	(0.006)	(0.010)	(0.010)	(0.006)	(0.129)	-	(0.006)
(5)	$S^* + x$	1.002	0.719	1.481	3.184	0.213	5.494	-	0.748
		(0.016)	(0.006)	(0.010)	(0.010)	(0.006)	(0.129)	-	(0.006)
(6)	$S^* + x + f(x)$	1.002	0.719	1.481	3.183	0.211	5.476	-	0.752
		(0.016)	(0.006)	(0.010)	(0.010)	(0.006)	(0.129)	-	(0.006)
(7)	$S^* + x + f(x) + r$	1.004	0.720	1.481	3.184	0.212	5.540	-	0.785
		(0.016)	(0.006)	(0.010)	(0.010)	(0.006)	(0.131)	-	(0.007)
(8)	$S^* + x + f(x) + r + f(\varepsilon)$	0.999	0.745	1.505	3.185	0.212	5.367	-	0.530
		(0.016)	(0.006)	(0.011)	(0.010)	(0.006)	(0.127)	-	(0.530)
(9)	$S^* + x + f(x) + r + f(\varepsilon) + \varepsilon$	1.087	0.758	1.509	3.177	0.208	5.307	-	0.471
		(0.016)	(0.006)	(0.011)	(0.010)	(0.006)	(0.126)	-	(0.005)
Misspecified models									
(10)	No mispercept. of x ($\hat{x} = x$)								
	(Gains in outcomes relative to baseline)	0.518	0.177	0.108	-0.030	-0.022	-	-	-
	(Gains in S^* relative to baseline)	0.133	0.060	0.014	-0.008	-0.008	1.818	-	0.084
	(Gains in (9) relative to baseline)	0.277	0.099	0.042	-0.019	-0.014	1.740	-	0.018
(11)	No mispercept. of x, e if $\pi > 0$								
	(Gains in outcomes relative to baseline)	0.138	0.058	0.027	-0.006	-0.008	-	-	-

Note: This table presents the counterfactuals. Columns: Welfare: EU according to fully informed payoffs. Place: probability of placement. (E(rank),Quality, VA): avg. (rank of placed school within ROL,quality,school value added (in student-level SD)), conditional on placement. (Single, Double, V(1st)): number of single clicks, double clicks, and value of the first pin click. Rows are as follows. Full model baseline: includes all possible misperceptions and biases. Full information: $\pi_{ijT} > 1$ for all (i, j), and $\hat{x} = x$. Gains: difference in outcomes between full information and baseline. Decomposition: sequential correction of beliefs and misperceptions. Better search (S^*): search is perfectly informative. $S^* + x$: S^*+ provides full information about price and quality of known schools. $S^* + x + f(x)$: $S^* + x+$ correct distribution of school characteristics of unknown schools. $S^* + x + f(x) + r$: $S^* + x + f(x)+$ correct misperceptions about rejection chances at known schools. $S^* + x + f(x) + r + f(\varepsilon)$: $S^* + x + f(x) + r+$ correct beliefs about the distribution of match value shocks of unknown schools. $S^* + x + f(x) + r + \varepsilon + f(\varepsilon)$: $S^* + x + f(x) + r + \varepsilon+$ correct misperceptions about the match value shocks of known schools. No mispercept. of x: gains from misspecified model assuming $\hat{x}_{ijt} = x_{ij}$ relative to baseline. No mispercept. of x, e if $\pi > 0$: gains in misspecified model assuming $\hat{x}_{ijt} = x_{ij}$ and perfect learning for all schools with $\pi_{ijT} > 0$ relative to baseline (as in data)

decisions.[45] Under full information, we endow households with full information

[45]We condition on the actions that households took in the data when drawing latent terms. We adjust w_{ijt} so that all households are assigned to the "control" arms in the feedback and search experiments.

about all schools and correct all misperceptions, setting $\pi_{ijT} = 2$ and $\hat{x}^2_{ij} = x_{ij}$ for all i and j. This scenario provides an upper bound on welfare gains.

Welfare would increase by the equivalent of .767 fewer kilometers traveled, placement probability would increase by 11.3 percentage points, expected quality by 0.2 points on a 4-point scale, and value added by 0.076 student standard deviations. We do not report search activity as it is irrelevant under full information.

Gains decomposition: We next decompose the gains from full information, quantifying the relative contribution of each misperception, bias, and friction in the search process that we consider. To aid interpretation, our first step (row 4) consists of improving and simplifying the search technology. In particular, we make "pin clicks" fully informative, taking α^{pin} sufficiently large to guarantee $\pi_{ijt} = 2$ in all subsequent periods for schools j that i clicks. Specifying a single fully-informative search action makes our setting simpler and closer to others.

Under this improved search technology, we then correct biases, misperceptions, and imperfect information in the following order. We correct misperceptions of observables, setting $\hat{x}^2_{ij} = \hat{x}^1_{ij} = x_{ij}$ for all ij (row 5). We then set beliefs about the distribution of unknown schools' price and quality to their objective values (row 6), provide rational expectations about rejection chances (row 7; set $\hat{r} = r$), provide rational expectations about the distribution of match-value shocks ε and measurement error (row 8), and finally remove measurement error on match-value shocks, setting $\hat{\varepsilon}^1_{ij} = \varepsilon_{ij}$ (row 9).

Improving the search technology would cause welfare to increase substantially, by the equivalent of 0.21 kilometers. Placement rates would rise from .73 to .77, with small positive effects on the quality and value added of assigned schools. Most of the remaining welfare gains would then come from fixing misperceptions of observables. Moving from row (4) to (5), welfare would increase by about a quarter of a kilometer per household. Quality and value added of assigned schools would increase as well, by .13 points and .06 student standard deviations respectively, to nearly the level that would be attained under full information. Secondar-

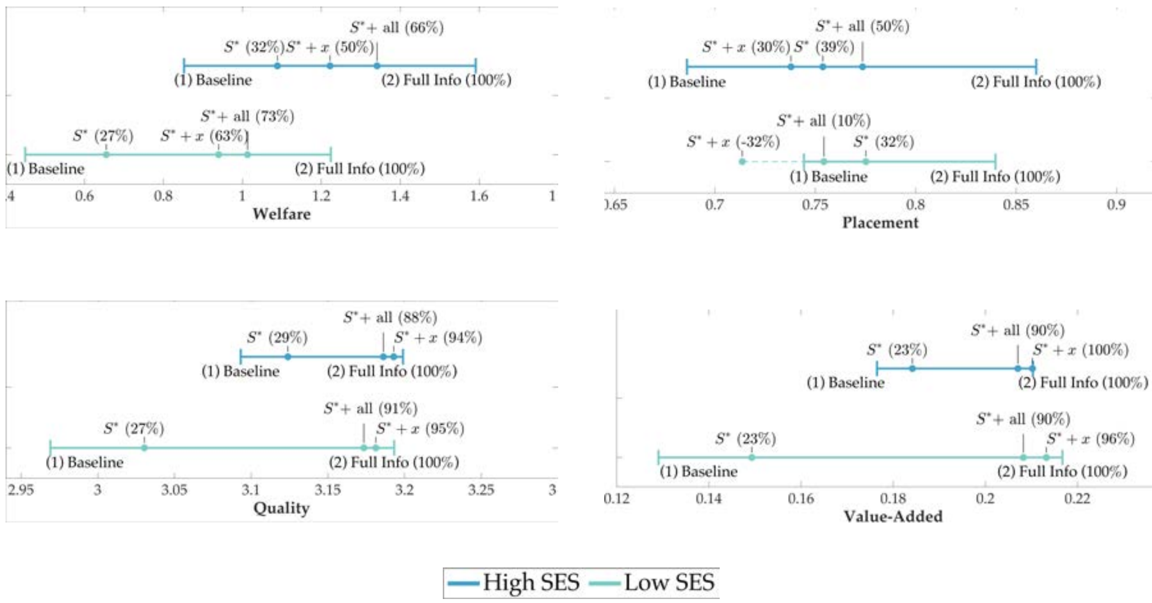


FIGURE 6.—Gains of full information. Notes: This figure shows the gains from moving from the *baseline case* to a *full information environment* using a solid horizontal line, in green for low SES and blue for high SES. Three dots show the outcomes that would be achieved under the *bettersearch* (S^*), the *bettersearch* (S^*) $+x$, and the *bettersearch* (S^*)$+x+f(x)+r+\varepsilon+f(\varepsilon)$ counterfactuals. The share of the gains under these three counterfactuals is displayed in parenthesis.

ily, providing full information about ε whenever $\pi_{ijt} > 0$ would raise welfare by about .09 kilometers.

Heterogeneity by SES: Figure 6 complements Table V, showing heterogeneity in a subset of these outcomes by SES.[46] For high SES households, improving search and fixing all information frictions achieves about 66% of the gains from full information. For low SES households, this figure is 73%. More importantly, gains in quality and value added are twice as large for low-SES households. At baseline, the quality of assigned schools is roughly .12 points lower for low-SES households than for high-SES households. Providing full information would entirely close this gap. Merely providing accurate perceptions of $\hat{x}$ would do so as well.

Search Activity: The final three columns of Table V show the number of single (pin) and double (profile) clicks as well as the average subjective value, given households' perceptions and beliefs, of the first on-platform search. Improving the search technology would raise the perceived returns and induce roughly 50%

[46]We provide further details in Appendix tables A.XV and A.XVI.

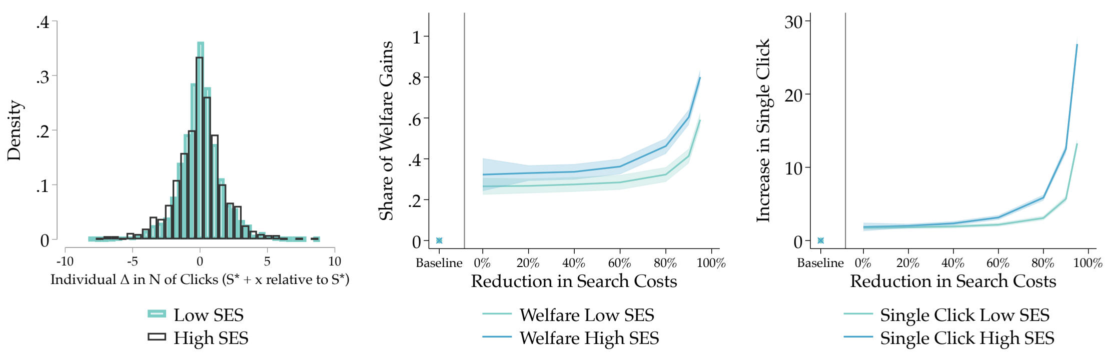
more search activity, from 3.7 clicks at baseline to 5.5 on average. However, further changes in mean search levels would be small or negative. As estimates indicate that households underestimate the mean and overestimate the variance of ε, correcting these beliefs can induce changes with either sign. We find that the variance dominates. Accurate information about the distribution of ε would reduce perceived returns to search on average (row 8). Removing measurement error on ε (row 9) would further reduce true and perceived returns to search.

These means mask important heterogeneity. Figure 7a shows the distribution of the individual change in the number of clicks when we provide correct information about x's of all the known schools on top of the better search technology. Absolute changes are large. Some families will search more once they learn the true characteristics of the schools known in the baseline, and others will search less.

Search Costs: To understand the role of search costs, we run a final sequence of counterfactuals starting from the *better search* scenario in which we gradually reduce the search costs. The main results are presented in Figure 7b and 7c. We provide more details in Table A.XVII. Consistent with the results already shown, improving the search technology achieves 30% of the welfare gains, and increases the number of schools clicked by almost two. However, once the technology is improved, large cost reductions are necessary to move families to search more and submit a portafolio with higher expected welfare. The combination of improving the search technology and reducing the search costs by 80% achieves 50% (35%) of the gains from full information for high (low) SES families. As the costs approach 0, families search all schools and their expected welfare converges to the full information benchmark.

Misspecified models: Finally, to assess the importance of modeling biased beliefs and misperceptions, we simulate counterfactuals in simpler misspecified models. In the first misspecified model, we estimate assuming $\hat{x}^2_{ij} = \hat{x}^1_{ij} = x_{ij}$ for all (i, j), as if we had not collected data on misperceptions of observables. We call this counterfactual *no misperception of x* ($\hat{x} = x$). The second misspecified model, *no misperception of x, e if* $\pi > 0$, assumes in addition that there is no measurement error on ε and no distinction between $\pi_{ijt} = 1$ and $\pi_{ijt} = 2$, as if we had pursued a standard "consideration set" approach. We compare a full-information counterfactual, simulated under these assumptions, to baseline estimates. Had we ignored misperceptions about schools' price and quality, we would have reversed the sign of estimated impacts on school quality. Providing full information about all schools would have led households to sort into schools with quality 2.97 on average, lower than at baseline.

We have estimated search costs and simulated a "better search" counterfactual under this first misspecifed model as well, further assuming that households have rational expectations over unknown x's and shocks. Compared to our main specification, this model underestimates the impact of "better search" relative to baseline. By construction, providing information about x has zero effect.



(a) Change N. of Clicks **(b)** Search Cost Reduction **(c)** Search Cost Reduction

FIGURE 7.—Panel (a) shows the individual change in number of clicks between the counterfactuals *bettersearch* $+x$ and *bettersearch*. Panel (b) shows welfare by the distribution of better search technology. Panel (c) shows single clicks by the distribution of better search technology. The shaded area indicates the 95% confidence interval.

9. CONCLUSION

This paper investigates the interactions between parents' biases and misperceptions and their information-acquisition efforts, applications under uncertainty, and assignments in a "school choice" market. We estimate a novel model of search, information, and demand for schools, using new data on parents' search activities, awareness of schools, (mis)perceptions of their characteristics and admissions chances, and beliefs about the distribution of local schools' characteristics, together with variation induced by randomized information experiments.

Our experiments and counterfactuals show that providing information before families make search decisions complements their efforts, raising welfare and assigning them to higher-quality schools. Consistent with theory, providing accurate information would have heterogeneous effects on families' search efforts, with some families engaging in greater search effort while others stop sooner.

We consider counterfactuals that provide information broadly. It may be valuable to policymakers to understand whom to treat, with what information, and at what time, however, given that information provided at the wrong time may be ignored in practice. We provide a unified framework for analyzing the impacts of information provision about strategic behavior, admissions chances, and schools' characteristics, which may be useful for addressing these questions. In addition, our model may serve an input for future research on equilibrium and supply-side behavior in these markets.

References

ABALUCK, JASON AND JONATHAN GRUBER (2016): "Evolving choice inconsistencies in choice of prescription drug insurance," American Economic Review, 106 (8), 2145–2184. [2]

ABDULKADIROGLU, ATILA, PARAG A. PATHAK, ALVIN E. ROTH, AND TAYFUN SONMEZ (2005): "The Boston Public School Match," American Economic Review, Papers and Proceedings, 95, 368–371. [5]

AGARWAL, NIKHIL AND PAOLO SOMAINI (2016): "Demand Analysis using Strategic Reports: An application to a school choice mechanisms," Working paper. [5]

AGARWAL, NIKHIL AND PAULO SOMAINI (2018): "Demand Analysis using Strategic Reports: An application to a school choice mechanism," Econometrica, 86 (2), 391–444. [5]

AGARWAL, NIKHIL AND PAULO J SOMAINI (2022): "Demand Analysis under Latent Choice Constraints," Working Paper 29993, National Bureau of Economic Research. [5]

AGARWAL, SUMIT, JOHN GRIGSBY, ALI HORTAÇSU, GREGOR MATVOS, AMIT SERU, AND VINCENT YAO (2020): "Searching for approval," Tech. rep., National Bureau of Economic Research. [2, 6]

AJAYI, KEHINDE AND MODIBO SIDIBE (2020): "School choice under imperfect information," Economic Research Initiatives at Duke (ERID) Working Paper, (294). [2]

AJAYI, KEHINDE F, WILLA H FRIEDMAN, AND ADRIENNE M LUCAS (2020): "When Information is Not Enough: Evidence from a Centralized School Choice System," Working Paper 27887, National Bureau of Economic Research. [6]

AKRESH, RICHARD, DANIEL HALIM, AND MARIEKE KLEEMANS (2023): "Long-term and intergenerational effects of education: Evidence from school construction in Indonesia," The Economic Journal, 133 (650), 582–612. [20]

ALLENDE, CLAUDIA, FRANCISCO GALLEGO, AND CHRISTOPHER NEILSON (2019): "Approximating the equilibrium effects of informed school choice," Working Paper. [6]

ANDRABI, TAHIR, JISHNU DAS, AND ASIM IJAZ KHWAJA (2017): "Report cards: The impact of providing school and child test scores on educational markets," American Economic Review, 107 (6), 1535–1563. [6]

ARTEAGA, FELIPE, ADAM J KAPOR, CHRISTOPHER A NEILSON, AND SETH D ZIMMERMAN (2022): "Smart matching platforms and heterogeneous beliefs in centralized school choice," The Quarterly Journal of Economics, 137 (3), 1791–1848. [4, 5, 6, 10, 11]

ATTANASIO, ORAZIO, TEODORA BONEVA, AND CHRISTOPHER RAUH (2022): "Parental beliefs about returns to different types of investments in school children," Journal of Human Resources, 57 (6), 1789–1825. [20]

BANDIERA, ORIANA, VITTORIO BASSI, ROBIN BURGESS, IMRAN RASUL, MUNSHI SULAIMAN, AND ANNA VITALI (2023): "The Search for Good Jobs: Evidence from a Six-year Field Experiment in Uganda," Working Paper 31570, National Bureau of Economic Research. [6]

BELLONI, ALEXANDRE, VICTOR CHERNOZHUKOV, AND CHRISTIAN HANSEN (2014): "Inference on Treatment Effects after Selection among High-Dimensional Controls," Review of Economic Studies, 81 (2), 608–650. [22]

BELOT, MICHELE, PHILIPP KIRCHER, AND PAUL MULLER (2019): "Providing advice to jobseekers at low cost: An experimental study on online advice," The Review of Economic Studies, 86 (4), 1411–1447. [6]

BERGMAN, PETER, ERIC W CHAN, AND ADAM KAPOR (2020): "Housing search frictions: Evidence from detailed search data and a field experiment," Tech. rep., National Bureau of Economic Research. [6]

BHATTACHARYA, VIVEK, JOSÉ IGNACIO CUESTA, GASTÓN ILLANES, ANA MARÍA MONTOYA, RAIMUNDO UNDURRAGA, AND GABRIELA COVARRUBIAS (2024): "Information Frictions in Mortgage Refinancing," Working Paper. [2]

BLACK, SANDRA E, PAUL J DEVEREUX, AND KJELL G SALVANES (2005): "Why the apple doesn't fall far: Understanding intergenerational transmission of human capital," American Economic Review, 95 (1), 437–449. [20]

BOEHM, EDUARD (2023): "Intermediation, Choice Frictions, and Adverse Selection: Evidence from the Chilean Pension Market," . [2]

BROWN, ZACH Y AND JIHYE JEON (2023): "Endogenous information and simplifying insurance choice," Working Paper. [2, 7]

BURGESS, SIMON, ELLEN GREAVES, ANNA VIGNOLES, AND DEBORAH WILSON (2015): "What parents want: School preferences and school choice," The Economic Journal, 125 (587), 1262–1289. [20]

CARRANZA, ELIANA, ROBERT GARLICK, KATE ORKIN, AND NEIL RANKIN (2022): "Job search and hiring with limited information about workseekers' skills," American Economic Review, 112 (11), 3547–3583. [6]

COHODES, SARAH R, SEAN P CORCORAN, JENNIFER L JENNINGS, AND CAROLYN SATTIN-BAJAJ (2022): "When do informational interventions work? Experimental evidence from New York City high school choice," Educational Evaluation and Policy Analysis, 01623737231203293. [6]

CORCORAN, SEAN P, JENNIFER L JENNINGS, SARAH R COHODES, AND CAROLYN SATTIN-BAJAJ (2018): "Leveling the Playing Field for High School Choice: Results from a Field Experiment of Informational Interventions," Working Paper 24471, National Bureau of Economic Research. [6]

CORRADINI, VIOLA (2024): "Information and Access in School Choice Systems: Evidence from New York City," Working Paper. [6]

CORRADINI, VIOLA AND CLEMENCE IDOUX (2023): "Overcoming Racial Gaps in School Preferences: the Role of Peer Diversity on School Choice," Working Paper. [6]

CORREA, JOSE, RAFAEL EPSTEIN, JUAN ESCOBAR, IGNACIO RIOS, BASTIAN BAHAMONDES, CARLOS BONET, NATALIE EPSTEIN, NICOLAS ARAMAYO, MARTIN CASTILLO, ANDRES CRISTI, ET AL. (2019): "School choice in Chile," in Proceedings of the 2019 ACM Conference on Economics and Computation, 325–343. [5, 9, 10, 1]

CORTÉS, PATRICIA, JESSICA PAN, LAURA PILOSSOPH, ERNESTO REUBEN, AND BASIT ZAFAR (2023): "Gender differences in job search and the earnings gap: Evidence from the field and lab," The Quarterly Journal of Economics, 138 (4), 2069–2126. [6]

DE LOS SANTOS, BABUR, ALI HORTAÇSU, AND MATTHIJS R WILDENBEEST (2012): "Testing models of consumer search using data on web browsing and purchasing behavior," American Economic Review, 102 (6), 2955–2980. [6]

DE LOS SANTOS, BABUR, ALI HORTAÇSU, AND MATTHIJS R WILDENBEEST (2017): "Search with learning for differentiated products: Evidence from e-commerce," Journal of Business & Economic Statistics, 35 (4), 626–641. [6]

DINERSTEIN, MICHAEL, LIRAN EINAV, JONATHAN LEVIN, AND NEEL SUNDARESAN (2018): "Consumer price search and platform design in internet commerce," American Economic Review, 108 (7), 1820–1859. [6]

DIZON-ROSS, REBECCA (2019): "Parents' beliefs about their children's academic ability: Implications for educational investments," American Economic Review, 109 (8), 2728–2765. [20]

GOEREE, MICHELLE SOVINSKY (2008): "Limited information and advertising in the US personal computer industry," Econometrica, 76 (5), 1017–1074. [28, 29]

GURANTZ, ODED, JESSICA HOWELL, MICHAEL HURWITZ, CASSANDRA LARSON, MATEA PENDER, AND BROOKE WHITE (2021): "A national-level informational experiment to promote enrollment in selective colleges," Journal of Policy Analysis and Management, 40 (2), 453–479. [6]

HANDEL, BEN AND JONATHAN KOLSTAD (2015a): Getting the most from marketplaces: Smart policies on health insurance choices, Brookings Institution Washington, DC. [2]

HANDEL, BENJAMIN R AND JONATHAN T KOLSTAD (2015b): "Health insurance for "humans": Information frictions, plan choice, and consumer welfare," American Economic Review, 105 (8), 2449–2500. [2, 7]

HASTINGS, JUSTINE S AND JEFFREY M WEINSTEIN (2008): "Information, school choice, and academic achievement: Evidence from two experiments," The Quarterly Journal of Economics, 123 (4), 1373–1414. [6]

HE, YINGHUA, SHRUTI SINHA, AND XIAOTING SUN (2021): "Identification and Estimation in Many-to-one Two-sided Matching without Transfers," arXiv preprint arXiv:2104.02009. [5]

HODGSON, CHARLES AND GREGORY LEWIS (2023): "You can lead a horse to water: Spatial learning and path dependence in consumer search," Tech. rep., National Bureau of Economic Research. [6]

HOXBY, CAROLINE, SARAH TURNER, ET AL. (2013): "Expanding college opportunities for high-achieving, low income students," Stanford Institute for Economic Policy Research Discussion Paper, 12 (014), 7. [6]

HOXBY, CAROLINE M AND SARAH TURNER (2015): "What high-achieving low-income students know about college," American Economic Review, 105 (5), 514–517. [6]

KAPOR, ADAM, MOHIT KARNANI, AND CHRISTOPHER NEILSON (2022): "Aftermarket Frictions and the Cost of Off-Platform Options in Centralized Assignment Mechanisms," Working Paper 30257, National Bureau of Economic Research. [5]

KAPOR, ADAM J, CHRISTOPHER A NEILSON, AND SETH D ZIMMERMAN (2020): "Heterogeneous beliefs and school choice mechanisms," American Economic Review, 110 (5), 1274–1315. [5]

LUFLADE, MARGAUX (2017): "The value of information in centralized school choice systems," Working Paper. [6]

MCCALL, JOHN JOSEPH (1970): "Economics of information and job search," The Quarterly Journal of Economics, 84 (1), 113–126. [5]

MCCULLOCH, ROBERT AND PETER E ROSSI (1994): "An exact likelihood analysis of the multinomial probit model," Journal of Econometrics, 64 (1-2), 207–240. [5]

MIZALA, ALEJANDRA AND MIGUEL URQUIOLA (2013): "School markets: The impact of information approximating schools' effectiveness," Journal of Development Economics, 103, 313–335. [6]

MORAGA-GONZÁLEZ, JOSÉ LUIS, ZSOLT SÁNDOR, AND MATTHIJS R WILDENBEEST (2023): "Consumer search and prices in the automobile market," The Review of Economic Studies, 90 (3), 1394–1440. [6]

NEILSON, CHRISTOPHER (2021): "Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students," Working Paper. [14, 1]

OREOPOULOS, PHILIP (2006): "Estimating average and local average treatment effects of education when compulsory schooling laws really matter," American Economic Review, 96 (1), 152–175. [20]

RICHBURG-HAYES, LASHAWN, CAITLIN ANZELONE, AND NADINE DECHAUSAY (2017): "Nudging change in human services: Final report of the Behavioral Interventions to Advance Self-Sufficiency (BIAS) project," OPRE Report, 23. [27]

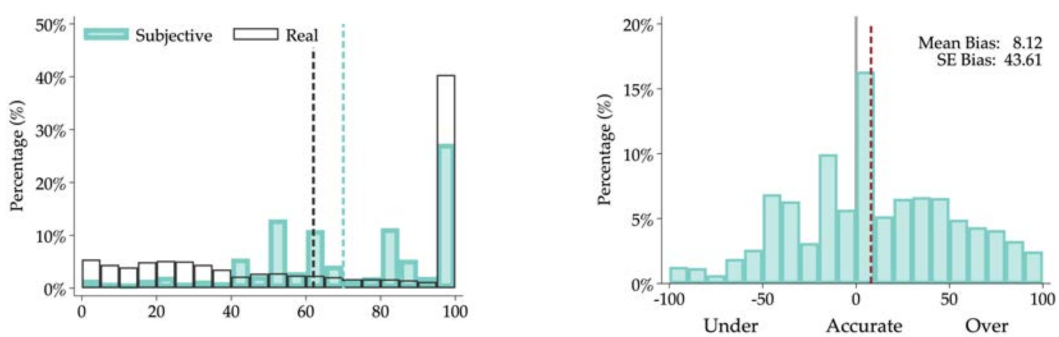
SORENSEN, ALAN T (2000): "Equilibrium price dispersion in retail markets for prescription drugs," Journal of Political Economy, 108 (4), 833–850. [2, 6]

WEITZMAN, MARTIN (1978): Optimal search for the best alternative, vol. 78, Department of Energy. [5]

ONLINE APPENDIX

A.1. ADDITIONAL DESCRIPTIVE RESULTS

In this section, we present additional descriptive results. Figure A.1 plots differences in perceived and true admission chances. We find that beliefs are biased upwards on average and exhibit compression. Table A.I shows that parental beliefs about the availability of schools affect search effort. Consistent with sequential search, we find in Table A.II that parents are more likely to stop searching when the last searched school is highlight-worthy. Tables A.III, A.IV, and A.V show that behavior on the school explorer platform affects the knowledge, beliefs, and perceptions of parents. Clicking on a school makes it more likely that a parent knows a school well (Table A.III) and that a parent's perceptions about school quality and admission chances are correct (Table A.IV). We also find that a parent's experience on the school explorer platform affects beliefs about the distribution of schools' characteristics. In Table A.V, we find that parents who clicked on more highlight-worthy schools in the explorer also report an increase in the perceived number of highlight-worthy schools in the midline survey.



(a) Distribution of Placement Chances **(b)** Bias of Placement Chances

FIGURE A.1.—Error in placement chances. Notes: Panel (A) shows the perceived and real distribution of placement chances for the school listed as first preference at baseline. If a school offers more than one program, placement chances are calculated according to the most common program offered by the school. The dotted lines indicate the mean values for each distribution. Panel (B) shows the bias on perceived placement chances of the first preference school at baseline, measured as perceived placement chances minus real placement chances. Positive values indicate that the parent responded a higher placement chance than real and negative values indicate that the parent responded a lower placement chance than real. The red dashed line indicates the mean bias and the grey solid line indicates the point of zero bias.

TABLE A.I

EFFECTS OF PERCEIVED VS REAL SCHOOL AVAILABILITY ON SEARCH EFFORT IN CONTROL GROUP

	Number of Pin Clicks		
	Mean	50th Pct.	75th Pct.
	(1)	(2)	(3)
Panel A: Number of High-Quality Schools			
Perceived Number	0.099	0.136**	0.440***
	(0.098)	(0.069)	(0.135)
True Number	0.306***	0.159***	0.480***
	(0.069)	(0.055)	(0.109)
Panel B: Number of Low-Price Schools			
Perceived Number	0.173*	0.145**	0.299**
	(0.094)	(0.064)	(0.128)
True Number	0.145***	0.081**	0.230***
	(0.044)	(0.033)	(0.067)
Mean/Percentile	8.01	3.00	10.00
Observations	1,027	1,027	1,027

Note: This table shows how perceived and true school availability affects search effort in the control group. Column 1 regresses perceived and real number of schools on mean number of pin clicks in the school explorer. Column 2 is a quantile regression for the 50th percentile of pin clicks and column 3 is a quantile regression for the 75th percentile of pin clicks.

TABLE A.II

EFFECTS OF SEARCH HISTORY ON STOPPING DECISION

Outcome:	Stopped Searching		
Sample Restriction:	Pooled	High SES	Low SES
	(1)	(2)	(3)
School is Highlight-worthy	0.031***	0.040**	0.028**
	(0.010)	(0.020)	(0.012)
Mean of Outcome	0.111	0.110	0.111
Observations	4210	1070	3122

Note: This table shows how search history affects the stopping decision. The sample consists of parent-school pin click observations in the control group. All columns regress an indicator variable for whether the parent stopped searching after this school pin click, the true number of total schools, the true number of highlight-worthy schools, and fixed effects for how many school pin clicks the parent made until this point. Column 1 uses all observations. Column 2 is restricted to high SES parents and column 3 is restricted to low SES parents.

TABLE A.III

PLATFORM BEHAVIOR AFFECTS SCHOOL KNOWLEDGE

Outcome:	Knows School Well at Midline		
Sample Restriction:	Unconditional	Unknown at Baseline	Known by Name at Baseline
	(1)	(2)	(3)
School Pin (Single) Click	0.048***	0.038*	0.053
	(0.019)	(0.020)	(0.051)
School Profile (Double) Click	0.124***	0.107**	0.217***
	(0.028)	(0.049)	(0.061)
Baseline - Knows By Name	0.285***		
	(0.022)		
Baseline - Knows Well	0.718***		
	(0.021)		
Mean of Outcome	0.270	0.050	0.369
Observations	2412	1429	499

Note: This table presents the regression of search behavior on school knowledge levels in the midline survey. Robust standard errors are reported in parentheses. Column 1 uses all observations. Column 2 is restricted to schools the parents did not know in the baseline survey. Column 3 is restricted to schools the parents only knew by name in the baseline survey. The sample is restricted to control group parents who opened the school explorer platform.

TABLE A.IV

SEARCH HISTORY AFFECTS THE PERCEPTION OF SCHOOL ATTRIBUTES

	Price		Quality		Pr Admission Chance	
	Absolute Value	Correct	Absolute Value	Correct	Absolute Value	Correct
	(1)	(2)	(3)	(4)	(5)	(6)
Single Click	-0.062	0.036	-0.095**	0.118***	-0.035	0.125**
	(0.056)	(0.047)	(0.047)	(0.041)	(0.028)	(0.052)
Double Click	-0.050	0.054	-0.108**	0.074*	-0.001	-0.056
	(0.059)	(0.050)	(0.050)	(0.043)	(0.027)	(0.048)
Outcome at Baseline	0.226***	0.305***	0.396***	0.343***	0.385***	0.257***
	(0.038)	(0.041)	(0.034)	(0.037)	(0.050)	(0.058)
Truth	0.050*	-0.033	-0.219***	0.147***	-0.095**	0.129**
	(0.027)	(0.023)	(0.030)	(0.024)	(0.038)	(0.065)
Mean of Outcome	0.349	0.681	0.521	0.548	0.328	0.237
Observations	470	470	599	599	355	355

Note: This table shows how clicking a school in the explorer affects the perceptions of price (Columns 1-2), quality (Columns 3-4), and probability of being admitted (Columns 5-6). Robust standard errors are reported in parentheses. Columns 1, 3 and 5 represent the absolute difference between the perceived and actual value. Columns 2, 4 and 6 are indicators if the parent's perceptions are correct. For admission chances, we consider the answer to be correct if the absolute difference between the perceived and actual value is not more than 10 percentage groups. Each regression also controls for the outcome variable measured at baseline as well as the true school attribute. The sample is restricted to parent-school observations in the control group who opened the school explorer platform.

TABLE A.V

PLATFORM BEHAVIOR AFFECTS BELIEFS

	Midline Survey	
	Perceived Number of Schools	Perceived Number of Highlight-worthy Schools
	(1)	(2)
Explorer Experience	0.113**	0.100**
	(0.051)	(0.041)
Baseline Perception	0.268***	0.158***
	(0.057)	(0.031)
Actual Value	0.069***	0.033*
	(0.018)	(0.017)
Mean of Outcome	6.242	1.907
Observations	550	537

Note: This table shows how explorer experiences affect perceptions in the midline survey. Robust standard errors are reported in parentheses. Column (1) regresses the perceived number of schools within 2km in the midline survey on the number of school profile clicks in the explorer, the perceived number of schools within 2km in the baseline survey, and the actual number of schools within 2km. Column (2) regresses the perceived number of highlight-worthy schools within 2km in the midline survey on the number of highlight-worthy school profile clicks in the explorer, the perceived number of highlight-worthy schools within 2km in the baseline survey, and the actual number of highlight-worthy schools within 2km. The sample is restricted to control group parents who opened the school explorer platform and completed the midline survey (N=616).

A.2. ADDITIONAL EXPERIMENTAL RESULTS

In this section, we present additional results from the two experiments. Table A.VI examines selection into the study sample by showing differences in household characteristics between the universe of applicants and the study sample. Table A.VII shows that there was no differential attrition between treatment groups. Tables A.VIII and A.IX present balance checks for the search aid interventions and the feedback intervention, respectively. We examine the treatment effects of the search aid interventions on application outcomes in Table A.X. We observe that the second treatment arm increases (i) the share of parents who submitted an application through the SAE, (ii) the likelihood that the second-ranked school is highlight-worthy, and (iii) the likelihood that the child enrolls in the school to which the child was assigned. We investigate heterogeneity in treatment effects by perceived returns to search in the baseline survey in Figure A.2 and Table A.XI and find that the treatment effects are concentrated among parents who said that they were unlikely to search for additional information on schools. Table A.XII further presents the effects of the feedback intervention on assignment and enrollment outcomes.

TABLE A.VI

DESCRIPTIVE STATISTICS FOR THE UNIVERSE OF APPLICANTS AND THE STUDY SAMPLE

	Universe	Control Group Sample		
		All	Low SES	High SES
	(1)	(2)	(3)	(4)
N	207,578	917	695	220
Panel A: Demographics				
SEP Household	0.51	0.43	0.52	0.15
Female	0.49	0.51	0.53	0.45
Panel B: Application Behavior				
Length initial attempt	2.93	3.59	3.51	3.83
Length final attempt	2.97	3.67	3.59	3.92
Total attempts	1.05	1.08	1.08	1.10
Panel C: Placement				
Placed in pref.	0.88	0.93	0.94	0.91
Placed 1st pref.	0.64	0.61	0.63	0.51
Partic. in 2nd round	0.07	0.08	0.07	0.10
Panel D: Enrolled School				
Enrolled at some school	0.97	0.98	0.98	0.95
Enrolled at placed	0.71	0.72	0.73	0.69
Free Tuition	0.75	0.75	0.80	0.59
Insufficient Quality	0.03	0.02	0.02	0.01
Mid-Low Quality	0.19	0.16	0.17	0.13
Mid Quality	0.60	0.55	0.54	0.60
High Quality	0.18	0.27	0.27	0.26
Highlight Worthy	0.45	0.48	0.52	0.36

Note: The table shows summary statistics for the universe of applicants and control group parents in the study sample. Column 1 consists of all students who either applied to prekindergarten, kindergarten, or first grade in 2021. Column 2 consists of control group children who submitted an application and entered the explorer platform. Column 3 consists of control group children whose mother did not complete college and Column 4 consists of control group children whose mother completed college. *Length of initial/final attempt* is the number of programs on an applicant's first and final choice application. *Total attempts* is the number of times an applicant submitted an application to the centralized system. *Placed in pref/1st* are indicators for any placement or for the school ranked first. *2nd round* variables describe participation in the second centralized placement round.

TABLE A.VII

ATTRITION CHECK

	Midline Survey (1)	Endline Survey (2)	Endline Survey (3)
Treatment 1	0.024	0.002	
	(0.022)	(0.016)	
Treatment 2	-0.021	0.006	
	(0.022)	(0.016)	
Feedback Treatment			-0.002
			(0.014)
Control Group Mean	0.532	0.147	0.154
Observations	3,111	3,111	3,055

Note: This table shows attrition rates by treatment status. The outcome in Column 1 is an indicator variable for whether the respondent completed the midline survey and the outcomes in Columns 2 and 3 are indicator variables for whether the respondent completed the endline survey. The sample is restricted to parents who opened the school explorer platform.

TABLE A.VIII

BALANCE CHECKS FOR SEARCH INTERVENTIONS

	Control		Treatment 1		Treatment 2		
	Mean	St. Dev.	Coeff.	St. Err.	Coeff.	St. Err.	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Choice Environment							
Number of available schools	16.222	[9.169]	-0.142	(0.279)	-0.076	(0.279)	3948
Number of available highlight-worthy schools	8.640	[5.014]	-0.110	(0.166)	0.031	(0.164)	3948
Panel B: Parent/Child Characteristics							
Child is female	0.495	[0.500]	0.016	(0.019)	0.023	(0.019)	3948
Child's birthyear	2017.096	[0.550]	0.012	(0.021)	0.014	(0.021)	3948
Mother completed college	0.220	[0.414]	0.018	(0.013)	-0.003	(0.013)	3945
Number of younger siblings	1.146	[0.387]	0.017	(0.015)	0.008	(0.015)	3948
Child has a disability (belief)	0.070	[0.255]	0.007	(0.011)	-0.004	(0.010)	3528
Parent works in a school	0.066	[0.249]	-0.002	(0.010)	-0.007	(0.009)	3885
SEP household	0.450	[0.498]	-0.015	(0.014)	-0.016	(0.014)	3908
Panel C: Initial Knowlege and Beliefs							
Expected satisfaction with process	5.235	[1.395]	0.049	(0.056)	-0.018	(0.057)	3689
Listed any school as first preference	0.909	[0.288]	0.001	(0.011)	0.008	(0.011)	3948
First-preference school is highlight-worthy	0.609	[0.488]	0.038*	(0.021)	0.053*	(0.020)	3245
Perceived admission change for first-preference school	0.684	[0.272]	0.013	(0.011)	0.022	(0.011)	3689
Number of schools known by name	3.301	[2.684]	-0.078	(0.103)	0.027	(0.102)	3948
Number of schools known well	1.874	[2.046]	-0.006	(0.078)	0.049	(0.079)	3948
Perceived number of available schools	7.444	[6.936]	0.016	(0.264)	-0.373	(0.261)	3948
Perceived number of available highlight-worthy schools	3.671	[3.615]	0.045	(0.136)	-0.123	(0.133)	3948
Parent believed to be SEP eligible	0.172	[0.378]	-0.000	(0.015)	-0.008	(0.014)	3948
SEP did not know about SEP status	0.665	[0.472]	-0.014	(0.018)	0.024	(0.018)	3948

Panel D: Treatment Summary	Control	Treatment 1	Treatment 2
Observations	1318	1313	1317
Whatsapp Reminder + SEP Status + Explorer	X	X	X
School Distribution		X	X
Highlight-worthy School			X

Note: This table shows balance for baseline covariates for the search aid interventions. Column 1 reports the control mean of the dependent variable for each relevant subgroup (standard deviations in brackets). Columns 3 and 5 report the difference in the dependent variable from OLS regressions of each outcome on indicator variables for treatment assignments and stratification dummies. Robust standard errors are reported in parentheses. Variables in Panel A come from administrative data. Variables in Panels B and C come from the baseline survey.

TABLE A.IX

BALANCE CHECK FOR FEEDBACK INTERVENTION

	Control		Feedback Treatment		
	Mean (1)	St. Dev. (2)	Coeff. (3)	St. Err. (4)	N (5)
Panel A: Choice Environment					
Number of available schools	15.778	[9.398]	-0.740	(0.860)	2581
Number of available highlight-worthy schools	8.355	[5.020]	0.187	(0.525)	2581
Panel B: Parent/Child Characteristics					
Child is female	0.516	[0.500]	-0.011	(0.025)	2581
Child's birthyear	2017.119	[0.500]	-0.017	(0.026)	2581
Mother completed college	0.206	[0.404]	0.023	(0.027)	2579
Number of younger siblings	1.126	[0.360]	0.032*	(0.019)	2581
Child has a disability (belief)	0.055	[0.229]	0.011	(0.013)	2300
Parent works in a school	0.063	[0.243]	0.001	(0.009)	2542
SEP household	0.455	[0.498]	-0.019	(0.026)	2581
Panel C: Initial Knowlege and Beliefs					
Expected satisfaction with process	5.273	[1.391]	-0.019	(0.064)	2435
Listed any school as first preference	0.937	[0.243]	-0.002	(0.019)	2581
First-preference school is highlight-worthy	0.647	[0.478]	0.041	(0.035)	2164
Perceived admission change for first-preference school	0.703	[0.264]	0.022*	(0.012)	2435
Number of schools known by name	3.352	[2.790]	0.095	(0.164)	2581
Number of schools known well	2.057	[2.117]	-0.023	(0.104)	2581
Perceived number of available schools	7.079	[6.207]	0.428	(0.320)	2581
Perceived number of available highlight-worthy schools	3.565	[3.086]	0.221	(0.174)	2581
Parent believed to be SEP eligible	0.165	[0.371]	-0.014	(0.015)	2581
SEP did not know about SEP status	0.676	[0.468]	0.022	(0.019)	2581
Panel D: Search Treatments					
Search Treatment 1	0.326	[0.469]	0.002	(0.023)	2581
Search Treatment 2	0.349	[0.477]	-0.004	(0.020)	2581
Observations	1395		1186		

Note: This table shows balance for baseline covariates for the feedback intervention. Column 1 reports the control mean of the dependent variable for each relevant subgroup (standard deviations in brackets). Column 3 reports the difference in the dependent variable from OLS regressions of each outcome on an indicator variable for feedback treatment assignments and market fixed effects. Standard errors clustered at the market cluster level are reported in parentheses. Variables in Panel A come from administrative data. Variables in Panels B and C come from the baseline survey.

TABLE A.X

TREATMENT EFFECTS OF SEARCH INTERVENTION ON APPLICATION OUTCOMES

			Rank 1				Rank 2					
	Submitted Application	Application Length	Highlight-worthy	Value Added	Distance	Knew School Well at Baseline	Highlight-worthy	Value Added	Distance	Knew School Well at Baseline	Not Assigned	Enrolled in Assigned School
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Pooled												
Treatment 1	0.010	-0.049	-0.011	0.007	-0.046	-0.021	0.010	0.012	-0.374	-0.034	-0.012	0.015
	(0.013)	(0.072)	(0.017)	(0.017)	(0.218)	(0.021)	(0.021)	(0.019)	(0.353)	(0.025)	(0.011)	(0.020)
Treatment 2	0.026**	0.015	0.005	0.001	-0.153	0.000	0.050**	0.016	0.174	-0.042*	-0.002	0.035*
	(0.012)	(0.071)	(0.017)	(0.017)	(0.209)	(0.020)	(0.021)	(0.019)	(0.393)	(0.026)	(0.011)	(0.019)
Control Group Mean	0.893	3.444	0.696	0.246	2.006	0.670	0.683	0.168	2.972	0.485	0.069	0.765
Observations	3111	2818	2757	2703	2818	2017	2568	2492	2633	1491	2817	2748
Panel B: Heterogeneity by SES Status												
Treatment 1 × High SES	0.007	0.025	-0.024	0.047	0.154	-0.029	0.091**	0.023	0.285	-0.021	-0.009	-0.028
	(0.023)	(0.157)	(0.037)	(0.035)	(0.458)	(0.042)	(0.046)	(0.037)	(0.725)	(0.049)	(0.027)	(0.041)
Treatment 1 × Low SES	0.012	-0.077	-0.006	-0.009	-0.117	-0.019	-0.015	0.006	-0.586	-0.036	-0.014	0.031
	(0.015)	(0.081)	(0.020)	(0.020)	(0.247)	(0.024)	(0.024)	(0.023)	(0.401)	(0.029)	(0.012)	(0.022)
Treatment 2 × High SES	-0.005	0.060	0.012	0.035	-0.076	0.019	0.166***	0.031	0.622	0.016	-0.009	0.012
	(0.025)	(0.151)	(0.036)	(0.038)	(0.455)	(0.041)	(0.046)	(0.038)	(0.753)	(0.052)	(0.028)	(0.041)
Treatment 2 × Low SES	0.036**	0.002	0.002	-0.010	-0.180	-0.006	0.014	0.010	0.037	-0.057*	-0.000	0.042*
	(0.014)	(0.080)	(0.020)	(0.019)	(0.238)	(0.023)	(0.023)	(0.022)	(0.468)	(0.029)	(0.012)	(0.022)
p-value: Treat 1 x High SES = Treat 1 x Low SES	0.853	0.567	0.662	0.172	0.602	0.835	0.042	0.696	0.292	0.796	0.887	0.206
p-value: Treat 2 x High SES = Treat 2 x Low SES	0.158	0.733	0.794	0.289	0.839	0.598	0.003	0.644	0.517	0.218	0.783	0.514
Control Group Mean (High SES)	0.909	3.695	0.585	0.262	2.081	0.691	0.552	0.229	2.766	0.504	0.091	0.762
Control Group Mean (Low SES)	0.888	3.368	0.732	0.242	1.986	0.663	0.724	0.150	3.036	0.478	0.062	0.766
Observations 1 (High SES)	732	670	659	639	670	471	609	585	620	361	669	643
Observations 2 (Low SES)	2376	2146	2096	2062	2146	1544	1957	1905	2011	1129	2146	2103

Note: This table presents the results of the search interventions on application outcomes. Columns 1-2 refer to an indicator variable of application submitted and application length, Columns 3-6 refer to characteristics of the first ranked school in the application, and Columns 7-10 refer to the second ranked school in the application. In Panel A, we regress each outcome on indicator variables for both treatment arms, stratification dummies and baseline controls selected by LASSO. In Panel B, we further include the fully interacted effects of treatments and SES status. SES status is proxied by whether the mother completed college. Continuous outcomes are top-coded at the 99th percentile. The sample is restricted to parents who opened the school explorer platform.

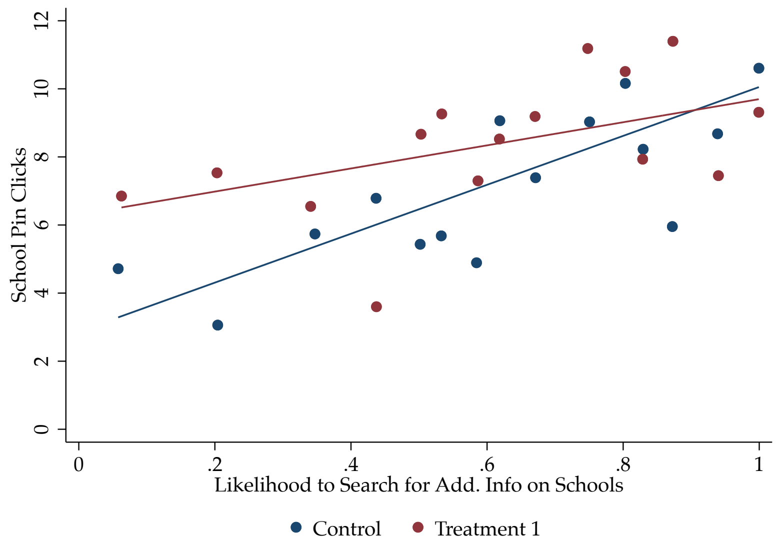


FIGURE A.2.—Treatment effects of search intervention by baseline likelihood to search. Notes: This figure plots the number of school pin clicks against the reported likelihood to search for additional information on schools in the baseline survey, separately for control group (blue) and treatment group 1 (red) parents.

TABLE A.XI

TREATMENT EFFECTS OF SEARCH INTERVENTION BY BASELINE LIKELIHOOD TO SEARCH

	Perceived Number of Schools		Number of Pin Clicks		Number of Schools Known	Enrolled School		
	All	Highlight-worthy	All	Highlight-worthy	At Least by Name	Highlight-worthy	Value Added	Distance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel B: Heterogeneity by SES Status								
Treatment 1 × Unlikely to Search	1.131***	0.549***	1.732***	0.767***	0.076	-0.017	0.018	-0.244
	(0.426)	(0.177)	(0.663)	(0.283)	(0.295)	(0.029)	(0.028)	(0.375)
Treatment 1 × Likely to Search	0.589	0.368**	-0.195	-0.172	0.347	0.003	-0.015	-0.077
	(0.420)	(0.173)	(0.829)	(0.346)	(0.296)	(0.030)	(0.027)	(0.376)
Treatment 2 × Unlikely to Search	0.781**	0.375**	0.413	0.455*	-0.097	0.006	-0.029	-0.827**
	(0.398)	(0.168)	(0.586)	(0.260)	(0.261)	(0.028)	(0.027)	(0.338)
Treatment 2 × Likely to Search	0.519	0.234	-0.732	0.181	0.111	0.029	-0.005	-0.278
	(0.433)	(0.180)	(0.777)	(0.345)	(0.287)	(0.029)	(0.027)	(0.339)
p-value: Treat 1 x Unlikely to Search = Treat 1 x Likely to Search	0.364	0.467	0.070	0.036	0.517	0.641	0.400	0.755
p-value: Treat 2 x Unlikely to Search = Treat 2 x Likely to Search	0.653	0.572	0.242	0.528	0.589	0.577	0.523	0.251
Control Group Mean (Unlikely to Search)	5.652	1.870	6.054	2.694	3.657	0.681	0.161	2.297
Control Group Mean (Likely to Search)	6.851	1.992	9.761	4.154	4.000	0.684	0.204	2.337
Observations 1 (Unlikely to Search)	779	761	1460	1460	496	1126	1097	1251
Observations 2 (Likely to Search)	789	772	1456	1456	518	1132	1098	1241

Note: This table presents the results of the search interventions on beliefs (Columns 1-2), search (Columns 3-4), knowledge (Column 5), and final school enrollment (Columns 6-8). In Panel A, we regress each outcome on indicator variables for both treatment arms, stratification dummies and baseline controls selected by LASSO. In Panel B, we further include the fully interacted effects of treatments and baseline likelihood to search. *Likely to Search* is a dummy variable that is equal to one if the maximum of the stated likelihoods to either search for additional information on (i) known or (ii) unknown schools at baseline is above the sample median. Continuous outcomes are top-coded at the 99th percentile. The sample is restricted to parents who opened the school explorer platform.

TABLE A.XII

TREATMENT EFFECTS OF FEEDBACK INTERVENTION ON ASSIGNMENT AND ENROLLMENT OUTCOMES

	Assignment			Enrollment			
	Not Assigned	Assigned to Added School	Assigned to Highlight-worthy School	Enrolled in Assigned School	Highlight-worthy	Value Added	Distance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Pooled							
Open Feedback	0.010	0.015*	0.010	-0.022	0.009	-0.039	0.676
	(0.015)	(0.008)	(0.053)	(0.031)	(0.051)	(0.050)	(1.034)
Control Group Mean	0.062	0.005	0.605	0.796	0.592	0.172	3.676
Observations	2116	2116	2116	2066	2066	1818	2063
Panel B: Heterogeneity by Parental Education							
Open Feedback × College Mother	0.029	0.014	0.028	0.024	0.028	-0.138*	2.079
	(0.037)	(0.009)	(0.068)	(0.055)	(0.078)	(0.067)	(2.062)
Open Feedback × Non-College Mother	0.006	0.016*	0.002	-0.037	0.002	-0.012	0.281
	(0.016)	(0.009)	(0.067)	(0.040)	(0.062)	(0.055)	(1.252)
p-value: Open Feedback x College Mother = Open Feedback x Non-College Mother	0.584	0.887	0.789	0.409	0.783	0.079	0.479
Control Group Mean (College Mother)	0.086	0.000	0.605	0.755	0.597	0.249	3.382
Control Group Mean (Non-College Mother)	0.055	0.006	0.605	0.806	0.591	0.152	3.751
Observations 1 (College Mother)	460	460	460	440	440	396	438
Observations 2 (Non-College Mother)	1656	1656	1656	1626	1626	1422	1625

Note: This table presents the results of the feedback intervention on assignment and enrollment outcomes. In Column 1 in Panel A, we regress the outcome on an indicator variable for whether the parent was assigned to the feedback treatment group. In Columns 2-7 in Panel A, we regress each outcome on an indicator variable for whether the respondent opened the feedback information instrumented by whether the parent was assigned to the feedback treatment group. In Panel B, we further include the fully interacted effects of the treatment and SES status. All regressions control for market fixed effects and application risk groups. SES status is proxied by whether the mother completed college.

A.3. MODEL APPENDIX

Beliefs about match quality: Households in our model use Equation (13) for three purposes: (1) to compute posterior means given their signal in the event $\pi_{ijt} = 1$ in order to form rank-order lists; (2) to form beliefs over the distribution of ε in this case, which are relevant for the subjective expected benefit of increasing π_{ijt} from 1 to 2; and (3) to form beliefs over the distribution of ε and $\hat{\varepsilon}^1$ at schools j with $\pi_{ijt} = 0$, in order to calculate the value of discovering more about currently unknown ($\pi_{ijt} = 0$) schools.

Applying the formula for Multivariate Normal conditional distributions, the household's posterior given its "low-information" ($\pi_{ijt} = 1$) signal is:

$$(\varepsilon_{ij}|\varepsilon_{ij} + e_{ij}) \sim N\left(\tilde{\mu} + \tilde{\rho}(\varepsilon_{ij} + e_{ij} - \tilde{\mu}), (1 - \tilde{\rho})\tilde{\sigma}_{\varepsilon}^2\right), \tag{16}$$

where $\tilde{\rho} = \tilde{\sigma}_\varepsilon^2 / \left(\tilde{\sigma}_\varepsilon^2 + \tilde{\sigma}_e^2\right)$ denotes the subjective informativeness of the "low-information" signal. From Equation (16) we have $\hat{E}(\varepsilon_{ij}|\varepsilon_{ij}+e_{ij}) = \tilde{\mu} + \tilde{\rho}(\varepsilon_{ij}+e_{ij}-\tilde{\mu})$. Hence, for j with $\pi_{ijt}=1$, household i believes $\varepsilon_{ijt} \sim N(\hat{\varepsilon}_{ij}^1, (1-\tilde{\rho})\tilde{\sigma}_\varepsilon^2)$. Observe that this subjective belief is biased whenever $\tilde{\mu} \neq 0$. An analogous calculation shows that, for unknown schools j, households believe (unconditionally) that $\varepsilon_{ij} \sim N(\tilde{\mu}, \tilde{\sigma}_\varepsilon^2)$ and $\hat{\varepsilon}_{ij}^1 \sim N(\tilde{\mu}, \frac{\tilde{\sigma}_\varepsilon^2}{\sqrt{\tilde{\sigma}_\varepsilon^2+\tilde{\sigma}_e^2}})$.

To solve for the reduced-form parameters, we substitute the objective joint distribution into Equation (16), obtaining:

$$\begin{pmatrix} E(\varepsilon_{ij}|\varepsilon_{ij}+e_{ij}) \\ \varepsilon_{ij} \end{pmatrix} \sim N\left(\begin{pmatrix} \tilde{\mu}(1-\tilde{\rho}) \\ 0 \end{pmatrix}, \begin{pmatrix} \tilde{\rho}^2(\sigma_\varepsilon^2+\sigma_e^2) & \tilde{\rho}\sigma_\varepsilon^2 \\ \tilde{\rho}\sigma_\varepsilon^2 & \sigma_\varepsilon^2 \end{pmatrix}\right).$$

A.4. ESTIMATION APPENDIX

A.4.1. Identification and Estimation of Second-Stage Parameters

Identification and Estimation of Beliefs about Match Quality: The terms $(\tilde{\mu}, \tilde{\rho}, \sigma_\varepsilon^2, \sigma_e^2)$ are identified from the reduced-form parameters estimated in step 1. We have

$$\sigma_\varepsilon^2 = \Sigma_{[2,2]}^\varepsilon, \;\; \tilde{\rho} = \frac{\Sigma_{[1,2]}^\varepsilon}{\Sigma_{[2,2]}^\varepsilon}, \;\; \tilde{\mu} = \mu^\ell/(1-\tilde{\rho}), \;\; \text{and } \sigma_e^2 = \frac{\Sigma_{[1,1]}^\varepsilon - \tilde{\rho}^2\sigma_\varepsilon^2}{\tilde{\rho}^2}. \tag{17}$$

To separately identify the subjective utility-shock variance $\tilde{\sigma}_\varepsilon^2$ and subjective measurement error variance $\tilde{\varepsilon}_e^2$ we need additional data. We use two baseline survey questions designed for this purpose. The first asks, if the parents were to discover an additional school with quality 4, zero price, and a distance of 2 kilometers from their house, with what probability would they add it to the top two places in their rank-order list. The second asks the same question, but the hypothetical school's characteristics are randomized: either it is made more expensive, or the quality is reduced by an increment. We elicit probabilities in $[0,1]$.

To map these questions to the model, we use the household's current state and utility draws; we assume that the household is reporting the subjective probability, denoted $\hat{P}r(u > u_{i21})$, that this random school will give higher utility than the second-highest currently known school, letting u_{i21} be the second-highest utility among known ($\pi_{ijt} > 0$) schools at time $t = 1$. For intuition, if households overestimate the variance of the shocks, they will tend to overestimate this probability when u_{i21} is high.

Letting (x, z) denote the hypothetical school's characteristics, we have:

$$\hat{Pr}(u > u_{i21}|x, z, u_{i21}) = 1 - \Phi(u_{i21}, \tilde{\mu} - z + x^{rc}\beta_i^x + x\overline{\beta} + x\gamma, \tilde{\sigma}_\varepsilon^2 + \Sigma_{[1,1]}^{\delta\eta}),$$

where $\Phi(a, \mu, \sigma^2)$ is the CDF of a Normal(μ, σ^2) distribution evaluated at a. The terms $x\overline{\beta}$ and $\Sigma_{[1,1]}^{\delta\eta}$ come from the distribution of mean utilities δ. The remaining terms come from substituting x, z, and the parameters of the subjective distribution of ε into expression (3), the expression for high-information subjective expected utility.

Because we have two survey questions, we are able to account for measurement error. For the two questions $m = 1, 2$, we allow measurement error $\nu^{\text{survey},\varepsilon,m} \sim N(0, \sigma_{\nu,\varepsilon}^2)$, iid across people and questions. Parents report

$$\hat{\hat{P}}_{ix} \equiv \max\{0, \min\{1, \hat{Pr}(u > u_{i21}|x_{im}, z, u_{i21}) + \nu^{\text{survey},\varepsilon,m}\}\}.$$

To estimate, we first obtain estimates of $(\tilde{\mu}, \tilde{\rho}, \sigma_\varepsilon^2, \sigma_e^2)$ by substituting our first-step estimates $\hat{\Sigma}^\varepsilon$ and $\hat{\mu}^\ell$ into Equation (17). Let $\hat{\tilde{\mu}}, \hat{\tilde{\rho}}, \hat{\sigma}_\varepsilon^2, \hat{\sigma}_e^2$ denote these estimates. With them in hand, as well as draws of u_{i21} from our first-stage MCMC procedure, we estimate the remaining parameters, and the variance of the measurement error, by maximizing the likelihood of reported beliefs $\hat{\hat{P}}_{ix}$ subject to the constraint that $\hat{\tilde{\rho}} = \frac{\hat{\tilde{\sigma}}_\varepsilon^2}{\hat{\tilde{\sigma}}_\varepsilon^2 + \hat{\tilde{\sigma}}_e^2}$.

Identification and Estimation of Beliefs about x: In modeling search decisions, the relevant beliefs are over characteristics of unknown schools within 5 kilometers. However, our survey questions elicit beliefs over the characteristics of all nearby schools, known and unknown. Our approach is to assume that households may probabilistically recall known schools—in which case they think of their subjective perceptions of these schools' characteristics—in addition to unknown schools. We capture survey measurement error via imperfect recall of "known" schools, and by letting households stochastically draw a set of unknown schools given their beliefs about the number of such schools and their characteristics.

In particular, households believe that there are N_{it}^{unknown} not-yet-discovered schools in their neighborhood at time t. We assume $N_{it}^{\text{unknown}} \sim \text{Poisson}(\lambda_{it}^{\text{total}})$, where $\lambda_{it}^{\text{total}} = \lambda_0^{\text{total}} + \lambda_1^{\text{total}} * (N^{\text{true}} - |\{j : \pi_{ijt} > 0\}| + \lambda_2^{\text{total}} * \text{treat}_{it} * (N^{\text{true}} - |\{j : \pi_{ijt} > 0\}|)$. The term treat_{it} is an indicator for having received search treatment 1 or 2 by time t. This expression allows beliefs about the number of schools to respond to the truth, and to allow our "search" treatments to provide information about it.

Survey s elicits the perceived number of schools by price-quality cell, $N_{is}^{\text{survey}} \in \mathbb{N}^K$. We collect this data at baseline and midline.[47] Let $t_i(s)$ denote the period in which household i takes survey s.[48] Let $N_{is} \in \mathbb{N}^K$ denote the number of "known" ($\pi_{ijt_i(s)} > 0$) schools with perceived characteristics $\hat{x}_{ij}^{\pi_{ijt_i(s)}}$ equal to the value of the kth cell.

We assume that households recall schools with probability p^{survey}, independently across schools. The number of reported "known" schools in cell k on survey s is therefore distributed $N_{is}^{\text{survey,known}} \sim \text{Binomial}(N_{isk}, p^{\text{survey}})$, independently across k conditional on the set of known schools.

The number of reported "unknown" schools is distributed $N^{\text{survey,unknown,total}} \sim \text{Binomial}(N_{it}^{\text{unknown}}, p^{\text{survey}})$. Conditional on this number, on survey s household i reports $N^{\text{survey,unknown}} \sim \text{Dirichlet}(N^{\text{survey,unknown,total}}, \alpha_{it_i(s)})$.

On survey s, we observe household i's report $N_{is}^{\text{survey}} = N_{is}^{\text{survey,unknown}} + N_{is}^{\text{survey,known}}$ schools. We estimate belief parameters and p^{survey} by maximum likelihood given this data, integrating over draws of π and $\hat{x}$ obtained in our MCMC procedure.

A.4.2. Step 1 Estimation Details

Let $\nu_{ijs}^{\text{survey}}$ denote survey measurement error in π_{ijt}^* on survey s. The household, if asked about knowledge of j on survey s at time $t_i(s)$, reports $\pi_{ijs}^{\text{survey}} = 1(\pi_{ijt_i(s)}^* + \nu_{ijs}^{\text{survey}} > 0) + 1(\pi_{ijt_i(s)}^* + \nu_{ijs}^{\text{survey}} > 1)$. Similarly, e_{ijt}^{survey} denote measurement error

[47]In the baseline survey, we elicit the number of schools in each of the 16 cells. At midline, the partition is coarser. See supplementary material for details.

[48]For the baseline survey, $t_i(s) = 1$. The timing of the midline survey varies.

on payoffs reported at $t < T$. Households rank schools that have $\pi_{ijt} > 0$ and $u_{ij}^{\pi_{ijt}} + e_{ijt}^{\text{survey}} > 0$ truthfully in order of $u_{ij}^{\pi_{ijt}} + e_{ijt}^{\text{survey}}$. There is no measurement error on the final rank-order list, which we take as reflecting the "true" preferences.

We augment the data with random coefficients $\{\alpha_i, \beta_i\}$ for all i, match-level terms $\{u_{ij}^1, u_{ij}^2, \hat{x}_{ij}^1, \hat{x}_{ij}^2, \pi_{ij1}^*, \ldots, \pi_{ijT}^*, e_{ij1}^{\text{survey}}, e_{ij2}^{\text{survey}}, \nu_{ij1}^{\text{survey}}, \ldots, \nu_{ijT}^{\text{survey}}\}$ for all i and all $j \in J_i$, and mean utilities and discoverabilities (δ_j, η_j) for all $j \in J$. In addition we track an indicator, $1(\text{misreport x})_{ijs}$, equal to 1 if household i reports $\hat{x}$ with error on survey s.

We first pick values of (u, π^*) consistent with reported rank-ordered lists. We pick starting values for $\hat{x}$, given our values π^*, setting $\hat{x}$ equal to survey responses where possible.[49] We then construct feasible measurement-error terms $e^{\text{survey}}, \nu^{\text{survey}}$, zero if possible, that are consistent with (u, π^*) and survey responses.

We then use a Gibbs sampler, iterating through the following steps:

1. Update $\overline{\alpha}, \overline{\beta} | x, \delta, \eta, \Sigma^{\delta\eta}$
2. Update $\Sigma^{\delta}\eta | x, \overline{\alpha}, \overline{\beta}\delta, \eta, \Sigma^{\delta\eta}$
3. Update $\delta, \eta | x, \hat{u}, \pi^*, \Sigma^{\delta}\eta, \overline{\alpha}, \overline{\beta}, \Sigma^{\delta\eta}$
4. Update random coefficients in utility β_i^x given $\hat{u}^1, \hat{u}^2, x, hatx, \Sigma^{rc}$
5. Update random coefficients in information α_i^{rc} given π, x, z
6. Update $\Sigma^{rc} | \beta_i^x$.
7. Update $\Sigma^{\text{rc},\pi} | \alpha_i^{rc}$
8. Update $\hat{x}^1, \hat{x}^2 | x, \hat{u}^1, \hat{u}^2, p^h$, survey responses.
9. Update $1(\text{misreport x})_{ijs} | pr(\text{misreport}), \hat{x}^1, \hat{x}^2, \hat{x}^{\text{survey}}$.
10. Update misreport-perceived-x probability $pr(\text{misreport}) | 1(\text{misreport x})$.
11. Update learning parameter $p^h | \hat{x}^1, \hat{x}^2, x$
12. Update "distortion functions" $\Gamma(\hat{x}|x) : \hat{x}, x$.
13. Update $\Sigma^{\varepsilon} | \hat{u}^1, \hat{u}^2, \beta_i, \gamma, \hat{x}, \delta$
14. Update $\Sigma^{\pi} | \pi_1^*, \ldots, \pi_T^*, \alpha, \eta, \alpha_i^{rc}$
15. Update linear information terms $\alpha^z, \alpha^w | \pi^*, \alpha_i^{rc}, \eta$.

[49] If both the midline and endline surveys take place for person i at time $t = 3$, and i reports different values of $\hat{x}_{ij}$ for some j in these two surveys, at least one of these must be due to survey measurement error.

16. Update linear utility terms $\mu^l, \gamma | \hat{u}^1, z, x, \hat{x}^1, \beta_i, \delta$.
17. Update $\hat{u}$ given surveys, other variables, measurement error, ROLs.
18. Update π^* given surveys, other variables, measurement error, $\hat{u}$, ROLs.
19. Update measurement-error terms e^{survey} given $\hat{u}$, ROLs.
20. Update ν^{survey} given π^*, survey responses.
21. Update variances of measurement-error terms $\sigma^2_{e,\text{survey}}, \sigma^2_{\nu,\text{survey}}$.

Updating linear parameters, variances, and covariance matrices are standard.

To describe how we update utilities, knowledge, and measurement errors on these objects, we must first describe the constraints imposed by the data. To state these constraints without special cases, we need the following notation: Let $e_{ijT} = 0$ for all i, j.[50] Optimality of the observed rank-order lists require information, utilities, and preference measurement error to satisfy the following constraints:

1. If j is not listed on the final rank-order list, then $\hat{u}_{ij}^{\pi_{ijt}} + e_{ijt} < 0$. This condition holds when either $\pi^*_{ijT} < 0$, $\pi^*_{ijt} \in [0, 1)$ and $\hat{u}^1_{ij} + e_{ijt} < 0$, or $\pi^*_{ijt} > 1$ and $\hat{u}^2_{ij} + e_{ijt} < 0$.
2. If j is ranked rth on the final ROL, then $\hat{u}_{ij'}^{\pi_{ij'}} + e_{ij't} > \hat{u}_{ij}^{\pi_{ij}} + e_{ijt} > \hat{u}_{ij''}^{\pi_{ij''}} + e_{ij't}$, where j' is the school ranked $(r-1)$th and j'' is the school ranked $(r+1)$th, provided that these exist.
   (a) If j is ranked first, then let $\hat{u}_{ij'}^{\pi_{ij'}} + e_{ij't} = \infty$.
   (b) If the ROL is of length r, so that j is the final option, then let $\hat{u}_{ij'}^{\pi_{ij'}} + e_{ij''t} = 0$.

If these conditions hold for t, we say that the latent variables are consistent with rank-order lists at time t.

Measurements of awareness provide the following additional constraint. If we elicit a measurement $\pi^{\text{survey}}_{ijs}$, on survey $s \in \{1, 2, 3\}$, then we must have

$$\pi^{\text{survey}}_{ijs} = 1(\pi^*_{ijt_i(s)} + \nu_{ijs} > 0) + 1(\pi^*_{ijt_i(s)} + \nu_{ijs} > 1), \tag{18}$$

where $t_i(s)$ is the time at which i takes survey s.

[50] We have measurement error e_{ijt} on baseline surveyed preferences, and on the preferences underlying the "just-before feedback" ROL, if any, that households submit at time $t = 2$, but not on final rank-order lists.

To update π^*, we loop through each (i, j, t), updating π_{ijt} conditional on the other elements of π_i and the other variables. Conditional on $\pi^*_{ij,-t}$, the variable π^*_{ijt} is normally distributed, with the mean and variance given by the (standard) formula. π^*_{ijt} is drawn from a normal distribution subject to optimality and survey-measurement constraints described above. Equivalently:

- If the high-information utility $\hat{u}^2_{ij} + e_{ijt}$ is not consistent with rank-order lists at time t, then $\pi^*_{ijt} < 1$.
- If the low-information utility $\hat{u}^1_{ij} + e_{ijt}$ is not consistent with rank-order lists at time t, then $\pi^*_{ijt} \notin [0, 1)$.
- If the no-information utility $\hat{u}^0_{ij} + e_{ijt}$ is not consistent with rank-order lists at time t, then $\pi^*_{ijt} \notin [0, 1)$.
- If there is a measurement $\pi^{\text{survey}}_{ijs}$ then $\pi^*_{ijt_i(s)} + e_{ijs}$ must satisfy equation 18, imposing upper and/or lower bounds on $\pi^*_{ijt_i(s)}$.

Constraints on utilities $\hat{u}$ and on measurement-error terms are analogous.

We use 5000 iterations, throwing out the first 2500 as burn-in. We choose relatively uninformative conjugate priors: variances are Gamma$(1, 1)$, regression coefficients and means are $N(0, 100)$, and covariance matrices of size (k, k) are $IW(k + 1, I_k)$.

To compute draws of latent variables at point estimates, we iterate through a similar Gibbs sampler, holding parameters fixed at their means along the chain estimated above, updating latent utilities, information π^*, measurement-error terms, $\hat{x}$'s, and random coefficients.

A.5. ADDITIONAL MODEL ESTIMATES

In this section, we present additional results from the model estimates and counterfactuals. Table A.XIII presents additional model parameters that describe parents' information and preferences. These parameters are estimated in the first step of the estimation procedure. Panel A shows the variance-covariance matrix of the subjective and true individual utility shocks. Panels B and E show the distortion functions for quality and price. Panel C shows the variance-covariance matrix for the individual information shocks over time. Panel D shows the random coefficients for the time effects (periods $t = 1, 2, 3$) over π. Panel F shows the coefficients for the knowledge shifters, and Panel G shows the estimated measurement error for our surveys. Finally, we report the probability of misreporting the subjective x's in the survey.

Table A.XIV presents additional model parameters that describe parents' beliefs and search behavior. We first report the mean estimated probability of each of the three latent types that define the heterogeneity in the Dirichlet parameters Λ_{it}. We also report the search technology parameters, which describe the probability of finding certain schools based on the school characteristics and our interventions. Finally, we present the unobserved match value shock primitives.

Figure A.3 shows the estimates for each school's unobservables, the mean utility, and "discoverability". Consistent with the positive discoverability-mean utility covariance parameter shown in Table IV, there is a clear correlation between schools that families prefer more (higher mean utility) and the ones that they are more likely to know (higher discoverability).

Figure A.4 shows the estimated distortion functions. Each panel shows the probability distribution of the *perceived values* for a school attribute (quality or price), conditional on the *true value*. Panels (A) and (B) show the distortion functions for low SES families, and panels (C) and (D) show the distortion functions for high SES families.

Figure A.5a shows the distribution of the deterministic component of the single click cost ($x_i^c \gamma^{\text{cost}}$), and Figure A.5b shows the distribution of the ratio of the value of the first search over the cost of the first search.

Figure A.6 shows the model fit on a series of expected school characteristics, behavior, and latent variables.

Figure A.XVI and Figure A.XV show the simulated counterfactuals separating by high and low SES families. Figure A.7 shows the distributional effects of a subset of the counterfactuals, relative to the base simulation.

Finally, Figure A.XVII presents a sequence of counterfactuals starting from the *better search* scenario in which we gradually reduce the search costs up to 5% of the original value.

TABLE A.XIII

ADDITIONAL PARAMETERS: INFORMATION AND PREFERENCES

	Param.	Low SES Coeff.	Low SES Std Err.	High SES Coeff.	High SES Std Err.
Panel A: Variance Covariance of Subjective and True Shocks (Σe)					
1	1	0.190	(0.026)	0.880	(0.153)
	2	0.045	(0.009)	0.189	(0.020)
2	2	0.618	(0.057)	1.414	(0.149)
Panel C: Variance Covariance of shocks v_{ijt} over time (Σ^v)					
1	1	0.359	(0.125)	0.209	(0.007)
	2	-0.210	(0.079)	-0.165	(0.009)
	3	-0.096	(0.039)	-0.121	(0.013)
2	2	0.387	(0.140)	0.352	(0.018)
	3	0.319	(0.120)	0.314	(0.020)
3	3	0.569	(0.197)	0.584	(0.034)
Panel D: Random Coefficients of time effects ($\Sigma rc\pi$)					
1	1	0.485	(0.067)	0.529	(0.054)
	2	0.057	(0.057)	-0.190	(0.052)
	3	-0.027	(0.009)	0.020	(0.008)
2	2	2.113	(0.278)	2.556	(0.267)
	3	-0.142	(0.028)	-0.155	(0.038)
3	3	0.015	(0.003)	0.020	(0.004)
Panel F: Knowledge Shifters (αzw)					
At least t2	α_w	1.096	(0.081)	0.928	(0.179)
At least t1	α_w	-0.208	(0.046)	-0.129	(0.026)
Distance	α_z	-0.146	(0.029)	-0.116	(0.018)
Treatment 1	α_w	0.139	(0.121)	0.448	(0.128)
Treatment 2	α_w	0.195	(0.060)	0.198	(0.237)
Highlight-worthy	α_w	0.362	(0.102)	0.258	(0.048)
Highlighted	α_w	-0.248	(0.052)	-0.065	(0.071)
Single Click	α_w	0.933	(0.180)	0.433	(0.059)
Double Click	α_w	1.369	(0.228)	0.772	(0.072)
Feedback	α_w	-0.699	(0.155)	-0.612	(0.172)
Feedback Pre	α_w	0.066	(0.057)	0.036	(0.063)
Feedback Known	α_w	1.126	(0.154)	0.986	(0.161)
Feedback Rec	α_w	-0.034	(0.181)	0.241	(0.185)
Panel G: Measurement Error					
Base Survey Utility	σ^2_ϵ	0.115	(0.012)	0.018	(0.029)
Survey Awareness	$\sigma^2_{\eta_s}$	0.385	(0.119)	0.017	(0.003)
Pr(misreport x's)	p^s	0.219	(0.019)	0.207	(0.039)

True	Subjective	Low SES Coeff.	Low SES Std Err.	High SES Coeff.	High SES Std Err.
Panel B: Distortion Function - Quality (Γ quality)					
1	1	0.127	(0.028)	0.129	(0.079)
	2	0.375	(0.055)	0.449	(0.115)
	3	0.415	(0.051)	0.358	(0.117)
	4	0.083	(0.015)	0.063	(0.048)
2	1	0.024	(0.007)	0.051	(0.026)
	2	0.323	(0.031)	0.449	(0.038)
	3	0.533	(0.033)	0.458	(0.027)
	4	0.121	(0.009)	0.042	(0.021)
3	1	0.001	(0.001)	0.009	(0.005)
	2	0.112	(0.018)	0.186	(0.021)
	3	0.592	(0.021)	0.613	(0.015)
	4	0.295	(0.015)	0.192	(0.022)
4	1	0.002	(0.001)	0.007	(0.005)
	2	0.059	(0.016)	0.075	(0.019)
	3	0.437	(0.029)	0.426	(0.036)
	4	0.502	(0.026)	0.492	(0.031)
Panel E: Distortion Function - Price (Γ price)					
1	1	0.594	(0.014)	0.657	(0.028)
	2	0.396	(0.014)	0.325	(0.026)
	3	0.009	(0.003)	0.018	(0.009)
	4	0.000	(0.000)	0.001	(0.001)
2	1	0.106	(0.021)	0.110	(0.047)
	2	0.782	(0.027)	0.674	(0.052)
	3	0.111	(0.026)	0.205	(0.034)
	4	0.001	(0.001)	0.011	(0.004)
3	1	0.037	(0.015)	0.035	(0.023)
	2	0.483	(0.031)	0.337	(0.050)
	3	0.474	(0.026)	0.600	(0.041)
	4	0.006	(0.003)	0.028	(0.019)
4	1	0.084	(0.037)	0.031	(0.029)
	2	0.289	(0.049)	0.049	(0.045)
	3	0.540	(0.058)	0.616	(0.071)
	4	0.087	(0.036)	0.304	(0.070)

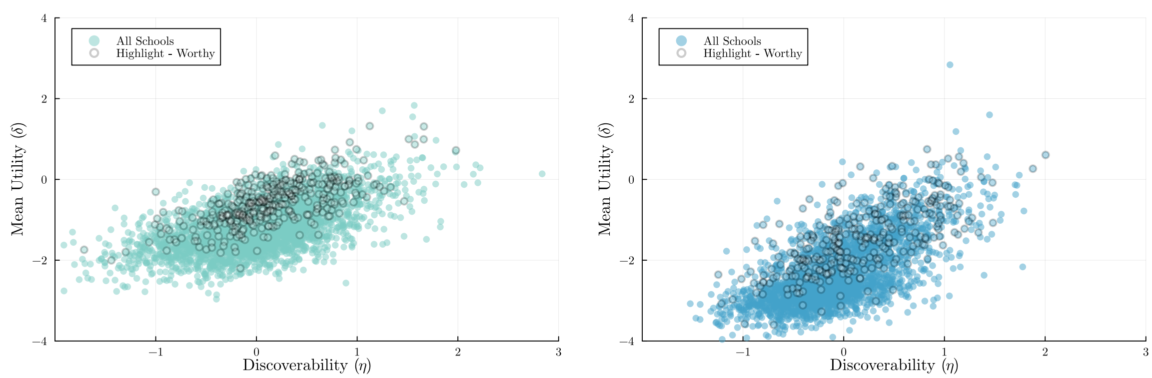
Note: This table presents additional estimates from step 1 of the model.

TABLE A.XIV

ADDITIONAL PARAMETERS: BELIEFS AND SEARCH COSTS

	Parameter	Low SES Coeff	Low SES Std Err.	High SES Coeff	High SES Std Err.
Mean Prob of each type					
Λ Type 1		0.614	-	0.654	-
Λ Type 2		0.329	-	0.251	-
Λ Type 3		0.057	-	0.095	-
Search Technology (γ click)					
Distance	1	-1.294	(0.040)	-1.072	(0.059)
Price = 1	2	-0.175	(0.089)	-0.215	(0.095)
Price = 2	3	-0.022	(0.109)	-0.048	(0.102)
Price = 3	4	0.162	(0.102)	0.193	(0.064)
Price = 4	5	0.035	(0.132)	0.069	(0.150)
Quality = 1	6	-0.193	(0.084)	-0.212	(0.098)
Quality = 2	7	-0.147	(0.095)	-0.066	(0.079)
Quality = 3	8	0.027	(0.088)	-0.023	(0.094)
Quality = 4	9	0.312	(0.089)	0.301	(0.083)
Highlightworthy	10	-0.048	(0.192)	0.089	(0.139)
Highlightworthy $\times$ Treat 2	11	0.340	(0.071)	0.200	(0.076)
Double Click (θ cost)					
Mean Cost		1.768	(0.540)	0.692	(0.134)
Log-Variance		0.555	(0.471)	-0.529	(0.284)
Match Value Shocks ε_{ij} Primitives					
Unobserved match-value shock	$\tilde{\mu}$	-0.445	(0.047)	-0.683	(0.096)
	$\tilde{\sigma}_{\varepsilon}$	3.316	(0.549)	2.221	(0.557)
	σ_{ε}	0.786	(0.044)	1.189	(0.097)

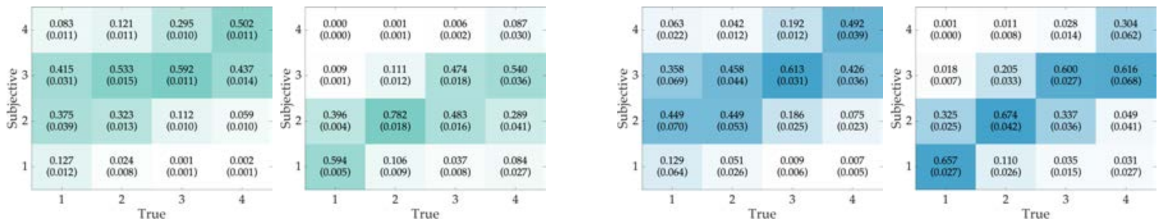
Note: This table presents additional estimates from step 2 of the model.



(a) Low SES

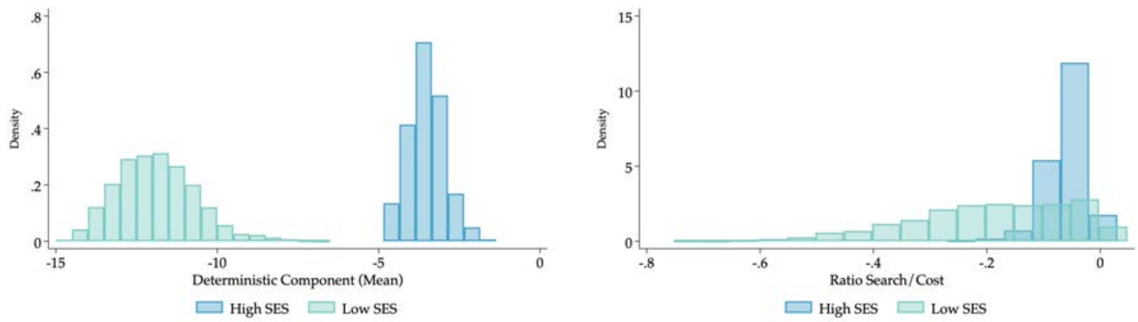
(b) High SES

FIGURE A.3.—Estimates of school unobservables. Notes: Each panel shows the estimated discoverability (X-axis) and the estimated mean utility (Y-axis) for each school. The left panel shows the estimates for low SES-families, and the right panel for high-SES families. The "highlight-worthy" schools are shown as white circles.



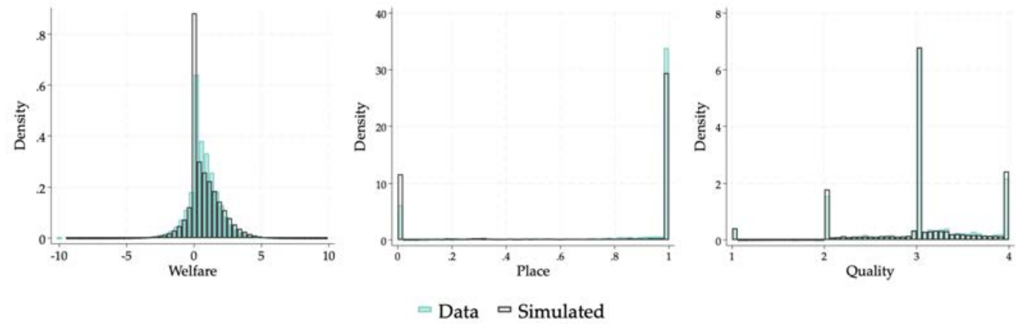
(a) Quality Low SES **(b)** Price Low SES **(c)** Quality High SES **(d)** Price High SES

FIGURE A.4.—Distortion Functions. Notes: Each panel shows estimated distortion functions for school attributes. Panels (A) and (B) show the estimates for low SES families, and Panels (C) and (D) for high SES families. Panels (A) and (C) show the quality distortion functions, and Panels (B) and (D) show the price distortion function.

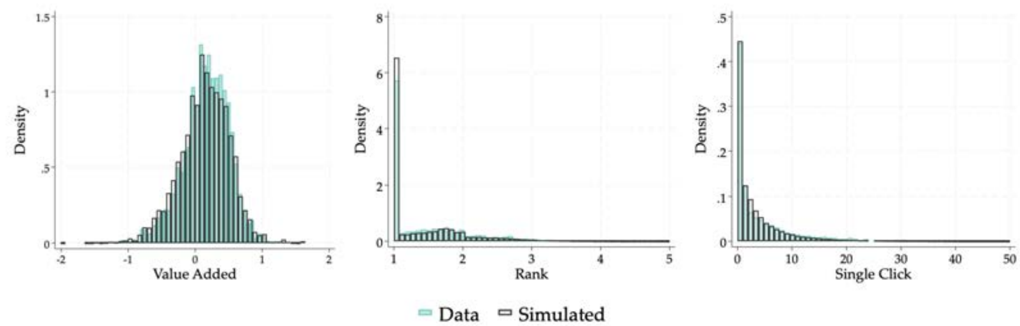


(a) Distribution of $x_i^c \gamma^{\text{cost}}$ **(b)** Ratio of Value 1st Search over Cost

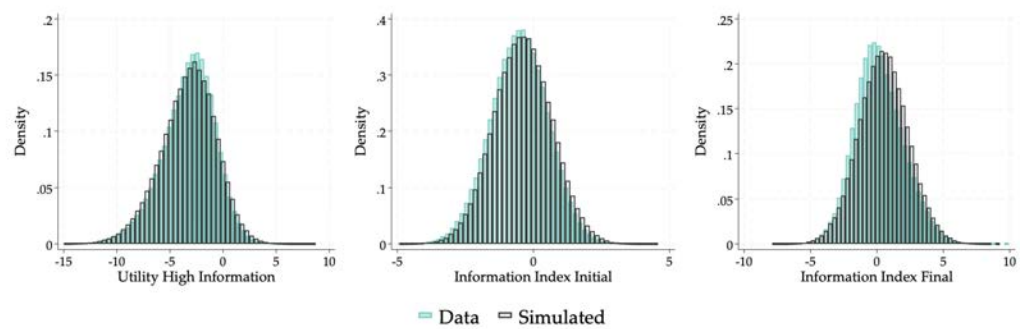
FIGURE A.5.—Panel (a) shows the distribution of the deterministic component of the single click cost ($x_i^c \gamma^{\text{cost}}$). Panel (b) shows the distribution of the ratio of the value of the first search over the cost of the first search.



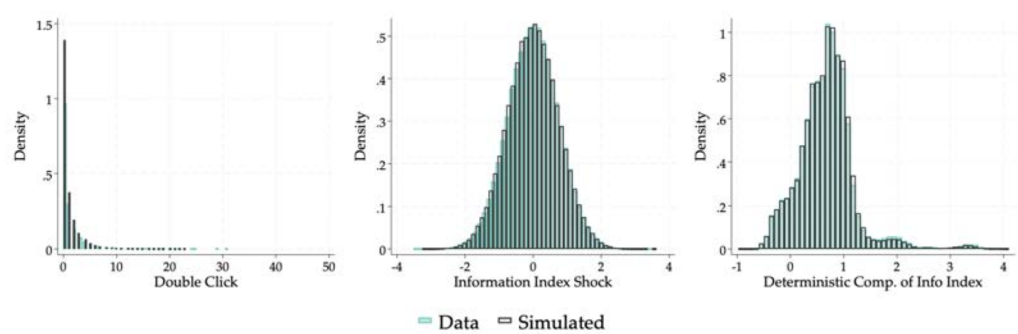
(a)



(b)



(c)



(d)

FIGURE A.6.—Model Fit. Panel (a) shows the model fit of school attributes (welfare, placement, and quality). Panel (b) shows the model fit of value-added, distance, and price. Panel (c) shows the model fit of utility of high information, information index initial, and information index final. Panel (d) shows the model fit of double click, information index shock and the deterministic component of the information index.

TABLE A.XV

MAIN RESULTS FOR LOW SES HOUSEHOLDS

		Welfare	Placement		Expected School Characteristics				Search (N.Clicks)		
			Place	E(rank)	Distance	Price	Quality	VA	Single	Double	V(1st)
Gains from Full Information											
(1)	Full model baseline	0.450	0.744	1.386	1.554	1.241	2.969	0.129	3.403	0.975	0.677
		(0.021)	(0.007)	(0.011)	(0.031)	(0.011)	(0.012)	(0.007)	(0.120)	(0.037)	(0.005)
(2)	Full information	1.224	0.840	1.552	1.603	1.371	3.193	0.217	-	-	-
		(0.018)	(0.006)	(0.013)	(0.030)	(0.014)	(0.012)	(0.007)	-	-	-
(3)	Gains (difference (2)-(1))	0.774	0.096	0.166	0.049	0.130	0.224	0.088	-	-	-
		(0.028)	(0.009)	(0.017)	(0.043)	(0.018)	(0.017)	(0.010)	-	-	-
	(% Change)	172.00%	12.90%	11.98%	3.15%	10.48%	7.54%	68.22%	-	-	-
Decomposition: sequential correction of beliefs and misperceptions											
(4)	Better Search (S^*)	0.655	0.775	1.411	1.541	1.263	3.030	0.149	5.152	-	0.859
		(0.021)	(0.007)	(0.011)	(0.031)	(0.012)	(0.012)	(0.007)	(0.142)	-	(0.007)
(5)	(4) $+x$	0.940	0.714	1.437	1.491	1.331	3.181	0.213	5.177	-	0.849
		(0.017)	(0.007)	(0.011)	(0.030)	(0.014)	(0.012)	(0.007)	(0.142)	-	(0.007)
(6)	(5) $+f(x)$	0.938	0.713	1.437	1.490	1.329	3.180	0.212	5.142	-	0.852
		(0.017)	(0.007)	(0.011)	(0.030)	(0.014)	(0.012)	(0.007)	(0.141)	-	(0.007)
(7)	(6) $+r$	0.941	0.715	1.438	1.490	1.330	3.182	0.213	5.202	-	0.894
		(0.017)	(0.007)	(0.011)	(0.030)	(0.014)	(0.012)	(0.007)	(0.145)	-	(0.007)
(8)	(7) $+f(\varepsilon)$	0.940	0.739	1.459	1.518	1.334	3.184	0.214	5.028	-	0.586
		(0.017)	(0.007)	(0.011)	(0.030)	(0.014)	(0.012)	(0.007)	(0.139)	-	(0.586)
(9)	(8) $+\varepsilon$	1.013	0.754	1.461	1.533	1.328	3.173	0.208	5.023	-	0.535
		(0.017)	(0.007)	(0.011)	(0.029)	(0.013)	(0.012)	(0.007)	(0.139)	-	(0.005)
Misspecified models											
(10)	No mispercept. of x ($\hat{x} = x$)										
	(Gains in outcomes relative to baseline)	0.473	0.175	0.095	0.021	0.003	-0.028	-0.020	-	-	-
	(Gains in S^* relative to baseline)	0.115	0.059	0.011	-0.062	-0.003	-0.006	-0.007	1.832	-	0.067
	(Gains in (9) relative to baseline)	0.245	0.099	0.038	-0.034	-0.005	-0.018	-0.013	1.800	-	0.014
(11)	No mispercept. of x, e if $\pi > 0$										
	(Gains in outcomes relative to baseline)	0.135	0.058	0.020	0.018	0.003	-0.007	-0.008	-	-	-

Note: This table presents the counterfactuals for Low SES. Columns: Welfare: EU according to fully informed payoffs. Place: probability of placement. (E(rank),Distance,Price,Quality, VA): avg. (rank of placed school within ROL,distance,price,quality,school value added (in student-level SD)), conditional on placement. (Single, Double, V(1st)): number of single clicks, double clicks, and value of the first pin click. Rows are as follows. Full model baseline: includes all possible misperceptions and biases. Full information: $\pi_{ijT} > 1$ for all (i, j), and $\hat{x} = x$. Gains: difference in outcomes between full information and baseline. Decomposition: sequential correction of beliefs and misperceptions. Better search (S^*): search is perfectly informative. $S^* + x$: S^*+ provides full information about price and quality of known schools. $S^* + x + f(x)$: $S^* + x+$ correct distribution of school characteristics of unknown schools. $S^* + x + f(x) + r$: $S^* + x + f(x)+$ correct misperceptions about rejection chances at known schools. $S^* + x + f(x) + r + f(\varepsilon)$: $S^* + x + f(x) + r+$ correct beliefs about the distribution of match value shocks of unknown schools. $S^* + x + f(x) + r + \varepsilon + f(\varepsilon)$: $S^* + x + f(x) + r + \varepsilon+$ correct misperceptions about the match value shocks of known schools. No mispercept. of x: gains from misspecified model assuming $\hat{x}_{ijt} = x_{ij}$ relative to baseline. No mispercept. of x, e if $\pi > 0$: gains in misspecified model assuming $\hat{x}_{ijt} = x_{ij}$ and perfect learning for all schools with $\pi_{ijT} > 0$ relative to baseline (as in data)

TABLE A.XVI

MAIN RESULTS FOR HIGH SES HOUSEHOLDS

		Welfare	Placement		Expected School Characteristics				Search (N.Clicks)		
			Place	E(rank)	Distance	Price	Quality	VA	Single	Double	V(1st)
Gains from Full Information											
(1)	Full model baseline	0.853	0.687	1.583	1.741	1.539	3.092	0.177	4.797	1.344	0.229
		(0.041)	(0.013)	(0.024)	(0.059)	(0.028)	(0.020)	(0.012)	(0.257)	(0.077)	(0.005)
(2)	Full information	1.591	0.860	1.798	1.812	1.527	3.199	0.210	-	-	-
		(0.038)	(0.009)	(0.031)	(0.057)	(0.028)	(0.020)	(0.011)	-	-	-
(3)	Gains (difference (2)-(1))	0.738	0.173	0.215	0.071	-0.012	0.107	0.033	-	-	-
		(0.056)	(0.016)	(0.039)	(0.082)	(0.040)	(0.028)	(0.016)	-	-	-
	(% Change)	86.52%	25.18%	13.58%	4.08%	-0.78%	3.46%	18.64%	-	-	-
Decomposition: sequential correction of beliefs and misperceptions											
(4)	Better Search (S^*)	1.091	0.753	1.615	1.685	1.525	3.125	0.185	6.658	-	0.415
		(0.041)	(0.012)	(0.025)	(0.057)	(0.028)	(0.020)	(0.012)	(0.299)	-	(0.008)
(5)	(4) $+x$	1.217	0.736	1.627	1.673	1.534	3.194	0.211	6.585	-	0.403
		(0.038)	(0.012)	(0.025)	(0.056)	(0.029)	(0.019)	(0.011)	(0.296)	-	(0.008)
(6)	(5) $+f(x)$	1.221	0.737	1.626	1.669	1.532	3.191	0.209	6.626	-	0.406
		(0.038)	(0.012)	(0.025)	(0.056)	(0.029)	(0.020)	(0.011)	(0.299)	-	(0.008)
(7)	(6) $+r$	1.221	0.736	1.626	1.671	1.532	3.191	0.208	6.704	-	0.411
		(0.038)	(0.012)	(0.025)	(0.056)	(0.029)	(0.020)	(0.011)	(0.299)	-	(0.008)
(8)	(7) $+f(\varepsilon)$	1.203	0.764	1.660	1.702	1.533	3.188	0.207	6.536	-	0.336
		(0.040)	(0.012)	(0.026)	(0.056)	(0.029)	(0.020)	(0.011)	(0.294)	-	(0.336)
(9)	(8) $+\varepsilon$	1.341	0.773	1.668	1.734	1.533	3.188	0.208	6.284	-	0.249
		(0.037)	(0.012)	(0.026)	(0.056)	(0.028)	(0.019)	(0.011)	(0.286)	-	(0.007)
Misspecified models											
(10)	No mispercept. of x ($\hat{x}=x$)										
	(Gains in outcomes relative to baseline)	0.673	0.181	0.154	0.042	-0.042	-0.041	-0.029	-	-	-
	(Gains in S^* relative to baseline)	0.193	0.064	0.021	-0.079	-0.020	-0.015	-0.012	1.770	-	0.143
	(Gains in (9) relative to baseline)	0.389	0.100	0.058	-0.021	-0.028	-0.024	-0.017	1.537	-	0.033
(11)	No mispercept. of x, e if $\pi > 0$										
	(Gains in outcomes relative to baseline)	0.151	0.058	0.050	0.005	-0.004	-0.004	-0.005	-	-	-

Note: This table presents the counterfactuals for High SES. Columns: Welfare: EU according to fully informed payoffs. Place: probability of placement. (E(rank),Distance,Price,Quality, VA): avg. (rank of placed school within ROL,distance,price,quality,school value added (in student-level SD)), conditional on placement. (Single, Double, V(1st)): number of single clicks, double clicks, and value of the first pin click. Rows are as follows. Full model baseline: includes all possible misperceptions and biases. Full information: $\pi_{ijT} > 1$ for all (i,j), and $\hat{x} = x$. Gains: difference in outcomes between full information and baseline. Decomposition: sequential correction of beliefs and misperceptions. Better search (S^*): search is perfectly informative. $S^* + x$: S^*+ provides full information about price and quality of known schools. $S^* + x + f(x)$: $S^* + x+$ correct distribution of school characteristics of unknown schools. $S^* + x + f(x) + r$: $S^* + x + f(x)+$ correct misperceptions about rejection chances at known schools. $S^* + x + f(x) + r + f(\varepsilon)$: $S^* + x + f(x) + r+$ correct beliefs about the distribution of match value shocks of unknown schools. $S^* + x + f(x) + r + \varepsilon + f(\varepsilon)$: $S^* + x + f(x) + r + \varepsilon+$ correct misperceptions about the match value shocks of known schools. No mispercept. of x: gains from misspecified model assuming $\hat{x}_{ijt} = x_{ij}$ relative to baseline. No mispercept. of x, e if $\pi > 0$: gains in misspecified model assuming $\hat{x}_{ijt} = x_{ij}$ and perfect learning for all schools with $\pi_{ijT} > 0$ relative to baseline (as in data)

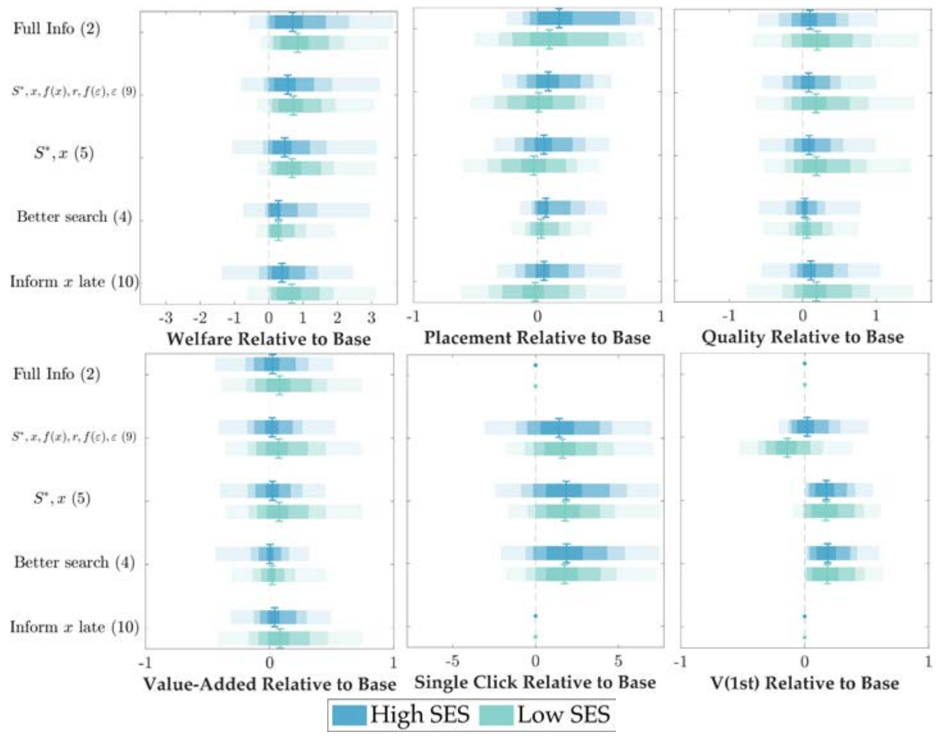


FIGURE A.7.—Distributional effects of counterfactual policies. Notes: This figure displays heterogeneity in the effects of the counterfactual policies described in Table V on a subset of the outcomes. Each panel displays the mean and distribution of individual-level changes in welfare, placement, quality, value-added of the school placed, single click, and value of 1st search for both low SES and high SES parents. The mean is represented by a dark line and the distribution by a bar with different shades. The dark center contains percentiles 25 to 75(50% of the students). The slightly lighter area contains percentiles 10 to 25 and 75 to 90 (30% of the students). The next area contains percentiles 5 to 10 and 90 to 95 (10% of the students).

TABLE A.XVII

SEARCH COST REDUCTION COUNTERFACTUAL

		Welfare	Placement		E(School Charact)		Search (N.Clicks)		
			Place	E(rank)	Quality	VA	Single	Double	V(1st)
(1)	Better Search 100%	0.753	0.770	1.456	3.051	0.157	5.490	-	-
		(0.019)	(0.006)	(0.010)	(0.010)	(0.006)	(0.129)	-	-
(2)	Better Search 80%	0.756	0.771	1.456	3.052	0.158	5.580	-	-
		(0.019)	(0.006)	(0.010)	(0.010)	(0.006)	(0.131)	-	-
(3)	Better Search 60%	0.761	0.772	1.460	3.053	0.157	5.727	-	-
		(0.019)	(0.006)	(0.010)	(0.010)	(0.006)	(0.133)	-	-
(4)	Better Search 40%	0.771	0.774	1.462	3.054	0.158	6.097	-	-
		(0.019)	(0.006)	(0.010)	(0.010)	(0.006)	(0.141)	-	-
(5)	Better Search 20%	0.811	0.782	1.473	3.062	0.161	7.414	-	-
		(0.019)	(0.006)	(0.010)	(0.010)	(0.006)	(0.167)	-	-
(6)	Better Search 10%	0.889	0.797	1.494	3.081	0.169	10.945	-	-
		(0.019)	(0.005)	(0.011)	(0.010)	(0.006)	(0.238)	-	-
(7)	Better Search 5%	1.028	0.819	1.532	3.112	0.181	20.030	-	-
		(0.019)	(0.005)	(0.011)	(0.010)	(0.006)	(0.396)	-	-

Note: This table presents additional counterfactuals from the model.

SUPPLEMENTARY MATERIAL

S.1. ADDITIONAL INFORMATION ON THE SCHOOL CHOICE SYSTEM IN CHILE

We conduct our intervention within the Chilean School Admission System (SAE). The SAE is a centralized system that allows students to apply to multiple schools, and rank them in order of preference. The SAE is administered by the Ministry of Education (Mineduc), and is the main mechanism for assigning students to schools in Chile.

There are three types of educational provisions in Chile: public schools owned and managed by the state mainly through municipalities, privately owned and managed schools subsidized by the state (voucher schools), and private schools owned and managed by the private sector. Voucher schools account for 55.58% of the total enrollment, and can charge out-of-pocket fees while receiving subsidies for each student, depending on the grade.[51] If a voucher school holds a *Subvención Escolar Preferencial* (SEP) agreement, students from low socioeconomic status that enroll carry a larger subsidy but do not pay any fee (for more details, see Neilson (2021)).[52]

Chile holds a student-proposed deferred acceptance system for centralized assignment. On a single nationwide online platform, parents with children from all levels, from Pre-K to 12th grade, apply to public and voucher schools. Almost all public and voucher schools participate in the platform. Off-platform options consist of fully private schools in all grades, as well as some publicly-funded preschools which may offer Kindergarten and/or 1st grade. In addition, there are a handful of schools in specialized settings such as hospitals, and schools exclusively for students with disabilities, which do not participate. For a detailed description of how the school admission system is implemented, see Correa et al. (2019).

There are three main stages in the SAE: the regular stage, the complementary stage, and the aftermarket:

[51] In 2015, the School Inclusion Law froze the co-payment, which will gradually fade out while subsidized funds increase.

[52] Note that low socioeconomic status for voucher eligibility is different from what we refer as low socioeconomic status in this paper (i.e. non-college educated mothers).

- **Main application stage:** The SAE application platform receives applications for roughly one month starting in early August. Parents may list as many schools as they like, in any order. There is no constraint on list length. Parents can update their application as often as they like during this stage. We focus on this stage in this paper.
- **Main assignment stage:** The SAE assigns students to schools based on their preferences and the priorities and quotas established by the Ministry of Education. Parents are notified of the results, and have a few days to accept or reject the assignment. Students who are not assigned or reject their assignment can participate in the complementary stage.
- **Complementary stage:** The process is the same as the main stage, but only schools with remaining vacancies are available. The platform receives applications for roughly one week during November.
- **Final results and aftermarket:** Final results are announced in early December. Parents assigned through the complementary process decide whether to accept or reject their assignment. Unassigned students are assigned to a default school, which is the closest school to their home with available slots that is not in the "insufficient" quality category. From late December to early January parents can enroll in their assigned school. After this period, students may change schools by enrolling in undersubscribed schools. This process is decentralized.

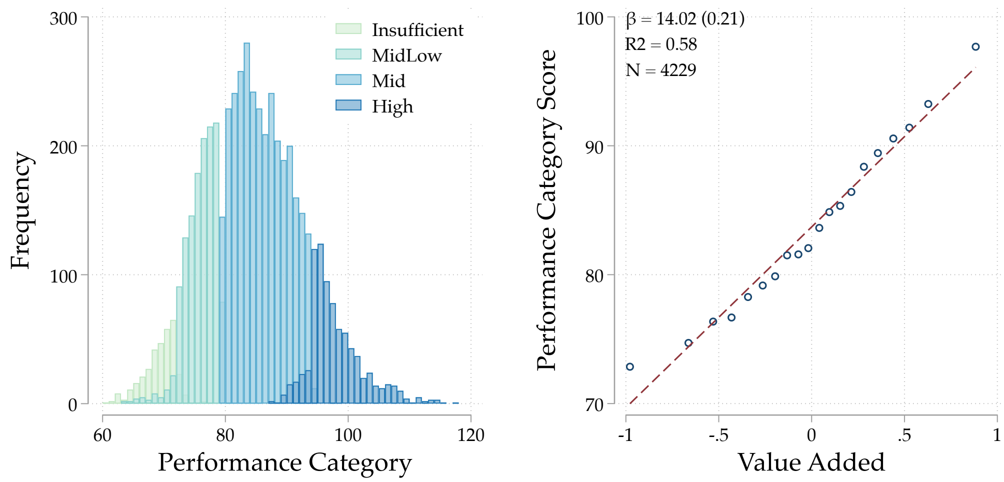
In 2021, 207,578 students applied to entry grades. Out of these students, 71% enrolled in the school that was assigned to them in the regular SAE process, 3% enrolled in the school that was assigned to them in the complementary process, 3% of students don't enroll in a school for 2022 and 13% enrolled in a school through the aftermarket for schools with SAE slots. The remaining 2% of students enrolled in private schools, and 7% of students enroll in a public or voucher option that didn't participate in SAE.

S.2. ADDITIONAL INFORMATION ON THE SCHOOL QUALITY CATEGORIES

Our quality measure comes from the Education Quality Agency (Agencia de Calidad de la Educación), which classifies schools into four categories (high, medium, medium-low, and insufficient performance). [53] The categories are based on a continuous performance score that uses the distribution of students in learning levels, indicators of personal development, and results from the SIMCE test, adjusted for student characteristics at the school level. Figure S.1a plots the distribution of the performance score. The different colors indicate the four discrete quality categories. The quality categories are not only based on different cutoff points of the performance score but also use additional criteria (such as the performance of selected student groups), resulting in an overlap of the performance score across quality categories.

Figure S.1b shows that the continuous performance score is highly correlated with the school value added measure. In our baseline survey, 95% of parents also reported that obtaining information on a school from the Education Quality Agency is a necessary step for them before adding a school to their application.

[53] We use the following criteria for these categories: High = 4, Medium = 3, Medium-Low = 2, Insufficient = 1. The mean quality category is 2.79 with standard deviation of 0.76



(a) **(b)**

FIGURE S.1.—School performance measure. Panel (A) shows the continuous performance of schools with colors for the discrete categories (Insufficient, MidLow, Mid, High). Panel (B) shows the correlation between the value added measure and the performance category score.

S.3. ADDITIONAL INFORMATION ON DATA COLLECTION AND SCHOOL EXPLORER PLATFORM

Surveys: All sample parents were invited to complete four rounds of surveys. Table S.I summarizes the timing and content of each survey round.

School Explorer: All sample parents received access to a school explorer platform that was developed by an EdTech NGO. Parents could use the explorer to learn about the characteristics of schools in their neighborhood. Figure S.2 shows the potential search path of a parent in the control group or in the first treatment group.[54] After receiving an initial set of instructions, parents see a map of schools around their home, which is indicated by the red pin on the map (Panel A). Each primary school is shown as a grey circle on the map. By clicking on one of the circles, a popup with basic information on the school is shown, including the distance to the parent's home, the quality, the price, and the admission probability (Panel B). Parents can then click again to view a detailed profile of the school, which contains additional information like the availability of infrastructure or a virtual tour of the school (Panels C-D). Figure S.3 plots explorer usage patterns over time.

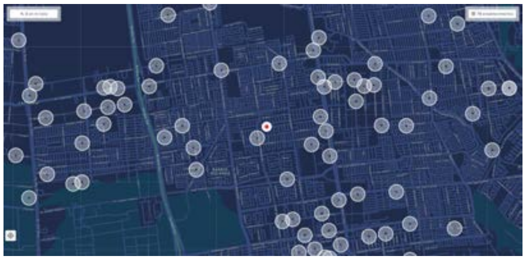
[54]For parents in the second treatment group, schools that are free and have high quality are highlighted in green. For more information on the difference between the treatment groups, see section S.4 in the supplementary material.

TABLE S.I

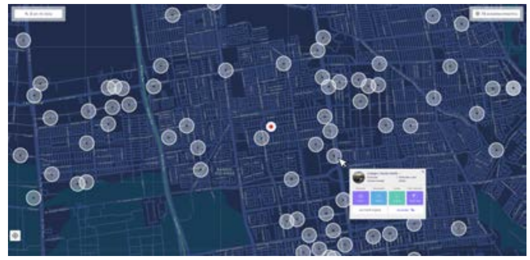
SUMMARY OF SURVEYS

Survey	Sections	Questions
Registration Form N = 13,721 May 25 - Jul 2	Respondent	SEP belief. Number of children
	Student Roster/Info	Student contact information. Mother education. Interest to apply to SAE
	Map + Beliefs	Address. Distribution of schools in neighborhood
Baseline N = 3,948 Jul 7 - Jul 16	Awareness	Knowledge level (know by name, know well, don't know) of 8 random schools within 2km (2 fake)
	Perceptions (x's)	Perceived price and academic performance of: (a) 1st ranked, (b) random school in application, (c) random known school not in application. Perceived admission chance of 1st ranked school
	Beliefs	Distribution of school characteristics in neighborhood (2kms around their house): (a) number of schools, (b) number of schools in each performance category (x4), (c) number of schools in each performance category-price cell (x16)
	ROL	Partial ROL (ranking to date). Perception on overall non placement risk of application.
	Other Questions	
	Own priority perception	Perception on whether their child is elegible for SEP. Parent staff priority in the application
	Search behavior	Probability of engaging in search and adding schools to application
Midline N = 1,669 Aug 24 - Oct 25	Awareness	Knowledge level of 4 schools not known in baseline (1 fake)
	Perceptions (x's)	Perceived price, academic performance and admission chances for up to 5 schools
	Beliefs	Final ROL (final ranking). Distribution of school characteristics in neighborhood (2kms around their house)
	Other Questions	
	Explorer Usage	Satisfaction with explorer. Was able to find new schools.
	Application + Report Card	If applied to SAE. Satisfaction with report card, if changed application after report card
Endline N = 540 Oct 21 - Oct 25	Awareness	Knowledge level of 5 random schools (1 fake) + top 5 schools in application.
	Perception (x's)	Perceived academic perfomance + price of top 4 schools in application. Perceived admission chance of top 3 schools in application. Perceived academic performance of 5 random schools (1 fake).
	Beliefs	Distribution of schools in neighborhood. SEP belief. Risk of application
	Other Questions	
	Behavior	Was report card useful and for what purpose. Likelihood of adding a school.
	Siblings	Siblings application, likelihood of rejecting assignment based on siblings

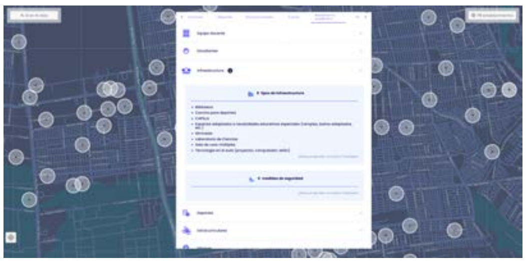
Note: This table presents a description of all surveys used in this study. SEP belief refers to the perceived eligibility for school vouchers. Distribution of schools in neighborhood refers to questions about the number of schools in each price and academic performance category, and the number of schools with primary education within 2km from the respondent's home. There are three knowledge levels of schools: I don't know it; know it by name; know it well. The midline survey was implemented through a call center. All other surveys were done through an online questionnaire distributed via email and WhatsApp.



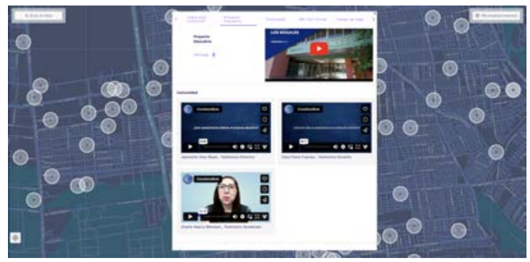
(a)



(b)



(c)



(d)

FIGURE S.2.—Example search path. This figure shows the potential search path of a parent in the control or first treatment arm.

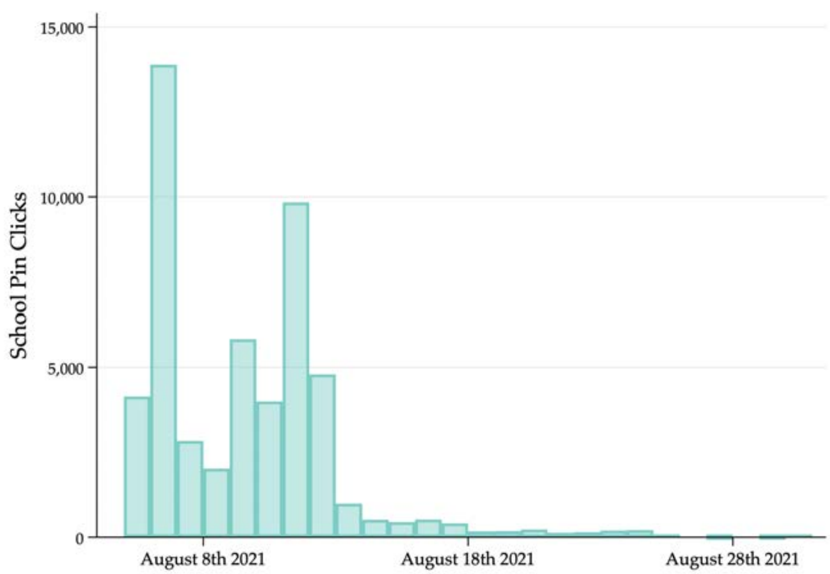


FIGURE S.3.—Search timing. Notes: This figure shows the timing of the search process. Access to the explorer platform was granted on August 5th and school applications closed on September 8th. Over 95% of school pin clicks occurred in the first two weeks.

S.4. ADDITIONAL INFORMATION ON SEARCH AID INTERVENTION

Recruitment: In total, 33,341 parents across approximately 2,700 kindergartens received the survey invitation. 13,721 of these parents (41%) completed the registration and pre-baseline form. Among parents who completed the pre-baseline form 9,062 parents (66%) met the eligibility criteria of the study. All eligible parents received then an invitation to the baseline survey, which was completed by 3,948 parents (43%).

Randomization: Eligible parents who completed the baseline survey were randomly assigned to one of three treatment arms. We later excluded 14 parents from the sample who were part of the research pilot. The randomization was done separately for parents who met the following three conditions: no older siblings, at least five primary schools within two kilometers of the home, and at least one highlight-worthy school within two kilometers of the home. Within both samples, we also stratified by region. For larger regions, we further stratified by perceived SEP status at baseline and maternal education. Tables S.II and S.III show balance checks separately for high and low SES households.

Intervention Details: Figures S.4 to S.7 show example screenshots of the information that was shown to parents as part of the search aid interventions. In this example, the household has access to 18 schools in total within 2km of the home (Figure S.4). Seven of these schools cost less than 50k CLP per month and seven schools have medium or high quality. The fourth panel shows the joint distribution, indicating that there are five highlight-worthy schools, defined as schools that cost less than 50k CLP per month and have at least medium quality. The second treatment group received the same information but was additionally shown where these schools are located on the map (Figure S.5). Both treatment groups further received a detailed table that shows the distribution of schools in each price and quality category (Figure S.6). The control group also received the school explorer platform but did not receive any information about the distribution of schools or their characteristics. After this information, parents entered the main part of the school explorer platform that allowed them to click on individual schools to obtain additional information. While all schools on the map were shown in grey for control and treatment group 1 parents, highlight-worthy schools were shown in green on the map for treatment group 2 parents (Figure S.7).

TABLE S.II

BALANCE CHECKS FOR SEARCH INTERVENTIONS FOR HIGH SES HOUSEHOLDS

	Control		Treatment 1		Treatment 2		
	Mean (1)	St. Dev. (2)	Coeff. (3)	St. Err. (4)	Coeff. (5)	St. Err. (6)	N (7)
Panel A: Choice Environment							
Number of available schools	15.512	[7.894]	0.124	(0.550)	-0.459	(0.578)	888
Number of available highlight-worthy schools	8.491	[4.526]	0.166	(0.345)	-0.075	(0.348)	888
Panel B: Parent/Child Characteristics							
Child is female	0.453	[0.499]	0.040	(0.041)	0.061	(0.043)	888
Child's birthyear	2017.087	[0.543]	-0.014	(0.046)	0.030	(0.046)	888
Number of younger siblings	1.163	[0.397]	0.019	(0.033)	0.033	(0.036)	888
Child has a disability (belief)	0.085	[0.279]	-0.019	(0.022)	-0.022	(0.023)	818
Parent works in a school	0.211	[0.409]	-0.047	(0.032)	-0.082	(0.031)	884
SEP household	0.161	[0.368]	0.003	(0.023)	-0.001	(0.023)	879
Panel C: Initial Knowlege and Beliefs							
Expected satisfaction with process	5.037	[1.403]	0.121	(0.128)	0.135	(0.128)	828
Listed any school as first preference	0.917	[0.276]	-0.022	(0.025)	-0.001	(0.024)	888
First-preference school is highlight-worthy	0.517	[0.501]	-0.022	(0.047)	0.094	(0.048)	700
Perceived admission change for first-preference school	0.679	[0.284]	0.013	(0.024)	0.028	(0.024)	828
Number of schools known by name	3.391	[2.735]	-0.103	(0.224)	0.115	(0.231)	888
Number of schools known well	2.028	[2.173]	-0.050	(0.175)	-0.024	(0.185)	888
Perceived number of available schools	7.758	[6.426]	0.873	(0.571)	-0.101	(0.524)	888
Perceived number of available highlight-worthy schools	3.339	[3.102]	0.699**	(0.301)	0.372**	(0.270)	888
Parent believed to be SEP eligible	0.104	[0.306]	0.016	(0.027)	0.016	(0.028)	888
SEP did not know about SEP status	0.654	[0.477]	-0.057	(0.040)	0.021	(0.040)	888

Panel D: Treatment Summary	Control	Treatment 1	Treatment 2
Observations	289	313	286
Whatsapp Reminder + SEP Status + Explorer	X	X	X
School Distribution		X	X
Highlight-worthy School			X

Note: This table shows balance for baseline covariates for the search aid interventions for high SES households. Column 1 reports the control mean of the dependent variable for each relevant subgroup (standard deviations in brackets). Columns 3 and 5 report the difference in the dependent variable from OLS regressions of each outcome on indicator variables for treatment assignments and stratification dummies. Robust standard errors are reported in parentheses. Variables in Panel A come from administrative data. Variables in Panels B and C come from the baseline survey.

TABLE S.III

BALANCE CHECKS FOR SEARCH INTERVENTIONS FOR LOW SES HOUSEHOLDS

	Control		Treatment 1		Treatment 2		
	Mean (1)	St. Dev. (2)	Coeff. (3)	St. Err. (4)	Coeff. (5)	St. Err. (6)	N (7)
Panel A: Choice Environment							
Number of available schools	16.416	[9.482]	-0.120	(0.327)	0.087	(0.322)	3057
Number of available highlight-worthy schools	8.679	[5.143]	-0.147	(0.193)	0.093	(0.189)	3057
Panel B: Parent/Child Characteristics							
Child is female	0.506	[0.500]	0.010	(0.022)	0.014	(0.022)	3057
Child's birthyear	2017.097	[0.552]	0.016	(0.023)	0.012	(0.024)	3057
Number of younger siblings	1.141	[0.383]	0.016	(0.017)	0.004	(0.017)	3057
Child has a disability (belief)	0.066	[0.248]	0.015	(0.012)	0.002	(0.012)	2707
Parent works in a school	0.025	[0.156]	0.011	(0.008)	0.015	(0.008)	2998
SEP household	0.533	[0.499]	-0.008	(0.016)	-0.019	(0.016)	3026
Panel C: Initial Knowlege and Beliefs							
Expected satisfaction with process	5.295	[1.387]	0.030	(0.063)	-0.063	(0.064)	2858
Listed any school as first preference	0.907	[0.291]	0.007	(0.013)	0.009	(0.012)	3057
First-preference school is highlight-worthy	0.635	[0.482]	0.057**	(0.023)	0.040**	(0.023)	2542
Perceived admission change for first-preference school	0.685	[0.268]	0.012	(0.012)	0.020	(0.012)	2858
Number of schools known by name	3.276	[2.672]	-0.065	(0.117)	0.019	(0.114)	3057
Number of schools known well	1.830	[2.010]	0.027	(0.088)	0.072	(0.088)	3057
Perceived number of available schools	7.355	[7.079]	-0.194	(0.302)	-0.395	(0.306)	3057
Perceived number of available highlight-worthy schools	3.759	[3.744]	-0.101	(0.154)	-0.213	(0.155)	3057
Parent believed to be SEP eligible	0.192	[0.394]	-0.004	(0.017)	-0.015	(0.017)	3057
SEP did not know about SEP status	0.669	[0.471]	-0.004	(0.021)	0.022	(0.020)	3057

Panel D: Treatment Summary	Control	Treatment 1	Treatment 2
Observations	1027	999	1031
Whatsapp Reminder + SEP Status + Explorer	X	X	X
School Distribution		X	X
Highlight-worthy School			X

Note: This table shows balance for baseline covariates for the search aid interventions for low SES households. Column 1 reports the control mean of the dependent variable for each relevant subgroup (standard deviations in brackets). Columns 3 and 5 report the difference in the dependent variable from OLS regressions of each outcome on indicator variables for treatment assignments and stratification dummies. Robust standard errors are reported in parentheses. Variables in Panel A come from administrative data. Variables in Panels B and C come from the baseline survey.

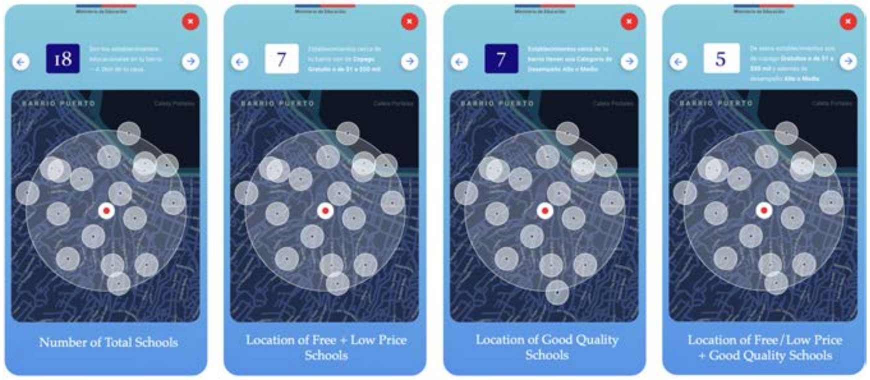


FIGURE S.4.—Search aid treatment 1

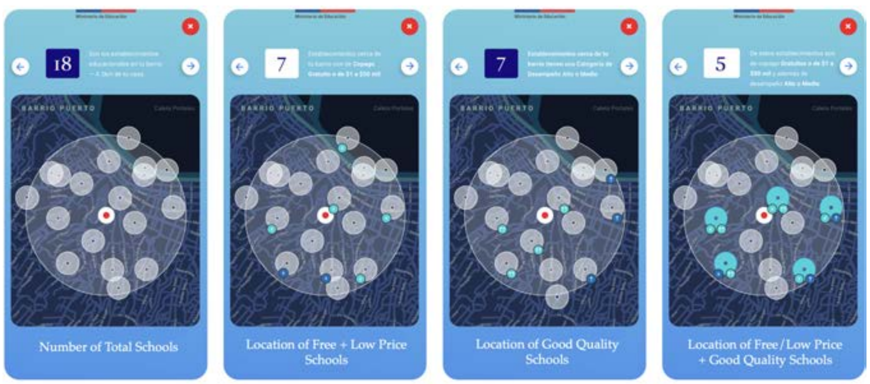


FIGURE S.5.—Search aid treatment 2

FIGURE S.6.—Additional distribution information for treatment groups 1 & 2

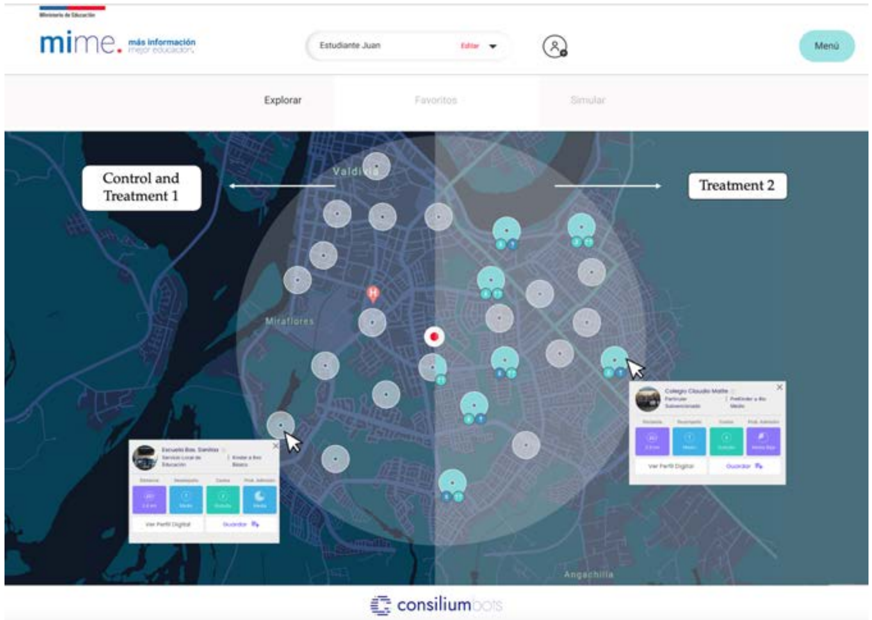


FIGURE S.7.—School explorer by treatment status

S.5. ADDITIONAL INFORMATION ON FEEDBACK INTERVENTION

Randomization: 318,520 applicants to entry grades with an application before the last week of the regular stage were randomized to receive feedback on their application. The research design was based on a geographical assignment where markets were divided into clusters. The assignment was stratified on the share of voucher eligible students, share of schools in high quality category, and the share of unassigned students in the main period of the previous year. Tables S.IV and S.V show balace checks for the high and low SES subsamples.

Intervention Details: Figure S.8 shows example screenshots of the information that was shown to parents as part of the feedback intervention. Panel (A) presents the student's current application with the option to view the applicants to date (Panel A.i) and the school characteristics (Panel A.ii). Panel (B) is a warning message if the current application is considered risky. Panel (C) and (D) present alternative schools not yet included in the current application. Panel (E) provides a detailed view of the alternatives offered. Panel (F) invites applicant to explore more schools and Panel (G) is a link to modify the current application.

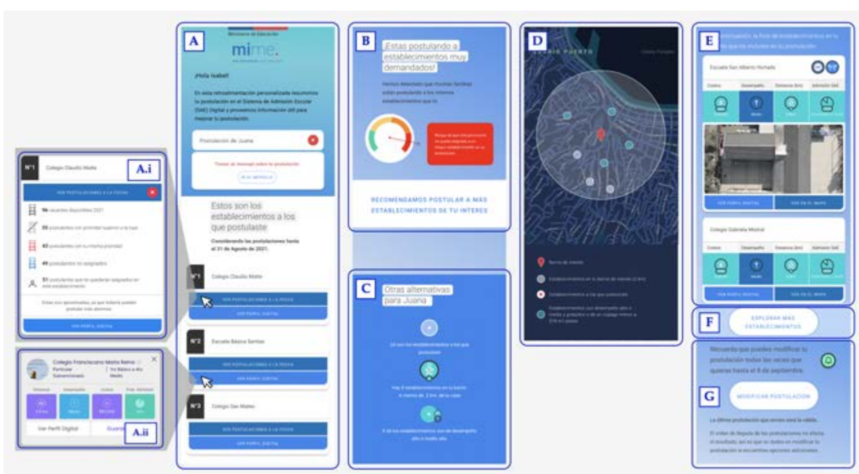


FIGURE S.8.—Feedback treatment

TABLE S.IV

BALANCE CHECK FOR FEEDBACK INTERVENTION FOR HIGH SES HOUSEHOLDS

	Control		Feedback Treatment		
	Mean	St. Dev.	Coeff.	St. Err.	N
	(1)	(2)	(3)	(4)	(5)
Panel A: Choice Environment					
Number of available schools	15.269	[8.410]	-1.561	(1.053)	521
Number of available highlight-worthy schools	8.456	[4.901]	-0.434	(0.602)	521
Panel B: Parent/Child Characteristics					
Child is female	0.507	[0.501]	-0.047	(0.043)	521
Child's birthyear	2017.146	[0.512]	-0.029	(0.046)	521
Number of younger siblings	1.160	[0.394]	0.015	(0.043)	521
Child has a disability (belief)	0.049	[0.215]	0.031	(0.025)	473
Parent works in a school	0.161	[0.368]	0.059	(0.038)	519
SEP household	0.173	[0.379]	-0.042	(0.036)	521
Panel C: Initial Knowlege and Beliefs					
Expected satisfaction with process	5.072	[1.477]	0.091	(0.155)	492
Listed any school as first preference	0.935	[0.246]	0.017	(0.031)	521
First-preference school is highlight-worthy	0.565	[0.497]	0.062	(0.069)	419
Perceived admission change for first-preference school	0.706	[0.270]	0.016	(0.026)	492
Number of schools known by name	3.432	[2.730]	0.251	(0.250)	521
Number of schools known well	2.204	[2.265]	0.163	(0.211)	521
Perceived number of available schools	7.514	[5.991]	0.970*	(0.567)	521
Perceived number of available highlight-worthy schools	3.391	[3.151]	0.559*	(0.325)	521
Parent believed to be SEP eligible	0.112	[0.316]	-0.018	(0.034)	521
SEP did not know about SEP status	0.656	[0.476]	0.035	(0.044)	521
Panel D: Search Treatments					
Search Treatment 1	0.340	[0.475]	-0.020	(0.044)	521
Search Treatment 2	0.323	[0.468]	0.007	(0.045)	521
Observations	272		249		

Note: This table shows balance for baseline covariates for the feedback intervention for high SES households. Column 1 reports the control mean of the dependent variable for each relevant subgroup (standard deviations in brackets). Column 3 reports the difference in the dependent variable from OLS regressions of each outcome on an indicator variable for feedback treatment assignments and market fixed effects. Standard errors clustered at the market cluster level are reported in parentheses. Variables in Panel A come from administrative data. Variables in Panels B and C come from the baseline survey.

TABLE S.V

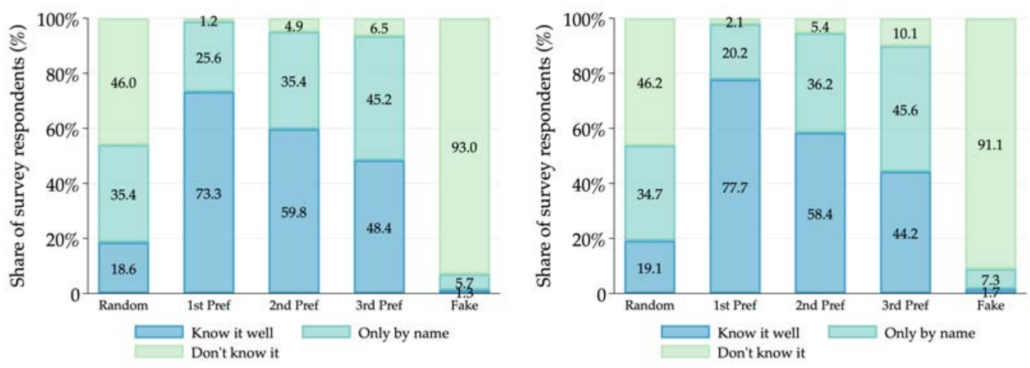
BALANCE CHECK FOR FEEDBACK INTERVENTION FOR LOW SES HOUSEHOLDS

	Control		Feedback Treatment		
	Mean (1)	St. Dev. (2)	Coeff. (3)	St. Err. (4)	N (5)
Panel A: Choice Environment					
Number of available schools	15.904	[9.628]	-0.472	(0.907)	2033
Number of available highlight-worthy schools	8.325	[5.051]	0.410	(0.561)	2033
Panel B: Parent/Child Characteristics					
Child is female	0.518	[0.500]	-0.004	(0.029)	2033
Child's birthyear	2017.111	[0.497]	-0.010	(0.028)	2033
Number of younger siblings	1.116	[0.350]	0.035	(0.022)	2033
Child has a disability (belief)	0.057	[0.233]	0.007	(0.014)	1793
Parent works in a school	0.038	[0.191]	-0.015**	(0.007)	1996
SEP household	0.529	[0.499]	-0.006	(0.028)	2033
Panel C: Initial Knowlege and Beliefs					
Expected satisfaction with process	5.328	[1.363]	-0.041	(0.073)	1918
Listed any school as first preference	0.937	[0.242]	-0.007	(0.020)	2033
First-preference school is highlight-worthy	0.667	[0.472]	0.046	(0.038)	1720
Perceived admission change for first-preference school	0.703	[0.263]	0.023*	(0.013)	1918
Number of schools known by name	3.330	[2.807]	0.103	(0.173)	2033
Number of schools known well	2.018	[2.078]	-0.096	(0.120)	2033
Perceived number of available schools	6.966	[6.263]	0.296	(0.349)	2033
Perceived number of available highlight-worthy schools	3.605	[3.068]	0.155	(0.182)	2033
Parent believed to be SEP eligible	0.179	[0.383]	-0.013	(0.018)	2033
SEP did not know about SEP status	0.682	[0.466]	0.021	(0.023)	2033
Panel D: Search Treatments					
Search Treatment 1	0.323	[0.468]	0.004	(0.026)	2033
Search Treatment 2	0.357	[0.479]	-0.007	(0.025)	2033
Observations	1102		931		

Note: This table shows balance for baseline covariates for the feedback intervention for low SES households. Column 1 reports the control mean of the dependent variable for each relevant subgroup (standard deviations in brackets). Column 3 reports the difference in the dependent variable from OLS regressions of each outcome on an indicator variable for feedback treatment assignments and market fixed effects. Standard errors clustered at the market cluster level are reported in parentheses. Variables in Panel A come from administrative data. Variables in Panels B and C come from the baseline survey.

S.6. ADDITIONAL DESCRIPTIVE RESULTS BY SES STATUS

In this section, we replicate our descriptive results separately for high and low SES parents. We observe little differences in school knowledge between high and low SES parents (Figure S.9). Figure S.10 shows that the beliefs about the distribution of school attributes of high SES parents tend to be more accurate than the beliefs of low SES parents. High SES parents also tend to have more accurate perceptions of school quality and prices (Figures S.11 and S.12), but low SES parents have slightly more accurate beliefs about placement chances (Figure S.13).



(a) High SES

(b) Low SES

FIGURE S.9.—Knowledge by SES status. Notes: Panel (A) plots the stated knowledge levels for five school categories: a random school within 2km of the respondent's home, the top three schools in the application, and a fake school for high SES households (N = 888). Panel (B) plots the same stated knowledge for five school categories for low SES households (N = 3057).

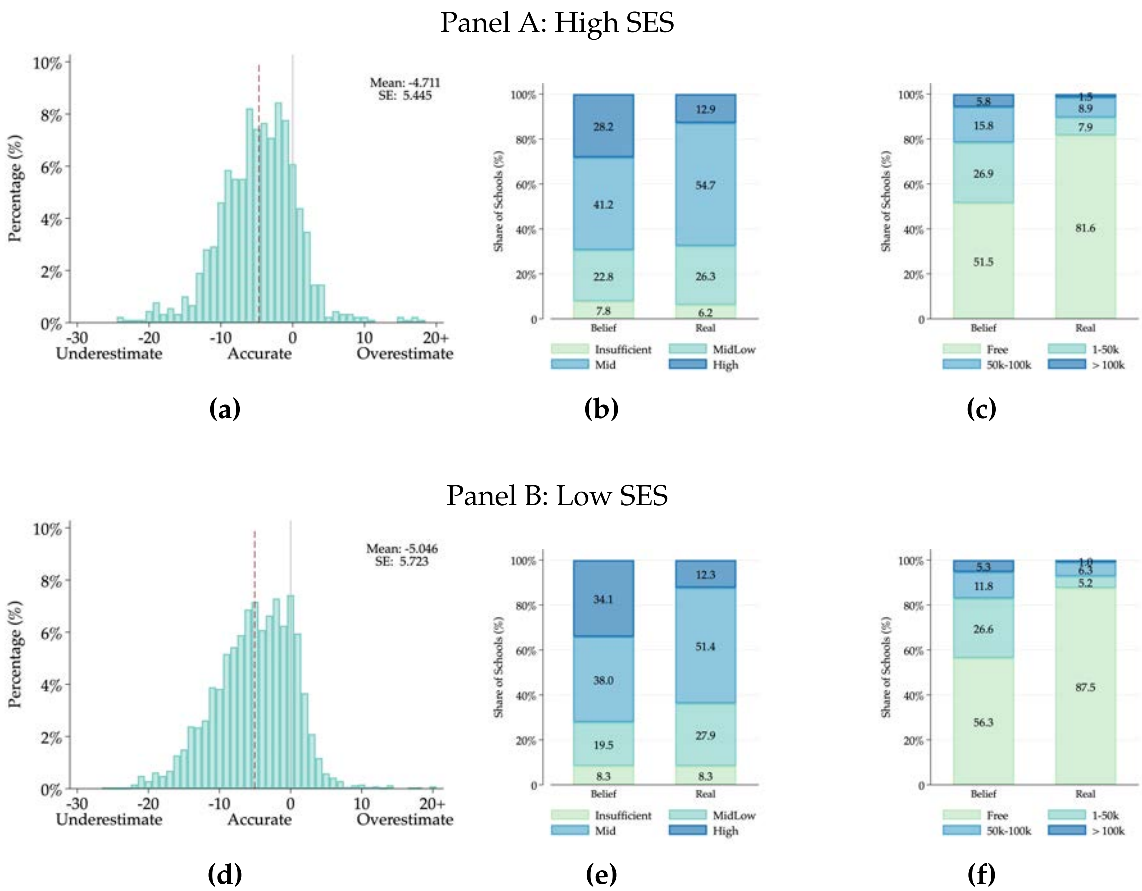
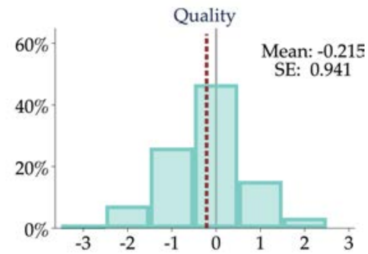
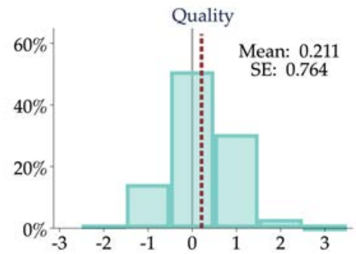


FIGURE S.10.—Beliefs about the Distribution of School Attributes by SES Status. Notes: Panels (A) and (D) show the bias in the beliefs of the number of highlight-worthy schools within 2km of the parent's home. Panel (B) and (E) show the perceived (left) and actual (right) share of schools in each of the four school quality categories. Panels (C) and (F) show the perceived (left) and actual (right) share of schools in each of the four school price categories. Panels (A-C) represent high SES households (N = 888) and Panels (D-F) represent low SES households (N = 3057).

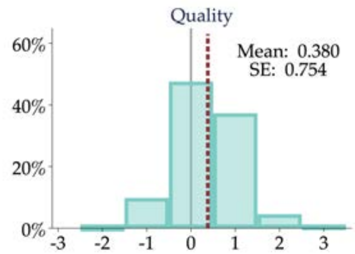
Panel A: High SES



(a) Random School Known

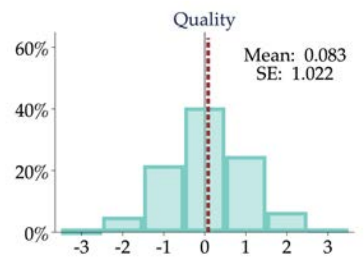


(b) Random School in App

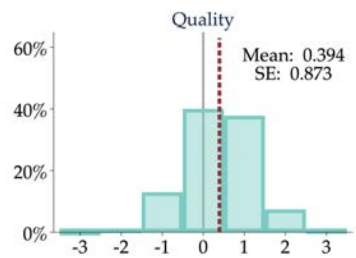


(c) First Preference

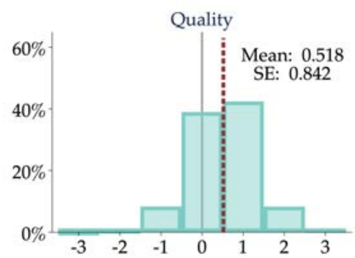
Panel B: Low SES



(d) Random School Known

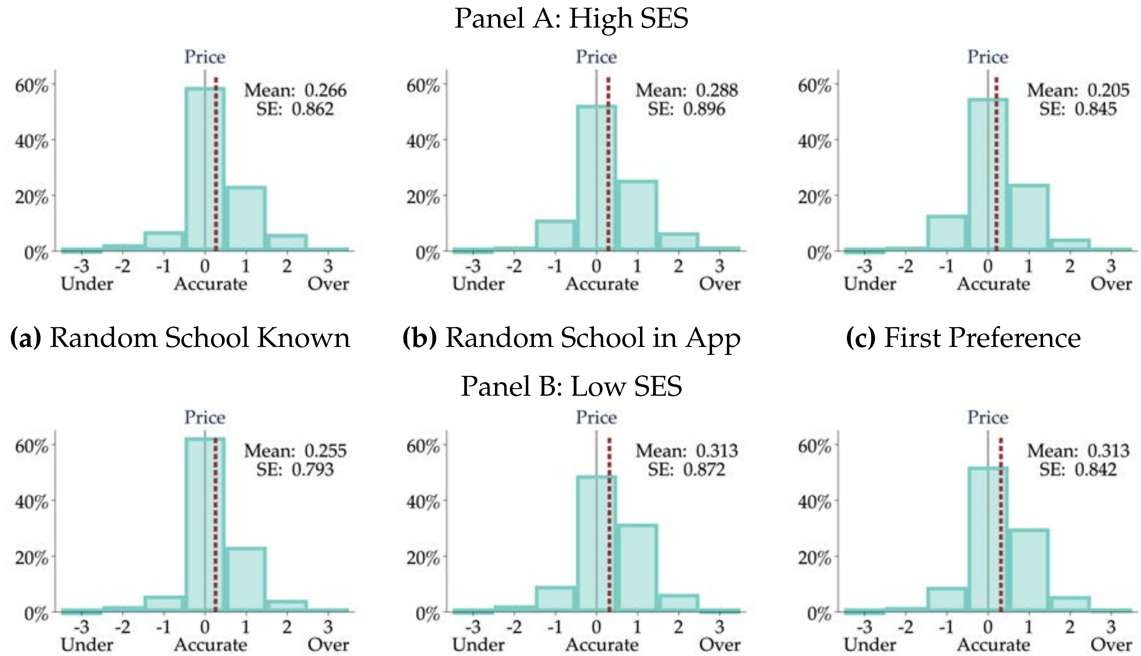


(e) Random School in App



(f) First Preference

FIGURE S.11.—Error in Quality by SES status. Notes: Panels (A) and (D) show the bias on perceived quality of a known random school asked in baseline. Panels (B) and (E) show the bias on perceived quality of a random school in the application list, excluding the first ranked school. Panels (C) and (F) show the bias on perceived quality of the first preference school at baseline. All biases are measured as perceived quality minus true quality. Positive values indicate that the parent perceived quality to be higher than the truth and negative values indicate that the parent perceived quality to be lower than the truth. Panels (A-C) represent high SES households (N = 888) and Panels (D-F) represent low SES households (N = 3057).



(a) Random School Known **(b)** Random School in App **(c)** First Preference

(d) Random School Known **(e)** Random School in App **(f)** First Preference

FIGURE S.12.—Error in Price by SES status. Notes: Panels (A) and (D) show the bias on perceived price of a known random school asked in baseline. Panels (B) and (E) show the bias on perceived price of a random school in the application list, excluding the first ranked school. Panels (C) and (F) show the bias on perceived price of the first preference school at baseline. All biases are measured as perceived price minus true price. Positive values indicate that the parent perceived price to be higher than the truth and negative values indicate that the parent perceived price to be lower than the truth. Panels (A-C) represent high SES households (N = 888) and Panels (D-F) represent low SES households (N = 3057).

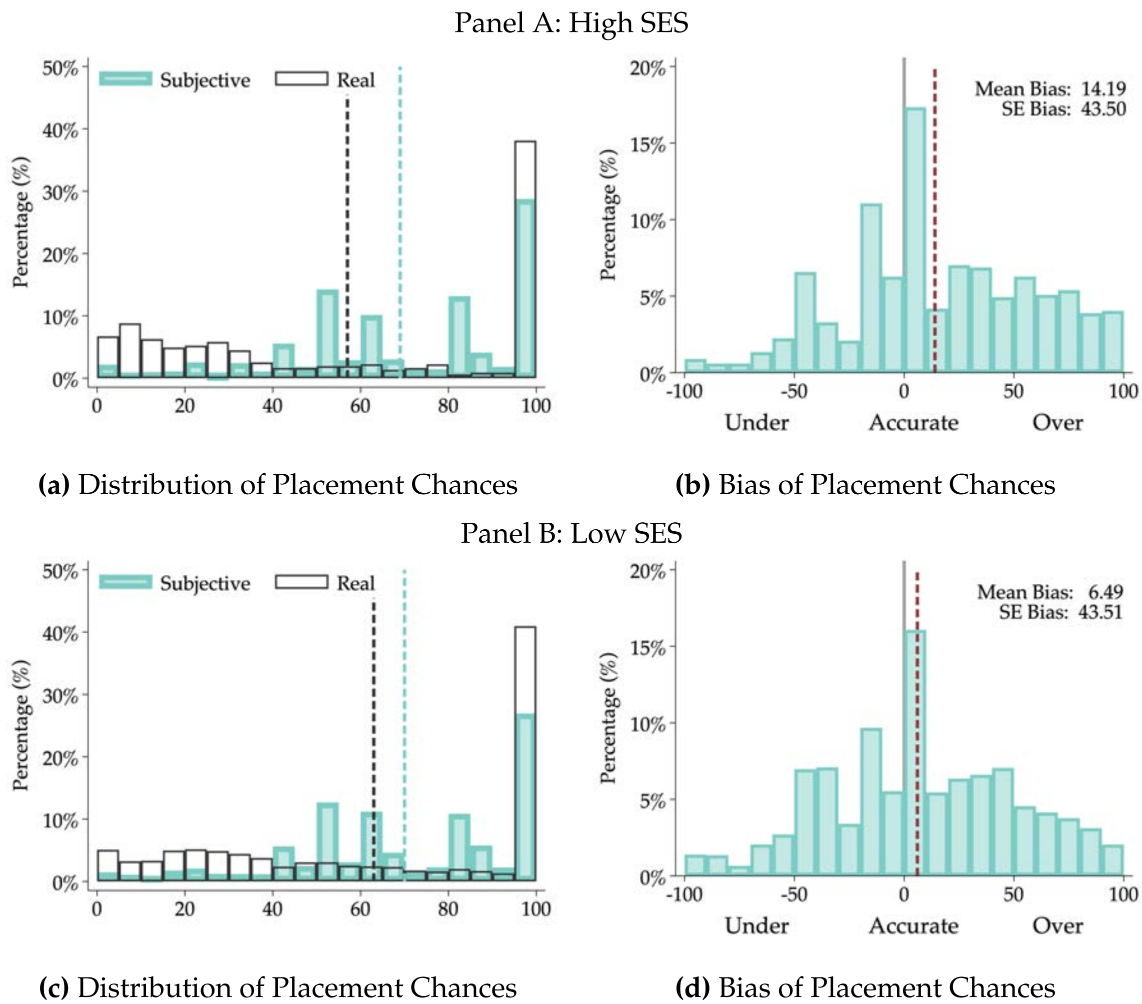


FIGURE S.13.—Error in Placement Chances by SES Status. Notes: Panels (A) and (C) show the perceived and true distribution of placement chances for first preference at baseline for high and low SES households respectively. Placement chances are calculated according to the most common program the school has if they have more than one program in the application process. Panels (B) and (D) show the bias on perceived placement chances of the first preference school at baseline, measured as perceived placement chances minus true placement chances. Positive values indicate that the parent perceived admission chances to be higher than the truth and negative values indicate that the parent perceived admission chances to be lower than the truth.