SEARCH AND BIASED BELIEFS IN EDUCATION MARKETS

PATRICK AGTE
Department of Economics, Stockholm School of Economics

CLAUDIA ALLENDE
Graduate School of Business, Stanford University

ADAM KAPOR
Department of Economics, Princeton University

CHRISTOPHER NEILSON
Department of Economics, Yale University

FERNANDO OCHOA
Department of Economics, New York University

First version: June 28, 2024

This version: February 25, 2026

This paper asks how search costs, limited awareness of schools, misperceptions of schools' attributes, and inaccurate beliefs over unknown schools affect parents' search and application decisions in Chile's nationwide 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. Parents do not know all schools, misperceive quality ratings of the schools they know and like, and underestimate the number of available schools. Addressing these frictions would raise welfare and cause students to match to schools with higher quality and value added, but search cost reductions alone would lead to lower quality. The effects of information and search-cost reductions are complementary.

Contact information: patrick.agte@hhs.se; callende@stanford.edu; akapor@princeton.edu; christopher.neilson@yale.edu; nano.ochoa@nyu.edu. We thank JPAL-LAC and the implementation team of data scientists and developers at ConsiliumBots for their help throughout the project. We thank the government partners that made this research possible, in particular the leadership at the Sistema de Admission Escolar (SAE) of the Ministry of Education in Chile. We thank Sofía Valdivia, María Constanza Muñoz, Isabel Jacas, and María Elena Guerrero for excellent research assistance, and Mariel Beyoda and Eduardo Vargas at J-PAL for their help during fieldwork. This project was pre-registered under AEA registry ID AEARCTR-0007998. This research was partially funded by NSF #1919504.

A model without misperceptions would incorrectly predict that providing full information reduces the quality of chosen schools.

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 a 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 is needed to rationalize low levels of search, and 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, (noisy) information, and (biased) beliefs interact with search costs, distorting their information-acquisition efforts, application decisions, and school assignments?

To address this question, we develop and estimate a model of parents' search and application decisions, collect novel data on parents' search activity, knowledge of schools, preferences, perceptions and beliefs, and implement two randomized field experiments to vary their information about known and unknown schools. Our model incorporates heterogeneous preferences and awareness of schools, noisy signals of schools' characteristics and admission chances, and biased beliefs over the distribution of signals and true values. Parents are initially aware of a lim-

[1]See e.g. Sorensen (2000), Handel and Kolstad (2015b), Agarwal et al. (2024), 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, 2024, Abaluck and Gruber, 2016), providing default options (Handel and Kolstad, 2015a), and encouraging delegation to intermediaries (Boehm, 2023).

ited set of schools. They may learn more via costly sequential search before submitting a rank-order list to a school choice mechanism. We measure parents' search activity with novel data from a "school explorer" platform that we designed and implemented in collaboration with the Chilean Ministry of Education. We measure parents' forecasts, perceptions, preferences, and awareness of schools via a panel of household surveys, which we link to explorer activity and to administrative application data. To generate variation, we conduct a "search aid" intervention in the explorer providing distributional information about nearby schools and a "feedback" intervention providing information about specific schools' prices, qualities, and admission chances.

We find that parents underestimate the number of available nearby schools and overestimate the quality of schools that they like and know. These biases reduce their search effort and cause their children to attend lower-quality schools. Providing full information and knowledge of all nearby schools—an infeasible benchmark—would achieve welfare gains of the equivalent of 0.765 fewer kilometers traveled, and cause students to attend schools with higher quality scores and value added. Correcting parents' perceptions of known schools' price, quality, and admissions chances, and their beliefs about the number and characteristics of nearby schools, would achieve two thirds of these welfare gains and essentially all of the gains in quality and value added. In contrast, driving search costs to zero without providing information would achieve about a third of the possible welfare gains and lead students to lower-quality schools than at baseline. Costs and (mis)information interact, with gains from jointly removing search costs and eliminating information frictions larger than the sum of each policy's effect separately. Had we lacked data on parents' perceptions and incorrectly assumed parents perceive known schools' price and quality accurately, we would have reversed the sign of the quality and value-added impacts.

We establish these findings through the following steps. First, we present descriptive analyses using our baseline survey, subsequent "midline" and "endline" surveys, and administrative application data. Parents know at least by name fewer than 60% of random nearby schools asked about at baseline, several months before applications are due, and know fewer than 20% of schools well (self-reported).

Thus, there is a role for search effort. However, parents underestimate the number of nearby schools that they have not discovered, and the number of nearby high-quality low-price schools, reducing perceived returns to search. Moreover, parents perceive the price, quality, and admissions chance of "known" schools with noise, leading them to overestimate the quality of their first-choice school by about half a category on a four-point scale, on average (Figure 3b). Administrative preference rankings over schools are sensitive to perceived quality score, suggesting that parents value this characteristic.

Second, we analyze our experiments. We embedded the "search aid" intervention in the school explorer platform, and shared it with parents just after our baseline survey. Its goal was to provide information relevant for the returns to searching for schools. Its first treatment arm provided personalized information about the joint distribution of nearby schools' prices and quality scores. A second arm provided the same information and additionally highlighted nearby low-cost high-quality schools, making them salient. These treatments caused changes in beliefs, search activity, knowledge of schools, and placements, with the effects driven by parents with college-educated mothers, a proxy for high SES. Effects are larger for parents whose initial beliefs were farther from the truth.

Our "feedback" intervention used administrative application data, roughly a week before the deadline, to provide tailored feedback on the initial application parents submitted, targeting misperceptions about known schools.[3] Its goal was to provide information about specific, known schools. This intervention corrected perceptions about school characteristics and had large impacts on the application decisions of low-SES parents.

Finally, we use the data, descriptive analyses, and experiments to identify and estimate 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

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

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). We extend prior empirical school-choice models(Agarwal and Somaini, 2018, Kapor et al., 2020, 2024) by incorporating misspecified beliefs and by modeling search, allowing us to simulate endogenous information acquisition decisions when beliefs, perceptions, or search costs change.

Our empirical strategy exploits repeated measurements within person and school as well as shifters of information—experimental variation and households' search outcomes—that are excluded from utility. 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, perceptions of known schools, and forecasts of the returns to search, from up to three surveys per family.

We estimate the model in three 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, 2018, Kapor et al., 2020) with limited availability (Kapor et al., 2024) to our panel-data context.[4] Second, we estimate parents' beliefs over unknown schools, using repeated survey measures of their beliefs about the number and characteristics of nearby schools. Third, we impose optimality of their search decisions to recover search costs and simulate counterfactuals.

Our results provide a unified treatment of the impacts of providing information about admissions chances (Arteaga et al., 2022, Ajayi et al., 2025, Gurantz et al., 2021, Hoxby et al., 2013, Hoxby and Turner, 2015, Luflade, 2019) 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

[4]For choice with limited availability and/or awareness, see also Agarwal and Somaini (2025) and He et al. (2024). We differ by observing rank-order lists and measures of parents' awareness of schools.

et al., 2017, Allende et al., 2019, Bergman et al., 2020). The estimated impacts of information interventions depend on both the importance and size of the friction being targeted and the effectiveness or "takeup" of the intervention. Our model and direct data on updating of beliefs and information let us see which frictions matter and extrapolate to hypothetical interventions with full policy take-up, correcting multiple information margins simultaneously. We validate the model against experimental effects it was not estimated to match.

Our paper is related to contemporaneous work estimating households' perceptions of school quality (Corradini, 2024), eliciting perceptions of schools' characteristics from household surveys (Corradini and Idoux, 2023), and providing information to school choice participants (Campos, 2024, Ainsworth et al., 2023). We differ by modeling search and information acquisition, allowing counterfactuals in which participants endogenously acquire information.

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., 2024). 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., 2025, 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, 2024). Our model is closely tied to the institutional details of our setting and the features of the explorer platform.

A limitation of this paper is that we conduct "single-agent" counterfactuals, holding schools' characteristics and admissions chances fixed. In our setting 41% of first-choice schools had excess capacity. Entry was heavily regulated and prices were fixed,[5] but other aspects of schools such as quality "markdowns" may ad-

[5]See our supplementary material S.1.

just 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 describes the setting. Section 3 presents the model and provides a motivating example. Section 4 presents our study design. Section 5 provides descriptive analysis. Section 6 evaluates the experiments. Section 7 provides assumptions for estimation, Section 8 describes estimation, and Section 9 presents results. Section 10 concludes.

2. EMPIRICAL SETTING

Our study takes place within the Chilean centralized School Admission System (SAE). The SAE accounts for 89% of primary school matriculation in the country, including almost all public and voucher-accepting private schools.

The SAE assigns applicants to schools using a student-proposing deferred acceptance algorithm (Correa et al., 2019). Seats at oversubscribed schools are rationed through quotas, coarse priorities, and school-specific lottery-based tiebreakers. Therefore, placement probabilities are student-school specific. Parents form school portfolios and submit a rank-ordered-list (ROL) through a centralized platform. There are no restrictions on ROL length, making the mechanism strategy-proof.[6] The main application round starts in early August. Applications can be submitted and freely edited for roughly one month (Arteaga et al., 2022).[7]

We focus on parents applying to entry grades (pre-K, kindergarten, and first grade).[8] 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

[6]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.

[7]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.

[8]In Chile, pre-kindergarten and kindergarten are not mandatory. Nonetheless, most schools in the country offer both pre-kindergarten and kindergarten.

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 the official academic quality classification produced by Chile's Education Quality Agency.[9] In the Online Appendix, we provide additional details on the construction of this measure and show that the underlying continuous score is highly correlated with our school value-added measure (Appendix Section A.1, Figure A.1b).[10] Consistent with its salience, 95% of parents in our baseline survey report that obtaining Education Quality Agency information is a necessary step before adding a school to their application (Appendix Section A.1). Related experimental evidence suggests that such information is relevant for students' outcomes: Allende et al. (2019) study a randomized information intervention that reports test-score-based school quality information from the same Agency (distinct from the composite index we use here) and find sizable medium-run gains in students' own test scores.

Most schools have excess capacity. As shown in Table A.II, 77% of school programs at entry grades have excess capacity, defined as having more available seats than students assigned through the centralized system. Among programs that actively participate in the centralized assignment, there are on average 2.1 applicants

[9]The Education Quality Agency (Agencia de Calidad de la Educación) classifies schools into four categories (high, medium, medium-low, and insufficient) based on a continuous performance score that places substantial weight on standardized learning outcomes (SIMCE) and is adjusted for student socioeconomic context. The SAE platform makes parents familiar with this classification as it matters 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.

[10]Appendix Table A.I also provides details about how the quality measure is correlated with peer composition.

per available seat. Table A.III shows that lower-quality schools are more likely to be undersubscribed.

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 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.

Motivated by this evidence, 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, in collaboration with this paper's research team, and was made available to the participants through the government website.

The school explorer platform aggregated all public information on all schools in the country. Two features were key for our study. First, the explorer used data on students' location and characteristics to provide individualized information, which we experimentally varied across participants. Second, to generate useful search data, the interface was simple, consisting of a map centered on the prospective student's home and indicating the locations (but not names) of nearby schools. Accordingly, the explorer allowed parents to discover and compare nearby schools that they had not been aware of, or whose names they had not known.

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. [11]

[11] Information on admission probabilities were displayed in quartiles to make the results easily understandable for parents. We treat the information as a coarse but informative signal about the continuous underlying admission probabilities. Arteaga et al. (2022) use a similar approach to communicate admission chances.

The explorer allowed parents to perform two main activities. Clicking on a school location on the map ("school pin click") provided the name of the school, 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.[12]

3. FRAMEWORK

We consider the problem of an imperfectly-informed parent participating in the school choice process. The parent may be unaware of some schools, may hold inaccurate beliefs about the number and characteristics of unknown nearby schools, may receive noisy signals of the characteristics of schools it "knows", and may interpret these signals incorrectly. To discover schools and learn more about them, the parent may engage in costly sequential search. Once done, the parent submits a rank-order list to a student-proposing deferred acceptance algorithm.

The parent's misperceptions and imperfect information may distort its search and application decisions, possibly causing it to be assigned to a less-desirable school. To understand which frictions matter and how they interact with search costs, we now develop a model of this parent, which motivates our research design, described in the next section. We first present the framework, then provide a motivating example, and finally summarize comparative statics that guide our design and analysis. We provide the full parametric model in Section 7.

3.1. Model

We consider parents $i \in I$ choosing schools $j \in J$.[13] Parent i has a maximal choice set $J_i \subseteq J$. If placed in $j \in J_i$, household i receives a true payoff $u_{ij}(x_{ij}, \varepsilon_{ij})$ that depends on researcher-observed school characteristics x_{ij} and a match-specific term ε_{ij}. However, j will reject i with probability $r_{ij} \in [0, 1]$, independently across schools. The utility of being unplaced is $u_{i0} = 0$. We treat $(u_{ij}(\cdot), x_{ij}, \varepsilon_{ij}, r_{ij})$ as prim-

[12]The detailed view provides information on nine sections, including available school infrastructure and extracurricular activities. Some schools also provide a virtual tour and testimonies by the school principal and teachers. See supplementary material for screenshots and additional details on the explorer platform.

[13]We treat parents as the relevant decision-makers, as our analysis focuses on primary school choice.

itives of the environment, fixed under counterfactuals. Where no ambiguity arises, we will write $u_{ij} = u_{ij}(x_{ij}, \varepsilon_{ij})$, suppressing the (true) arguments.

Timing: Time is discrete: $t \in \{0,1,2,3\}$. At $t=0$, parents are endowed with awareness of a subset of schools and imperfect information about "known" schools. At $t=1$, they engage in sequential search using the explorer. At $t=2$, parents submit rank-order lists to a student-proposing deferred acceptance algorithm. At $t=3$, they may revise their rank-order lists in response to new information before the application deadline. This timing is motivated by the timing of our interventions and the school choice process.

Before presenting the application and search decisions, we introduce the primitives relating to information.

1. Limited awareness of schools: Let $\pi_{ijt} \in \{0,1,2,3\}$ denote the level of knowledge of school $j \in J_i$ held by parent i at time t. If $\pi_{ijt} = 0$ then parent i does not know j well enough to apply to it. School j is "known by name" if $\pi_{ijt} = 1$, "known well" if $\pi_{ijt} = 2$, and known perfectly when $\pi_{ijt} = 3$.[14] Four levels are not essential, but are motivated by the design of our survey measures.

2. Noisy signals: If school j is known but not perfectly so, then parent i receives noisy signals of school j's payoff-relevant characteristics and rejection chance. Formally, when $\pi_{ijt} \in \{1,2\}$ the parent receives signals ($\hat{x}_{ij}^{(\pi_{ijt})}$, $\hat{\varepsilon}_{ij}^{(\pi_{ijt})}$, $\hat{r}_{ij}^{(\pi_{ijt})}$), drawn from joint distributions $G^y(\hat{y}^{(1)}, \hat{y}^{(2)}, y)$ over signals and true values for $y \in \{x, \varepsilon, r\}$ respectively. It is a key simplifying assumption that the same index, π_{ijt}, governs awareness of school j and the accuracy of information about it.

3. Biased beliefs and misperceptions: We allow parents to update incorrectly given the signals they see—that is, to misperceive the signals—and to make inaccurate forecasts of how they will update with more information. The former errors will distort rank-order lists, while the latter will affect the perceived returns to search. The true distributions G^y have densities that factor into true-value densities and conditional densities of signals: $g^y(\hat{y}^{(1)}, \hat{y}^{(2)}, y) = f^y(y) g^y(\hat{y}^{(1)}, \hat{y}^{(2)} | y)$ for $y \in \{x, \varepsilon, r\}$. Parents believe signals are produced as $\hat{g}^y(\hat{y}^{(1)}, \hat{y}^{(2)} | y)$ and believe the truth is distributed as $\hat{f}_{it}^y(y)$.

[14]The state $\pi_{ijt} = 3$ is used to define full-information environments. It occurs in counterfactual benchmarks but not in our data.

4. Subjective expected payoffs: The subjective expected utility of a placement in j, given the information available at state $\pi_{ijt} > 0$, is:

$$\hat{u}_{ij}^{\pi_{ijt}} = E_{\hat{g}}(u_{ij}(x_{ij}, \varepsilon_{ij}) | \hat{x}_{ij}^{(\pi_{ijt})}, \hat{\varepsilon}_{ij}^{(\pi_{ijt})}; \pi_{ijt}),$$

where the expectation is according to the subjective beliefs. Our model accommodates misperceptions. For instance, the parent may hold the beliefs $\hat{g}^y(\hat{y}|y) = 1(y = \hat{y})$, in which case it treats its (possibly inaccurate) signal as the truth.

5. Subjective forecasts: Search-relevant beliefs $\hat{f}_{it}^{y}(\cdot)$ at time t may be systematically biased as well. We flexibly model beliefs over x and allow for systematic optimism/pessimism and compression/overdispersion of match shocks ε and admission chances r. In addition, we consider misperceptions of the number of unknown schools. The true number is $N_i^{\text{unknown}}(\pi_{i \cdot t}) = |J_i \setminus C_i(\pi_{i \cdot t})|$, but parents believe there are $\hat{N}_{it}$ such schools. These distortions may affect the perceived returns to search.

6. Misspecified learning and updating: Given our data, we focus on learning about the number of unknown schools $\hat{N}_{it}$ and the distribution of unknown schools' observables $\hat{f}_{it}^{x}(\cdot)$. These beliefs may update over time as a function of parents' search outcomes as well as the treatments that we provide. We model this updating flexibly.

We now present the rank-order portfolio and search decisions.

Student assignment: A rank-order list L_i is an ordered list $(j_1, \ldots, j_k)$ with $j_1, \ldots, j_k \in J_i$. The payoff of such a list is:

$$U(L_i, u_i, r_i) = \sum_{n=1}^{k} \left(\prod_{m=1}^{n-1} r_{ij_m} \right) (1 - r_{ij_n}) u_{ij_n}. \tag{1}$$

That is, EU-maximizing parent i is considered for the nth ranked school if they have been rejected by the $n-1$ higher-ranked schools. If so, i is admitted with probability $1 - r_{ij_n}$, receiving payoff u_{ij_n} in this case. The final list, submitted at $t = 3$, determines i's placement and payoff.

Application decision: The expected payoff term (1) is maximized by ranking all acceptable options in descending order of utility $u_{ij}(x_{ij}, \varepsilon_{ij})$.[15] Moreover, because expression (1) is linear in u_{ij} for each j, the subjectively expected utility maximizing rank-order list $L_i^*(\pi_{i\cdot t}) = (j_1^*, \ldots, j_k^*)(\pi_{i\cdot t})$ at time t ranks all schools $\{j \in J_i : \pi_{ijt} > 0, \hat{u}_{ij}^{\pi_{ijt}} > 0\}$ in descending order of subjective expected utilities $\hat{u}_{ij}^{\pi_{ijt}}$. We adopt the convention that $\hat{u}_{ij}^{(0)} < 0$ for all i, j.[16] Parent i's subjective expected payoff is thus given by:

$$\hat{U}_i^*(\pi_{i\cdot t}) = U(L_i^*(\pi_{i\cdot t}), \hat{u}_i^{\pi_{i\cdot t}}, \hat{r}_i),$$

where without loss we suppose $E_{\hat{g}}(r_{ij}|\hat{r}_{ij}) = \hat{r}_{ij}$. When we observe a rank-order list, it is the "truthful" list $L_i^*(\pi_{i\cdot t})$.

Search decision: Consistent with descriptive evidence, we model search as a sequence of stop-or-continue decisions. Initially, parent i has conducted $s = 0$ pin clicks. To sample the sth school, i incurs a cost c_{is} and draws a school $j \in J_i$. Each school has a salience level, $q_j(d_{ij}, x_{ij}, \tau_i)$, which depends on distance d_{ij}, characteristics x, and the search technology, τ_i. The probability of "clicking" a particular school, conditional on search, is given by

$$p_{ij}^{\text{click}} = q_j(d_{ij}, x_{ij}, \tau_i) / \sum_{j \in J_i} q_j(d_{ij}, x_{ij}, \tau_i).$$

If parent i underestimates the number of unknown schools, it underestimates the probability of finding one. In particular, i perceives the salience of unknown schools—those for which $\pi_{ijt} = 0$—to be

$$\hat{q}_{ij}(\hat{N}_{it}) = \frac{\hat{N}_{it}}{N_{it}^{\text{unknown}}} q_j(d_{ij}, x_{ij}, \tau_i), \tag{2}$$

[15]For simplicity we take r_{ij} as invariant to the parent's rank-order list. In principle, in school choice mechanisms, these chances may vary. However, it is well known that, in the student-proposing deferred acceptance mechanism, the payoff is maximized by ranking all acceptable options in descending order of payoff.

[16]An alternative interpretation of the state $\pi_{ijt} = 0$ is that, given the limited information available, the expected payoff is lower than the outside option.

where $\hat{N}_{it}$ is the subjective number of schools.

Clicking on a school at time $t = 1$ stochastically increases the information level $\pi_{ijt'}$ for $t' > 1$. It also affects future-period beliefs over the number and characteristics of unknown schools. For instance, if the parent finds an unknown high-quality school, it will update about the number of schools and the share that are of high quality.

The value of a new school: We now turn to the search decision at $t = 1$. We begin by considering the value of discovering a new school. Fix a vector $\pi_{i\cdot t}$ and its associated optimal rank-order list $(j_1^*, \ldots, j_k^*)(\pi_{i\cdot t})$. Suppose a new school becomes known, i.e. $\pi_{ij't}$ increases from 0 to $q > 0$ for some j', with all other information held fixed. Let $\pi'_{i\cdot t} = \pi_{i\cdot t} + q \cdot e_{j'}$ denote the new state.

Let n^* denote the number of elements of $\{\hat{u}_{ij}^{\pi_{ijt}} : \pi_{ijt} > 0\}$ that are greater than $\hat{u}_{ij'}^{q}$. The agent's expected payoff gain from adding such a school is given by

$$\hat{U}_i^*(\pi'_{i\cdot t}) - \hat{U}_i^*(\pi_{i\cdot t}) = \left(\prod_{m=1}^{n^*} r_{ij_m}\right)(1 - \hat{r}_{ij'})\left(\hat{u}_{ij'}^{q} - U((j_{n^*+1}, \ldots, j_k), \hat{u}_i^{\pi_{i\cdot t}}, \hat{r}_i)\right). \quad (3)$$

Expression (3) is linear in the new school's rejection chance $\hat{r}_{ij'}$, and is a piecewise linear nonnegative nondecreasing convex function of the new school's expected payoff given the drawn signals, $\hat{u}_{ij'}^{q}$. [17]

Simplifying search: We assume that parents use a one-step-lookahead heuristic. That is, having conducted $s - 1$ clicks so far, they continue when the gain from conducting one further click is larger than the search cost c_{is}. The value of the sth search is given by

$$\hat{E}(\hat{U}_i^*(\pi'_{i\cdot 1}) - \hat{U}_i^*(\pi_{i\cdot 1})), \quad (4)$$

where π is the current state, and π' is the state in the event the parent conducts its sth search, with the expectation over the identity of the found school j, the amount of information gained $(\pi'_{ij1} - \pi_{ij1})$, and the characteristics and signals drawn from $\hat{f}$ and $\hat{g}$ in the event $\pi'_{ij1} > \pi_{ij1}$.

[17] See Appendix A.2.

We view the one-step-lookahead heuristic as a simple model of actual decision-making. Making fully optimal stopping decisions in this environment is difficult, as agents must integrate over their future beliefs over unknown schools. For example, a parent may want to search, even when the immediate gain (expression 4) is low, if it believes there is a set of possible signals that will cause it to update positively about the distribution of school payoffs. Such reasoning would require a high degree of sophistication from parents and would compel us to take a stand on subjective beliefs over future forecasts.

Information shifters: We end by briefly describing the role of search and experimental variation. Parents' information state π_{ijt} depends on a set of time-varying observables, w_{ijt}, that are excluded from preferences. We experimentally manipulate elements of w_{ijt} via an intervention at $t = 3$ providing "feedback" on the rank-order list submitted at $t = 2$. At the beginning of $t = 1$, we provide a search-aid intervention that shifts beliefs about the number and observable characteristics of schools. We parameterize the perceived number of schools $\hat{N}_{it}$ and the subjective distribution $\hat{f}^x_{it}$ as functions of parameters λ_{it} and φ_{it}, respectively, which in turn vary with treatment and search history; details are provided in the Appendix. An additional treatment arm at time $t = 1$ varies the search technology τ_i, making high-quality low-cost schools more salient.

3.2. *A Motivating Example*

We now show a simple illustrative example of the forces that distort the returns to search. Our example takes place at $t = 1$. For brevity, we omit t subscripts.

Suppose parent i has two possible schools, $J_i = \{1, 2\}$, of which it knows one: $\pi_{i1} = 1$ but $\pi_{i2} = 0$. The known school provides true payoff $u_{i1} = 1$ if the child is admitted, but rejects with probability $r_{i1} = 0.25$. Given its knowledge level, the parent forms perceptions of payoff and admissions: $\hat{u}_{i1}^{(\pi_{i1})} = E_{\hat{g}}(u_{i1} \mid \hat{x}_{i1}^{\pi_{i1}}, \hat{\varepsilon}_{i1}^{\pi_{i1}})$ and $\hat{r}_{i1}^{\pi_{i1}}$.

A single search incurs cost c_{i1} and reveals school 2 with probability $p_{i2}(\tau_i)$, causing π_{i2} to equal 1. However, the parent does not know whether there is a "school 2". It believes its choice set is $J_i = \{1\}$ with probability $1 - \lambda_i$ and $J_i = \{1, 2\}$ with probability λ_i, for $\lambda_i \in [0, 1]$. Hence it believes the probability of discovering a second school is $\lambda_i p_{i2}(\tau_i)$. If search does not reveal school 2, it revisits school 1. In this example we take the gain from revisiting school 1 to be zero.

The parent expects the payoff of school 2, if found, to be drawn from a belief distribution $\hat{f}^u(\hat{u})$, which summarizes beliefs over x, ε, and signals thereof. For simplicity, school 2 admits with probability one. Allowing admission risk would not affect the core ranking-reversal mechanism described below.

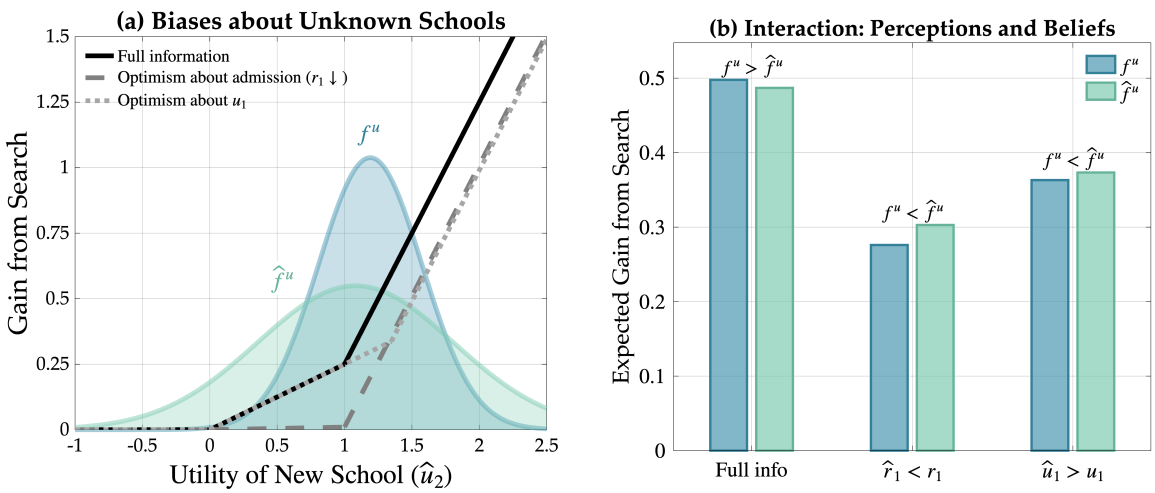


FIGURE 1.—Motivating example: return to search under misperceptions and biased beliefs. Panel (a) overlays misperceptions about the known school (three gain-from-search curves) and beliefs about unknown schools (f^u and $\hat{f}^u$). Panel (b) summarizes expected gains for all six combinations and highlights where the ranking between f^u and $\hat{f}^u$ flips. The y-axis in Panel (b) is the expected gain, i.e., the integral of each gain-from-search curve in Panel (a) against each distribution. The three groups in Panel (b) correspond to accurate perceptions, admission optimism ($\hat{r}_{i1} < r_{i1}$), and quality optimism ($\hat{u}_{i1} > u_{i1}$).

The gain from discovering school 2: Panel (a) of Figure 1 plots the perceived gain as a function of the new school's payoff (expression 3), together with the true and subjective payoff distributions over unknown schools. When $\hat{u}_{i2} < 0$, the new school is worse than the outside option and contributes nothing. When $0 < \hat{u}_{i2} < \hat{u}_{i1}^{\pi_{i1}}$, the new school is ranked second and is relevant with probability $\hat{r}_{i1}$. When $\hat{u}_{i2} > \hat{u}_{i1}^{\pi_{i1}}$, the new school is ranked first. The solid line shows the case of accurate perceptions, i.e. $\hat{u}_{i1} = u_{i1}, \hat{u}_{i2} = u_{i2}, \hat{r}_{i1} = r_{i1}, \hat{r}_{i2} = r_{i2}$.

Misperceptions about known schools. The dashed line shows a parent that underestimates its rejection probability, perceiving $\hat{r}_{i1} \approx 0$. Because admission to school 1 is perceived as nearly certain, a second-ranked school is perceived as almost irrelevant. The dotted line shows a parent that overestimates school 1's quality, perceiving $\hat{u}_{i1} > u_{i1}$ and raising the threshold for the new school to be ranked first. Both types of misperception reduce the perceived return to search.

Beliefs about unknown schools and the interaction: Panel (a) illustrates the effects of inaccurate beliefs over unknown schools' payoffs, which may operate through observables $\hat{f}^x(\cdot|\varphi_i)$, unobservables $\hat{f}^\varepsilon(\cdot)$, or the number of schools λ_i. The expected return to search (expression 3) is the integral of the gain curve weighted by the distribution of payoffs. The true and subjective distributions of new-school payoffs, $f^u(\hat{u})$ and $\hat{f}^u(\hat{u})$, are overlaid on the gain curve.[18] Subjective beliefs $\hat{f}^u$ exhibit pessimism (lower mean) but greater uncertainty (higher variance) than the true distribution.

Under accurate perceptions of the known school (solid gain curve), f^u yields a higher expected return. However, both types of misperceptions in Panel (a) give higher expected payoffs under the higher-variance subjective distribution $\hat{f}^u$. Panel (b) shows expected payoffs under each distribution and type of misperception. An implication is that information interventions that move beliefs from $\hat{f}^u$ to f^u, reducing pessimism, need not induce more search. In contrast, the returns to search increase when the probability of a second school (λ_i) rises or the search technology τ_i raises the probability of finding it.

3.3. *Comparative Statics*

We now state comparative statics that guide our experimental design and analysis, generalizing the previous example. Throughout, fix a parent i and period t and suppress the i and t subscripts. Let $(j_1, \ldots, j_k)$ be the known schools—those with $\pi_j > 0$—ranked in descending order of perceived payoff, with perceived payoffs $\hat{u}_{j_1}^{(\pi_{j_1})} \geq \cdots \geq \hat{u}_{j_k}^{(\pi_{j_k})} > 0$ and perceived rejection chances $\hat{r}_{j_1}, \ldots, \hat{r}_{j_k}$. Let v^{known} denote the expected gain (Eq (3)) conditional on the drawn school being a known school $j \in \{j_1, \ldots, j_k\}$.

[18] For illustration, we depict these as gaussian densities with distinct means and variances; Section 7 specifies a richer parametric structure for beliefs. The unconditional distribution of search outcomes also includes the event of not discovering a school. We depict conditional-on-discovery densities for clarity.

A distribution F dominates G in the *increasing convex order*, written $F \geq_{\text{icx}} G$, if $E[\phi(X)] \geq E[\phi(Y)]$ for every increasing convex function $\phi : \mathbb{R} \to \mathbb{R}$, where $X \sim F$ and $Y \sim G$. Equivalently, $F \geq_{\text{icx}} G$ if and only if $\int_x^\infty 1 - F(u)\mathrm{d}u \geq \int_x^\infty 1 - G(u)\mathrm{d}u$ for all x.[19]

PROPOSITION 1—Comparative statics: *Suppose the gain v^{known} is zero, and that perceived rejection chances of unknown schools $\hat{r}_{j'} \sim g^r(\hat{r})$ are independent of perceived utilities $\hat{u}_{j'}$.*

(a) The expected gain from discovering a new school (expression (3)) is decreasing in each perceived payoff $\hat{u}_{j_m}$, $m = 1, \ldots, k$, and increasing in each perceived rejection chance $\hat{r}_{j_m}$, $m = 1, \ldots, k$.

(b) Let f^u and $\hat{f}^u$ be two distributions of the expected payoff of a newly discovered school. The expected gain (Expression (3)) is higher under f^u than under $\hat{f}^u$ at every information state π_i if and only if f^u dominates $\hat{f}^u$ in the increasing convex order.

The proof is given in Appendix A.2. Part (a) says that a parent that overestimates known schools' payoffs or underestimates their rejection chances will perceive a lower return to search. Part (b) says stringent conditions are required for one distribution of unknown payoffs to dominate another. In particular, as our example illustrates, "reducing pessimism" (raising the mean of $\hat{f}^u$ toward that of f^u) need not raise the perceived return to search. f^u must have both a weakly higher mean and be more dispersed. A plausible information intervention may reduce pessimism but also reduce the variance of beliefs. The effect of such an intervention is an empirical question.

The signs of the following impacts are also ambiguous. They depend on parents' current state. We construct examples in Appendix A.2.

Increasing the perceived number of unknown schools: If the expected gain from revisiting a "known" ($\pi_j > 0$) school is smaller than the expected gain conditional on clicking an unknown ($\pi_j = 0$) school, then returns are increasing in $\hat{N}_i$. Otherwise, they are decreasing.

[19] Shaked and Shanthikumar (2007, Theorem 4.A.1 and eq. 4.A.5).

Uniform bias in rejection chances at known and unknown schools: The return to discovering a first school is lower when the perceived rejection chance is higher. In contrast, when many schools are known, pessimism about all rejection chances may raise the perceived returns to search.

Making desirable schools more salient: The effects of changes in τ that make certain schools more salient depend on the share of these schools that is already known. For instance, highlighting high-payoff schools can reduce search if parents believe most such schools are already known.

Finally, we observe that interactions between search costs and misperceptions are an empirical question. For instance, misperceiving rejection chances $\hat{r}$ distorts search decisions but not rankings; hence when search costs are sufficiently small, the impacts of these misperceptions on true welfare (Equation (1)) vanish. In contrast, misperceptions of payoffs $\hat{u}$ distort rankings as well as search decisions, and may grow more costly as search costs shrink and more schools become known.

4. RESEARCH DESIGN

We conducted multiple survey rounds and two field experiments in partnership with an ed-tech NGO and the Ministry of Education of Chile over an approximately three-month period. We link these surveys and experiments to parents' search activity in the explorer, and to administrative data on school applications and placements.

Our study design follows a structured timeline with four periods, $t \in \{0, 1, 2, 3\}$, corresponding to the periods in our model. These correspond to approximately three months before the application deadline, approximately two months before the deadline, the penultimate week, and the final week and post-application period, respectively.

Within each period, events occur in the following order. First, experimental interventions (if any) take place. Second, parents' information updates, both passively and as a function of the intervention that was received. Third, parents make search or application decisions. In the case of search, their decisions further update their information. Fourth, we collect measurements through surveys, platform tracking, or administrative data.

Figure 2 provides an overview of our study, outlining the timeline, interventions, data collection, and model objects. Within a period (row), time progresses through the four phases (columns) from left to right: interventions (black), learning (blue), decisions (blue), and finally measurements (green). The remainder of this section unpacks this figure. Subsection 4.1 describes recruitment and the construction of the study sample. Subsection 4.2 presents our experimental interventions. Subsection 4.3 describes data sources and measurement.

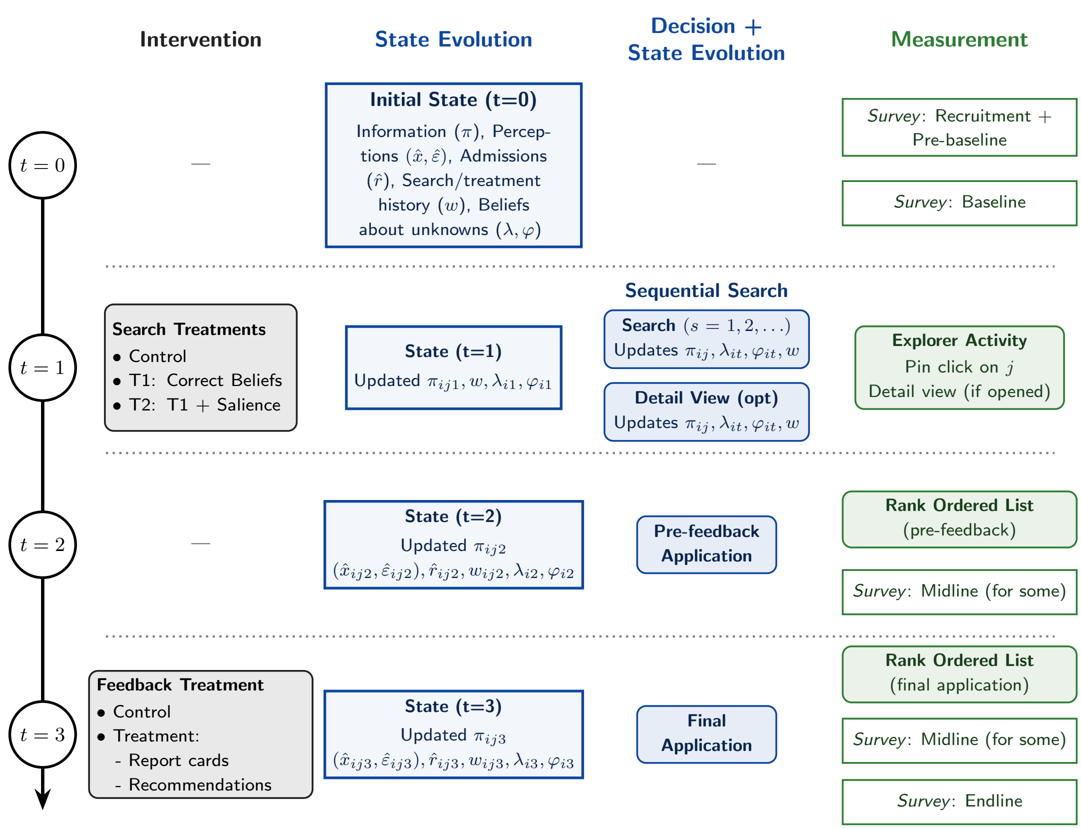


FIGURE 2.—Timeline for the Model, School Choice Process, Data Collection, and Interventions. The figure shows four periods ($t = 0, 1, 2, 3$) organized by columns: Intervention (black), State Evolution (blue), Decision + State Evolution (blue), and Measurement (green). Notation: π_{ijt} is information level about school j; perceptions are $(\hat{x}_{ijt}, \hat{\varepsilon}_{ijt}, \hat{r}_{ijt})$; beliefs about unknowns are summarized by $(\lambda_{it}, \varphi_{it})$; and w records treatments and search history. See Section 7 for formal definitions.

4.1. *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 children and collect location and demographic information. To be eligible for the study, the focal child must not have previously applied to a school through the SAE system. This means that we exclude children who have already received a school assignment through SAE and are reapplying to transfer to another school. We allow children to have older siblings who are already enrolled in school.

A total of 3,948 parents signed up for the study and completed the baseline survey. 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.[20] Table A.IV 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.

4.2. *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. Parents were randomly assigned to one of two treatment arms or a control group after completing the baseline survey. The treatment was embedded in a "training" that was required to access the explorer interface. After completing the treatment, parents gained access to the explorer.

The first treatment informed parents about number of schools within 2km of a respondent's home within each of 16 price and Education Quality Agency quality

[20]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 explorer platform.

category bins.[21] We classified good inexpensive schools—those that cost less than 50k CLP per month and were of either high (4) or medium (3) performance—as "*highlight-worthy*". The second treatment additionally highlighted these "highlight-worthy" schools on the map in green and displayed their price and quality (Figure S.3), making it easier for parents to find them. The control group received access to the school explorer platform but did not receive any information about the distribution of schools or their characteristics.

We interpret the first treatment as inducing variation in beliefs (parameters λ, φ) about the number of nearby schools and the distribution of their characteristics. The second treatment in addition is intended to vary the search technology (τ), raising the probability of finding highlight-worthy schools. Within our framework, if parents systematically underestimate the number of highlight-worthy schools, the first treatment should increase search effort and improve enrollment outcomes. The second treatment should likewise improve enrollment outcomes, though its effect on search intensity is theoretically ambiguous.

Feedback Intervention: We implement our second randomized information intervention 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 with a link providing tailored feedback on their applications (Figure S.6). 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.[22]

We interpret this treatment as providing information about schools, especially those that are already part of the family's application. In our framework, it is a shifter of w and should improve enrollment outcomes.

[21] Figure S.2-S.5 in the supplementary material show an example with screenshots of the platform.

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

4.3. 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, parents 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 practice, we construct two application snapshots: the final submitted application, and the application as of approximately ten days before the deadline. The latter was an input for the feedback treatment, for those who received it.
- **Past-year administrative data:** We use administrative data from the 2020 application process to construct a predicted probability of non-assignment at each school for each applicant, which we provide to applicants in the explorer and as part of the feedback treatment discussed in Section 4.2.
- **School Characteristics:** This data contains relevant information on schools, including their location, monthly school fees, and Education Quality Agency quality category.[23] In our analysis, we further report results using the value-added measure from Neilson (2021), expressed in standard deviations for the full nationwide student population.

Explorer: The explorer was made available to parents approximately two months before the application deadline, just after the "search" treatments described in the following section. We tracked all activities the parents performed in the explorer platform. From this data we construct parents' sequences of pin and profile clicks. Essentially all explorer use occurred in the two weeks after the explorer was provided, consistent with the timing in Figure A.3.

[23]Around half of our sample families 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.

Surveys: We conducted four surveys to collect information on family 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 parents had to apply for schools. It was sent to eligible parents and collected a detailed list of schools that the parents knew, 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.[24] To measure knowledge of schools, we used a three-point scale: a parent could report not knowing a school, knowing it by name, or knowing it well.[25] We also included questions about the rank-ordered list of schools to which parents were planning to apply, and about the perceived returns to search.[26]
- **Midline Survey:** This survey was conducted by phone in the final weeks of the application period, with timing varying across parents.[27] 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 perceptions of the appli-

[24]The survey elicited the perceived number of schools in 16 distinct price-quality categories within two kilometers of their home. Respondents were prompted with Education Quality Agency quality scores, and informed about the national distribution of these scores, to ensure a common understanding of the quality definition.

[25]The question did not provide additional guidance on what qualifies as "knowing a school well."

[26]Of particular interest, we elicited the probability that schools of given characteristics, if discovered, would be added to the top of the rank-order list.

[27]In our framework, if the midline survey took place before the final week, we say it occurred at time $t = 2$. Otherwise, it took place after the feedback intervention at time $t = 3$.

cation process, knowledge of schools in the neighborhood, and perceptions of schools' price, quality, and admissions chances.

Table S.I provides details about the timing and the content of each survey round. 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.V). However, even in the absence of differential attrition across treatment arms, selection into the midline survey could still affect some of our model estimates if parents who learned more through the app were also more likely to respond. To address this concern, Appendix Tables A.VI and A.VII show that baseline characteristics in the control group are very similar between attrited and non-attrited parents, suggesting no systematic selection into the surveys.

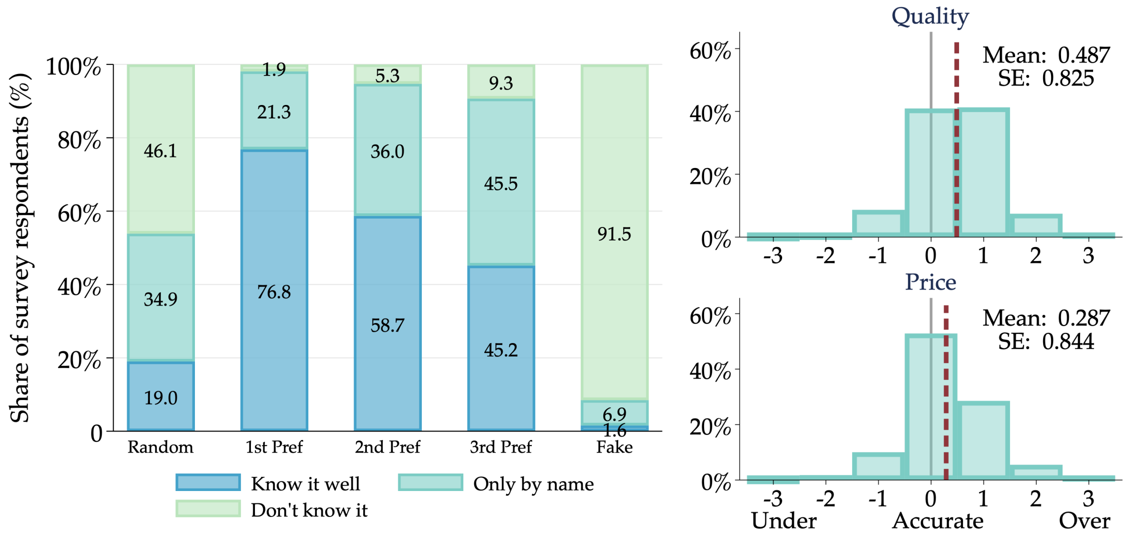
Table A.VIII provides additional details on the parents in our survey sample. Panel A describes their choice environment, Panel B presents parent and child characteristics, and Panel C reports initial knowledge and beliefs. Parents have, on average, 16.2 schools within 2km of their home. Of these, 8.6 schools (53%) meet the "highlight-worthy" criteria. As we discuss in more detail in the next section, this contrasts sharply with baseline beliefs: the average parent believes that there are only 7.3 schools in the neighborhood and that 3.6 schools are highlight-worthy. The average focal child is 3.9 years old and 44% of parents are eligible for the school voucher program.

5. DESCRIPTIVE ANALYSIS

We combine our administrative and survey data to document empirical patterns consistent with the key features of our framework.

5.1. *Knowledge, Perceptions, and Beliefs*

Parents have limited knowledge of nearby schools: At baseline, we asked respondents how well they knew eight randomly selected schools within 2km of their home, plus two fake schools to validate the question. At midline, we asked about schools on the application list, partially overlapping with this set. Figure 3a shows responses by school type. 19% of parents report knowing a randomly-selected school well and 35% by name only. As expected, respondents report knowing schools on



(a) Knowledge Levels By Baseline Ranking **(b)** Beliefs for First Preference School

FIGURE 3.—Knowledge of Schools. 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 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. Since application data come from the midline survey, we restrict the sample to the control group (N=1,318). Panel (b) shows the bias on perceived quality and price of the first-ranked school at baseline for the full sample (N=2,998 for quality and N=2,523 for price). All biases are measured as the perceived value minus the true value. Positive values indicate that the parent overestimates the quality or price of the school and negative values indicate that the parents underestimate the quality or price of the school. Quality is measured in four categories based on the classification of the Education Quality Agency. Prices are also categorized into four categories (free, 1-50k CLP, 50k-100k CLP, 100k+ CLP). Red dashed lines indicate the mean bias. The solid grey lines indicate the point of zero bias.

their applications, especially those at the top, better than random neighborhood schools. 77% know their reported first preference well versus 59% for their second. Reassuringly, over 91% correctly report not knowing the nonexistent schools. Figure A.4 shows that parents are more likely to know higher-quality and closer schools, indicating that unknown schools are negatively selected on these dimensions, but a substantial fraction of nearby high-quality schools nonetheless remain unknown, suggesting that search may be useful.

Parents misperceive "known" schools' price, quality, and admissions chance: Figure 3b plots perception errors for quality and price of the baseline first-preference school. Parents systematically overestimate the quality category of schools they intend to apply to; by contrast, perception errors are centered around zero for random schools (Figure A.5), suggesting that application decisions respond to perceived quality. Departures from true values are not pure level-shift biases: parents are also overoptimistic about the relative quality rank of their first-preference school (Figure A.6a).[28] In addition, parents overestimate prices.

Comparing baseline beliefs about admission probability to objective placement chances, Figure A.7 shows that beliefs are biased upward on average and exhibit compression, with parents underestimating both the share of schools with chances below 40% and the share where admission is nearly certain.[29]

Final rank-order lists respond to price and quality perceptions: Table A.X presents a rank-ordered logit of application rankings on actual and perceived school attributes, showing that the survey measures capture decision-relevant perceptions. Using endline data (Column 1), perceived price and quality predict rankings, while true values are irrelevant conditional on perceptions. Pooling all survey rounds (Column 2), true quality becomes predictive, but coefficients remain substantially larger for perceptions.[30]

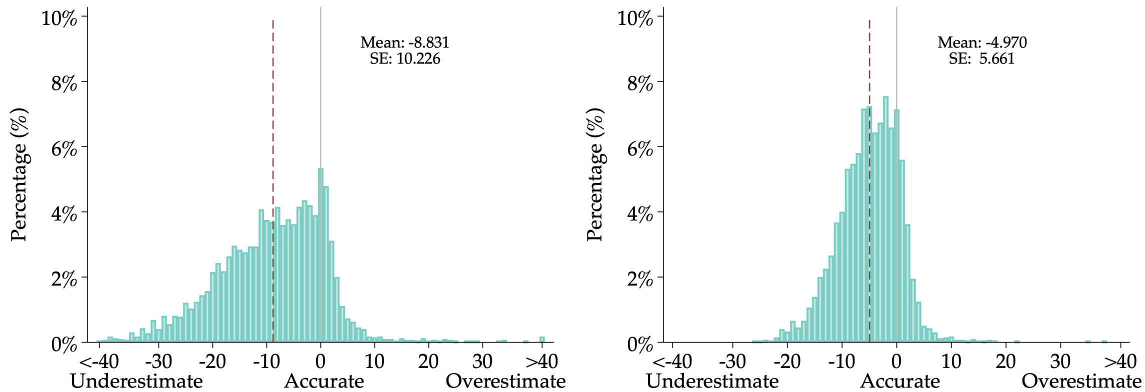
Better knowledge predicts more accurate perceptions of price and quality but not admission chance: Table A.IX regresses beliefs about school characteristics on a "know well" indicator with parent $\times$ school fixed effects, comparing the same parent–school pairs across waves as knowledge changes.[31] Within school, parents are 10.8 percentage points more likely to know the correct price and 4.7 percentage points more likely to report the correct quality when they report knowing the school "well" than when they knew it "by name". By contrast, we observe no differences in perceived admission chances.

[28]Figure A.6b shows that 23% of parents believe their first-preference school has higher quality than a random school not in their application, even when both schools have equal quality.

[29]This figure shows baseline beliefs. We also elicited perceived admission chances at midline for a larger set of schools, including repeated measurements, and at endline for the top three ranked schools.

[30]True quality may correlate with rankings due to measurement error or learning between survey rounds; perceptions at application time (midline and endline surveys) are more accurate than at baseline.

[31]Schools the respondent did not know at all are excluded, since no perception data was collected.



(a) Number of Schools **(b)** Number of Highlight-worthy Schools

FIGURE 4.—Beliefs about the Distribution of School Attributes. Notes: The first figure shows the bias in the beliefs of the total schools within 2km of the parent's home and the second figure shows the bias in the beliefs of the number of highlight-worthy schools. All biases are measured as the perceived value minus the true value at baseline. The sample consists of all parents (N=3,948).

Parents underestimate the number of available schools: We elicited beliefs about the number of schools within 2km and asked parents to allocate them across sixteen price $\times$ quality cells. Figure 4 shows that counts are systematically biased downward. On average, parents underestimate the total number of schools by almost nine and the number of highlight-worthy schools by five. Beliefs about relative shares are noisy.[32]

5.2. *Search Behavior*

We next use data from the school explorer platform to describe search patterns.

Biases predict search intensity: Parents typically underestimate the number of relevant schools and overvalue the quality of their first-choice school. Consistent with our hypotheses, those with greater biases conduct fewer searches. Table A.XI shows that, conditional on the true values, beliefs about counts of all schools, high-quality schools, and low-price schools positively correlate with the number of "pin clicks". Moreover, columns 5–7 show a significant negative relationship between the number of pin clicks and overestimates of the first-choice school's highlight-

[32] Appendix Figure A.8 shows that parents on average underestimate the share of free schools and overestimate the share that are of high quality. Biases about shares are not large enough to lead parents to overestimate the number of high-quality schools.

worthiness and admission chances. Parents who are content with their first choice may perceive lower returns.

"Clicks" predict greater knowledge and more accurate perceptions: Table A.XII shows that, controlling for baseline knowledge, clicking on a school increases the likelihood of knowing it well at midline by 4.8 percentage points. Effects are larger for schools where parents double-clicked to open the full profile. Moreover, misperceptions are smaller at "clicked" schools. Table A.XIII shows that, conditional on baseline perceptions and true attributes, clicking is associated with a 14.5 percentage point increase in reporting the correct quality.[33] However, parents still hold inaccurate quality perceptions at 40% of "clicked' schools. Effects on perceived prices and admission chances are limited as well, suggesting that parents' perceptions update only partially.

Platform use predicts beliefs over unknown schools: Regressing midline beliefs on clicks among control group respondents, Table A.XIV shows that, conditional on baseline beliefs and actual school counts, clicking on more schools is associated with higher perceived counts at midline (Column 1). Similarly, clicking on more highlight-worthy schools is associated with higher perceived counts of highlight-worthy schools (Column 2).

More general search patterns: Table A.XV provides evidence consistent with sequential search: the decision to stop depends on search history. Figure A.3 shows that most search occurs immediately after receiving platform access, in line with our model's timing.

[33]These patterns are descriptive: selection into clicking is endogenous. The coefficient for price is positive but insignificant. Results are similar using absolute perception errors instead of a correctness dummy. A potential concern is the set of observed perceptions is selected (see Appendix Table S.I for additional information on how we select schools). To address this concern, Panels B–D show that our results are similar across different types of schools, including the baseline first-ranked school, for which perceptions are always observed the in baseline and midline surveys.

5.3. *Heterogeneity by SES Status*

Following the previous literature, we analyze results separately by SES, proxied by maternal college completion.[34] As shown in Table A.IV, high-SES parents apply to 0.33 more schools on average. Low-SES families are 37 percentage points more likely to be eligible for the targeted voucher program (*Subvención Escolar Preferencial*, SEP), which provides additional resources to schools, raises admission chances, and eliminates out-of-pocket fees for eligible students at participating voucher schools. Accordingly, low-SES parents are 21 percentage points more likely to enroll their children in a school that is free for them. Children of high- and low-SES parents are equally likely to attend a high-quality school (26% vs. 27%).

In the supplementary material, we show how knowledge, perceptions, and beliefs vary by SES status. We find little difference in school knowledge levels. By contrast, high-SES parents have more accurate beliefs about the distribution of nearby schools: they underestimate the number of highlight-worthy schools by 4.7 on average, compared to 5 for low-SES parents. High-SES parents also have lower mean errors in perceived school quality (0.38 vs 0.52) and price (0.21 vs 0.31) for their first-ranked school. Low-SES parents have slightly more accurate beliefs about placement chances. Figure A.9 shows that timing varies, with high-SES parents starting to search for schools 0.33 months earlier on average.

5.4. *Summary*

Taken together, the descriptive evidence paints a coherent picture of how biased beliefs distort the school search process. Parents possess incomplete information about nearby schools, systematically misperceive key school attributes, and underestimate the number of high-quality options in their neighborhoods. As shown in Section 3.3, both of these biases reduce the perceived returns to search: the former by understating the probability of discovering a good school, the latter by overstating the value of the current best option. Consistent with this, we present suggestive evidence that households with larger biases along these dimensions conduct

[34] Existing work has studied how school availability, preferences, and beliefs explain differences by SES in schooling decisions (Burgess et al., 2015, Dizon-Ross, 2019, Attanasio et al., 2022), and how parental education affects child attainment (Black et al., 2005, Oreopoulos, 2006, Akresh et al., 2023).

fewer searches. We next leverage our experimental variation to examine whether correcting beliefs alters search decisions and, ultimately, school choice outcomes.

6. EXPERIMENTAL ANALYSIS

We begin by examining the effects of the search aid interventions and then turn to the impacts of the feedback intervention.

6.1. Impact of Search Aid Interventions

We use the randomization of the search aid interventions to study how search behavior and application outcomes change when parents are provided with information about the distribution of schools in their neighborhood. The sample is well-balanced (Table A.XVI) and attrition rates are similar across treatment groups (Table A.V). For parent i, we estimate:

$$Y_i = \alpha^S + \beta_1^S T1_i + \beta_2^S T2_i + \zeta_i^S + \epsilon_i^S. \tag{5}$$

Y_i is the outcome variable and ζ_i^S are stratification dummies. We show results for the pooled sample and separately for high- and low-SES parents.

Search interventions affect beliefs: Both treatment arms increase the perceived number of total and highlight-worthy schools in their neighborhood (Table I, Panel A, Columns 1–2). Relative to their control group counterparts, parents in treatment group 1 believe that there are 24% more highlight-worthy schools in their neighborhood. High-SES parents update their beliefs about the number of total and highlight-worthy schools in their neighborhood, but low-SES parents update only their beliefs about the number of highlight-worthy schools. Appendix Table A.XVII further shows that treatment 2 increases the likelihood that high-SES know the correct quality of a school by 11.8 percentage points.[35]

Belief updates are partial: The treatment interventions do not fully eliminate belief errors. The first treatment reduces the absolute difference between the perceived

[35] A potential concern is selection in the schools for which we elicited perceptions, as this set partly depends on application decisions which in turn depend on the treatment. We find positive and statistically significant effects when restricting the sample to schools listed by parents in the baseline survey, which was determined prior to treatments. The results are also robust to controlling for parent–school characteristics, including school quality, distance, and baseline knowledge.

TABLE I

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	0.837***	0.466***	0.633	0.226	0.277	0.013	0.005	-1.778
	(0.313)	(0.124)	(0.535)	(0.225)	(0.220)	(0.023)	(0.019)	(1.166)
Treatment 2	0.564*	0.295**	-0.280	0.268	-0.016	0.033	-0.006	0.212
	(0.322)	(0.127)	(0.496)	(0.220)	(0.211)	(0.023)	(0.019)	(1.423)
Control Group Mean	6.242	1.907	8.020	3.443	3.710	0.671	0.170	6.144
Observations	1670	1632	3108	3108	1075	2446	2368	2746
Panel B: Heterogeneity by Parental Education								
Treatment 1 × High SES	2.248***	0.628**	3.542***	1.316**	1.657***	0.020	0.077**	1.869
	(0.701)	(0.270)	(1.280)	(0.518)	(0.490)	(0.049)	(0.039)	(1.956)
Treatment 1 × Low SES	0.427	0.420***	-0.313	-0.129	-0.162	0.012	-0.019	-2.925**
	(0.350)	(0.141)	(0.575)	(0.246)	(0.245)	(0.026)	(0.022)	(1.395)
Treatment 2 × High SES	1.415**	0.540*	0.167	0.454	1.033**	0.146***	0.030	5.114*
	(0.710)	(0.299)	(1.204)	(0.508)	(0.429)	(0.050)	(0.041)	(2.787)
Treatment 2 × Low SES	0.332	0.230	-0.439	0.203	-0.343	0.000	-0.018	-1.232
	(0.363)	(0.140)	(0.536)	(0.242)	(0.241)	(0.025)	(0.022)	(1.648)
p-value: Treat 1 x High SES = Treat 1 x Low SES	0.021	0.495	0.006	0.012	0.001	0.882	0.034	0.045
p-value: Treat 2 x High SES = Treat 2 x Low SES	0.176	0.350	0.645	0.657	0.005	0.009	0.308	0.050
Control Group Mean (High SES)	6.110	1.783	9.897	4.058	3.386	0.547	0.193	3.528
Control Group Mean (Low SES)	6.278	1.941	7.441	3.253	3.815	0.711	0.163	6.949
Observations 1 (High SES)	362	357	732	732	246	587	564	643
Observations 2 (Low SES)	1308	1275	2376	2376	829	1859	1804	2103

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 and stratification dummies. In Panel B, we further include a dummy for SES status and the full set of interactions between both 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.

and true number of highlight-worthy schools by 6%. Similarly, the 11.8 percentage point in high-SES treatment-2 parents' knowledge of school quality scores translates into a 26% reduction in the share of parents holding incorrect perceptions. This partial updating may reflect both inattention and survey measurement error.

Treatments affect high-SES search behavior: While we find limited effects for the pooled sample, we observe substantial increases in the number of school pin clicks among high-SES parents in the first treatment group (Columns 3–4). Consistent with increased search, we also observe knowledge gains in the midline survey (Column 5). High-SES parents in treatment group 1 report that they know 49% more schools at least by name. By contrast, we find null to negative effects for low-SES parents.

Treatments affect enrollment: We next study the effects of the search aid interventions on school enrollment (Columns 6–8). We again find limited effects on the pooled sample (Panel A). However, among high-SES parents, we find that the first treatment arm leads to a significant increase in the average value added of the enrolled school. High-SES parents in the second treatment arm are also 27% more likely to enroll their child in a highlight-worthy school.[36]

Effects are larger when biases are larger: We also examine whether treatment effects on enrollment vary with baseline beliefs. For each belief variable, we interact the treatment indicators with both the true value and the belief bias measure introduced in Section 5.2. Our framework predicts two sets of heterogeneous responses. While parents who are less pessimistic about the number of available schools should experience weaker treatment effects, parents who are more optimistic about the attributes of their first-ranked school at baseline should respond more to the interventions.

Table A.XIX shows suggestive results that are consistent with these predictions. First, Treatment 1 has significantly smaller effects on enrolling in a highlight-worthy school for parents who are less pessimistic about the number of highlight-worthy and low-price schools in their neighborhood at baseline (Columns 2 and 4). Second, Treatment 1 has significantly larger effects for parents who were overoptimistic about the attributes and admission chances of their first-ranked school at baseline (Column 5). For Treatment 2, we find larger effects for parents who were more overoptimistic about their admission chances.

6.2. *Impact of Feedback Treatment*

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

$$Y_i = \alpha^F + \beta^F F_i + \zeta_i^F + X_i' \psi^F + \varepsilon_i^F.$$

[36]Table A.XVIII 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. 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.

Y_i is the outcome variable, ζ_i^F 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.

TABLE II
TREATMENT EFFECTS OF FEEDBACK INTERVENTION

	First Stage	Perceptions		Application			Enrollment		
	Open Feedback	Correct Price	Correct Quality	Added School	Changed Rank 1 School	Deleted School	Highlight-worthy	Value Added	Distance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Pooled									
Feedback Treatment	0.578***								
	(0.018)								
Open Feedback		0.172*	0.227***	0.061***	-0.002	0.020*	-0.031	-0.039	2.081
		(0.092)	(0.081)	(0.018)	(0.005)	(0.012)	(0.057)	(0.050)	(1.758)
F-Stat	1000.016								
Control Group Mean	0.000	0.362	0.553	0.026	0.003	0.009	0.705	0.172	4.814
Observations	2116	765	753	2116	2116	2116	1850	1818	2066
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**	0.017	-0.000	-0.001	-0.059	-0.137**	4.529
		(0.150)	(0.119)	(0.030)	(0.002)	(0.016)	(0.087)	(0.067)	(3.711)
Open Feedback × Low SES		0.300**	0.199*	0.076***	-0.002	0.027*	-0.014	-0.012	1.373
		(0.116)	(0.105)	(0.022)	(0.007)	(0.015)	(0.062)	(0.055)	(2.013)
p-value: Open Feedback x High SES = Open Feedback x Low SES	0.289	0.034	0.731	0.124	0.747	0.187	0.628	0.080	0.462
F-Stat (High SES)	322.667								
F-Stat (Low SES)	616.052								
Control Group Mean (High SES)	0.000	0.441	0.667	0.029	0.000	0.008	0.618	0.249	4.412
Control Group Mean (Low SES)	0.000	0.343	0.528	0.025	0.003	0.009	0.729	0.153	4.926
Observations 1 (High SES)	460	165	160	460	460	460	406	396	440
Observations 2 (Low SES)	1654	600	593	1654	1654	1654	1442	1420	1624

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 2–9 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 a dummy for SES status and the full set of interactions between 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.

Column 1 of Table II presents the first-stage results. 58% of parents assigned to the treatment group opened the feedback intervention (Panel A). These rates are similar for high and low-SES parents (Panel B). We also examine the effect on school-level knowledge using information from the endline survey. The feedback intervention increases the likelihood that a parent correctly perceives the price and quality category of a school in their application by 17 and 23 percentage points, respectively. Columns 4 through 7 show that the feedback intervention also affects application behavior. Parents who receive feedback are 6 percentage points more likely to add a school and 2 percentage points more likely to delete a school from their application. The latter indicates that parents previously held incorrect beliefs about the attributes of some schools in their choice set. We find suggestive evidence that the treatment effects on perception and application behavior are concentrated among low-SES parents. However, we cannot rule out that the effects are similar across SES status for any of the application outcomes.

Finally, we find no effects on final enrollment outcomes. This likely reflects the fact that most parents do not change their first-ranked school, instead adding or removing schools in the middle or at the bottom of their rankings. One exception is a negative treatment effect on value added for parents with high socioeconomic status, which may in part reflect that the feedback intervention emphasizes school attributes other than value added.

While the absence of average enrollment effects suggests limited welfare gains for the typical parent in our sample, there may be substantial heterogeneity by baseline beliefs. Consistent with this interpretation, Table A.XXI provides suggestive evidence that, among parents assigned to the control group in the search aid experiment, enrollment outcomes improve more for parents who are more overoptimistic about their first-ranked school at baseline.

6.3. Discussion of Experimental Evidence

We find that both search aid interventions and the feedback intervention affected search and application outcomes, although the mechanisms through which they operate differ. The first arm of the search intervention provided aggregate information about the set of schools available to parents. Because the explorer in this arm did not make salient which schools had high quality and low prices, parents had to explicitly click on schools to learn about their attributes. By contrast, parents in the second search intervention could immediately identify which schools were highlight-worthy because these schools were displayed in green in the explorer. As a result, beliefs about known and unknown schools could improve even without a corresponding increase in pin clicks.

The feedback intervention also provided information about known schools (through report cards) and about unknown schools, (through recommendations for additional schools). However, this information may have been delivered too late in the school choice process to generate large changes in enrollment outcomes. Nonetheless, the fact that we observe changes in application outcomes across all treatment arms for some subgroups suggests that, consistent with our framework, incorrect beliefs about both known and unknown schools play an important role in shaping school search and choice behavior.

We also document substantial heterogeneity by SES status. We argue that there are two main explanations for why the effects of the search aid interventions are larger for high-SES parents. First, high-SES parents may be better able to understand the information provided due to higher digital literacy. Consistent with this interpretation, we find larger effects for these parents on the perceived number of available schools and the likelihood that they know the correct quality of a school.[37] Second, the timing of the search aid interventions may have been too early for low-SES parents. The importance of timing has been documented in several other contexts (Richburg-Hayes et al., 2017, De Groote and Rho, 2024), and, as discussed in Section 5.3, low-SES parents report starting their search later than high-SES parents. Because most parents used the explorer primarily when they first received access to it, a larger share of low-SES parents may not have benefited from the information if they had not yet begun the search process. Even if they used the explorer, low-SES parents may have been less focused or may have spent less time engaging deeply with it. These same mechanisms may also explain why the feedback intervention tends to be more beneficial for low-SES parents. The timing of the feedback intervention aligns more closely with their search process, and the information it provides may be easier to act upon.

[37] An explanation for the limited changes in perceptions of prices might also have been confusion about the school voucher system. Voucher eligibility is determined on a timeline that means many parents do not know whether they qualify at the time of their search. 67% of parents were unsure about their voucher eligibility status at baseline.

7. EMPIRICAL MODEL

We now estimate the model presented in Section 3. Our goal is to quantify the welfare and school-quality impacts of misperceptions, search costs, and their interaction. To understand whether information frictions matter, and how they interact with search costs, we need to extrapolate from our experiments: are there large potential gains to be realized from heavier-touch interventions that move beliefs by a greater amount? Do the gains from reducing information frictions shrink if search costs are reduced, or are information and cost reductions complementary? We address these questions with counterfactual simulations.

7.1. *Parametric Assumptions for Estimation*

For tractability, we make the following assumptions, specializing the model described in Section 3. We estimate the model separately by SES group. All parameters may differ arbitrarily by group.

1. Payoffs (u)*:* The choice set J_i consists of all available schools within 5km of i's home. The value of a placement in school $j \in J_i$ is given by:

$$u_{ij} = d_{ij}\beta_i^d + x_{ij}'\beta_i^x + \delta_j + \varepsilon_{ij}. \tag{6}$$

This value depends on distance d_{ij}, price and quality category x_{ij}, school effects δ_j, and a match-specific component ε_{ij}. School effects and quality categories are common across parents within an SES group; distance, price, and admissions chances are parent-school specific.[38] We allow for heterogeneity by placing four jointly normally distributed random coefficients on distance, price, quality, and a constant that enters all "inside" options.[39] As a scale normalization, we fix $E[\beta_i^d]$ (the mean distance coefficient) to -1.

2. Information (π)*:* We model awareness and information—$\pi_{ijt} \in \{0,1,2,3\}$, denoting zero, low, high, or full information about j at time t respectively—via a latent

[38] Out-of-pocket costs depend on family SEP eligibility and school participation in SEP.

[39] The vector x_{ij} includes a constant. We take the mean of the coefficient on this constant to be zero, as this mean is subsumed by δ_j.

index $\pi^*_{ijt} \in \mathbb{R}$ and a set of thresholds, extending models of consideration sets (Goeree, 2008) to tractably accommodate imperfect knowledge of "known" schools:

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

$$\pi^*_{ijt} = (d_{ij}, w_{ijt})'\alpha_i + \eta_j + \nu_{ijt} \tag{8}$$

Equation (7) fixes the thresholds for "low" and "high" information at 0 and 1, respectively, normalizing the index's scale and location. If $\pi^*_{ijt} > 1$, a school is known well; otherwise it is known by name $(\pi^*_{ijt} \in (0, 1])$ or not at all. The state $\pi_{ijt} = 3$ does not occur in the data. We use it for counterfactual benchmarks in which parents have full information about all payoff-relevant objects.

The vector w_{ijt} contains observed, time-varying shifters of information that are excluded from preferences: treatment indicators, on-platform pin clicks and profile views, and time $(t \in \{0, 1, 2, 3\})$ indicators. We write the corresponding coefficient vector as $\alpha_i = \overline{\alpha} + \tilde{\alpha}_i$, where $\tilde{\alpha}_i$ is a mean-zero random coefficient vector with covariance matrix Σ^α. We place multivariate-normal random coefficients on the time indicators to allow for off-platform learning with flexible timing. Other elements of α_i are zero.

The term η_j captures school-level "discoverability". It varies with schools' price and quality, and is jointly normal with the utility term δ_j, allowing for selection in the set of known schools that is consistent with our descriptive evidence. In particular, we assume

$$(\delta_j, \eta_j)' \sim N\left(\left(x_j\overline{\beta}, x_j\overline{\alpha}\right)', \Sigma^{\delta\eta}\right). \tag{9}$$

Finally, ν_{ijt} is an idiosyncratic information shock, independent across schools and parents, but correlated within parent-school over time. We assume $\nu_{ij} \sim N(0, \Sigma^\nu)$.

When $\pi_{ijt} = 3$, parents observe school j's true characteristics $(x_{ij}, \varepsilon_{ij}, r_{ij})$. We model signals and perceptions at $\pi_{ijt} \in \{1, 2\}$ as follows.

3. Observable attributes (x): The conditional distribution $g^x(\hat{x}_{ij}^{(1)}, \hat{x}_{ij}^{(2)}|x_{ij})$ is given by:

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

where $\Gamma(\cdot|x)$ are multinomial distributions whose values depend arbitrarily on the true value x. That is, "low-information" ($\pi = 1$) signals are drawn as flexible functions of the true characteristics. When a parent gains "high" information about a school, it learns the truth with probability p^h. Otherwise, it does not update about x.

Parents believe that their signals are correct: $\hat{g}^x(\hat{x}_{ij}^{\pi_{ijt}}|x_{ij}) = 1(\hat{x}_{ij}^{\pi_{ijt}} = x_{ij})$. This is a simplifying assumption that is not essential. It is motivated by our survey, which elicits perceptions in discrete price and quality categories, and the evidence in Table A.X.[40]

4. Match value (ε): We model both the true and subjective distributions of ε as Gaussian, but allow pessimism (or optimism) and over- or under-dispersion. These biases may distort search decisions and the rankings of schools that are known "by name". For instance, parents may be excessively pessimistic about the unobservables of schools they do not know, or do not know well, leading them to underestimate search returns and penalize schools known "by name" in their rank-order lists. We assume that ε_{ij} is observed when $\pi_{ijt} = 2$. Thus our full-information benchmark, $\pi_{ijt} = 3$, provides no further information about ε_{ij}.[41]

[40] An alternative approach would model parents' uncertainty over an underlying running variable, about which the quality and price bins provide partial information (Vatter, 2025).

[41] We interpret this assumption as part of the definition of the counterfactual full-information environment rather than a restriction on the model. It says that our full-information benchmark provides the information about match-specific unobservables that agents would have if they knew schools "well".

In the "low-information" event $\pi_{ijt} = 1$, parent i observes $\tilde{\varepsilon}_{ij} = \varepsilon_{ij} + e_{ij}$ with classical measurement error e_{ij}. Objectively,

$$\begin{pmatrix} \varepsilon_{ij} + e_{ij} \\ \varepsilon_{ij} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2 + \sigma_e^2 & \sigma_\varepsilon^2 \\ \sigma_\varepsilon^2 & \sigma_\varepsilon^2 \end{pmatrix} \right). \tag{10}$$

However, parents believe that

$$\begin{pmatrix} \varepsilon_{ij} + e_{ij} \\ \varepsilon_{ij} \end{pmatrix} \sim N \left(\begin{pmatrix} \tilde{\mu} \\ \tilde{\mu} \end{pmatrix}, \begin{pmatrix} \tilde{\sigma}_\varepsilon^2 + \tilde{\sigma}_e^2 & \tilde{\sigma}_\varepsilon^2 \\ \tilde{\sigma}_\varepsilon^2 & \tilde{\sigma}_\varepsilon^2 \end{pmatrix} \right). \tag{11}$$

The parameter $\tilde{\mu}$ captures bias. If $\tilde{\mu} < 0$, parents are systematically pessimistic about their payoffs at unknown ($\pi_{ijt} = 0$) schools. Parents may also hold inaccurate beliefs about the prior variance or the informativeness of the signal. Parents apply Bayes' rule and Equation (11) to form a subjective expectation, $\hat{\varepsilon}_{ij}^{(1)} = \hat{E}(\varepsilon_{ij} | \tilde{\varepsilon}_{ij})$, see Appendix A.6.

Our model focuses on misperceptions of match shocks ε rather than mean utilities δ. As payoffs (Eq. (6)) depend on their sum, however, one could equivalently interpret the "pessimism" parameter $\tilde{\mu}$ as pessimism about mean utilities δ.

5. Rejection chances (r): The true rejection probability r_{ij} is parent–school specific (e.g., it varies with family eligibility such as SEP and with school-level priorities/capacities). Signals for $\pi_{ijt} \in \{1,2\}$ are distributed as:

$$\hat{r}_{ij}^{\pi_{ijt}} = \min\{1, \max\{0, o_{i0} + o_{i1}(r_{ij} - o_{i0})\}\}, \;\; \pi_{ijt} \in \{1,2\}, \tag{12}$$

where $(o_{i0}, o_{i1}) \sim N(\mu_o, \Sigma_o)$ are parent-level random terms. This specification allows for optimism or pessimism (via o_{i0}) and compression (via o_{i1}), consistent with our survey evidence. As with x, parents take their signals to be the true value. Following the descriptive evidence, signals are invariant to the level of information $\pi_{ijt} \in \{1,2\}$.

6. Beliefs, learning, and updating: Parent i believes the number of not-yet discovered schools $\hat{N}_{it}$ is Poisson with mean λ_{it}, and the probabilities of each of the 16 price-quality cells are distributed Dirichlet with parameter φ_{it}. These beliefs update in response to information treatments and knowledge $\pi_{i \cdot t}$. Parents belong to discrete types that differ in how initial belief parameters $(\lambda_{i0}, \varphi_{i0})$ are formed as a function of the true choice set (J_i) and initial knowledge ($\pi_{i \cdot 0}$), and in how they update as information is revealed. Our specification nests Bayesian updating over the probabilities of price-quality cells, but we also allow parents to under-react to information. We present this part of the model in Appendix A.6.

Beliefs over shocks ε are given by Equation (11). Beliefs over r are given by $f^r(r) = \hat{f}^r(\min\{1, \max\{0, o_{i0} + o_{i1}(r_{ij} - o_{i0})\}\})$. That is, parent correctly anticipate the distribution of perceptions $\hat{r}_{ij}$. We adopt this assumption given the absence of evidence on the subjective distribution of rejection chances at unknown schools. Among known schools, perceptions are no more accurate at those known well (Table A.IX).

7. Subjective expected payoffs $(\hat{u})$*:* Subjective expected payoffs at information level $\pi_{ijt} \in \{1, 2\}$ are given by

$$\hat{u}_{ij}^{(\pi_{ijt})} = \hat{E}(u_{ij} | \hat{x}_{ij}^{\pi_{ijt}}, \hat{\varepsilon}_{ij}^{\pi_{ijt}}) = d_{ij}\beta_i^d + \hat{x}_{ij}^{\pi_{ijt}}\beta_i^x + \delta_j + \hat{\varepsilon}_{ij}^{\pi_{ijt}}.$$

As the true payoff (Expression (6)) is linear in x_{ij} and ε_{ij}, the subjective expected payoff is formed by substituting in the subjective means of these variables. The weights β_i reflect preference heterogeneity and are invariant.[42]

8. Search: At each $s = 0, 1, \ldots$, parents choose whether to conduct their sth pin click or move on. Conditional on clicking school j's pin, they learn a detail-view cost and may optionally conduct a detail view at j. The cost of the sth pin click is:

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

where $\bar{c}_i^{\text{pin}} \sim N(0, \sigma_{\bar{c}}^2)$. If instead the parent stops, it pays $c_{i0s}^{\text{pin}} = -\sigma^{\text{pin}} \varepsilon_{i0s}^{\text{pin}}$. The shocks $\varepsilon_{i0s}^{\text{pin}}$ and $\varepsilon_{i1s}^{\text{pin}}$ are independently standard Gumbel (T1EV) distributed. If the parent 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.

[42]We do not model uncertainty over the mapping from characteristics to payoffs.

If a parent chooses to conduct a pin click, the next clicked school is $j \in J_i$ with probability

$$Pr_{ij} = Pr(\text{view } j|\text{continue}) = \exp(x_{ij}^{\text{click}}\gamma^{\text{click}}) \,/ \sum_{j\in J_i} \exp(x_{ij}^{\text{click}}\gamma^{\text{click}}). \tag{13}$$

The variables x_{ij}^{click} consist of schools' price, quality, distance, a highlight-worthiness indicator, and an interaction between highlight-worthiness and treatment 2. Thus certain schools (nearby schools; high-quality schools) may be found more easily. We inlcude all characteristics that enter discoverability (η). Perceptions $\hat{N}_{it}$ distort the perceived returns as described in Section 3.

The "detail view" cost shocks are revealed only after the parent has conducted a pin click. When computing the "one-step-ahead" value of the next pin click, parents form an expected value anticipating the optimal "detail view" decision. We provide details in the Appendix.

Indicators for a "pin click" and "profile view" at j are elements of w_{ijt} for $t > 0$. Parents revisit previously-searched schools according to the probabilities in Equation (13), but additional clicks do not provide further information. Hence as more schools are clicked at least once, the return to further search falls.

8. ESTIMATION

All parameters are estimated separately for low-SES and high-SES parents. We allow measurement error on every survey variable, and on the "just-before-feedback" administrative rank-order list submitted at $t = 2$. We begin with an informal discussion of identification.

8.1. Identification

Identification comes from (1) repeated measurements of rank-order lists, awareness, subjective beliefs and perceptions of schools' characteristics, (2) observable variation in information, excluded from preferences, that is induced by our treatment assignments and parents' search activity, (3) a type of "parallel trends" assumption that we place on the evolution of information, and (4) in the case of beliefs $\hat{f}^{\varepsilon}$, survey questions that elicit relevant forecasts. We now discuss these in turn.

Panel data: We obtain repeated measurements of rank-order lists $\{L_{it} : t \in \{0, 2, 3\}\}$, surveyed knowledge ($\pi_{ijt}^{\text{survey}}$), perceptions ($\hat{x}_{ijt}^{\text{survey}}, \hat{r}_{ijt}^{\text{survey}}$), and beliefs over the number and characteristics of unknown schools.[43] Repeated measurements within people and person-school pairs allow us to accommodate measurement error in perceived characteristics, reported knowledge levels, and non-final rank-order lists.[44] Correlations between information and perceptions identify updating. For instance, if perceptions of x_{ij} become more accurate when parents know j "well," we conclude that the updating parameter p^h is large.

Variation in information: Experiments and search activity induce observable variation in w over time that is excluded from payoffs but shifts information. Changes in reported knowledge, perceptions, and rankings within person and school identify the effects of treatments and clicks on π^* and perceptions $\hat{x}$. For instance, if almost all schools receiving "feedback" become known well, we conclude that the effect of feedback is large.

"Parallel trends": A possible concern is that some elements of w, such as search decisions, are endogenous, depending on baseline preferences and information. For instance, suppose person i conducts a pin click at time $t = 1$ at school j but stops before clicking on school k. The key assumption, implicit in equation (8), is that information levels π_{ijt}^* and π_{ikt}^* would have evolved in parallel absent i's decision to click on j, up to shocks η that average out. Thus greater growth in information at j than at k can be ascribed to search.

Alternatively, we could have exclusively used variation in w that is independent of all unobservables. For instance, w could have consisted only of randomized treatment assignments or other valid instruments for treatment or search. Such an approach would not need panel data: one rank-order list and one survey wave would suffice. While one could simulate "full information" counterfactuals, how-

[43]We use administrative rank-order lists at times $t = 2$ and $t = 3$ and a list elicited on our baseline survey. All three surveys provide knowledge and perceptions measures. We use beliefs over N_{it}^{unknown} and the distribution of x from the baseline and midline surveys.

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

ever, it is not obvious how to model endogenous information acquisition in our setting, and simulate parents' search decisions under counterfactual information and costs, without using the search data.

Takeup and importance: We observe the effects of treatments on applications and on the relevant intermediate outcomes (beliefs, perceptions, awareness). This data lets us extrapolate to larger belief shifts. For instance, the search treatments have small effects on low-SES parents' applications, but they also have small effects on beliefs—imperfect takeup—despite inaccurate baseline beliefs. Without intermediate outcomes, we might wrongly conclude that correcting these beliefs is not useful.

Payoffs and unobserved characteristics: To identify misspecified beliefs $\hat{f}^{\varepsilon}(\cdot)$, we use repeated rank-order list measures and two baseline survey questions. The first question asks with what probability parents would add a hypothetical school with specified characteristics to the top two places in their rank-order list, conditional on discovering it. The second varies the hypothetical school's characteristics. If parents overestimate the variance of unobservables, they will overestimate this probability.

We first estimate a reduced form (equation (17), see Appendix A.6) that is sufficient for application decisions conditional on π^*, and is identified by ranking data. In the presence of measurement error, a Bayesian will shrink noisy signals of ε toward their mean, but parents might have the wrong mean or fail to shrink correctly. In particular, the extent to which parents penalize schools known "by name" relative to those known well, these schools' dispersion in parents' rank-order lists conditional on their observables, and the amount of movement within rank-order lists when they become "known well" reveal the objective distribution of match shocks ε, the subjective mean μ^l, and the subjective signal-to-noise ratio $\frac{\tilde{\sigma}_\varepsilon^2}{\tilde{\sigma}_\varepsilon^2+\tilde{\sigma}_e^2}$. The survey questions then pin down the subjective utility and measurement-error variances $\tilde{\sigma}_\varepsilon^2$ and $\tilde{\sigma}_e^2$, which are needed for forecasts. See Appendix A.6 for details.

Effect of true and perceived x's: Within-school-j variation in $\hat{x}_{ijt}$ across parents, and variation within parents over time, let us distinguish mean-utility parameters $\overline{\beta}$ on x from the impact of perceived attributes on choice behavior (through the preference coefficients β), consistently with the descriptive patterns described in Table A.X.

Estimation: We estimate the model in three steps. First, we estimate the distribution of $(\hat{u}^{\pi}, \pi^{*}, \hat{x})$, the parameters relevant for these objects, and a reduced form for the distribution of shocks (the parameters μ^{l} and Σ^{ε} of Equation (17)), via a Gibbs sampler, taking posterior means as our point estimates.

We use the baseline survey rank-order list at time $t = 0$, rank-order lists at $t \in \{2, 3\}$, explorer clicks, survey responses on knowledge and perceptions of x at baseline, midline, and endline. Reported awareness $\hat{\pi}^{\text{survey}}$ and non-final rank-order lists are constructed with additive Gaussian measurement error in the underlying indices (π^{*} and $\hat{u}^{\pi}$, respectively). Perceived $\hat{x}$ are misreported (drawn uniformly on $\{1, 2, 3, 4\}$) with a probability that we estimate.

Second, we estimate all remaining parameters, except those entering search costs, via maximum likelihood.

1. 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.
2. We estimate beliefs over the number of unknown schools, and the perceived distribution of x, using baseline and midline survey questions on the perceived number of schools within distinct price and quality cells, again allowing for measurement error. To compute the estimator, we use an EM algorithm. The structural beliefs are over unknown schools within 5km, but the survey elicits counts of all schools (known and unknown) within a smaller 2km radius. To address this measurement challenge, we build a model of survey responses, assuming that survey respondents probabilistically recall a subset of known schools and stochastically sample unknown schools according to their beliefs (see Appendix A.6).
3. With draws of $\hat{u}$ and the parameters of equation (17) in hand, we maximize the likelihood of survey responses to estimate the remaining parameters of equations (10) and (11).

These objects suffice for computing the value of search (Equation (3)). In the final step, we impose optimality of the search decision and estimate search costs via MLE. We provide details on all steps in the Online Appendix.

9. RESULTS AND COUNTERFACTUAL SIMULATIONS

We first discuss parameter estimates. The main results are given in Section 9.2.

9.1. Estimates

Table III presents selected parameter estimates from the model in Section 7, estimated as described in Section 8. The table is organized into five panels: preferences and information (Panel A), variance-covariance of random coefficients (Panel B), match value shock primitives (Panel C), school-level random effects (Panel D), and search costs (Panel E). All parameters are estimated separately by SES group. A full set of estimates is available in Tables A.XXII and A.XXIII.

Panel A reports preference and information parameters. The first rows show the coefficients on perceived price (β^p) and perceived quality (β^q) from the payoff equation (6), where all coefficients are relative to the mean distance coefficient, which is normalized to -1. Preferences have the expected signs: parents value school quality and dislike prices. Low-SES parents are relatively more responsive to prices and relatively less responsive to quality.

The knowledge shifters govern the information index in equation (8). Distance reduces knowledge, and highlight-worthy schools (low price, high quality) are more likely to be known, consistent with the descriptive evidence in Section 5. Schools that parents have explored on the platform are substantially more likely to be known well. Treatment 1 increases knowledge for both SES groups, while Treatment 2 increases knowledge for low-SES parents and is close to zero for high-SES parents. When a high-SES parent knows a school "well" ($\pi_{ijt} = 2$), they learn the true x's with probability $p^h = 32.4\%$. For low-SES parents, this probability is 4.7%.

The admission chance parameters show that 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 parents' beliefs more extreme than the truth.

The last rows of Panel A report estimates for the latent types that govern subjective beliefs about unknown schools. As described in Section 7, beliefs about the number of unknown schools follow a mixture model with three latent types $h \in \{1, 2, 3\}$, where each type has a Poisson parameter λ_{ih} and an estimated mixing probability $\hat{\pi}_h$. We report $\bar{\lambda}_h$, the mean of λ_{ih} across individuals within each type; estimated type probabilities are in Table A.XXIII. This parameter directly affects the perceived probability of discovering a new school when searching: a parent with a low λ believes few undiscovered schools exist, which reduces the perceived return to search even if search costs are low. The estimates reveal substantial heterogeneity across types. For low-SES parents, $\bar{\lambda}$ ranges from 4.5 (type 1, $\hat{\pi}_1 = 0.56$) to 8.1 (type 3, $\hat{\pi}_3 = 0.15$), with the majority of parents assigned to the most pessimistic type. For high-SES parents, the range is wider, from 4.6 ($\hat{\pi}_1 = 0.56$) to 10.8 ($\hat{\pi}_3 = 0.18$), with a similar concentration in the lowest-λ type. This heterogeneity in λ is a key driver of the interaction between information and search costs documented in the counterfactual analysis: parents with low λ underestimate the returns to search, so that reducing search costs alone has little effect on their school choices.

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 ($\sigma_{d,p} > 0$), implying that price-sensitive parents are also less willing to travel. The covariances of both price and distance with quality are positive ($\sigma_{d,q}, \sigma_{p,q} > 0$), implying that parents who value school quality tend to be less price- and distance-sensitive.

Panel C reports the match value shock primitives. The negative μ^l means that parents expect lower payoffs at schools known only by name than at schools known well, consistent with Bayesian shrinkage toward a pessimistic prior (Equation (17)).

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 SES groups, and more expensive and higher quality schools have higher expected mean utilities. Schools with higher prices are more easily discovered by both SES groups, while the effect of quality on discoverability is small and imprecisely estimated. 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 parents may find via search have lower mean utilities on average than those already known.[45]

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 low-SES parents (Figure A.12a), but variances are larger as well.[46]

We present additional estimation results in the Online Appendix alongside the full set of model coefficients. First, the distortion functions $\Gamma(\cdot|x)$ describe how parents at the lowest knowledge level ($\pi_{ijt} = 1$) perceive school price and quality given the true values (Section 7). Figure A.11 shows that perceived distributions are compressed relative to the truth, and that both SES groups systematically overestimate quality. Low-SES parents also overestimate prices, possibly because some are unaware of their targeted voucher eligibility, which makes most schools free for them. Second, Figure A.13 shows that the model fits the data well, comparing the distributions of observed and predicted school characteristics, placement probabilities, and search behavior.

9.2. *Counterfactuals*

Table IV presents our main counterfactual results, pooling across low- and high-SES parents. In each counterfactual, we change some aspect of the information environment or search costs and simulate search behavior, applications, allocations (holding admissions cutoffs fixed), and final enrollment. We report the expected

[45]We plot the distribution of mean utility and discoverability in Figure A.10, showing a positive relationship that is steeper for high-SES parents.

[46]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 parents' decisions are more responsive to perceived values while low-SES parents' decisions are more random.

TABLE III

SELECTED MODEL ESTIMATES

Panel A: Preferences and Information	Param.	Low SES Coeff.	Low SES Std Err.	High SES Coeff.	High SES Std Err.
Subjective Expected Utility Parameters					
Perceived Price	β^p	-0.474	(0.035)	-0.234	(0.060)
Perceived Quality	β^q	0.622	(0.052)	0.757	(0.093)
Knowledge Shifters (Selected)					
Distance	α_z	-0.148	(0.008)	-0.115	(0.006)
Treatment 1	α_w	0.125	(0.030)	0.157	(0.070)
Treatment 2	α_w	0.304	(0.060)	-0.060	(0.106)
Highlight-worthy	α_w	0.331	(0.075)	0.281	(0.068)
Single Click	α_w	0.896	(0.030)	0.456	(0.041)
Double Click	α_w	1.384	(0.043)	0.760	(0.063)
Pr(learn true x's | know well)	p^h	0.047	(0.008)	0.324	(0.025)
Subjective Admission Chances Parameters					
Optimism (mean)	μ_{o0}	0.685	(0.009)	0.683	(0.021)
Optimism (sd)	σ_{o0}	0.197	(0.010)	0.216	(0.029)
Compression (mean)	μ_{o1}	0.160	(0.019)	0.262	(0.042)
Compression (sd)	σ_{o1}	0.232	(0.029)	0.298	(0.037)
Subjective Beliefs of Unknown Schools					
Type $h=1$	$\bar{\lambda}_1$	4.505		4.613	
Type $h=2$	$\bar{\lambda}_2$	6.950		7.224	
Type $h=3$	$\bar{\lambda}_3$	8.096		10.796	

Panel B: Variance Covariance of Random Coefficients (Utility)	Param.	Low SES Coeff.	Low SES Std Err.	High SES Coeff.	High SES Std Err.
Variances					
Constant	σ_c	16.814	(2.279)	14.772	(2.516)
Distance	σ_d	0.535	(0.027)	0.667	(0.066)
Price	σ_p	1.835	(0.173)	1.217	(0.280)
Quality	σ_q	0.709	(0.109)	0.613	(0.116)
Covariances					
Constant-Distance	$\sigma_{c,d}$	-1.372	(0.098)	-1.857	(0.269)
Constant-Price	$\sigma_{c,p}$	-3.567	(0.490)	-2.526	(0.661)
Constant-Quality	$\sigma_{c,q}$	-2.692	(0.454)	-2.213	(0.457)
Distance-Price	$\sigma_{d,p}$	0.184	(0.031)	0.090	(0.070)
Distance-Quality	$\sigma_{d,q}$	0.083	(0.026)	0.170	(0.066)
Price-Quality	$\sigma_{p,q}$	0.140	(0.072)	0.074	(0.086)

Panel C: Match Value Shocks ε_{ij} Primitives	Param.	Low SES Coeff.	Low SES Std Err.	High SES Coeff.	High SES Std Err.
Mean Subjective Expectation of ε_{ij} given $\tilde{\varepsilon}_{ij}$	μ_l	-0.410	(0.031)	-0.581	(0.065)
Variance Covariance of Errors (Σe)					
1	1	0.199	(0.028)	0.831	(0.148)
	2	0.044	(0.017)	0.185	(0.078)
2	2	0.655	(0.044)	1.428	(0.217)

Panel D: Random Effects (δ and η)	Param.	Low SES Coeff.	Low SES Std Err.	High SES Coeff.	High SES Std Err.
Coefficients on Mean Elements (Mean Utility)					
Constant	$\bar{\beta}$	-2.349	(0.241)	-4.072	(0.556)
Price	$\bar{\beta}^p$	0.149	(0.027)	0.577	(0.063)
Quality	$\bar{\beta}^q$	0.366	(0.032)	0.423	(0.070)
Coefficients on Mean Elements (Discoverability)					
Constant	$\bar{\alpha}$	-0.208	(0.045)	-0.571	(0.092)
Price	$\bar{\alpha}^p$	0.190	(0.029)	0.337	(0.033)
Quality	$\bar{\alpha}^q$	0.007	(0.040)	0.039	(0.047)
Variances					
Mean Utility	σ_δ	0.511	(0.051)	0.971	(0.181)
Discoverability	σ_η	0.495	(0.024)	0.433	(0.021)
Covariances					
M. Utility - Discov.	$\sigma_{\delta,\eta}$	0.242	(0.020)	0.247	(0.037)

Panel E: Search Costs	Param.	Low SES Coeff.	Low SES Std Err.	High SES Coeff.	High SES Std Err.
Pin Click					
Std. dev. of $\bar{c}_i^{pin}$	$\sigma_{\bar{c}}$	2.690	(0.495)	0.260	(0.074)
Std. dev. of the shock	σ^{pin}	3.692	(0.687)	0.432	(0.189)
Intercept - first	$\bar{c}_0$	3.726	(0.911)	0.164	(0.141)
Std. dev. of the shock - first	σ_0^{pin}	1.198	(0.674)	1.594	(0.423)
Coefs on xc					
Constant	γ^{cost}	-4.782	(0.797)	-0.671	(0.288)
Mean π	γ^{cost}	1.077	(0.157)	0.174	(0.068)
Pr Place i	γ^{cost}	-0.714	(0.112)	-0.049	(0.050)
Profile Click (Detailed View)					
Mean	$\bar{c}^{det.}$	2.248	(0.501)	0.650	(0.084)
Standard Deviation	$\sigma^{det.}$	0.805	(0.553)	-0.607	(0.172)

Note: This table presents selected parameter estimates. All parameters are estimated separately by SES group. Standard errors are bootstrapped. Panel A: preferences over perceived school characteristics and information parameters, including knowledge shifters, subjective admission chance parameters, and subjective beliefs about unknown schools ($\bar{\lambda}_h$ is the mean of the Poisson parameter λ_{ith} across individuals within latent type h before search starts $t=0$). Panel B: variance-covariance matrix of random coefficients in utility. Panel C: match value shock primitives (μ^l is the mean subjective expectation of ε_{ij} given the signal $\tilde{\varepsilon}_{ij}$). Panel D: school-level random effects for mean utility (δ) and discoverability (η). Panel E: search cost parameters for pin clicks (deterministic component, variance, and first-click premium) and profile clicks.

utility of the final allocation under true preferences (welfare), placement probabilities, mean characteristics of placed schools, and search activity.

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. We condition on the actions that parents took in the data when drawing latent terms and adjust the information shifters

TABLE IV

COUNTERFACTUAL 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	0.546	0.739	1.437	2.998	0.141	4.342	1.243	0.712
		(0.026)	(0.011)	(0.012)	(0.009)	(0.006)	(0.197)	(0.065)	(0.158)
(2)	Full information	1.312	0.837	1.590	3.164	0.204	-	-	-
		(0.061)	(0.005)	(0.016)	(0.019)	(0.008)	-	-	-
(3)	Gains (difference (2)-(1))	0.765	0.097	0.154	0.165	0.063	-	-	-
		(0.051)	(0.013)	(0.017)	(0.018)	(0.007)	-	-	-
Gains from fixing perceptions or eliminating search costs (relative to baseline)									
(4)	Fix Information	0.508	-0.009	0.063	0.156	0.060	2.717	0.644	2.044
		(0.038)	(0.007)	(0.009)	(0.014)	(0.006)	(1.358)	(0.330)	(0.499)
(5)	No Search Cost	0.225	0.141	0.100	-0.018	-0.012	-	-	-
		(0.023)	(0.013)	(0.010)	(0.004)	(0.002)	-	-	-
(6)	Cost and Information Interaction	0.032	-0.035	-0.010	0.027	0.015	-	-	-
		(0.008)	(0.006)	(0.007)	(0.006)	(0.003)	-	-	-

Note: This table presents counterfactual simulations pooled across SES groups. Columns: Welfare is the expected utility of the final allocation under true preferences (in equivalent kilometers, as the distance coefficient is normalized to -1). Place: probability of placement. E(rank): expected rank of placed school within the submitted application. Quality: school quality index (1–4 scale). VA: school value added in student-level standard deviations. Single, Double: number of pin and profile clicks. V(1st): perceived value of first click. Rows (1)–(2) report levels for the baseline and full-information scenarios; row (3) reports the difference. Rows (4)–(6) report differences relative to baseline. Row (4): fix all information (set $\hat{x}_{ij} = x_{ij}$, correct beliefs about unknown schools' characteristics, admission chances, and match values). Row (5): eliminate search costs. Row (6): interaction, defined so that row (3) = row (4) + row (5) + row (6). Search activity columns are not reported when search is irrelevant (full information) or unbounded (no search costs).

w_{ijt} (treatment indicators and search history) so that all parents are assigned to the "control" arms in the feedback and search experiments. Under full information, we endow parents with complete knowledge of all schools and correct all misperceptions and biases, setting $\pi_{ijT} = 3$ (the full-information benchmark in Section 3, where payoff-relevant objects are observed without error) and $\hat{x}_{ij} = x_{ij}$ for all i and j. This scenario provides an upper bound on welfare gains. Row (3) reports the difference between the two scenarios, which we use as the reference for the counterfactuals that follow. Welfare would increase by the equivalent of 0.765 fewer kilometers traveled, placement probability would increase by 9.7 percentage points, expected quality by 0.165 points on a 4-point scale, and value added by 0.063 student standard deviations.

Fixing information and eliminating search costs: We next consider two counterfactuals that correspond to policies a planner could pursue. In the first (row 4), we fix all information but retain imperfect awareness of schools and search costs. We correct misperceptions of school characteristics ($\hat{x}_{ij} = x_{ij}$), set the subjective probability of discovering unknown schools equal to the objective probability (correcting λ_i and φ_i), remove bias and compression in admission chance beliefs ($o_{i0} = 0$, $o_{i1} = 1$, so that $\hat{r}_{ij} = r_{ij}$), and eliminate bias and measurement error in match value signals ($\mu^l = 0$, $\sigma_e^2 = 0$).[47] In the second (row 5), we eliminate search costs (i.e., take $c_{is}^{\text{pin}}, c_{is}^{\text{detail}} = 0$ for all i, s) while keeping all beliefs and misperceptions at their baseline values, so that parents can search freely but remain misinformed.

Fixing information alone would account for 66% of the welfare gains and would capture nearly all of the gains in school quality (0.156 of 0.165) and value added (0.060 of 0.063). Eliminating search costs alone would account for 29% of the welfare gains, but would reduce both school quality (by 0.018) and value added (by 0.012). Under biased beliefs, inducing parents to search more does not lead them to better schools.

Complementarity between information and search costs: Row (6) captures the interaction between the two counterfactuals described above. It is defined so that the total gain from full information (row 3) equals the sum of fixing information (row 4), eliminating search costs (row 5), and the interaction (row 6). A positive interaction means that fixing information and eliminating search costs together yields gains that exceed the sum of doing each separately. For welfare, this interaction is small, accounting for 4% of the total gain. In quality and value added, in contrast, the gains from fixing information are substantially larger when search costs are zero. For quality, the interaction would account for 16% of the total gain, and for value added it would account for 24%. An interpretation is that fixing misperceptions is more valuable when the set of known schools is larger.

[47] In the supplement, Table S.VI, we provide further results. We decompose this counterfactual, considering the effects of addressing each of these misperceptions or biases in sequence. In addition, we consider the effects of improving the search technology so that a pin click always causes high knowledge ($\pi_{ij3} = 2$), with and without addressing information frictions. Most gains come from addressing misperceptions of x, and, secondarily, from providing correct beliefs about the number and characteristics of unknown schools.

The interaction could have had the opposite sign. For instance, if parents were misinformed only about rejection chances r, then search decisions would have been distorted, but in the absence of search costs there would have been no distortion in the application decision; hence, reducing search costs and removing information frictions would have been substitutes.

Heterogeneity: We focus on two dimensions of heterogeneity, summarized in Figure 5. The first is by SES, the same grouping used throughout the paper. For both SES groups, the interaction between information and search costs accounts for less than 5% of welfare gains but between 23% and 25% of the gains in value added. Low-SES parents benefit more in absolute terms, and full information would close the baseline quality gap that exists between SES groups.[48]

The second dimension is by type of baseline bias. We split parents by whether they especially underestimate the number of highlight-worthy schools at baseline (Figure 4b). As most parents underestimate this count, we split at the 30th percentile. Parents who most underestimate the number of highlight-worthy schools (low λ_i) have larger welfare gains in the full-information benchmark, and would have substantially larger gains from either fixing information or removing search costs, than those parents who are less biased, consistent with the number-of-schools channel from Section 3.2. Moreover, among parents who are pessimistic about the number of highlight-worthy schools, the interaction would account for 54% of the total value-added gain (0.026 of 0.048).

Search activity and search cost reduction: To understand the mechanisms through which information and search costs affect outcomes, we examine how search behavior changes under these counterfactuals. The last three columns of Table IV show that fixing information raises the number of single clicks by 2.8 (from 4.3 at baseline to 7.1) and nearly quadruples the perceived value of the first search (from 0.71 to 2.76), confirming that correcting beliefs raises perceived returns and induces more search. Figure 6 shows the effects of gradually reducing search costs, with and without correcting information. Without correcting information, even an

[48]Full results by SES are in Panel A of Appendix Table A.XXIV.

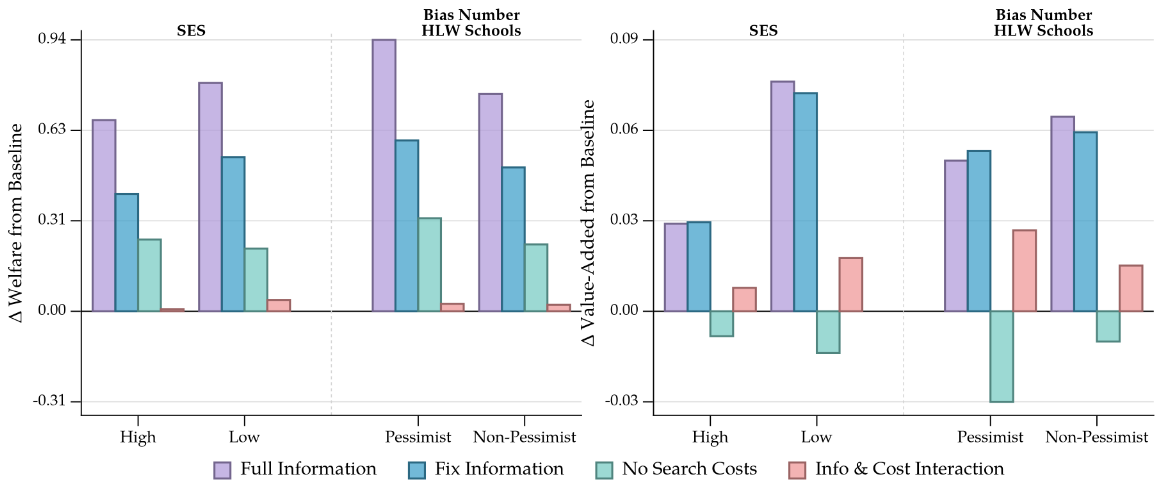


FIGURE 5.—Counterfactual gains by group. The figure shows the difference in welfare (left) and value added (right) relative to baseline achieved by each counterfactual, expressed as a percentage of the full-information gain. Welfare is the expected utility of the final allocation under true preferences (in equivalent kilometers). Value added is in student-level standard deviations. Groups: "High SES" and "Low SES" split by parental college education. "Pessimist" and "Non-Pessimist" split by whether parents underestimate the number of highlight-worthy schools at baseline ($\leq$ 30th percentile of the gap between perceived and actual count). Within each group, bars report gains from fixing information (set $\hat{x}_{ij} = x_{ij}$, correct all beliefs) and from eliminating search costs. The "Info & Cost Interaction" bar shows the fraction of the full-information gain attributable to the complementarity between information and search costs, defined so that the full-information gain equals the sum of fixing information, eliminating search costs, and the interaction.

80% reduction in search costs would achieve modest welfare gains and negligible or negative effects on school quality.

Implications for empirical work: Finally, to assess the importance of modeling biased beliefs, we consider what a researcher would conclude from counterfactual simulations based on simpler models that restrict the misperceptions our framework allows for. A researcher who had not collected survey data on perceptions of school characteristics would estimate the model under the assumption that perceived school characteristics equal their true values at all knowledge levels ($\hat{x}^k_{ij} = x_{ij}$ for $k = 1,2$ and $\tilde{f}^x(x;\varphi) = f^x(x)$). We call this the *no misperception of x* specification. We re-estimate the model under this restriction and simulate counterfactuals from the resulting estimates.

Figure 7 compares the counterfactual results across specifications. Under the full model, providing full information would increase school quality by 0.165 and

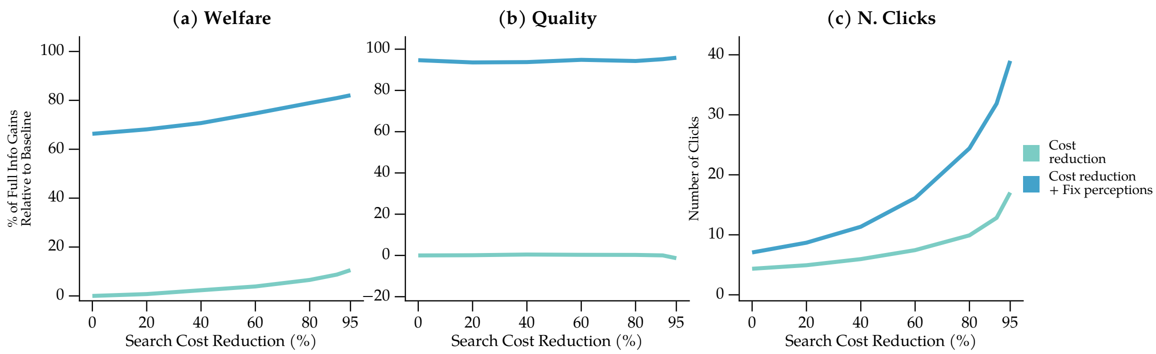


FIGURE 6.—Search cost reduction and information. The figure shows welfare (left), quality (middle), and number of clicks (right) as a function of the percentage reduction in search costs (horizontal axis, 0%–100%). Two series are plotted: "Cost Reduction" reduces search costs while keeping all beliefs and misperceptions at their baseline values; "Cost Reduction + Fix Perceptions" additionally sets $\hat{x}_{ij} = x_{ij}$ and corrects beliefs about admission chances, match values, and the number of undiscovered schools. Welfare is the expected utility of the final allocation under true preferences (in equivalent kilometers). Quality is the school quality index (1–4 scale). Clicks is the total number of pin clicks. Dashed horizontal lines mark the full-information benchmark. The gap between the two series at each cost reduction level reflects the complementarity between information and search costs. Full results are reported in Table S.VII.

value added by 0.063. Under the misspecified model, these signs reverse: full information would reduce school quality and value added relative to baseline. The reversal occurs because the misspecified model attributes the gap between perceived and true school characteristics to preference heterogeneity and school unobservables instead, incorrectly concluding that parents do not value quality.

We also simulate the elimination of search costs in the misspecified model. Parents maintain misperceptions of ε, and need not gain high information about every school, although every school is clicked. Under this specification, eliminating search costs would reduce quality by 0.025 and value added by 0.016.

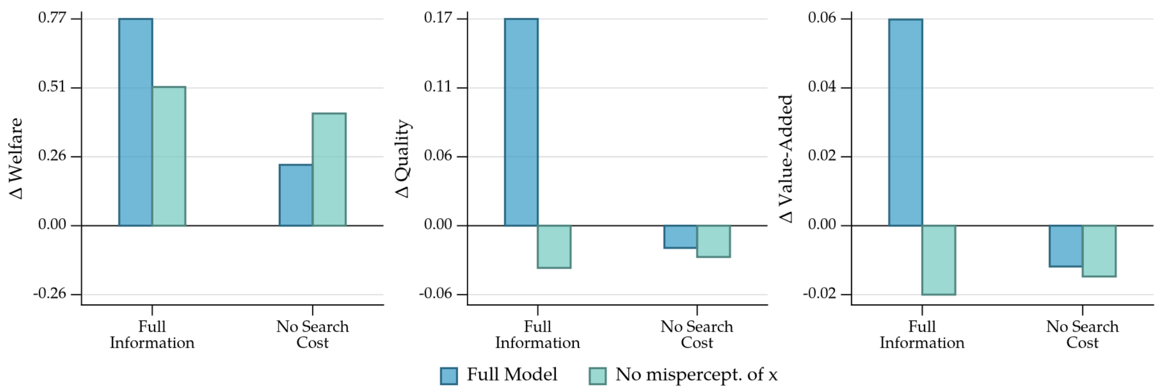


FIGURE 7.—Counterfactual results under misspecified models. The figure compares the change in welfare (left), quality (middle), and value added (right) relative to baseline under two counterfactuals—full information and elimination of search costs—across two model specifications. "Full Model" corresponds to our main specification. "No mispercept. of x" re-estimates the model imposing $\hat{x}_{ij} = x_{ij}$ at all knowledge levels, so that perceived and true school characteristics always coincide. Welfare is in equivalent kilometers, quality on a 1–4 scale, and value added in student-level standard deviations.

Interpreting our findings: To contextualize our results, we focus on value added. Campos (2024) and Ainsworth et al. (2023) experimentally provide value-added information in Los Angeles and Romania, respectively, finding effects of 0.13-0.25 SD in the value added of the top-ranked school (Campos) and 0.05 SD in enrolled-school value added with 0.20 SD among a subgroup rejected from their first choice (Ainsworth). We find 0.06 SD in enrolled-school VA in our perfect information-intervention counterfactuals, and 0.08 SD in our search experiment for high-SES parents (treatment 1).A key difference is that our experiment was not intended to maximally improve value added or other quality measures. Rather, its purpose was to vary beliefs over unknowns to identify our model.[49] In addition, we did not provide information about value added specifically.

Counterfactual gains exceed our experimental treatment effects in the full sample (Section 6.3). This pattern is expected: the experiments are designed to separately vary specific components of information, take-up was partial, and observed

[49]For this reason we did not provide a list of schools by value added, or similar information that might have been helpful, in the search intervention.

belief updating was modest. The counterfactuals extrapolate to cases of complete belief updating about all relevant objects.

Our counterfactuals show that biased beliefs and inaccurate perceptions, not search costs per se, are the primary driver of welfare and quality losses. Moreover, interactions between information and search costs are quantitatively important for school quality and value added, and effects are larger for parents with greater biases. Our findings suggest that interventions that have the goal of delivering useful school quality information may be especially effective when delivered before parents search and when complemented with efforts to improve parents' search technology or reduce search costs.

10. CONCLUSION

This paper investigates the relationship 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 parents make search decisions complements their efforts, raising welfare and assigning them to higher-quality schools. Consistent with theory, providing accurate information has heterogeneous effects on parents' search efforts, with larger effects among parents with greater initial pessimism about the availability of desirable schools.

We consider average effects of policies that provide information or reduce costs. It may be valuable to policymakers to understand whom to treat, with what information, and at what time, given that information provided at the wrong time may be ignored in practice. Our model, data, and design may be useful for questions of timing and targeting. In addition, while our counterfactuals hold schools' admission chances and characteristics fixed, our analysis may serve as 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. [4]

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

——— (2025): "Demand Analysis under Latent Choice Constraints," The Review of Economic Studies, rdaf093. [5]

AGARWAL, SUMIT, JOHN GRIGSBY, ALI HORTAÇSU, GREGOR MATVOS, AMIT SERU, AND VINCENT YAO (2024): "Searching for approval," Econometrica, 92 (4), 1195–1231. [2, 6]

AINSWORTH, ROBERT, RAJEEV DEHEJIA, CRISTIAN POP-ELECHES, AND MIGUEL URQUIOLA (2023): "Why Do Households Leave School Value Added on the Table? The Roles of Information and Preferences," American Economic Review, 113 (4), 1049–1082. [6, 55]

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 (2025): "When Information is Not Enough: Evidence from a Centralised School Choice System," The Economic Journal, 136 (673), 26–60. [5]

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. [30]

ALLENDE, CLAUDIA, FRANCISCO A. GALLEGO, AND CHRISTOPHER A. NEILSON (2019): "Approximating the Equilibrium Effects of Informed School Choice," Working Paper. [6, 8]

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. [5]

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, 7, 9]

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. [30]

BANDIERA, ORIANA, VITTORIO BASSI, ROBIN BURGESS, IMRAN RASUL, MUNSHI SULAIMAN, AND ANNA VITALI (2025): "The search for good jobs: evidence from a six-year field experiment in Uganda," Journal of Labor Economics, 43 (3), 885–935. [6]

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. [30]

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

BROWN, ZACH Y AND JIHYE JEON (2024): "Endogenous information and simplifying insurance choice," Econometrica, 92 (3), 881–911. [2, 6]

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. [30]

CAMPOS, CHRISTOPHER (2024): "Social Interactions, Information, and Preferences for Schools: Experimental Evidence from Los Angeles," Working Paper 33010, National Bureau of Economic Research. [6, 55]

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. [5]

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. [5]

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. [4, 7, 8, 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 GROOTE, OLIVIER AND MINYOUNG RHO (2024): "Matching disadvantaged children to day care: Evidence from a centralized platform," Working Paper. [36]

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. [30]

GOEREE, MICHELLE SOVINSKY (2008): "Limited information and advertising in the US personal computer industry," Econometrica, 76 (5), 1017–1074. [38]

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. [5]

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, 6]

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. [5]

HE, YINGHUA, SHRUTI SINHA, AND XIAOTING SUN (2024): "Identification and Estimation in Many-to-One Two-Sided Matching Without Transfers," Econometrica, 92 (3), 749–774. [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. [5]

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

KAPOR, ADAM, MOHIT KARNANI, AND CHRISTOPHER NEILSON (2024): "Aftermarket frictions and the cost of off-platform options in centralized assignment mechanisms," Journal of Political Economy, 132 (7), 2346–2395. [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 (2019): "The value of information in centralized school choice systems," Working Paper. [5]

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. [5]

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. [23, 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. [30]

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. [36]

SHAKED, MOSHE AND J GEORGE SHANTHIKUMAR (2007): Stochastic orders, Springer. [18]

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

VATTER, BENJAMIN (2025): "Quality Disclosure and Regulation: Scoring Design in Medicare Advantage," Econometrica, 93 (3), 959–1001. [39]

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

ONLINE APPENDIX

A.1. ADDITIONAL INFORMATION ON THE SCHOOL QUALITY CATEGORIES

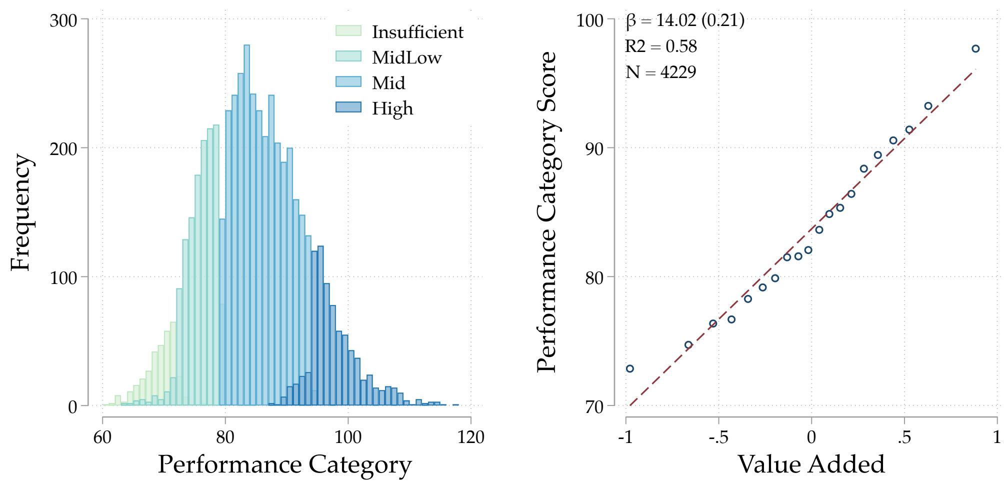
Our quality measure comes from the Education Quality Agency (Agencia de Calidad de la Educación), which classifies schools into four categories (high, medium, medium-low, and insufficient performance).[50] The categories are based on a continuous performance score constructed using three main components. First, learning outcomes (67% weight): the distribution of students across SIMCE learning level categories in 4th and 8th grade. Second, indicators of personal and social development (33% weight), which include school climate and coexistence, academic self-esteem and school motivation, citizenship participation and civic formation, healthy lifestyle habits, school retention, gender equity, attendance, and SIMCE performance trends. Third, student context characteristics: the performance score is adjusted for the socioeconomic context of the school's students using a context index that identifies whether a school operates in a more or less favorable context, incorporating a value-added approach. Figure A.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 A.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.

Table A.I shows the peer composition of schools for each of the four quality categories and Table A.II shows how school characteristics vary for over- and undersubscribed school programs. In Table A.III, we present information about excess

[50]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

capacity by school quality. Figure A.2 further plots the distribution of the number of applicants per available seat across programs.



(a) **(b)**

FIGURE A.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.

TABLE A.I

PEER COMPOSITION BY QUALITY CATEGORY

	Quality Category			
	Insufficient	Medium-Low	Medium	High
% SEP/Disadvantaged	61%	56%	50%	47%
Mother's Education Distribution:				
Less than high school (incomplete secondary or below)	52%	40%	29%	23%
High school graduate (complete secondary)	39%	45%	46%	47%
Postsecondary technical degree (complete; non-university)	6%	9%	14%	14%
University degree (complete)	3%	6%	11%	15%
N Students	5,592	33,830	108,251	31,737

Note: This table shows peer composition by school quality category. The unit of analysis is the applicant. Of the 207,578 applicants to entry grades (pre-K, kindergarten, or first grade) in 2021, 179,410 (86%) enrolled in 2022 and have complete quality measures and first-grade peer composition data for their enrolled school. This group constitutes the analysis sample. School quality categories are based on the performance category of the school where students enrolled in 2022. SEP/Disadvantaged status is measured at the time of the 2021 application and refers to students classified as prioritario in the SAE system. Mother's education composition reflects peer composition at each school, measured among first-grade students in 2016, the most recent year available. Mother's education categories are defined as: (1) University degree (complete), (2) Postsecondary technical degree (complete; non-university), (3) High school graduate (complete secondary), and (4) Less than high school (incomplete secondary or below).

TABLE A.II

SCHOOL CHARACTERISTICS BY EXCESS CAPACITY

	Oversubscribed	Undersubscribed
Quality Category (mean)	3.02 (0.63)	2.65 (0.76)
Insufficient (%)	1%	8%
Medium-Low (%)	15%	28%
Medium (%)	64%	54%
High (%)	19%	10%
Price Category (mean)	1.84 (1.22)	1.11 (0.51)
Free (%)	65%	95%
Low (%)	3%	1%
Mid (%)	14%	2%
High (%)	18%	2%
Value Added (mean)	0.165	-0.141
Avg Distance from Applicants (km)‡	2.92 (3.45)	2.80 (3.26)
SAE-active programs (morning and afternoon shifts)		
% Seats Filled†	100 (0.0)	37.0 (27.5)
Applicants per Seat†	4.73	1.12
N Programs†	2,252	4,973
N Programs	3,927	13,004

Note: This table describes school-campus-grade programs at entry grades (pre-K through 1st grade). A program is oversubscribed if all available seats were filled by the deferred acceptance (DA) algorithm. In student-proposing DA, a school fills all seats only if it rejected at least one applicant during the algorithm. Quality categories range from 1 (Insufficient) to 4 (High), based on the Education Quality Agency classification. Price categories range from 1 (Free) to 4 (High, 100k+ CLP). Standard deviations in parentheses. † Restricted to morning and afternoon shift programs that actively participate in the centralized SAE assignment. Full-day programs, which largely fill through direct enrollment outside the centralized system, are included in all other rows. ‡ Restricted to urban schools; schools with average applicant distance greater than 50 km are excluded as geocoding outliers.

TABLE A.III

EXCESS CAPACITY BY QUALITY CATEGORY

Quality Category	% with Excess Capacity	Avg Excess Seats		N Programs	
		Level	% of Seats	Count	% of Total
Insufficient	94%	17	74%	900	6%
Medium-Low	83%	14	59%	3,520	25%
Medium	68%	10	43%	8,088	57%
High	55%	7	27%	1,749	12%
Total	72%	11	47%	14,257	100%

Note: This table shows the distribution of excess capacity across quality categories at the school-campus-grade level for entry grades (pre-K through 1st grade). Excess capacity is defined as the difference between available seats and students assigned by the DA algorithm. Programs with excess capacity have more available seats than assigned students. Avg Excess Seats (% of Seats) is the average excess seats as a percentage of average total seats per program in that category. N Programs (% of Total) is the share of total programs with non-missing quality data. 2,674 programs without a quality classification (ungraded or new schools) are excluded.

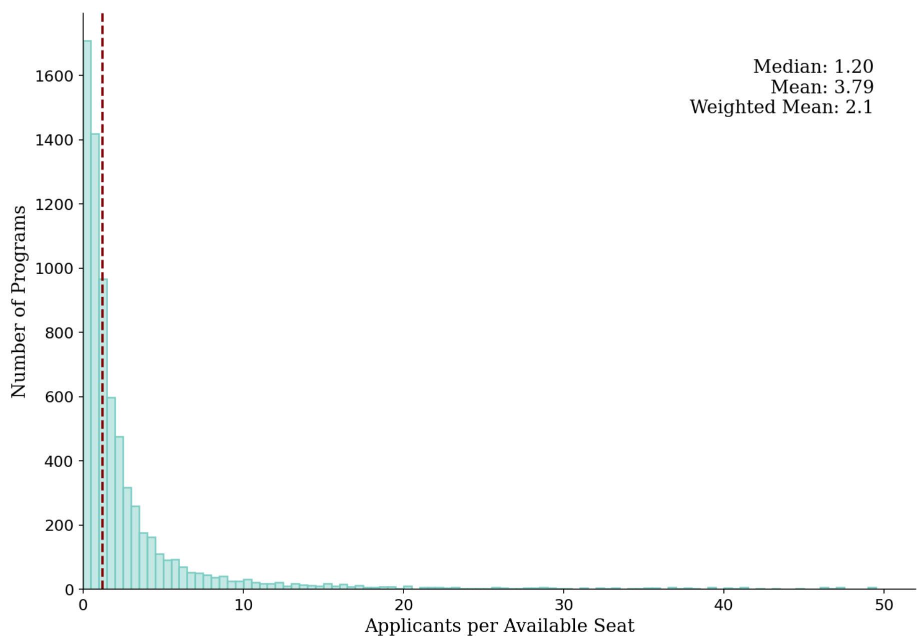


FIGURE A.2.—Distribution of Applicants per Available Seat. Each observation is a school-campus–grade program in SAE-active shifts (morning and afternoon) at entry grades. The mean is the unweighted average of applicants per seat across programs. The weighted mean is total applicants divided by total seats, weighting each program by its number of seats.

A.2. MODEL AND COMPARATIVE STATICS APPENDIX

This section provides additional information on Section 3.

Convexity of Search Returns (Expression (3)): Supposing school j' is placed in the (n^*+1)th position, the payoff gain is a linear function of $\hat{u}^q_{ij'}$ with slope $\left(\prod_{m=1}^{n^*} r_{ij_m}\right)(1-\hat{r}_{ij'}) \geq 0$. If j' is not ranked, the payoff gain is zero for all $\hat{u}^q_{ij'}$. Expression (3) is the upper envelope of these linear functions.

PROOF OF PROPOSITION 1: Part (a): Each of the linear functions produced when j' is ranked in the (n^*+1)th position, for $n^* = 0, \ldots, k$, is weakly decreasing in $\hat{u}_{ij}$ and weakly increasing in $\hat{r}_{ij}$ for each j. The result follows. Part (b) ($\Rightarrow$): Expression (3) is an increasing convex function of the discovered school's payoff. ($\Leftarrow$): functions of the form $h(u) = \max\{0, u - a\}$ are a basis for the set of increasing convex functions. Search gains (Expression (3)) take this form when there is a single known school, school 1, with expected payoff $\hat{u}_{i1} = a$ and rejection chance $\hat{r}_{i1} = 0$. *Q.E.D.*

Uniform misperceptions of r: Suppose $J_i = \{1, 2\}$, $r_1 = r_2 = 0$, $\hat{r}_1 = \hat{r}_2 = \alpha$, $\hat{u}_1 = u_1 = 1$ almost surely, and $\hat{u}_2 = u_2 \sim U[0,1]$. That is, true rejection chances at known and unknown schools are zero, but they are perceived to be α. Let p_2 be the chance of clicking school 2. If school 1 is known, the expected returns to search are $p_2\alpha(1-\alpha)/2$, which is increasing in α for $\alpha < \frac{1}{2}$. In contrast, if neither school is known, the returns are $p_2(1-\alpha)/2 + (1-p_2)(1-\alpha)$, which is decreasing in α.

Making desirable schools more salient: Consider our example in Section 3.2. Consider a change in τ that makes high-quality schools easier to find. If school 1 has a high quality score but school 2 does not, then the probability of finding school 2 falls. On the other hand, if only school 2 has a high quality score, the probability of discovering it rises.

A.3. ADDITIONAL SAMPLE INFORMATION

In this section, we present additional information about our study sample. Figure A.3 plots explorer usage patterns over time. Table A.IV examines selection into the study sample by showing differences in family characteristics between the universe of applicants and the study sample. Table A.V shows that there was no differential attrition across treatment groups. Tables A.VI and A.VII present comparisons between parents who completed the midline and endline surveys and those who did not.

TABLE A.IV

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.

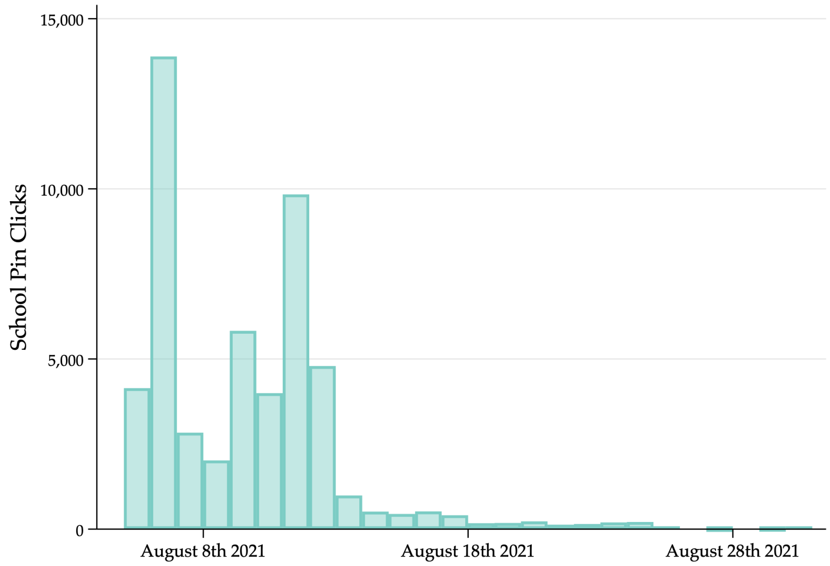


FIGURE A.3.—Search timing. Notes: This figure shows the timing of the search process. Access to the explorer platform was granted on August 5th and school applications closed on September 8th. Over 95% of school pin clicks occurred in the first two weeks.

TABLE A.V

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.VI

MIDLINE ATTRITION COMPARISON WITHIN CONTROL GROUP

	Did Not Complete Midline		Completed Midline		
	Mean (1)	St. Dev. (2)	Coeff. (3)	St. Err. (4)	N (5)
Panel A: Choice Environment					
Number of available schools 2km	16.222	[9.169]	0.615	(0.512)	1318
Number of available highlight-worthy schools	8.640	[5.014]	0.431	(0.279)	1318
Number of available high quality schools	9.769	[5.316]	0.459	(0.297)	1318
Number of available low price schools	14.761	[8.858]	0.553	(0.494)	1318
Panel B: Parent/Child Characteristics					
Child is female	0.495	[0.500]	0.008	(0.028)	1318
Mother completed college	0.220	[0.414]	-0.018	(0.023)	1316
Number of younger siblings	1.146	[0.387]	-0.016	(0.022)	1318
Child has a disability (belief)	0.070	[0.255]	-0.014	(0.015)	1185
Parent works in a school	0.066	[0.249]	0.009	(0.014)	1295
SEP household	0.450	[0.498]	0.018	(0.028)	1302
Child's age	3.904	[0.550]	0.032	(0.031)	1318
Panel C: Initial Knowledge and Beliefs					
Expected satisfaction with process	5.235	[1.395]	0.080	(0.080)	1224
Listed any school as first preference	0.904	[0.295]	0.005	(0.016)	1318
First-preference school is highlight-worthy	0.633	[0.482]	0.036	(0.030)	1053
Perceived admission chance for first-preference school	0.684	[0.272]	0.032**	(0.016)	1224
Number of schools known by name	3.301	[2.684]	0.117	(0.148)	1318
Number of schools known well	1.874	[2.046]	0.042	(0.114)	1318
Perceived number of available schools	7.444	[6.936]	0.065	(0.375)	1318
Perceived number of available highlight-worthy schools	3.671	[3.615]	0.184	(0.194)	1318
Perceived number of available high quality schools	5.011	[4.654]	0.015	(0.251)	1318
Perceived number of available low price schools	5.838	[5.051]	0.377	(0.274)	1318
Parent believes to be SEP eligible	0.172	[0.378]	-0.037*	(0.021)	1318
Parent is unsure about SEP status	0.665	[0.472]	0.050*	(0.026)	1318

Note: This table compares baseline covariates for parents who completed the midline survey with those parents that did not in the control group. Column 1 reports the mean of the dependent variable for parents who did not complete the midlne survey 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 whether the parent completed the midline survey. 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.VII

ENDLINE ATTRITION COMPARISON WITHIN CONTROL GROUP

	Did Not Complete Endline		Completed Endline		
	Mean	St. Dev.	Coeff.	St. Err.	N
	(1)	(2)	(3)	(4)	(5)
Panel A: Choice Environment					
Number of available schools 2km	16.222	[9.169]	1.012	(0.659)	1318
Number of available highlight-worthy schools	8.640	[5.014]	0.633*	(0.369)	1318
Number of available high quality schools	9.769	[5.316]	0.517	(0.385)	1318
Number of available low price schools	14.761	[8.858]	1.031	(0.641)	1318
Panel B: Parent/Child Characteristics					
Child is female	0.495	[0.500]	0.006	(0.036)	1318
Mother completed college	0.220	[0.414]	0.018	(0.030)	1316
Number of younger siblings	1.146	[0.387]	-0.006	(0.026)	1318
Child has a disability (belief)	0.070	[0.255]	-0.021	(0.017)	1185
Parent works in a school	0.066	[0.249]	-0.024	(0.016)	1295
SEP household	0.450	[0.498]	0.036	(0.036)	1302
Child's age	3.904	[0.550]	-0.061	(0.039)	1318
Panel C: Initial Knowledge and Beliefs					
Expected satisfaction with process	5.235	[1.395]	0.067	(0.104)	1224
Listed any school as first preference	0.904	[0.295]	0.006	(0.021)	1318
First-preference school is highlight-worthy	0.633	[0.482]	0.049	(0.037)	1053
Perceived admission chance for first-preference school	0.684	[0.272]	-0.026	(0.020)	1224
Number of schools known by name	3.301	[2.684]	0.049	(0.198)	1318
Number of schools known well	1.874	[2.046]	-0.024	(0.140)	1318
Perceived number of available schools	7.444	[6.936]	0.914	(0.612)	1318
Perceived number of available highlight-worthy schools	3.671	[3.615]	-0.042	(0.230)	1318
Perceived number of available high quality schools	5.011	[4.654]	0.211	(0.341)	1318
Perceived number of available low price schools	5.838	[5.051]	0.254	(0.336)	1318
Parent believes to be SEP eligible	0.172	[0.378]	0.065**	(0.029)	1318
Parent is unsure about SEP status	0.665	[0.472]	-0.049	(0.034)	1318

Note: This table compares baseline covariates for parents who completed the endline survey with those parents that did not in the control group. Column 1 reports the mean of the dependent variable for parents who did not complete the endline survey 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 whether the parent completed the endline survey. 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.VIII

ADDITIONAL DESCRIPTIVE STATISTICS

	Mean (1)	St. Dev. (2)	*N* (3)
Panel A: Choice Environment			
Number of available schools 2km	16.154	[9.289]	3948
Number of available highlight-worthy schools	8.617	[5.100]	3948
Number of available high quality schools	9.741	[5.423]	3948
Number of available low price schools	14.691	[8.934]	3948
Panel B: Parent/Child Characteristics			
Child is female	0.509	[0.500]	3948
Mother completed college	0.225	[0.418]	3945
Number of younger siblings	1.155	[0.397]	3948
Child has a disability (belief)	0.071	[0.257]	3528
Parent works in a school	0.063	[0.243]	3885
SEP household	0.441	[0.497]	3908
Child's age	3.896	[0.530]	3948
Panel C: Initial Knowledge and Beliefs			
Expected satisfaction with process	5.246	[1.399]	3689
Listed any school as first preference	0.905	[0.294]	3948
First-preference school is highlight-worthy	0.666	[0.472]	3149
Perceived admission chance for first-preference school	0.696	[0.265]	3689
Number of schools known by name	3.284	[2.724]	3948
Number of schools known well	1.888	[2.056]	3948
Perceived number of available schools	7.323	[6.763]	3948
Perceived number of available highlight-worthy schools	3.647	[3.425]	3948
Perceived number of available high quality schools	4.952	[4.342]	3948
Perceived number of available low price schools	5.789	[5.249]	3948
Parent believes to be SEP eligible	0.170	[0.375]	3948
Parent is unsure about SEP status	0.669	[0.471]	3948

Note: This table summary statistics for the full survey sample. Variables in Panel A come from administrative data. Variables in Panels B and C come from the baseline survey.

A.4. ADDITIONAL DESCRIPTIVE RESULTS

In this section, we present additional descriptive results. Figure A.4 plots differences in school knowledge by distance and school quality. Figure A.5 shows 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 first in the application. Figure A.6 shows that parents also have incorrect perceptions about the relative quality of schools. In Table A.IX, we exploit the panel structure of our data to document that better school knowledge is associated with more accurate perceptions about the price and quality of a school. Table A.X present results of a rank-ordered logit choice model with actual and perceived school attributes.

Figure A.7 plots differences in perceived and true admission chances. We find that beliefs are biased upwards on average and exhibit compression. Figure A.8 compares the actual and perceived share of schools in each quality and price category. Table A.XI shows that parental beliefs about the availability of schools affect search effort. Tables A.XII, A.XIII, and A.XIV 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.XII) and that a parent's perceptions about school quality and admission chances are correct (Table A.XIII). These relationships should be interpreted as descriptive, since deeper engagement (e.g., opening a full profile after a pin click) is endogenous. We also find that a parent's experience on the school explorer platform affects beliefs about the distribution of schools' characteristics. In Table A.XIV, 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. Consistent with sequential search, we find in Table A.XV that parents are more likely to stop searching when the last searched school is highlight-worthy. Figure A.9 documents differences in search timing by SES status.

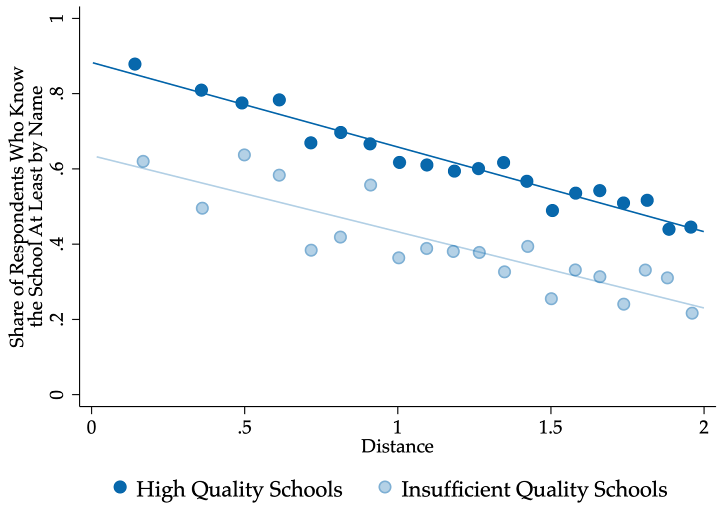
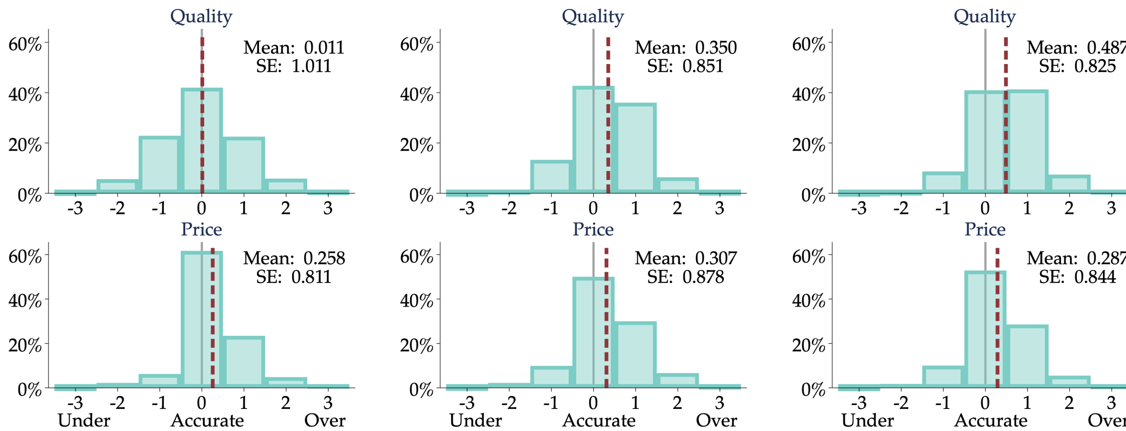
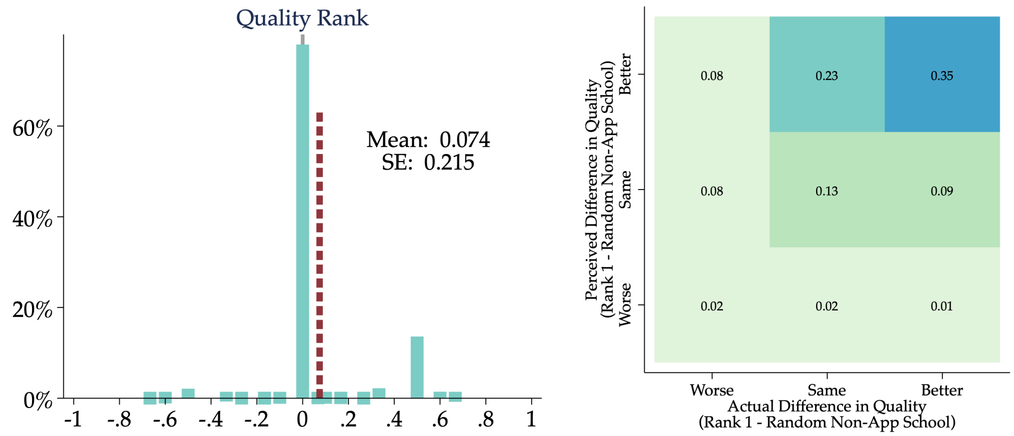


FIGURE A.4.—Knowledge of Schools by Distance and Quality. Notes: The figure uses baseline survey information for the full sample (N=3,823) to plot the share of parents who know a school at least by name by the distance of the school to the respondent's home (in km), separately for schools with high quality (dark blue) and insufficient quality (light blue).



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

FIGURE A.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=2,297 for quality and N=2,066 for price). 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=1,744 for quality and N=1,534 for price). Panel (c) shows the bias on perceived quality and price of the first preference school at baseline (N=2,998 for quality N=2,523 for price). All biases are measured as the perceived value minus the true value. Positive values indicate that the parent overestimates the quality or price of the school and negative values indicate that the parents underestimate the quality or price of the school. Quality is measured in four categories based on the classification of the Education Quality Agency. Prices are also categorized into four categories (free, 1-50k CLP, 50k-100k CLP, 100k+ CLP). Red dashed lines indicate the mean bias. The solid grey lines indicate the point of zero bias.



(a) Relative Quality of Rank 1 School **(b)** Difference in Bias

FIGURE A.6.—Relative belief measures. Panel (a) shows the bias on perceived relative quality of the first-ranked school at baseline for the full sample (N=2,998 for quality). Relative quality is defined as the quality percentile rank of the school relative to all other schools within 2km of the respondent's home. Bias is measured as the perceived value minus the true value. Positive values indicate that the parent overestimates the relative quality of the school and negative values indicate that the parents underestimate the relative quality or price of the school. Quality is measured in four categories based on the classification of the Education Quality Agency. Red dashed lines indicate the mean bias. The solid grey lines indicate the point of zero bias. Panel (b) plots parents' perceived difference in quality between their first-ranked school and a randomly selected school to which they did intend to apply (y-axis)) against the corresponding actual difference in quality (x-axis). Each cell reports the share of observations falling into the corresponding bin. Data from the baseline survey and the sample consists of all parents for which the relevant information is available (N= 2,040).

TABLE A.IX

SCHOOL KNOWLEDGE AND PERCEIVED SCHOOL ATTRIBUTES

	Price		Quality		Pr Admission Chance	
	Absolute Value (1)	Correct (2)	Absolute Value (3)	Correct (4)	Absolute Value (5)	Correct (6)
Knows Well	-0.144***	0.108***	-0.044*	0.047*	-0.004	0.002
	(0.037)	(0.027)	(0.026)	(0.024)	(0.013)	(0.025)
Mean of Outcome	0.444	0.627	0.563	0.514	0.341	0.199
Observations	3631	3631	4528	4528	2961	2961

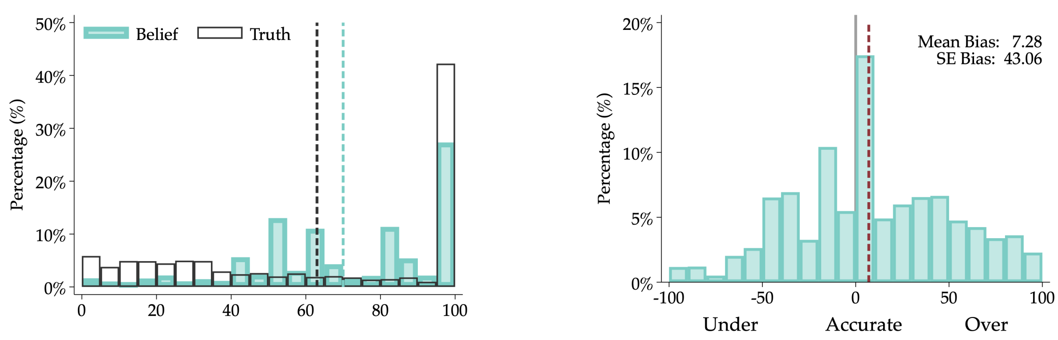
Note: This table show how school knowledge affects the perceptions of price (Columns 1-2), quality (Columns 3-4), and probability of being admitted (Columns 5-6) for the full sample. Columns 1, 3 and 5 represent the absolute difference between the perceived and actual value. Columns 2, 4 and 6 are indicators if the parent's perceptions are correct. For admission chances, we consider the answer to be correct if the difference between the perceived and actual value is not more than 10 percentage groups. Each regression pools information from the baseline, midline, and endline surveys and includes a dummy for whether the parent knows the school well and parent × school fixed effects. Robust standard errors are reported in parentheses.

TABLE A.X

EFFECTS OF PERCEIVED VS REAL CHARACTERISTICS ON RANKING

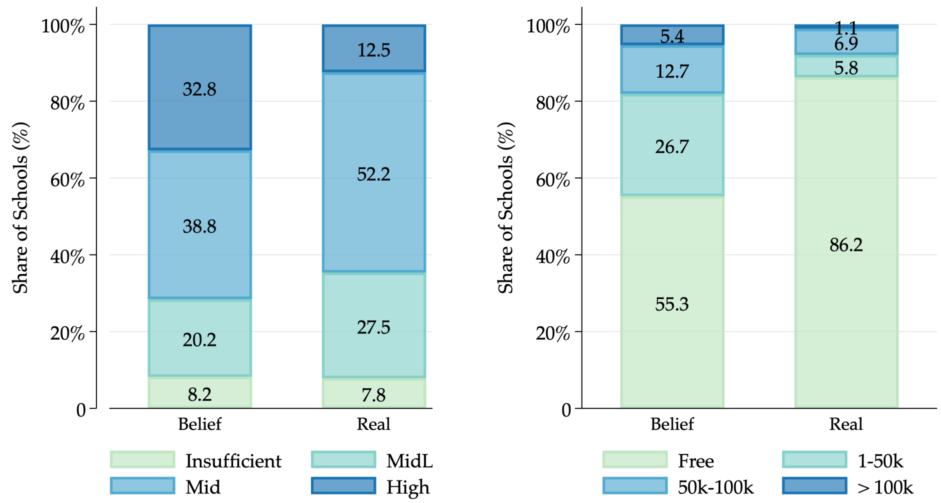
		Only Endline Survey Responses (1)		Combining All Survey Responses (2)	
Distance		-0.024*	(0.014)	-0.026***	(0.008)
Perceived Price Category	Free	0.311*	(0.182)	0.204**	(0.090)
	50k-100k CLP	-0.016	(0.205)	0.005	(0.111)
	100k+ CLP	0.283	(0.513)	0.097	(0.239)
True Price Category	Free	0.094	(0.214)	-0.111	(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.



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

FIGURE A.7.—Error in placement chances. Notes: Panel (a) shows the perceived and true 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 true placement chances. Positive values indicate that the parent responded a higher placement chance than the truth. The red dashed line indicates the mean bias and the grey solid line indicates the point of zero bias.



(a) Quality **(b)** Price

FIGURE A.8.—Beliefs about the Relative Distribution of School Attributes. Notes: Panel (a) shows the perceived (left) and actual (right) share of schools in each of the four school quality categories. Panel (b) 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).

TABLE A.XI

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

Outcome:	Number of Clicks						
Belief Measure:	Number of Schools	Number of Highlight-worthy Schools	Number of High Quality Schools	Number of Low Price Schools	1(1st Choice Is Highlight-worthy) × Admission Chance for 1st Choice	1(1st Choice Is Highlight-worthy)	Admission Chance for 1st Choice
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: OLS Regression							
Bias (Belief - Truth)	0.167**	-0.102	0.099	0.155*	-4.612***	-4.318***	-3.904***
	(0.074)	(0.097)	(0.098)	(0.090)	(1.476)	(1.444)	(1.502)
Truth	0.348***	0.151	0.405***	0.302***	-9.548***	-6.039***	-9.281***
	(0.076)	(0.107)	(0.103)	(0.093)	(1.729)	(1.658)	(1.817)
Control Group Mean	8.008	8.008	8.008	8.008	8.690	8.719	7.603
Observations	1,027	1,027	1,027	1,027	552	565	833
Panel B: Quantile Regression (50th Percentile)							
Bias (Belief - Truth)	0.150***	0.000	0.136**	0.145**	-4.000***	-5.000***	-2.350**
	(0.048)	(0.082)	(0.069)	(0.062)	(1.052)	(0.897)	(1.071)
Truth	0.237***	0.167*	0.295***	0.224***	-5.000***	-4.000***	-5.950***
	(0.052)	(0.091)	(0.080)	(0.064)	(1.212)	(0.981)	(1.256)
Control Group Percentile	3.00	3.00	3.00	3.00	3.00	3.00	3.00
Observations	1,027	1,027	1,027	1,027	552	565	833
Panel C: Quantile Regression (75th Percentile)							
Bias (Belief - Truth)	0.490***	-0.016	0.440***	0.292**	-7.870***	-9.000***	-5.810**
	(0.090)	(0.164)	(0.135)	(0.123)	(2.737)	(2.647)	(2.824)
Truth	0.729***	0.374**	0.920***	0.514***	-15.574***	-10.000***	-12.632***
	(0.097)	(0.182)	(0.158)	(0.126)	(3.152)	(2.896)	(3.311)
Control Group Percentile	10.00	10.00	10.00	10.00	10.00	10.00	10.00
Observations	1,027	1,027	1,027	1,027	552	565	833

Note: This table shows how perceived and true school availability and school attributes affect search effort in the control group. We regress each outcome on our bias measure, defined as beliefs minus the true value, and on the true value itself. Panel A reports estimates from OLS regressions, while Panels B and C report quantile regression estimates for the 50th and 75th percentiles, respectively.

TABLE A.XII

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.XIII

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)
Panel A: All Schools						
Single Click	-0.047	0.024	-0.135***	0.145***	2.442	0.030
	(0.053)	(0.045)	(0.047)	(0.039)	(4.378)	(0.044)
Double Click	-0.057	0.048	-0.065	0.040	0.239	-0.045
	(0.053)	(0.047)	(0.048)	(0.040)	(4.135)	(0.041)
Outcome at Baseline	0.247***	0.310***	0.392***	0.331***	0.528	0.217***
	(0.039)	(0.039)	(0.040)	(0.038)	(7.131)	(0.053)
Truth	0.059*	-0.045*	-0.193***	0.141***	-12.662**	0.036
	(0.032)	(0.024)	(0.034)	(0.022)	(5.037)	(0.050)
Mean of Outcome	0.357	0.674	0.503	0.559	20.376	0.180
Observations	524	524	676	676	416	416
Panel B: Schools Ranked First at Baseline						
Single Click	-0.035	0.013	-0.182***	0.134**	2.442	0.030
	(0.075)	(0.061)	(0.064)	(0.053)	(4.378)	(0.044)
Double Click	-0.044	0.048	-0.039	0.027	0.239	-0.045
	(0.067)	(0.057)	(0.056)	(0.048)	(4.135)	(0.041)
Outcome at Baseline	0.214***	0.293***	0.368***	0.288***	0.528	0.217***
	(0.055)	(0.051)	(0.057)	(0.056)	(7.131)	(0.053)
Truth	0.014	-0.027	-0.237***	0.182***	-12.662**	0.036
	(0.035)	(0.031)	(0.051)	(0.034)	(5.037)	(0.050)
Mean of Outcome	0.332	0.692	0.428	0.631	20.376	0.180
Observations	289	289	374	374	416	416
Panel C: Listed in Baseline						
Single Click	-0.039	0.017	-0.162***	0.126***		
	(0.064)	(0.054)	(0.054)	(0.045)		
Double Click	-0.051	0.044	-0.030	0.020		
	(0.059)	(0.052)	(0.049)	(0.043)		
Outcome at Baseline	0.240***	0.302***	0.360***	0.278***		
	(0.047)	(0.045)	(0.046)	(0.046)		
Truth	0.007	-0.025	-0.231***	0.172***		
	(0.029)	(0.026)	(0.042)	(0.028)		
Mean of Outcome	0.365	0.661	0.443	0.611		
Observations	381	381	499	499		
Panel D: Not Listed in Baseline						
Single Click	-0.052	0.030	-0.006	0.129		
	(0.097)	(0.091)	(0.113)	(0.081)		
Double Click	-0.060	0.049	-0.147	0.036		
	(0.130)	(0.126)	(0.142)	(0.125)		
Truth	0.246**	-0.116*	-0.111*	0.066*		
	(0.098)	(0.059)	(0.057)	(0.037)		
Outcome at Baseline	0.257***	0.322***	0.406***	0.432***		
	(0.058)	(0.085)	(0.069)	(0.070)		
Mean of Outcome	0.336	0.706	0.672	0.412		
Observations	143	143	177	177		

Note: This table shows how clicking a school in the explorer is associated with changes in perceptions of price (Columns 1-2), quality (Columns 3-4), and probability of being admitted (Columns 5-6). Outcomes are obtained from the midline and endline surveys. Endline responses are used in cases where both are available. 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. Robust standard errors are reported in parentheses. The sample is restricted to parent-school observations in the control group who opened the school explorer platform. Because selection into single vs. double clicking is not random, these relationships should be interpreted as descriptive evidence consistent with learning rather than as causal effects. Panel A consists of all schools, Panel B only consists of schools that the parent ranked first at baseline, Panel C consists of schools that the parent included in the application at baseline, and Panel D consists of schools that the parent did not include in the application at baseline.

TABLE A.XIV

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).

TABLE A.XV

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**	0.027**
	(0.010)	(0.020)	(0.012)
Mean of Outcome	0.111	0.110	0.111
Observations	4221	1070	3133

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.

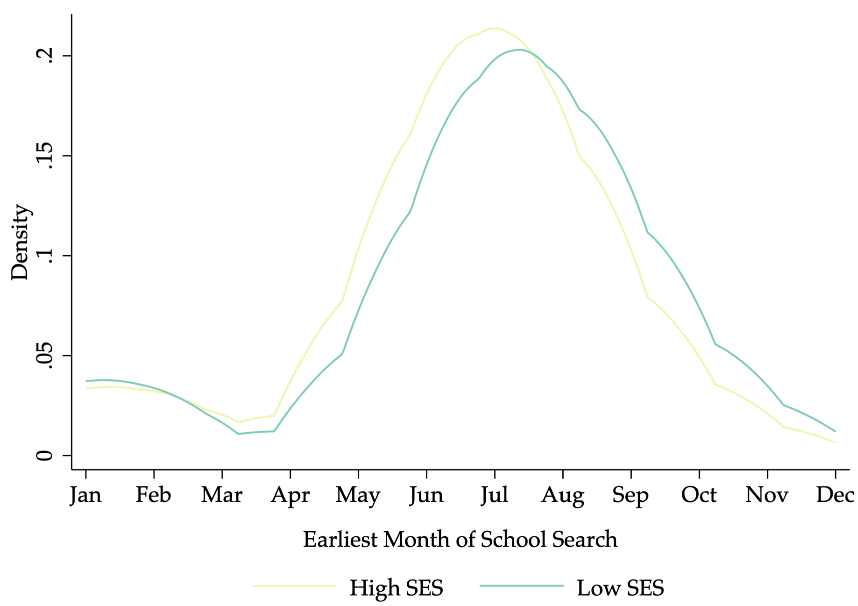


FIGURE A.9.—This figure shows the distribution of when parents plan to start searching for schools, by SES status. Information on search plans comes from the pre-baseline survey (N = 3,215). SES status is proxied by whether the mother completed college.

A.5. ADDITIONAL EXPERIMENTAL RESULTS

In this section, we present additional results from the two experiments. Table A.XVI shows that the sample is well-balanced. We examine the treatment effects of the search aid interventions on school perceptions in Table A.XVII and on application outcomes in Table A.XVIII. We observe that the second treatment arm increases the likelihood that parents have correct perceptions about the quality of the school. The second search intervention also 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. In Table A.XIX, we study heterogeneous treatment effects by baseline beliefs. Table A.XX presents the balance check for the feedback intervention and A.XXI reports heterogeneous treatment effects for the feedback intervention.

TABLE A.XVI

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 2km	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
Number of available high quality schools	9.769	[5.316]	-0.133	(0.176)	0.044	(0.175)	3948
Number of available low price schools	14.761	[8.858]	-0.126	(0.273)	-0.100	(0.269)	3948
Panel B: Parent/Child Characteristics							
Child is female	0.495	[0.500]	0.016	(0.019)	0.023	(0.019)	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
Child's age	3.904	[0.550]	-0.012	(0.021)	-0.014	(0.021)	3948
Panel C: Initial Knowledge 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	0.904	[0.295]	-0.005	(0.012)	0.007	(0.011)	3948
First-preference school is highlight-worthy	0.633	[0.482]	0.040**	(0.020)	0.051**	(0.020)	3149
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
Perceived number of available high quality schools	5.011	[4.654]	0.062	(0.175)	-0.239	(0.168)	3948
Perceived number of available low price schools	5.838	[5.051]	0.058	(0.199)	-0.206	(0.204)	3948
Parent believes to be SEP eligible	0.172	[0.378]	-0.000	(0.015)	-0.008	(0.014)	3948
Parent is unsure 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.XVII

EFFECTS OF SEARCH AID INTERVENTIONS ON SCHOOL PERCEPTIONS

	Price		Quality		Pr Admission Chance	
	Absolute Value	Correct	Absolute Value	Correct	Absolute Value	Correct
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pooled						
Treatment 1	0.026	-0.012	0.020	-0.010	-2.415	0.031**
	(0.027)	(0.023)	(0.025)	(0.019)	(1.578)	(0.015)
Treatment 2	0.011	0.001	0.017	0.004	-0.158	0.020
	(0.027)	(0.023)	(0.025)	(0.019)	(1.668)	(0.015)
Control Group Mean	0.372	0.668	0.527	0.538	16.494	0.155
Observations	4168	4168	5171	5171	4414	4414
Panel B: Heterogeneity by Parental Education						
Treatment 1 × High SES	0.054	-0.029	-0.022	0.058	-4.365	0.049*
	(0.050)	(0.043)	(0.051)	(0.039)	(3.185)	(0.030)
Treatment 1 × Low SES	0.024	-0.011	0.035	-0.028	-1.821	0.025
	(0.032)	(0.027)	(0.029)	(0.022)	(1.811)	(0.017)
Treatment 2 × High SES	0.073	-0.041	-0.098*	0.118***	-1.633	0.011
	(0.054)	(0.046)	(0.051)	(0.041)	(3.518)	(0.029)
Treatment 2 × Low SES	-0.007	0.013	0.047	-0.027	0.251	0.022
	(0.031)	(0.026)	(0.029)	(0.022)	(1.900)	(0.017)
p-value: Treat 1 x High SES = Treat 1 x Low SES	0.607	0.726	0.343	0.055	0.486	0.483
p-value: Treat 2 x High SES = Treat 2 x Low SES	0.203	0.308	0.014	0.002	0.638	0.729
Control Group Mean (High SES)	0.293	0.732	0.495	0.550	18.022	0.146
Control Group Mean (Low SES)	0.394	0.651	0.536	0.535	16.070	0.157
Observations 1 (High SES)	956	956	1204	1204	990	990
Observations 2 (Low SES)	3210	3210	3965	3965	3421	3421

Note: This table presents the results of the search interventions on school perceptions at the school-parent level. 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. Outcomes are obtained from the midline and endline surveys. Endline responses are used in cases where both are available. In Panel A, we regress each outcome on indicator variables for both treatment arms, and stratification dummies. In Panel B, we further include a dummy for SES status and the fully interacted effects of treatments and SES status. SES status is proxied by whether the mother completed college. Standard errors are clustered at the parent level. The sample is restricted to parents who opened the school explorer platform.

TABLE A.XVIII

TREATMENT EFFECTS OF SEARCH INTERVENTION ON APPLICATION OUTCOMES

			Rank 1				Rank 2					
	Submitted Applica-tion	App. 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	0.007	-0.050	0.001	0.010	-0.035	-0.029	0.016	0.012	-0.374	-0.054*	-0.012	0.015
	(0.013)	(0.075)	(0.021)	(0.018)	(0.221)	(0.026)	(0.022)	(0.020)	(0.357)	(0.032)	(0.011)	(0.020)
Treatment 2	0.030**	0.013	0.025	0.005	-0.156	-0.004	0.054**	0.016	0.174	-0.047	-0.002	0.035*
	(0.013)	(0.074)	(0.021)	(0.018)	(0.212)	(0.025)	(0.022)	(0.019)	(0.398)	(0.032)	(0.011)	(0.019)
Control Group Mean	0.893	3.444	0.691	0.246	2.006	0.670	0.679	0.168	2.972	0.485	0.069	0.765
Observations	3111	2818	2770	2703	2818	2017	2580	2492	2633	1491	2817	2748
Panel B: Heterogeneity by SES Status												
Treatment 1 × High SES	0.014	0.030	-0.017	0.050	0.214	-0.074	0.091*	0.027	0.285	-0.050	-0.009	-0.028
	(0.025)	(0.162)	(0.047)	(0.036)	(0.464)	(0.053)	(0.048)	(0.037)	(0.735)	(0.063)	(0.027)	(0.041)
Treatment 1 × Low SES	0.005	-0.086	0.009	-0.005	-0.124	-0.015	-0.006	0.004	-0.586	-0.054	-0.014	0.031
	(0.016)	(0.084)	(0.024)	(0.020)	(0.251)	(0.030)	(0.025)	(0.023)	(0.406)	(0.037)	(0.012)	(0.023)
Treatment 2 × High SES	0.009	0.051	0.068	0.045	-0.008	-0.030	0.183***	0.034	0.622	-0.006	-0.009	0.012
	(0.026)	(0.159)	(0.047)	(0.038)	(0.461)	(0.053)	(0.047)	(0.039)	(0.763)	(0.065)	(0.028)	(0.041)
Treatment 2 × Low SES	0.037**	0.002	0.009	-0.008	-0.205	0.004	0.014	0.010	0.037	-0.056	-0.000	0.042*
	(0.015)	(0.083)	(0.023)	(0.020)	(0.242)	(0.029)	(0.024)	(0.022)	(0.474)	(0.036)	(0.012)	(0.022)
p-value: Treat 1 x High SES = Treat 1 x Low SES	0.776	0.526	0.618	0.179	0.522	0.331	0.074	0.611	0.299	0.957	0.889	0.212
p-value: Treat 2 x High SES = Treat 2 x Low SES	0.350	0.785	0.258	0.227	0.708	0.568	0.001	0.599	0.522	0.500	0.786	0.520
Control Mean (High SES)	0.909	3.695	0.583	0.262	2.081	0.691	0.552	0.229	2.766	0.504	0.091	0.762
Control Mean (Low SES)	0.888	3.368	0.726	0.242	1.986	0.663	0.719	0.150	3.036	0.478	0.062	0.766
Observations 1 (High SES)	732	670	660	639	670	471	609	585	620	361	669	643
Observations 2 (Low SES)	2376	2146	2108	2062	2146	1544	1969	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, and stratification dummies. In Panel B, we further include a dummy for SES status and 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.

TABLE A.XIX

TREATMENT EFFECTS OF SEARCH INTERVENTION BY BASELINE BELIEFS

Outcome:	Enrolled School Is Highlight-worthy						
Heterogeneity Variable:	Number of Schools	Number of Highlight-worthy Schools	Number of High Quality Schools	Number of Low Price Schools	1(1st Choice Is Highlight-worthy) × Admission Chance for 1st Choice	1(1st Choice Is Highlight-worthy)	Admission Chance for 1st Choice
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment 1	0.118**	0.120**	0.102*	0.137***	-0.047	0.009	-0.073
	(0.050)	(0.053)	(0.054)	(0.049)	(0.052)	(0.059)	(0.079)
Treatment 1 × Bias (Belief - Truth)	-0.004	-0.019**	-0.008	-0.011**	0.134*	0.096	0.028
	(0.004)	(0.009)	(0.006)	(0.005)	(0.079)	(0.070)	(0.096)
Treatment 1 × Truth	-0.009**	-0.023**	-0.013*	-0.015***	0.073	-0.013	0.146
	(0.004)	(0.009)	(0.007)	(0.005)	(0.076)	(0.070)	(0.110)
Treatment 2	0.113**	0.137***	0.091*	0.147***	0.074	0.090	-0.026
	(0.048)	(0.052)	(0.053)	(0.048)	(0.054)	(0.060)	(0.081)
Treatment 2 × Bias (Belief - Truth)	0.002	-0.014	-0.000	-0.004	0.064	-0.024	0.217**
	(0.003)	(0.009)	(0.006)	(0.005)	(0.078)	(0.067)	(0.098)
Treatment 2 × Truth	-0.003	-0.019**	-0.005	-0.009**	-0.111	-0.069	0.079
	(0.004)	(0.009)	(0.006)	(0.005)	(0.076)	(0.070)	(0.112)
Bias (Belief - Truth)	0.003	0.023***	0.005	0.012***	0.058	0.050	-0.037
	(0.003)	(0.007)	(0.005)	(0.004)	(0.055)	(0.047)	(0.068)
Truth	0.005*	0.034***	0.010*	0.016***	0.600***	0.595***	0.034
	(0.003)	(0.007)	(0.005)	(0.004)	(0.055)	(0.049)	(0.078)
Control Group Mean	0.669	0.669	0.669	0.669	0.669	0.667	0.679
Observations	2,454	2,454	2,454	2,454	1,394	1,415	2,050

Note: This table presents the results of the search interventions on whether the child is enrolled in a highlight-worthy school by baseline beliefs. We regress each outcome on indicator variables for both treatment arms, stratification dummies, the bias measure and the true value of the heterogeneity variable, interactions between each treatment arm and the bias measure, and interactions between each treatment arm and the true value. The bias measure is defined as beliefs minus the true value. The sample is restricted to parents who opened the school explorer platform.

TABLE A.XX

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 2km	15.778	[9.398]	-0.740	(0.860)	2581
Number of available highlight-worthy schools	8.355	[5.020]	0.187	(0.525)	2581
Number of available high quality schools	9.440	[5.346]	0.257	(0.457)	2581
Number of available low price schools	14.345	[8.987]	-0.736	(0.913)	2581
Panel B: Parent/Child Characteristics					
Child is female	0.516	[0.500]	-0.011	(0.025)	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
Child's age	3.881	[0.500]	0.017	(0.026)	2581
Panel C: Initial Knowledge and Beliefs					
Expected satisfaction with process	5.273	[1.391]	-0.019	(0.064)	2435
Listed any school as first preference	0.929	[0.256]	-0.001	(0.019)	2581
First-preference school is highlight-worthy	0.673	[0.469]	0.038	(0.035)	2103
Perceived admission chance 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
Perceived number of available high quality schools	4.795	[4.046]	0.342*	(0.188)	2581
Perceived number of available low price schools	5.649	[4.831]	0.288	(0.282)	2581
Parent believes to be SEP eligible	0.165	[0.371]	-0.014	(0.015)	2581
Parent is unsure 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.XXI

TREATMENT EFFECTS OF FEEDBACK INTERVENTION BY BASELINE BELIEFS

Outcome:	Enrolled School Is Highlight-worthy	
Heterogeneity Variable:	1(1st Choice is Highlightworthy) × Admission Chance for 1st Choice	
	Full Sample	Control Group in Search Aid Interventions
	(1)	(2)
Open Feedback	-0.057	-0.297*
	(0.087)	(0.157)
Open Feedback × Bias (Belief - Truth)	0.017	0.463**
	(0.115)	(0.229)
Open Feedback × Truth	0.092	0.414*
	(0.119)	(0.229)
Bias (Belief - Truth)	0.145***	-0.018
	(0.047)	(0.101)
Truth	0.537***	0.462***
	(0.059)	(0.108)
Control Group Mean	0.713	0.667
Observations	1031	317

Note: This table reports the effects of the feedback intervention on whether the child is enrolled in a highlight-worthy school by baseline beliefs. The heterogeneity variable is defined as an indicator for whether the first-ranked school is highlight-worthy multiplied by the admission chance of getting into the first-ranked school. We regress each outcome on an indicator for whether the respondent opened the feedback information, the bias measure, and the true value of the heterogeneity variable, as well as interactions between the 'open feedback' indicator and the bias measure and between the 'open feedback' indicator and the true value. The 'open feedback' indicator and its corresponding interaction terms are instrumented using assignment to the feedback treatment group and interactions between treatment assignment, the bias measure, and the true value. All regressions include market fixed effects and application risk group fixed effects. Column 1 reports results for the full sample and Column 2 restricts the sample to parents in the control group of the search aid interventions.

A.6. MODEL APPENDIX

This section provides further details on our structural model.

Elements of w: We include indicators for search treatments 1 and 2, an interaction between search treatment 2 and highlight-worthiness, an indicator for opening the feedback treatment, an interaction between opening feedback and period 2 information to capture selection into opening the feedback, indicators for feedback on specific schools j, and indicators for a recommendation for school j.

A.6.1. Beliefs about x and the number of schools

Overview: Beliefs about the number of relevant schools and their observables are given by a mixture model with three latent types, $h \in \{1,2,3\}$. Types are persistent over time, and are unobserved by the econometrician. If parent i is of type h, it believes the number of unknown elements of J_i is distributed Poisson(λ_{iht}), and has Dirichlet-Multinomial beliefs over the probabilities of the sixteen (price,quality) cells $\{1,2,3,4\}^2$ with parameter vector φ_{iht}.

The subjective-belief parameters λ_{ith} and φ_{ith} in turn depend on the true values, on the number of schools that i knows at least by name at time t, and on i's perceptions of their characteristics. Our specification of φ nests a form of Bayesian updating, but we also allow parents to under-react to information provided by treatments or search.

It is natural to model parents' beliefs as Dirichlet when the data available to them consists of counts of schools within cells, as provided by our treatments and their search histories, since the Dirichlet is the conjugate prior for multinomial data.

Details: distribution of x: Prices and quality scores are divided into 16 cells. Parent i of type h has Dirichlet beliefs at time t over the probabilities of each cell, with parameter vector φ_{iht}.

Let x^k be the kth (price,quality) cell. The kth element of φ_{iht} is given by:

$$\varphi_{ihtk} = \overline{\varphi}_{h0k} + b^{\varphi}_{1iht} * |\{j \in J_i : x_{ij} = x_k\}| + b^{\varphi}_{2iht} * |\{j \in J_i : \pi_{ijt} > 0 \text{ and } \hat{x}_{ij}^{\pi_{ijt}} = x_k\}|. \tag{14}$$

That is, $\overline{\varphi}$ represents the initial prior of type h before it draws a choice set. The terms b^{φ}_{1iht} and b^{φ}_{2iht} capture updating in response to the true state and the signals that i has received so far, respectively.

We allow information about the distribution of nearby schools' characteristics (treatment 1) and information about specific schools (treatment 2) to affect beliefs. In particular, both coefficients vary with i's search treatment, provided that i has received the treatment by time t. We have $b^{\varphi}_{1iht} = \tilde{b}^{\varphi}_{1h0} + \tilde{b}^{\varphi}_{1h1} 1(\text{treatment}_i = 1)1(t \geq 1) + \tilde{b}^{\varphi}_{1h2} 1(\text{treatment}_i = 2)1(t \geq 1)$, where treatment$_i$ denotes i's search treatment assignment. Similarly, we have $b^{\varphi}_{2iht} = \tilde{b}^{\varphi}_{2h0} + \tilde{b}^{\varphi}_{2h1} 1(\text{treatment}_i = 1)1(t \geq 1) + \tilde{b}^{\varphi}_{2h2} 1(\text{treatment}_i = 2)1(t \geq 1)$.

Our model nests Bayesian updating in response to the information that the parent has gained. In particular, supposing the agent believes its signals $\hat{x}$ are equal to the true x, the Bayesian updating rule sets $b^{\varphi}_{2iht} = 1$ for all h, t. We allow under- or over-reaction relative to this benchmark.[51]

Details: beliefs about the number of schools: Let N^{unknown}_{it} denote the number of not-yet-discovered schools at time t, i.e. $|\{j \in J_i : \pi_{ijt} = 0\}|$. If i is of type h, as defined above, it believes $N^{\text{unknown}}_{it} \sim \text{Poisson}(\lambda_{iht})$, where

$$\lambda_{iht} = b^{\lambda}_{0h} + (b^{\lambda}_{1h} + b^{\lambda}_{2h} * \text{treat}_{it}) * |\{j \in J_i : \pi_{ijt} = 0\}|. \tag{15}$$

This expression allows beliefs about the number of schools to respond to the truth and to partially update in response to our "search" treatments.

A.6.2. Beliefs about match quality

Parents in our model use Equation (11) 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

[51] The Dirichlet is the conjugate prior for the multinomial distribution. The set of schools i knows, $\{j \in J_i : \pi_{ijt} > 0\}$, is a multinomial draw over $\hat{x}^{(\pi_{ijt})}_{ij}$ of size $|\{j \in J_i : \pi_{ijt} > 0\}|$. If prior to observing this set the parent's parameters are φ, then the Bayesian posterior is $\varphi' + |\{j \in J_i : \pi_{ijt} > 0 \text{ and } \hat{x}^{(\pi_{ijt})}_{ij} = x_k, \ k = 1, \ldots, 16\}|$.

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 parent's posterior given its "low-information" $(\pi_{ijt} = 1)$ signal is:

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

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 $\pi_{ijt} = 1$, parent 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, parents 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 obtain the implied joint distribution of the subjective expectation and the true match value, we substitute the objective joint distribution into Equation (16), obtaining:

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

For convenience, we let μ^ℓ and Σ^ε denote the terms above, which are functions of the primitives.

$$\begin{pmatrix} \hat{\varepsilon}_{ij}^{(1)} \\ \varepsilon_{ij} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu^\ell \\ 0 \end{pmatrix}, \Sigma^\varepsilon \right). \tag{17}$$

A.7. ESTIMATION APPENDIX

This section provides additional details on estimation. We begin with our models of beliefs over unobservables (ε) and observables (x). We then discuss step 3 estimation and inference. Finally, we provide details on the Gibbs Sampler used in step 1.

A.7.1. 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 (Eq. (17)) estimated in step 1. We have

$$\sigma_\varepsilon^2 = \Sigma^\varepsilon_{[2,2]}, \;\; \tilde{\rho} = \frac{\Sigma^\varepsilon_{[1,2]}}{\Sigma^\varepsilon_{[2,2]}}, \;\; \tilde{\mu} = \mu^\ell / (1 - \tilde{\rho}), \;\; \text{and } \sigma_e^2 = \frac{\Sigma^\varepsilon_{[1,1]} - \tilde{\rho}^2 \sigma_\varepsilon^2}{\tilde{\rho}^2}. \tag{18}$$

To separately identify the subjective utility-shock variance $\tilde{\sigma}_\varepsilon^2$ and subjective measurement error variance $\tilde{\sigma}_e^2$ 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 parent's current state and utility draws; we assume that the parent is reporting the subjective probability, denoted $\hat{Pr}(u > u_{i21})$, 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 parents overestimate the variance of the shocks, they will tend to overestimate this probability when u_{i21} is high.

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

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

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 $\Sigma^{\delta\eta}_{[1,1]}$ come from the distribution of mean utilities δ. The remaining terms come from substituting x, d, and the parameters of the subjective distribution of ε into the subjective expected utility under high information ($\pi = 2$), i.e., equation (6) evaluated at perceived inputs with $\pi = 2$.

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 $\nu^{\text{survey},\varepsilon,m} \sim N(0, \sigma^2_{\nu,\varepsilon})$, iid across people and questions. Parents report

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

To estimate, we first obtain estimates of $(\tilde{\mu}, \tilde{\rho}, \sigma^2_\varepsilon, \sigma^2_e)$ by substituting our first-step estimates $\hat{\Sigma}^\varepsilon$ and $\hat{\mu}^\ell$ into Equation (18). Let $\hat{\tilde{\mu}}, \hat{\tilde{\rho}}, \hat{\tilde{\sigma}}^2_\varepsilon, \hat{\tilde{\sigma}}^2_e$ 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 $\hat{\tilde{\rho}} = \frac{\hat{\tilde{\sigma}}^2_\varepsilon}{\hat{\tilde{\sigma}}^2_\varepsilon + \hat{\tilde{\sigma}}^2_e}$.

Identification and Estimation of Beliefs about x: In modeling search decisions, the relevant beliefs are over characteristics of unknown schools within five kilometers. However, our survey questions elicit beliefs over the characteristics of all schools, known and unknown, within two kilometers.

To relate the survey to the model, we assume that parents may probabilistically recall known schools—in which case they think of their subjective perceptions of these schools' characteristics—in addition to unknown schools. In particular, agents report the sum of two sets of schools: a set of probabilistically-recalled known schools, and a sample drawn according their beliefs over unknown schools.

Probabilistic recall captures both survey measurement error and parents' decisions not to report schools in J_i that they believe are outside of the 2km radius. We capture survey measurement error via imperfect recall of "known" schools, and by letting parents stochastically draw a set of unknown schools given their beliefs about the number of such schools and their characteristics.

Survey s elicits the perceived number of schools by price-quality cell, $N^{\text{survey}}_{is} \in \mathbb{N}^K$. We collect this data at baseline and midline.[52] Let $t_i(s)$ denote the period in which parent i takes survey s.[53] Let $N_{is} \in \mathbb{N}^K$ denote the number of "known"

[52]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.

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

$(\pi_{ijt_i(s)} > 0)$ schools with perceived characteristics $\hat{x}_{ij}^{\pi_{ijt_i(s)}}$ equal to the value of the kth cell.

We assume that parents 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 parent i reports $N^{\text{survey,unknown}} \sim \text{Dirichlet-Multinomial}(N^{\text{survey,unknown,total}}, \alpha_{it_i(s)})$.

On survey s, we observe parent 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.7.2. Step 3: Estimating Search Costs

The value of a detail view at school j in information state $\pi_{i\cdot 1}$ is given by

$$V_{ij}^{\text{detail}}(\pi_{i\cdot 1}) = \hat{E}(\hat{U}_i^*(\pi_{i\cdot 1} + e_j\alpha^{\text{detail}}) - \hat{U}_i^*(\pi_{i\cdot 1})),$$

where e_j is the vector with a 1 in the jth place, 0 elsewhere. That is, the value is the expected gain when π_{ij1} increases by an amount equal to the coefficient on "detail view" in Equation (8).

The value of a pin click alone is given by:

$$V_i^{\text{pin}}(\pi_{i\cdot 1}) = \hat{E}(\hat{U}_i^*(\pi'_{i\cdot 1}) - \hat{U}_i^*(\pi_{i\cdot 1}))$$

where the expectation is over the clicked school j, with $\pi'_{i\cdot 1} = \pi_{i\cdot 1} + e_j\alpha^{\text{pin}}$ where α^{pin} is the coefficient on "pin click" in Equation (8).

The expected value of a pin click and the option to conduct a detail view is given by

$$V_i(\pi_{i\cdot 1}) = V_i^{\text{pin}}(\pi_{i\cdot 1}) + E\max\{V_{ij}^{\text{detail}}(\pi'_{i\cdot 1}) - \bar{c}^{\text{detail}} + \sigma^{\text{detail}}\varepsilon_{i1s}^{\text{detail}}, \sigma^{\text{detail}}\varepsilon_{i0s}^{\text{detail}}\},$$

where the expectation is over the state π' that will be attained as a result of the pin click.

We first estimate the detail-view cost parameters via maximum likelihood of the observed detail-click decisions. Next, we calculate $V_i(\cdot)$ for all i over a grid of draws of i's latent variables at the point estimates, via simulation. Finally, we estimate the parameters of the pin-click cost via maximum likelihood of the sequence of pin-click decisions.

A.7.3. Inference

We conduct inference on parameters and counterfactuals by bootstrap, simulating 100 datasets resampling with replacement and re-running estimation steps 1-3 and counterfactual simulations.

A.7.4. Step 1: Estimation Details

Let $\nu_{ijs}^{\text{survey}}$ denote survey measurement error in π_{ijt}^* on survey s. The parent, 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, let e_{ijt}^{survey} denote measurement error on payoffs reported at $t < T$. Parents 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 parent 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.[54] 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.

[54]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.

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^d, β_i^x) given $\hat{u}^1, \hat{u}^2, x, \hat{x}$
5. Update information-coefficient vectors α_i given π, x, d, w
6. Update $\Sigma^{\beta} | \{(\beta_i^d, \beta_i^x)\}_i$.
7. Update $\Sigma^{\alpha} | \{\alpha_i\}_i$
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 $\Sigma^{\varepsilon} | \hat{u}^1, \hat{u}^2, \beta_i, \hat{x}, \delta$
14. Update $\Sigma^{\pi} | \pi_1^*, \ldots, \pi_T^*, \overline{\alpha}, \eta, \{\alpha_i\}_i$
15. Update mean information coefficients $\overline{\alpha} | \pi^*, \{\alpha_i\}_i, \eta$.
16. Update linear utility terms $\mu^l | \hat{u}^1, d, x, \hat{x}^1, \beta_i, \delta$.
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 $e_{ij3} = 0$ for all i, j.[55] Optimality of the observed rank-order lists require information, utilities, and preference measurement error to satisfy the following constraints:

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

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 the upper bound $\hat{u}_{ij'}^{\pi_{ij'}} + e_{ij't}$ is replaced by ∞.
   (b) If the ROL is of length r, so that j is the final option, then the lower bound $\hat{u}_{ij''}^{\pi_{ij''}} + e_{ij''t}$ is 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

$$\pi^{\text{survey}}_{ijs} = 1(\pi^*_{ijt_i(s)} + \nu_{ijs} > 0) + 1(\pi^*_{ijt_i(s)} + \nu_{ijs} > 1), \tag{19}$$

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

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 $\pi^*_{ijt} \geq 0$.
- If there is a measurement $\pi^{\text{survey}}_{ijs}$ then $\pi^*_{ijt_i(s)} + \nu_{ijs}$ must satisfy equation 19, 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 $\text{Gamma}(1,1)$, 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.8. ADDITIONAL MODEL ESTIMATES

In this section, we present additional results from the model estimates and counterfactuals. Table A.XXII 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.XXIII 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 beliefs over the number and characteristics of unknown schools. 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.10 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 III, there is a clear correlation between schools that parents prefer more (higher mean utility) and the ones that they are more likely to know (higher discoverability).

Figure A.11 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 parents, and panels (C) and (D) show the distortion functions for high-SES parents.

Figure A.12a shows the distribution of the deterministic component of the single click cost ($x_i^c \gamma^{\text{cost}}$), and Figure A.12b shows the distribution of the ratio of the value of the first search over the cost of the first search.

Figure A.13 shows the model fit on a series of expected school characteristics, behavior, and latent variables.

Tables A.XXIV present the simulated counterfactuals separately for high- and low-SES parents, and by pessimism about the number of highlight-worthy schools.

Finally, Table S.VII presents a sequence of counterfactuals starting from the *baseline* and *fix information* scenarios in which we gradually reduce the search costs up to 5% of the original value.

TABLE A.XXII

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	0.199	(0.028)	0.831	(0.148)
	2	0.044	(0.017)	0.185	(0.078)
2	2	0.655	(0.044)	1.428	(0.217)
Panel C: Variance Covariance of shocks v_{ijt} over time (Σ^v)					
1	1	0.360	(0.009)	0.208	(0.011)
	2	-0.206	(0.011)	-0.168	(0.012)
	3	-0.090	(0.012)	-0.122	(0.013)
2	2	0.383	(0.017)	0.367	(0.028)
	3	0.311	(0.015)	0.327	(0.024)
3	3	0.563	(0.019)	0.598	(0.033)
Panel D: Random Coefficients of time effects ($\Sigma rc\pi$)					
1	1	0.473	(0.016)	0.505	(0.030)
	2	0.074	(0.030)	-0.188	(0.047)
	3	-0.035	(0.009)	0.022	(0.015)
2	2	2.075	(0.120)	2.841	(0.250)
	3	-0.134	(0.024)	-0.233	(0.051)
3	3	0.015	(0.004)	0.028	(0.008)
Panel F: Knowledge Shifters (αzw)					
At least t2	α_w	1.062	(0.036)	1.065	(0.043)
At least t1	α_w	-0.209	(0.013)	-0.152	(0.031)
Distance	α_z	-0.148	(0.008)	-0.115	(0.006)
Treatment 1	α_w	0.125	(0.030)	0.157	(0.070)
Treatment 2	α_w	0.304	(0.060)	-0.060	(0.106)
Highlight-worthy	α_w	0.331	(0.075)	0.281	(0.068)
Highlighted	α_w	-0.275	(0.037)	-0.081	(0.054)
Single Click	α_w	0.896	(0.030)	0.456	(0.041)
Double Click	α_w	1.384	(0.043)	0.760	(0.063)
Feedback	α_w	-0.643	(0.073)	-0.613	(0.126)
Feedback Pre	α_w	0.011	(0.049)	0.028	(0.029)
Feedback Known	α_w	1.050	(0.039)	0.981	(0.142)
Feedback Rec	α_w	-0.160	(0.150)	0.303	(0.233)
Panel G: Measurement Error					
Base Survey Utility	σ^2_ϵ	0.359	(0.028)	0.059	(0.008)
Survey Awareness	$\sigma^2_{\eta_s}$	0.010	(0.001)	0.014	(0.002)
Pr(misreport x's)	p^s	0.279	(0.012)	0.293	(0.014)

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	0.074	(0.017)	0.157	(0.048)
	2	0.340	(0.021)	0.411	(0.112)
	3	0.495	(0.026)	0.366	(0.109)
	4	0.090	(0.009)	0.066	(0.025)
2	1	0.009	(0.005)	0.043	(0.015)
	2	0.277	(0.024)	0.441	(0.033)
	3	0.586	(0.026)	0.470	(0.029)
	4	0.127	(0.006)	0.046	(0.009)
3	1	0.001	(0.001)	0.010	(0.006)
	2	0.094	(0.008)	0.194	(0.015)
	3	0.597	(0.011)	0.600	(0.021)
	4	0.308	(0.007)	0.196	(0.013)
4	1	0.002	(0.002)	0.008	(0.005)
	2	0.039	(0.010)	0.080	(0.026)
	3	0.426	(0.016)	0.439	(0.027)
	4	0.533	(0.012)	0.473	(0.030)
Panel E: Distortion Function - Price (Γ price)					
1	1	0.608	(0.008)	0.673	(0.018)
	2	0.382	(0.008)	0.311	(0.018)
	3	0.010	(0.001)	0.015	(0.004)
	4	0.000	(0.000)	0.001	(0.000)
2	1	0.095	(0.011)	0.112	(0.030)
	2	0.774	(0.015)	0.679	(0.043)
	3	0.130	(0.012)	0.200	(0.031)
	4	0.001	(0.001)	0.009	(0.009)
3	1	0.054	(0.012)	0.027	(0.016)
	2	0.475	(0.015)	0.374	(0.040)
	3	0.461	(0.012)	0.574	(0.038)
	4	0.010	(0.003)	0.026	(0.013)
4	1	0.098	(0.030)	0.047	(0.034)
	2	0.317	(0.046)	0.065	(0.067)
	3	0.467	(0.045)	0.571	(0.074)
	4	0.117	(0.025)	0.317	(0.063)

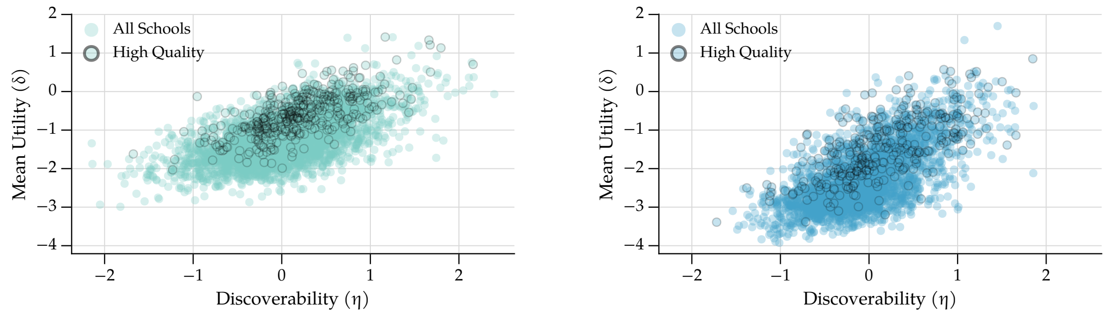
Note: This table presents additional step 1 parameter estimates by SES group. Panel A: variance-covariance of subjective and true match value shocks (Σe). Panel B: distortion function for quality (Γ quality), mapping true quality categories to subjective perceived quality. Panel C: variance-covariance of knowledge shocks v_{ijt} over time (Σ^v). Panel D: random coefficients on time effects ($\Sigma rc\pi$). Panel E: distortion function for price (Γ price). Panel F: full set of knowledge shifters (α_{zw}). Panel G: measurement error parameters. Standard errors are bootstrapped.

TABLE A.XXIII

ADDITIONAL PARAMETERS: BELIEFS AND SEARCH COSTS

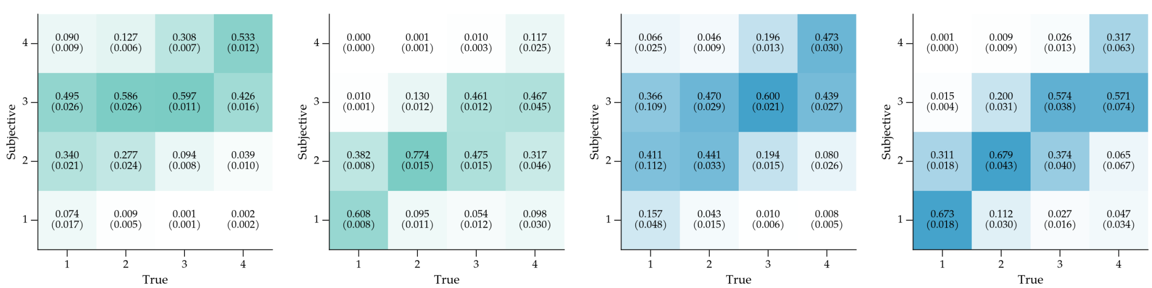
		Low SES		High SES	
	Parameter	Coeff	Std Err.	Coeff	Std Err.
Mean Prob of each type					
Λ Type 1		0.557	(0.007)	0.555	(0.035)
Λ Type 2		0.288	(0.012)	0.267	(0.023)
Λ Type 3		0.154	(0.010)	0.177	(0.024)
Search Technology (γ click)					
Distance	1	-1.294	(0.040)	-1.072	(0.039)
Price = 1	2	-0.175	(0.070)	-0.215	(0.085)
Price = 2	3	-0.022	(0.085)	-0.048	(0.085)
Price = 3	4	0.162	(0.085)	0.193	(0.089)
Price = 4	5	0.035	(0.119)	0.069	(0.103)
Quality = 1	6	-0.193	(0.074)	-0.212	(0.081)
Quality = 2	7	-0.147	(0.070)	-0.066	(0.058)
Quality = 3	8	0.027	(0.073)	-0.023	(0.054)
Quality = 4	9	0.312	(0.064)	0.301	(0.064)
Highlightworthy	10	-0.048	(0.145)	0.089	(0.110)
Highlightworthy × Treat 2	11	0.340	(0.054)	0.200	(0.110)
Double Click (θ cost)					
Mean Cost		2.248	(0.501)	0.650	(0.084)
Log-Variance		0.805	(0.553)	-0.607	(0.172)
Match Value Shocks ε_{ij} Primitives					
Unobserved match-value shock	$\tilde{\mu}$	-0.440	(0.034)	-0.668	(0.087)
	$\tilde{\sigma}_\varepsilon$	3.699	(0.614)	2.322	(0.412)
	σ_ε	0.809	(0.032)	1.195	(0.077)

Note: This table presents step 2 parameter estimates by SES group. Standard errors are bootstrapped. The first section reports mean type probabilities for the latent types governing $\bar{\lambda}_{0h}$ (the Poisson parameter for beliefs about unknown schools). The search technology section (γ click) reports logistic coefficients that map school characteristics (distance, price, quality, highlightworthiness, and their interaction with the search treatment) to pin-click signal quality. The double-click section (θ cost) reports the mean and log-variance of profile-click costs. The final section reports match value shock primitives: the subjective prior mean ($\tilde{\mu}$), subjective prior standard deviation ($\tilde{\sigma}_\varepsilon$), and true standard deviation (σ_ε) of unobserved match values.



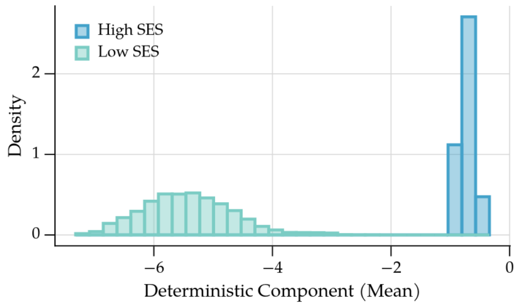
(a) Low SES **(b)** High SES

FIGURE A.10.—Estimates of school unobservables. Each dot represents one school (trimmed to the 1st–99th percentile on each axis for readability). The horizontal axis shows estimated discoverability (η) and the vertical axis shows estimated mean utility (δ). Circled dots mark high quality schools (quality = 4). A positive relationship indicates that schools with higher mean utility are more likely to be known at baseline.

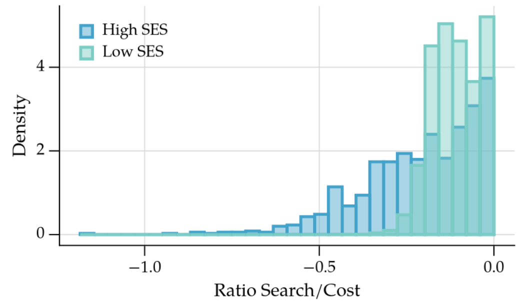


(a) Quality Low SES **(b)** Price Low SES **(c)** Quality High SES **(d)** Price High SES

FIGURE A.11.—Distortion Functions. Each panel shows a 4×4 estimated distortion function. Rows indicate the subjective (perceived) category and columns indicate the true category (1–4 scale). Cell entries report the conditional probability that a parent perceives the school in the row category given that the true category is the column value, with analytical standard errors in parentheses. Panels (a) and (c) show quality distortion functions and panels (b) and (d) show price distortion functions. Panels (a)–(b) correspond to low-SES parents and panels (c)–(d) to high-SES parents.

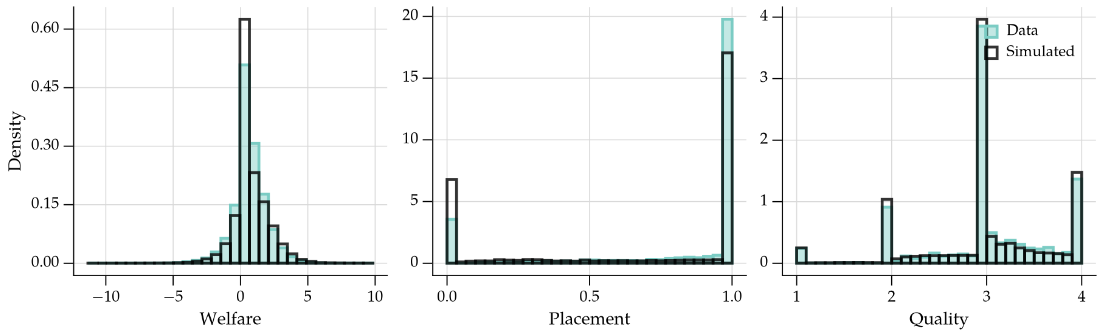


(a) Distribution of $x_i^c \gamma^{\text{cost}}$

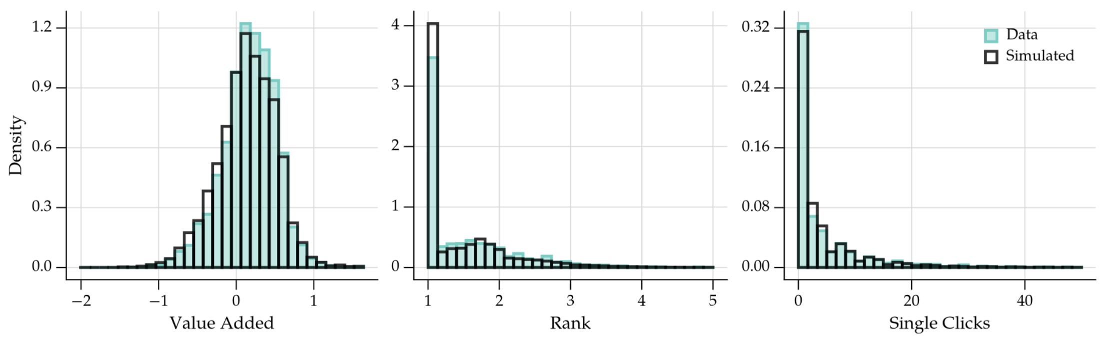


(b) Ratio of Value 1st Search over Cost

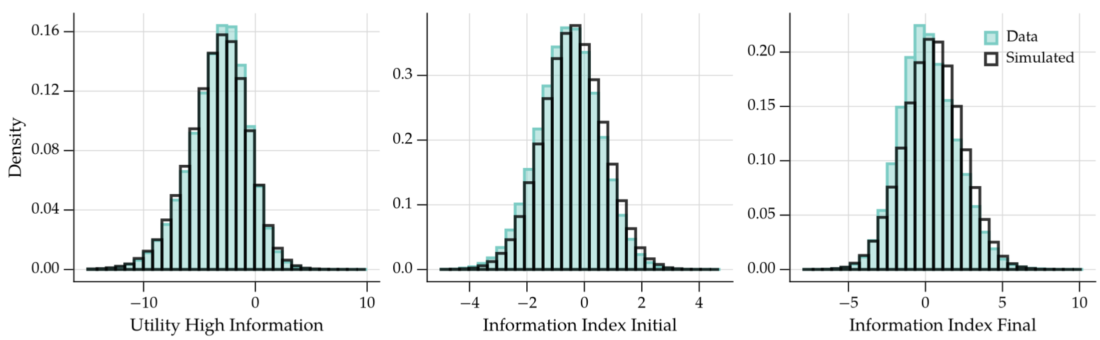
FIGURE A.12.—Panel (a) shows the distribution of the deterministic component of the single click cost ($x_i^c \gamma^{\text{cost}}$). Panel (b) shows the distribution of the ratio of the value of the first search over the cost of the first search.



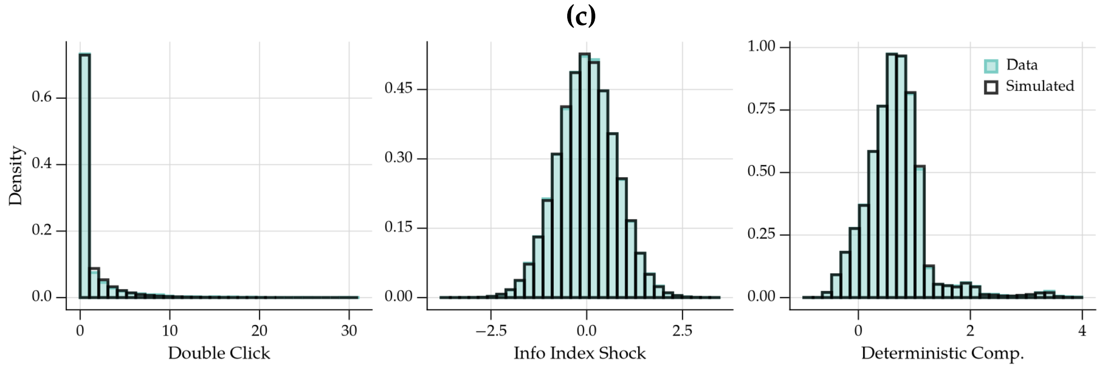
(a)



(b)



(c)



(d)

FIGURE A.13.—Model Fit. Panel (a) shows the model fit of school attributes (welfare, placement, and quality). Panel (b) shows the model fit of value-added, rank, and single clicks. 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.XXIV

MAIN RESULTS HETEROGENEITY BY SES AND PERCEIVED NUMBER OF HIGHLIGHT-WORTHY SCHOOLS

	Welfare	Placement		E(School Charact)		Search (N.Clicks)		
		Place	E(rank)	Quality	VA	Single	Double	V(1st)
Panel A: Heterogeneity by Parental Education								
High SES								
(1) Baseline	0.920	0.724	1.615	3.094	0.178	5.988	1.682	0.242
	(0.051)	(0.021)	(0.029)	(0.021)	(0.014)	(0.364)	(0.113)	(0.031)
(2) Full Information	1.586	0.859	1.795	3.185	0.206	-	-	-
	(0.084)	(0.011)	(0.035)	(0.026)	(0.013)	-	-	-
(3) Fix Information	1.328	0.765	1.688	3.179	0.207	12.103	3.119	1.119
	(0.068)	(0.016)	(0.030)	(0.024)	(0.013)	(3.548)	(0.931)	(0.137)
(4) No Search Cost	1.170	0.833	1.746	3.086	0.171	-	-	-
	(0.073)	(0.012)	(0.028)	(0.020)	(0.013)	-	-	-
(5) Cost and Info. Interaction	0.008	-0.015	-0.023	0.014	0.007	-	-	-
	(0.020)	(0.012)	(0.016)	(0.007)	(0.004)	-	-	-
Low SES								
(1) Baseline	0.438	0.744	1.386	2.971	0.130	3.864	1.116	0.849
	(0.026)	(0.011)	(0.010)	(0.009)	(0.005)	(0.166)	(0.061)	(0.202)
(2) Full Information	1.232	0.830	1.529	3.157	0.203	-	-	-
	(0.050)	(0.006)	(0.015)	(0.023)	(0.010)	-	-	-
(3) Fix Information	0.974	0.720	1.442	3.147	0.199	5.595	1.529	3.231
	(0.035)	(0.009)	(0.014)	(0.020)	(0.009)	(1.443)	(0.327)	(0.831)
(4) No Search Cost	0.656	0.895	1.481	2.952	0.117	-	-	-
	(0.040)	(0.008)	(0.010)	(0.008)	(0.006)	-	-	-
(5) Cost and Info. Interaction	0.040	-0.041	-0.007	0.029	0.017	-	-	-
	(0.010)	(0.008)	(0.007)	(0.006)	(0.003)	-	-	-
Panel B: Heterogeneity by difference in perceived vs real number of highlight-worthy schools								
Pessimist ($\leq$ 30th pctile)								
(1) Baseline	0.455	0.714	1.492	3.079	0.182	5.252	1.503	0.757
	(0.033)	(0.018)	(0.036)	(0.026)	(0.012)	(0.305)	(0.086)	(0.176)
(2) Full Information	1.400	0.883	1.706	3.220	0.230	-	-	-
	(0.082)	(0.009)	(0.037)	(0.032)	(0.014)	-	-	-
(3) Fix Information	1.050	0.749	1.585	3.222	0.233	10.543	2.810	4.366
	(0.060)	(0.015)	(0.035)	(0.030)	(0.016)	(2.814)	(0.688)	(1.054)
(4) No Search Cost	0.779	0.912	1.638	3.039	0.153	-	-	-
	(0.063)	(0.011)	(0.038)	(0.024)	(0.010)	-	-	-
(5) Cost and Info. Interaction	0.026	-0.065	-0.026	0.038	0.026	-	-	-
	(0.014)	(0.008)	(0.016)	(0.012)	(0.007)	-	-	-
Non-Pessimist ($>$ 30th pctile)								
(1) Baseline	0.583	0.742	1.440	2.990	0.137	4.128	1.185	0.696
	(0.027)	(0.013)	(0.013)	(0.016)	(0.007)	(0.174)	(0.060)	(0.150)
(2) Full Information	1.339	0.833	1.592	3.152	0.199	-	-	-
	(0.059)	(0.006)	(0.016)	(0.021)	(0.009)	-	-	-
(3) Fix Information	1.084	0.729	1.498	3.140	0.194	6.138	1.646	2.232
	(0.046)	(0.011)	(0.012)	(0.018)	(0.009)	(1.079)	(0.277)	(0.515)
(4) No Search Cost	0.816	0.880	1.546	2.975	0.127	-	-	-
	(0.042)	(0.005)	(0.011)	(0.015)	(0.007)	-	-	-
(5) Cost and Info. Interaction	0.023	-0.035	-0.013	0.028	0.015	-	-	-
	(0.009)	(0.007)	(0.006)	(0.006)	(0.003)	-	-	-

Note: This table presents counterfactual simulations by subgroup. All values are in levels. Columns: Welfare is the expected utility of the final allocation under true preferences (in equivalent kilometers, as the distance coefficient is normalized to -1). Place: probability of placement. E(rank): expected rank of placed school within the submitted application. Quality: school quality index (1–4 scale). VA: school value added in student-level standard deviations. Single, Double: number of pin and profile clicks. V(1st): perceived value of first click. Panel A splits by parental education (high SES = at least one college-educated parent; low SES = none). Panel B splits by whether parents underestimate the number of highlight-worthy schools at baseline (pessimist $\leq$ 30th percentile of the gap between perceived and actual count). Within each subgroup, rows report: (1) baseline simulation, (2) full information, (3) fix all information, (4) eliminate search costs, (5) cost and information interaction. Standard errors are bootstrapped. Search activity columns are not reported when search is irrelevant (full information) or unbounded (no search costs).

TABLE A.XXV

SEARCH COST REDUCTION COUNTERFACTUAL

		Welfare	Placement		E(School Charact)		Search (N.Clicks)		
			Place	E(rank)	Quality	VA	Single	Double	V(1st)
Panel A: Search Cost Reduction									
(1)	100%	0.546	0.739	1.437	2.998	0.141	4.342	1.243	0.712
		(0.022)	(0.003)	(0.005)	(0.006)	(0.003)	(0.142)	(0.042)	(0.003)
(2)	80%	0.552	0.743	1.438	2.999	0.141	4.932	1.412	0.712
		(0.019)	(0.003)	(0.004)	(0.006)	(0.003)	(0.117)	(0.039)	(0.003)
(3)	60%	0.564	0.752	1.441	2.999	0.140	5.952	1.701	0.712
		(0.021)	(0.003)	(0.004)	(0.006)	(0.003)	(0.128)	(0.039)	(0.003)
(4)	40%	0.576	0.762	1.444	2.999	0.140	7.448	2.131	0.712
		(0.019)	(0.003)	(0.005)	(0.006)	(0.002)	(0.153)	(0.048)	(0.004)
(5)	20%	0.596	0.777	1.449	2.999	0.139	9.913	2.833	0.712
		(0.020)	(0.003)	(0.004)	(0.007)	(0.003)	(0.134)	(0.037)	(0.003)
(6)	10%	0.613	0.789	1.455	2.998	0.139	12.828	3.655	0.712
		(0.022)	(0.003)	(0.005)	(0.005)	(0.003)	(0.175)	(0.053)	(0.003)
(7)	5%	0.627	0.803	1.463	2.996	0.137	17.044	4.833	0.712
		(0.020)	(0.004)	(0.006)	(0.006)	(0.003)	(0.184)	(0.057)	(0.003)
Panel B: Search Cost Reduction and Information Fix									
(8)	100%	1.054	0.730	1.500	3.155	0.201	7.059	1.887	2.756
		(0.009)	(0.006)	(0.009)	(0.008)	(0.005)	(0.177)	(0.048)	(0.010)
(9)	80%	1.068	0.737	1.502	3.153	0.200	8.688	2.335	2.756
		(0.010)	(0.006)	(0.009)	(0.009)	(0.005)	(0.218)	(0.061)	(0.010)
(10)	60%	1.087	0.746	1.507	3.153	0.200	11.346	3.037	2.756
		(0.010)	(0.007)	(0.008)	(0.009)	(0.005)	(0.175)	(0.053)	(0.010)
(11)	40%	1.118	0.761	1.515	3.155	0.201	16.135	4.315	2.756
		(0.009)	(0.006)	(0.009)	(0.008)	(0.005)	(0.226)	(0.066)	(0.010)
(12)	20%	1.150	0.777	1.526	3.154	0.201	24.414	6.514	2.756
		(0.010)	(0.007)	(0.009)	(0.008)	(0.005)	(0.201)	(0.057)	(0.010)
(13)	10%	1.166	0.785	1.534	3.156	0.201	31.870	8.476	2.756
		(0.010)	(0.007)	(0.009)	(0.008)	(0.005)	(0.190)	(0.067)	(0.010)
(14)	5%	1.174	0.789	1.539	3.157	0.202	38.995	10.377	2.756
		(0.009)	(0.006)	(0.009)	(0.008)	(0.005)	(0.169)	(0.069)	(0.010)

Note: This table reports outcome levels under search cost reduction at various cost levels, pooled across SES groups. Panel A reduces search costs while keeping beliefs at their baseline values. Panel B additionally sets $\hat{x}_{ij} = x_{ij}$ and corrects beliefs about admission chances, match values, and the number of undiscovered schools. Each row corresponds to a fraction of the original search cost: 100% is the baseline (or baseline with information fix), 5% retains only 5% of the cost. Columns: Welfare is the expected utility of the final allocation under true preferences (in equivalent kilometers). Place: probability of placement. E(rank): expected rank of placed school. Quality: school quality index (1–4). VA: school value added (student-level s.d.). Single: number of pin clicks. Double: number of profile clicks. V(1st): perceived value of first click. Standard deviations across simulation draws are reported in parentheses.

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.[56] 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)).[57]

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:

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

[57] 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 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.1 shows the potential search path of a parent in the control group or in the first treatment group.[58] 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).

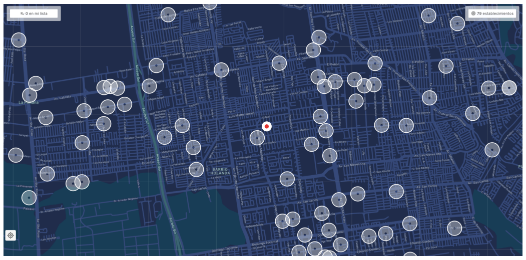
[58]For parents in the second treatment group, schools that are free and have high quality are highlighted in green. For more information on the difference between the treatment groups, see section S.3 in the supplementary material.

TABLE S.I

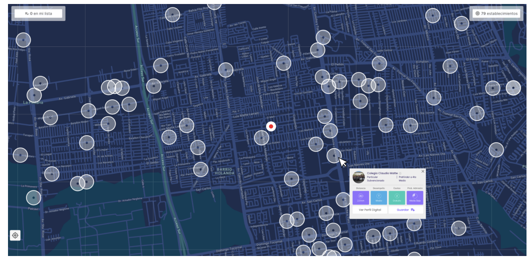
SUMMARY OF SURVEYS

Survey	Sections	Questions
Pre-Baseline + 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. Search plans.
	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 hypothetical application, (c) random known school not in hypothetical 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	Perception on whether their child is eligible for SEP. Parent staff priority in the application. 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 and not listed in application (+1 fake)
	Perceptions (x's)	Updated school knowledge level, perceived price, academic performance and admission chances for up to 5 schools selected based on the following priority criteria: (a) 1st-ranked school in application, (b) 2nd-ranked school in application, (c) last-ranked school in application, (d) 1st-ranked school in baseline (if not already mentioned in a-c), (e) random school that they learned about since baseline, (f) random highlight-worthy school that they didn't add to their application set but that they either knew at baseline or clicked in the explorer, (g) 2nd school for which we asked beliefs questions at baseline (if not already mentioned in a-c)
	Beliefs	Distribution of school characteristics in neighborhood
	Other Questions	Satisfaction with explorer. Found new schools due to explorer. Applied to SAE. Satisfaction with report card. Changed application after report card
Endline N = 540 Oct 21 - Oct 25	Awareness	Knowledge level of 4 random schools (+1 fake) + top 5 schools in application.
	Perception (x's)	Perceived academic performance + price of top 4 schools in application. Perceived admission chance of top 3 schools in application. Perceived academic performance of 4 random schools (+1 fake).
	Beliefs	Distribution of schools in neighborhood. SEP belief. Perceived risk of application
	Other Questions	Was report card useful and for what purpose. Likelihood of adding a school. Siblings application. Likelihood of rejecting assignment based on siblings

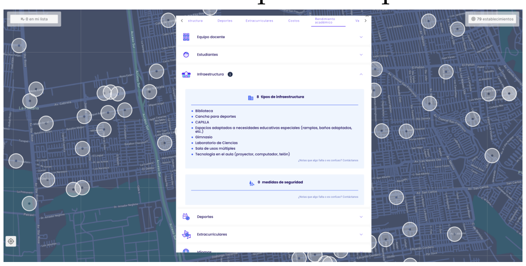
Note: This table presents a description of all surveys used in this study. SEP belief refers to the perceived eligibility for school vouchers. Distribution of schools in neighborhood refers to questions about the number of schools in each price and academic performance category, and the number of schools 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.



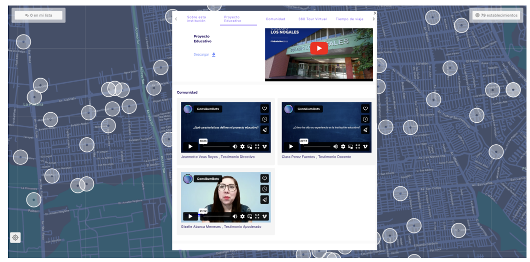
(a) Explorer map



(b) Pin click



(c) Profile (overview)



(d) Profile (details)

FIGURE S.1.—Example search path. This figure shows the potential search path of a parent in the control or first treatment arm.

S.3. 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 parents.

Intervention Details: Figures S.2 to S.5 show example screenshots of the information that was shown to parents as part of the search aid interventions. In this example, the family has access to 18 schools in total within 2km of the home (Figure S.2). 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.3). Both treatment groups further received a detailed table that shows the distribution of schools in each price and quality category (Figure S.4). 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.5).

TABLE S.II

BALANCE CHECKS FOR SEARCH INTERVENTIONS FOR HIGH-SES PARENTS

	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 2km	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
Number of available high quality schools	9.903	[4.748]	0.064	(0.375)	-0.163	(0.371)	888
Number of available low price schools	13.830	[7.438]	0.248	(0.518)	-0.324	(0.548)	888
Panel B: Parent/Child Characteristics							
Child is female	0.453	[0.499]	0.040	(0.041)	0.061	(0.043)	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
Child's age	3.913	[0.543]	0.014	(0.046)	-0.030	(0.046)	888
Panel C: Initial Knowledge 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]	-0.030	(0.025)	-0.011	(0.025)	888
First-preference school is highlight-worthy	0.531	[0.500]	-0.003	(0.047)	0.094	(0.048)	686
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
Perceived number of available high quality schools	5.076	[4.396]	0.697*	(0.390)	-0.018*	(0.344)	888
Perceived number of available low price schools	5.754	[4.863]	0.940**	(0.452)	0.332**	(0.433)	888
Parent believes to be SEP eligible	0.104	[0.306]	0.016	(0.027)	0.016	(0.028)	888
Parent is unsure 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 parents. 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 PARENTS

	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 2km	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
Number of available high quality schools	9.729	[5.467]	-0.167	(0.203)	0.121	(0.202)	3057
Number of available low price schools	15.015	[9.193]	-0.115	(0.322)	0.038	(0.312)	3057
Panel B: Parent/Child Characteristics							
Child is female	0.506	[0.500]	0.010	(0.022)	0.014	(0.022)	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
Child's age	3.903	[0.552]	-0.016	(0.023)	-0.012	(0.024)	3057
Panel C: Initial Knowledge 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	0.900	[0.301]	0.002	(0.013)	0.011	(0.013)	3057
First-preference school is highlight-worthy	0.662	[0.473]	0.054**	(0.023)	0.037**	(0.023)	2460
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
Perceived number of available high quality schools	4.989	[4.730]	-0.102	(0.197)	-0.258	(0.195)	3057
Perceived number of available low price schools	5.859	[5.108]	-0.134	(0.225)	-0.292	(0.236)	3057
Parent believes to be SEP eligible	0.192	[0.394]	-0.004	(0.017)	-0.015	(0.017)	3057
Parent is unsure 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 parents. 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.

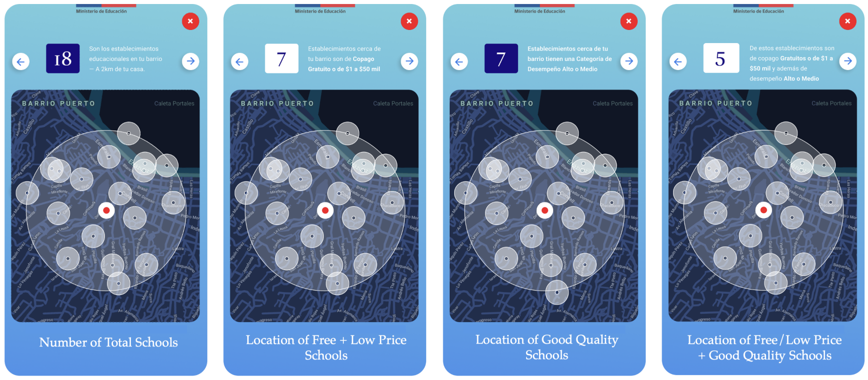


FIGURE S.2.—Search aid treatment 1

FIGURE S.3.—Search aid treatment 2

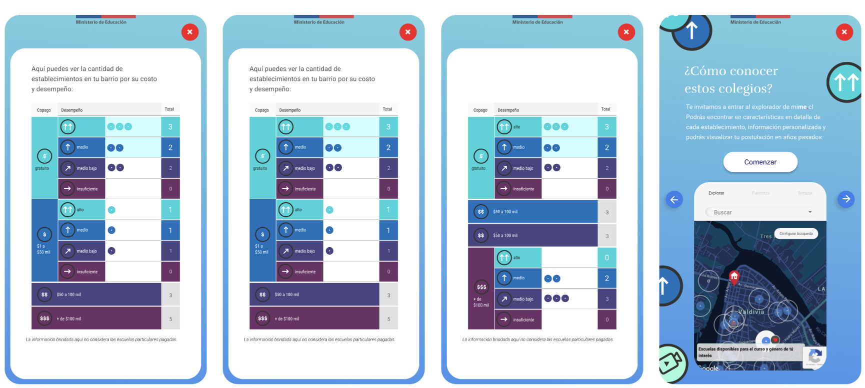


FIGURE S.4.—Additional distribution information for treatment groups 1 & 2

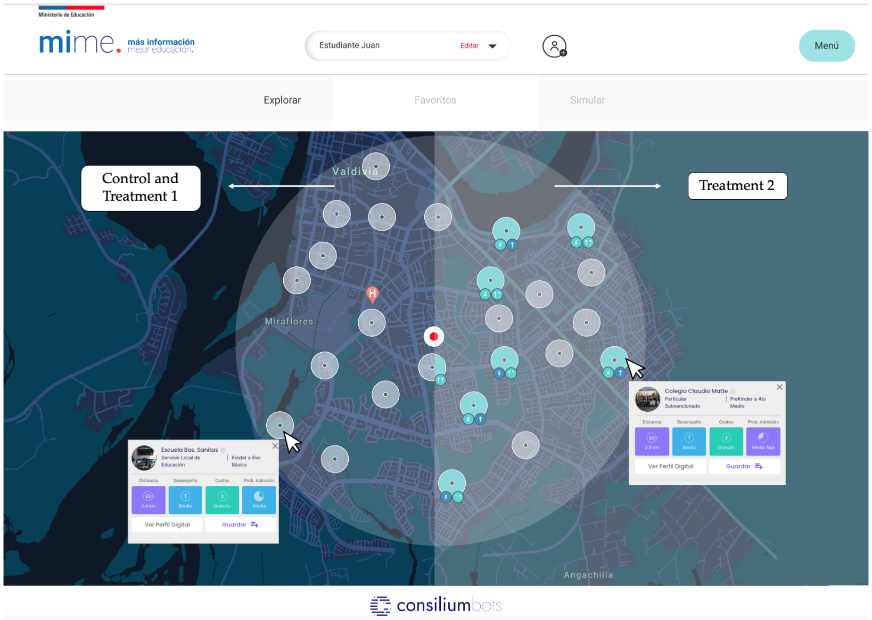


FIGURE S.5.—School explorer by treatment status

S.4. 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 balance checks for the high- and low-SES subsamples.

Intervention Details: Figure S.6 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.

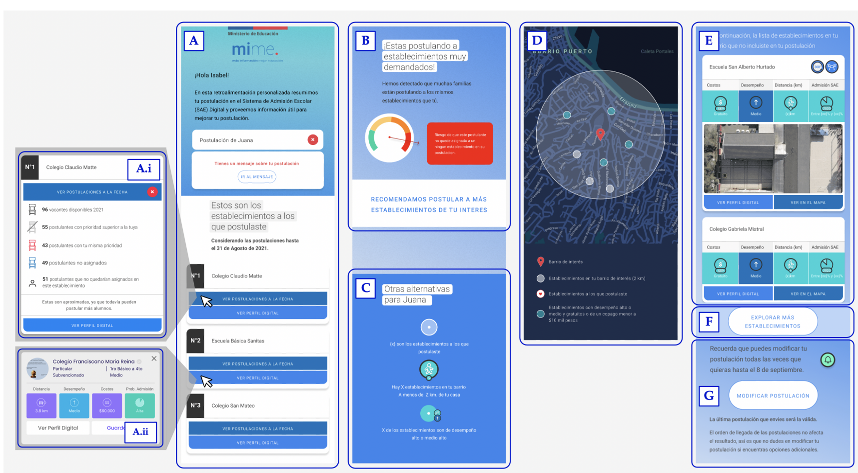


FIGURE S.6.—Feedback treatment

TABLE S.IV

BALANCE CHECK FOR FEEDBACK INTERVENTION FOR HIGH-SES PARENTS

	Control		Feedback Treatment		
	Mean (1)	St. Dev. (2)	Coeff. (3)	St. Err. (4)	N (5)
Panel A: Choice Environment					
Number of available schools 2km	15.269	[8.410]	-1.561	(1.053)	521
Number of available highlight-worthy schools	8.456	[4.901]	-0.434	(0.602)	521
Number of available high quality schools	9.765	[5.322]	-0.456	(0.583)	521
Number of available low price schools	13.697	[7.845]	-1.424	(1.039)	521
Panel B: Parent/Child Characteristics					
Child is female	0.507	[0.501]	-0.047	(0.043)	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
Child's age	3.854	[0.512]	0.029	(0.046)	521
Panel C: Initial Knowledge and Beliefs					
Expected satisfaction with process	5.072	[1.477]	0.091	(0.155)	492
Listed any school as first preference	0.925	[0.264]	0.021	(0.032)	521
First-preference school is highlight-worthy	0.585	[0.494]	0.070	(0.069)	407
Perceived admission chance 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
Perceived number of available high quality schools	4.833	[4.204]	0.942**	(0.398)	521
Perceived number of available low price schools	5.891	[4.743]	0.434	(0.433)	521
Parent believes to be SEP eligible	0.112	[0.316]	-0.018	(0.034)	521
Parent is unsure 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 parents. 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

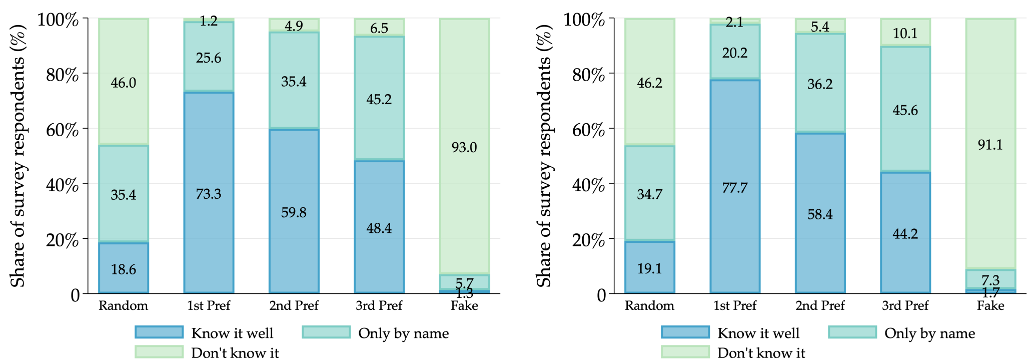
BALANCE CHECK FOR FEEDBACK INTERVENTION FOR LOW-SES PARENTS

	Control		Feedback Treatment		
	Mean (1)	St. Dev. (2)	Coeff. (3)	St. Err. (4)	N (5)
Panel A: Choice Environment					
Number of available schools 2km	15.904	[9.628]	-0.472	(0.907)	2033
Number of available highlight-worthy schools	8.325	[5.051]	0.410	(0.561)	2033
Number of available high quality schools	9.353	[5.350]	0.498	(0.486)	2033
Number of available low price schools	14.505	[9.245]	-0.492	(0.985)	2033
Panel B: Parent/Child Characteristics					
Child is female	0.518	[0.500]	-0.004	(0.029)	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
Child's age	3.889	[0.497]	0.010	(0.028)	2033
Panel C: Initial Knowledge and Beliefs					
Expected satisfaction with process	5.328	[1.363]	-0.041	(0.073)	1918
Listed any school as first preference	0.930	[0.255]	-0.007	(0.022)	2033
First-preference school is highlight-worthy	0.694	[0.461]	0.041	(0.037)	1672
Perceived admission chance 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
Perceived number of available high quality schools	4.781	[4.009]	0.198	(0.196)	2033
Perceived number of available low price schools	5.583	[4.856]	0.263	(0.312)	2033
Parent believes to be SEP eligible	0.179	[0.383]	-0.013	(0.018)	2033
Parent is unsure 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 parents. 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.5. 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.7). Figure S.8 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.9 and S.10), but low-SES parents have slightly more accurate beliefs about placement chances (Figure S.11).

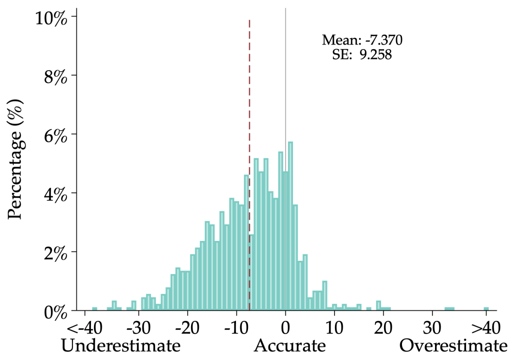


(a) High SES

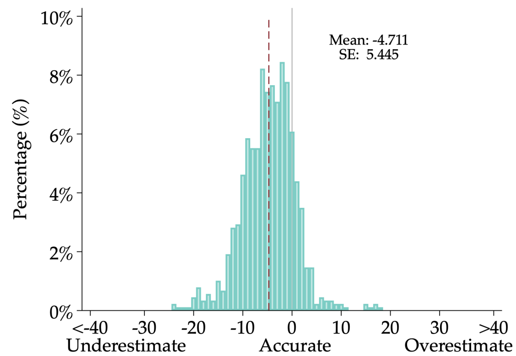
(b) Low SES

FIGURE S.7.—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 parents (N = 888). Panel bB) plots the same stated knowledge for five school categories for low-SES parents (N = 3057).

Panel A: High SES

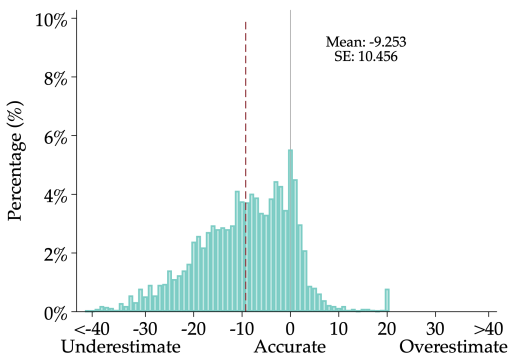


(a) Number of Schools

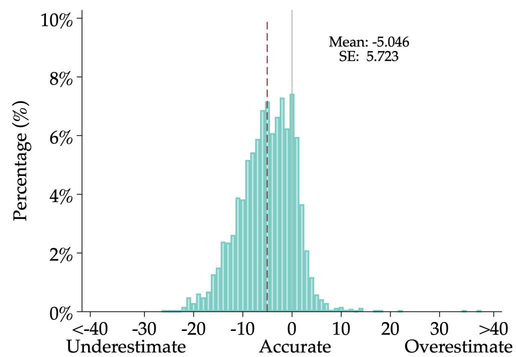


(b) Number of Highlight-worthy Schools

Panel B: Low SES



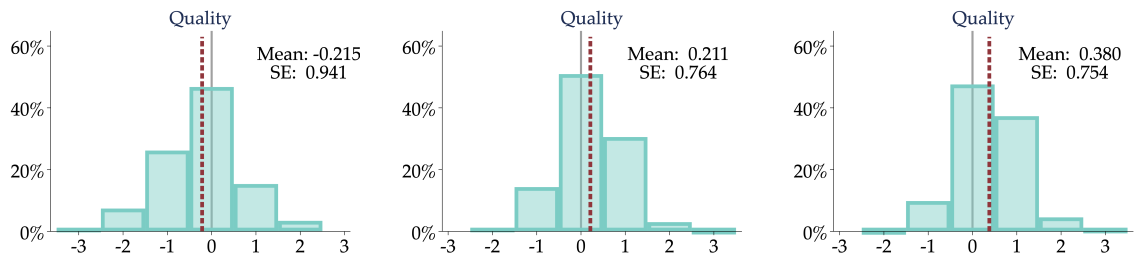
(c) Number of Schools



(d) Number of Highlight-worthy Schools

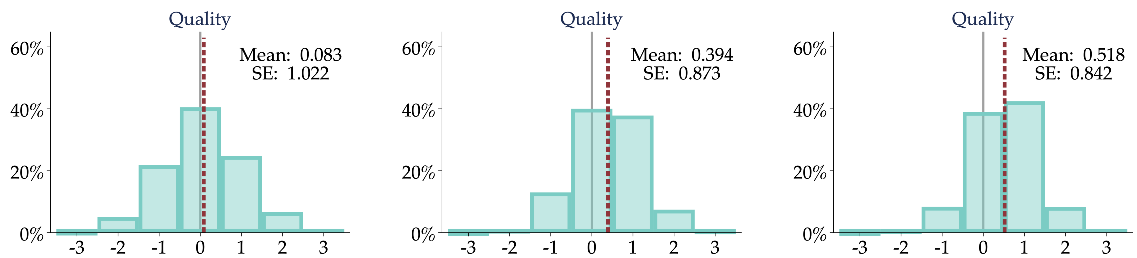
FIGURE S.8.—Beliefs about the Distribution of School Attributes by SES Status. Notes: Panels (a) and (b) show the bias in the beliefs of the number of total and highlight-worthy schools within 2km of the parent's home for high-SES parents (N = 888). Panels (c) and (d) show the bias in the beliefs of the number of total and highlight-worthy schools within 2km of the parent's home for low-SES parents (N = 3057).

Panel A: High SES



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

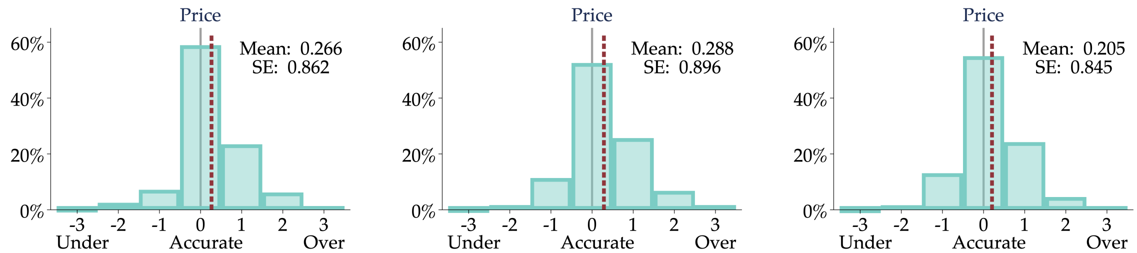
Panel B: Low SES



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

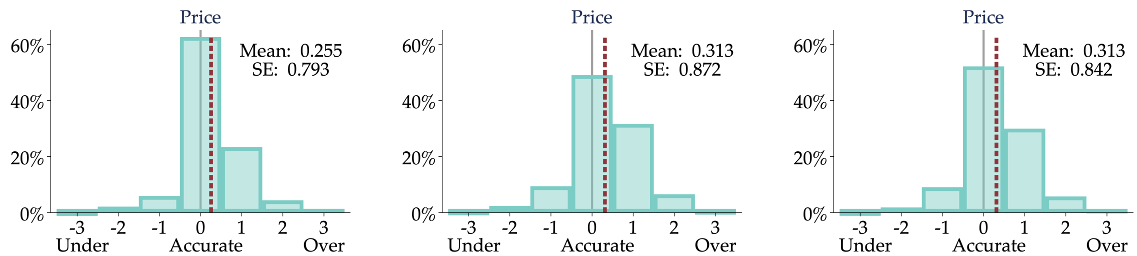
FIGURE S.9.—Error in Quality by SES status. Notes: Panels (a) and (d) show the bias on perceived quality of a known random school asked in baseline. Panels (b) and (e) show the bias on perceived quality of a random school in the application list, excluding the first ranked school. Panels (c) and (f) 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 parents (N = 888) and Panels (d-f) represent low-SES parents (N = 3057).

Panel A: High SES



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

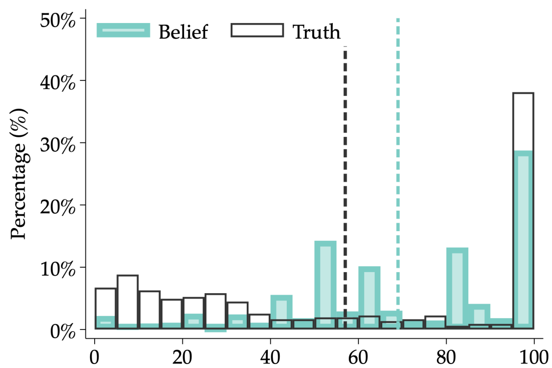
Panel B: Low SES



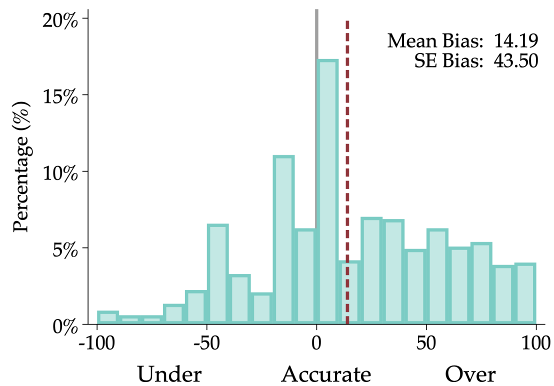
(d) Random School Known **(e)** Random School in App **(f)** First Preference

FIGURE S.10.—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 parents (N = 888) and Panels (d-f) represent low-SES parents (N = 3057).

Panel A: High SES

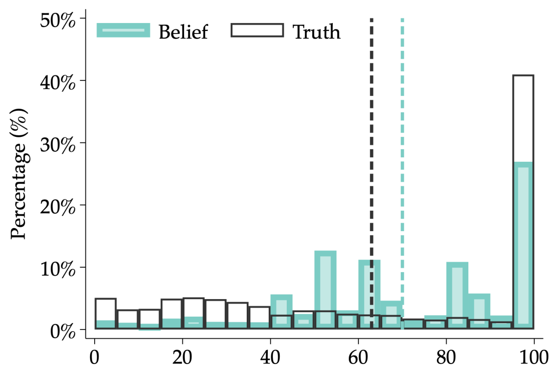


(a) Distribution of Placement Chances

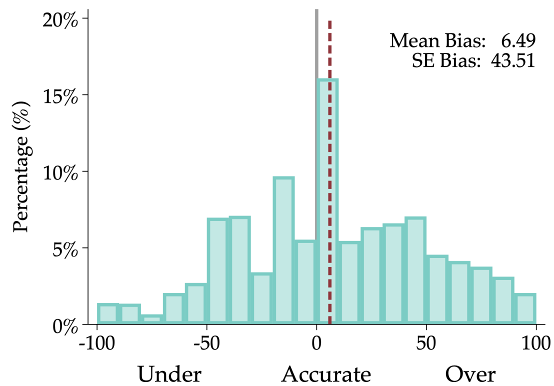


(b) Bias of Placement Chances

Panel B: Low SES



(c) Distribution of Placement Chances



(d) Bias of Placement Chances

FIGURE S.11.—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 parents 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.

S.6. ADDITIONAL COUNTERFACTUAL RESULTS

This section presents additional counterfactual results that supplement the analysis in Section 9.2. Table S.VI reports the full set of counterfactual simulations, including the scenarios in Table IV and the following additional counterfactuals:

1. *Partial information correction (row 4).* Corrects only misperceptions of school characteristics and beliefs about the distribution of unknown schools (x and $f(x)$), without correcting admission chance beliefs or match value signals. Comparison with row (5), which fixes all information, isolates the contribution of school characteristic misperceptions.
2. *RCT effects (row 6).* Reports the simulated results keeping the search aid information treatments RCT, changes relative to the baseline are due to the effects of information provided in each treatment arm.
3. *Learn x late (row 7).* Provides parents with the true characteristics ($\hat{x}_{ij} = x_{ij}$) of all known schools after search has concluded, just before applications are due. Unlike the information corrections in rows (4)–(5), this intervention does not affect search decisions.
4. *Better search technology (rows 9–10).* Row (9) improves the search technology by making pin clicks fully informative after a single click, without correcting baseline misperceptions or reducing search costs. Row (10) combines the better search technology with the elimination of search costs.
5. *Sequential decomposition with better search (rows 11–13).* Starting from the better search technology (S^*), we sequentially correct misperceptions of observables (x, row 11), beliefs about unknown schools' characteristics ($f(x)$, row 12), and all remaining information frictions (r, $f(\varepsilon)$, ε, row 13). This decomposition quantifies the relative contribution of each bias conditional on an improved search technology.
6. *Additional misspecified model (row 15).* In addition to the "no misperception of x" model discussed in the main text (row 14), row (15) reports results under a specification that further eliminates the distinction between knowing a school "by name" and knowing it "well," removing measurement error on match values for all known schools ($\pi_{ijt} = 1$ vs. $\pi_{ijt} = 2$).

Table S.VII reports outcome levels under the better search technology (S^*) at progressively lower search costs, from 100% to 5% of the baseline search cost.

TABLE S.VI

ADDITIONAL COUNTERFACTUAL 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	0.546	0.739	1.437	2.998	0.141	4.342	1.243	0.712
		(0.026)	(0.011)	(0.012)	(0.009)	(0.006)	(0.197)	(0.065)	(0.158)
(2)	Full information	1.312	0.837	1.590	3.164	0.204	-	-	-
		(0.061)	(0.005)	(0.016)	(0.019)	(0.008)	-	-	-
(3)	Gains (difference (2)-(1))	0.765	0.097	0.154	0.165	0.063	-	-	-
		(0.051)	(0.013)	(0.017)	(0.018)	(0.007)	-	-	-
	(% Change)	140.14%	13.15%	10.69%	5.51%	44.80%	-	-	-
Fix perceptions and information interventions									
(4)	$x+f(x)$	0.395	-0.059	0.032	0.165	0.068	3.410	0.945	2.900
		(0.030)	(0.006)	(0.009)	(0.015)	(0.006)	(1.677)	(0.459)	(0.649)
(5)	$x+f(x)+r+f(\varepsilon)+\varepsilon$	0.508	-0.009	0.063	0.156	0.060	2.717	0.644	2.044
		(0.038)	(0.007)	(0.009)	(0.014)	(0.006)	(1.358)	(0.330)	(0.499)
(6)	RCT effects	0.095	0.067	0.063	0.009	0.009	0.226	0.152	-0.051
		(0.011)	(0.009)	(0.007)	(0.005)	(0.003)	(0.084)	(0.026)	(0.012)
(7)	Learn x late	0.474	-0.016	0.063	0.168	0.073	-	-	-
		(0.032)	(0.011)	(0.011)	(0.014)	(0.006)	-	-	-
Changes in search cost and technology									
(8)	No cost	0.225	0.141	0.100	-0.018	-0.012	-	-	-
		(0.023)	(0.013)	(0.010)	(0.004)	(0.002)	-	-	-
(9)	Better Search (S^*)	0.097	0.066	0.021	-0.007	-0.008	2.381	-	0.140
		(0.016)	(0.011)	(0.006)	(0.006)	(0.003)	(0.863)	-	(0.062)
(10)	No cost $+S^*$	0.288	0.165	0.109	-0.026	-0.019	-	-	-
		(0.025)	(0.012)	(0.013)	(0.009)	(0.005)	-	-	-
Decomposition: sequential correction with better search									
(11)	S^*+x	0.497	-0.008	0.044	0.158	0.062	2.368	-	0.128
		(0.041)	(0.012)	(0.011)	(0.016)	(0.006)	(0.845)	-	(0.062)
(12)	$S^*+x+f(x)$	0.608	0.038	0.086	0.159	0.061	10.638	-	4.152
		(0.053)	(0.019)	(0.021)	(0.017)	(0.007)	(5.034)	-	(0.772)
(13)	$S^*+x+f(x)+r+f(\varepsilon)+\varepsilon$	0.632	0.047	0.093	0.155	0.059	7.342	-	2.340
		(0.048)	(0.014)	(0.016)	(0.016)	(0.006)	(2.984)	-	(0.528)
Misspecified models									
(14)	No mispercept. of x ($\hat{x}=x$)								
	Full information	0.514	0.170	0.103	-0.034	-0.021	-	-	-
	No search cost	0.415	0.148	0.095	-0.025	-0.016	-	-	-
	S^*	0.152	0.065	0.017	-0.010	-0.008	2.236	-	0.057
	No cost $+S^*$	0.514	0.170	0.103	-0.034	-0.021	-	-	-
	$S^*+x+f(x)+r+f(\varepsilon)+\varepsilon$	0.314	0.109	0.054	-0.020	-0.014	5.412	-	2.463
(15)	No mispercept. of x, e if $\pi>0$								
	Full information	17.776	0.175	0.137	-0.055	-0.050	-	-	-
	No search cost	14.107	0.139	0.143	-0.049	-0.043	-	-	-
	No cost $+S^*$	17.776	0.175	0.137	-0.055	-0.050	-	-	-

Note: This table presents all counterfactual simulations pooled across SES groups. Columns: Welfare is the expected utility of the final allocation under true preferences (in equivalent kilometers, as the distance coefficient is normalized to -1). Place: probability of placement. E(rank): expected rank of placed school within the submitted application. Quality: school quality index (1–4 scale). VA: school value added in student-level standard deviations. Single, Double: number of pin and profile clicks. V(1st): perceived value of first click. Panel 1 (rows (1)–(3)): levels for the baseline and full-information scenarios plus gains. Panel 2 (rows (4)–(7)): information interventions reported as differences relative to baseline—fixing x and $f(x)$, fixing all information, RCT effects, and learning true x after search. Panel 3 (rows (8)–(10)): search cost interventions—eliminating search costs, adopting better search technology (S^*), and their combination. Panel 4 (rows (11)–(13)): sequential corrections starting from better search (S^*) and progressively adding x, $f(x)$, and the remaining information components. Panel 5 (rows (14)–(15)): gains under misspecified models that eliminate misperceptions of x or of both x and ε, reporting full-information, no-search-cost, and better-search gains relative to baseline. Search activity columns are not reported when search is irrelevant (full information) or unbounded (no search costs).

TABLE S.VII

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%	0.643	0.805	1.457	2.991	0.133	6.723	0.000	0.852
		(0.020)	(0.004)	(0.005)	(0.008)	(0.004)	(0.142)	(0.000)	(0.003)
(2)	Better Search 80%	0.646	0.807	1.459	2.992	0.133	6.958	0.000	0.852
		(0.020)	(0.003)	(0.006)	(0.007)	(0.004)	(0.130)	(0.000)	(0.003)
(3)	Better Search 60%	0.656	0.812	1.463	2.992	0.133	7.383	0.000	0.852
		(0.021)	(0.003)	(0.006)	(0.007)	(0.003)	(0.110)	(0.000)	(0.003)
(4)	Better Search 40%	0.667	0.820	1.465	2.990	0.132	8.217	0.000	0.852
		(0.023)	(0.004)	(0.006)	(0.007)	(0.004)	(0.175)	(0.000)	(0.003)
(5)	Better Search 20%	0.692	0.836	1.476	2.990	0.131	10.223	0.000	0.852
		(0.022)	(0.003)	(0.004)	(0.007)	(0.004)	(0.158)	(0.000)	(0.003)
(6)	Better Search 10%	0.724	0.856	1.488	2.986	0.129	13.349	0.000	0.852
		(0.021)	(0.003)	(0.006)	(0.007)	(0.004)	(0.152)	(0.000)	(0.003)
(7)	Better Search 5%	0.750	0.873	1.504	2.983	0.126	17.644	0.000	0.852
		(0.022)	(0.003)	(0.005)	(0.007)	(0.004)	(0.171)	(0.000)	(0.003)

Note: This table reports outcome levels under the better search technology (S^*) at various search cost levels. Each row corresponds to a fraction of the better-search cost: 100% is the baseline better-search scenario, 5% retains only 5% of that cost. Columns: Welfare is the expected utility of the final allocation under true preferences (in equivalent kilometers). Place: probability of placement. E(rank): expected rank of placed school. Quality: school quality index (1–4). VA: school value added (student-level s.d.). Single: number of pin clicks. Standard deviations across simulation draws are reported in parentheses.