Search and Biased Beliefs in Education Markets

Patrick Agte[1] Claudia Allende[2] Adam Kapor[3] Chris Neilson[1] Fernando Ochoa[4]

[1]Yale University [2]Stanford GSB [3]Princeton University [4]NYU

NBER SI IO

July 2024

Motivation: Search and Biased Beliefs

- Search frictions: popular explanation for demand for expensive / low-quality options:

Motivation: Search and Biased Beliefs

- Search frictions: popular explanation for demand for expensive / low-quality options:
  - Price dispersion in homogeneous-good markets (Sorensen 2000).

Motivation: Search and Biased Beliefs

- Search frictions: popular explanation for demand for expensive / low-quality options:
  - Price dispersion in homogeneous-good markets (Sorensen 2000).
  - Perhaps more severe with differentiated goods / many options:
    - health plans (Handel and Kolstad 2015); mortgages (AGHMSY 2024, Bhattacharya et al 2024); schools (Ajayi and Sidibe 2022; AKNZ 2022).

Motivation: Search and Biased Beliefs

- Search frictions: popular explanation for demand for expensive / low-quality options:
  - Price dispersion in homogeneous-good markets (Sorensen 2000).
  - Perhaps more severe with differentiated goods / many options:
    - health plans (Handel and Kolstad 2015); mortgages (AGHMSY 2024, Bhattacharya et al 2024); schools (Ajayi and Sidibe 2022; AKNZ 2022).
- Policy solutions: high "search costs" $\Longrightarrow$ rely less on search

Motivation: Search and Biased Beliefs

- Search frictions: popular explanation for demand for expensive / low-quality options:
  - Price dispersion in homogeneous-good markets (Sorensen 2000).
  - Perhaps more severe with differentiated goods / many options:
    - health plans (Handel and Kolstad 2015); mortgages (AGHMSY 2024, Bhattacharya et al 2024); schools (Ajayi and Sidibe 2022; AKNZ 2022).
- Policy solutions: high "search costs" $\implies$ rely less on search
  - simplify choice sets (Brown / Jeon 2023); add default options (HK2015); use intermediaries (Boehm 2023).

Motivation: Search and Biased Beliefs

- Search frictions: popular explanation for demand for expensive / low-quality options:
  - Price dispersion in homogeneous-good markets (Sorensen 2000).
  - Perhaps more severe with differentiated goods / many options:
    - health plans (Handel and Kolstad 2015); mortgages (AGHMSY 2024, Bhattacharya et al 2024); schools (Ajayi and Sidibe 2022; AKNZ 2022).
- Policy solutions: high "search costs" $\implies$ rely less on search
  - simplify choice sets (Brown / Jeon 2023); add default options (HK2015); use intermediaries (Boehm 2023).
- Alternative explanation: wrong info/beliefs can distort perceived returns to search:

Motivation: Search and Biased Beliefs

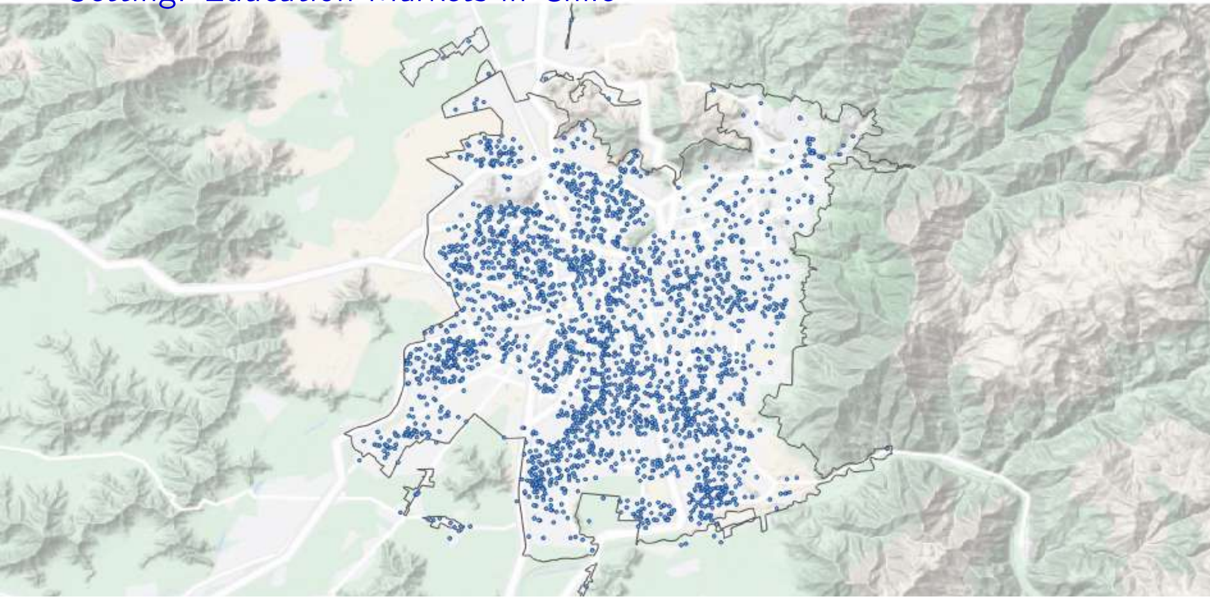
- Search frictions: popular explanation for demand for expensive / low-quality options:
  - Price dispersion in homogeneous-good markets (Sorensen 2000).
  - Perhaps more severe with differentiated goods / many options:
    - health plans (Handel and Kolstad 2015); mortgages (AGHMSY 2024, Bhattacharya et al 2024); schools (Ajayi and Sidibe 2022; AKNZ 2022).
- Policy solutions: high "search costs" $\implies$ rely less on search
  - simplify choice sets (Brown / Jeon 2023); add default options (HK2015); use intermediaries (Boehm 2023).
- Alternative explanation: wrong info/beliefs can distort perceived returns to search:
  - Overestimate quality of "known" options or underestimate unknowns $\Rightarrow$ need smaller search cost to rationalize data

Motivation: Search and Biased Beliefs

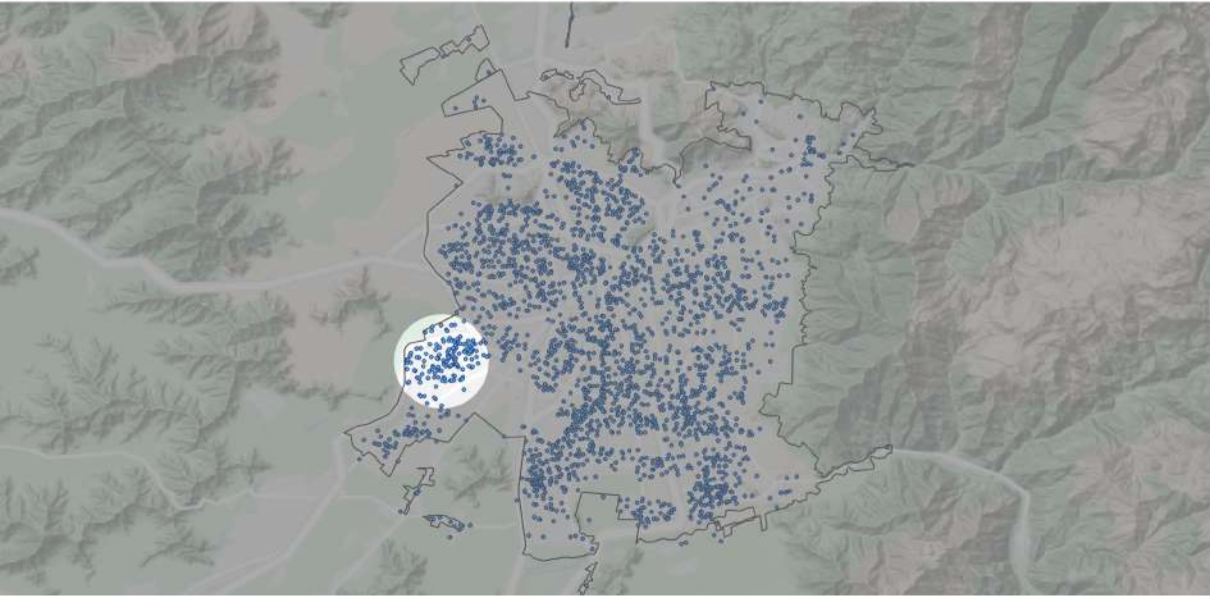
- Search frictions: popular explanation for demand for expensive / low-quality options:
  - Price dispersion in homogeneous-good markets (Sorensen 2000).
  - Perhaps more severe with differentiated goods / many options:
    - health plans (Handel and Kolstad 2015); mortgages (AGHMSY 2024, Bhattacharya et al 2024); schools (Ajayi and Sidibe 2022; AKNZ 2022).
- Policy solutions: high "search costs" $\implies$ rely less on search
  - simplify choice sets (Brown / Jeon 2023); add default options (HK2015); use intermediaries (Boehm 2023).
- Alternative explanation: wrong info/beliefs can distort perceived returns to search:
  - Overestimate quality of "known" options or underestimate unknowns
    $\Rightarrow$ need smaller search cost to rationalize data

$\implies$ **Research Question**: How do families' limited awareness of options, misperceptions and inaccurate beliefs about characteristics interact with preferences and search costs to distort their information-acquisition efforts, choices, and outcomes in a complex, high-stakes decision?

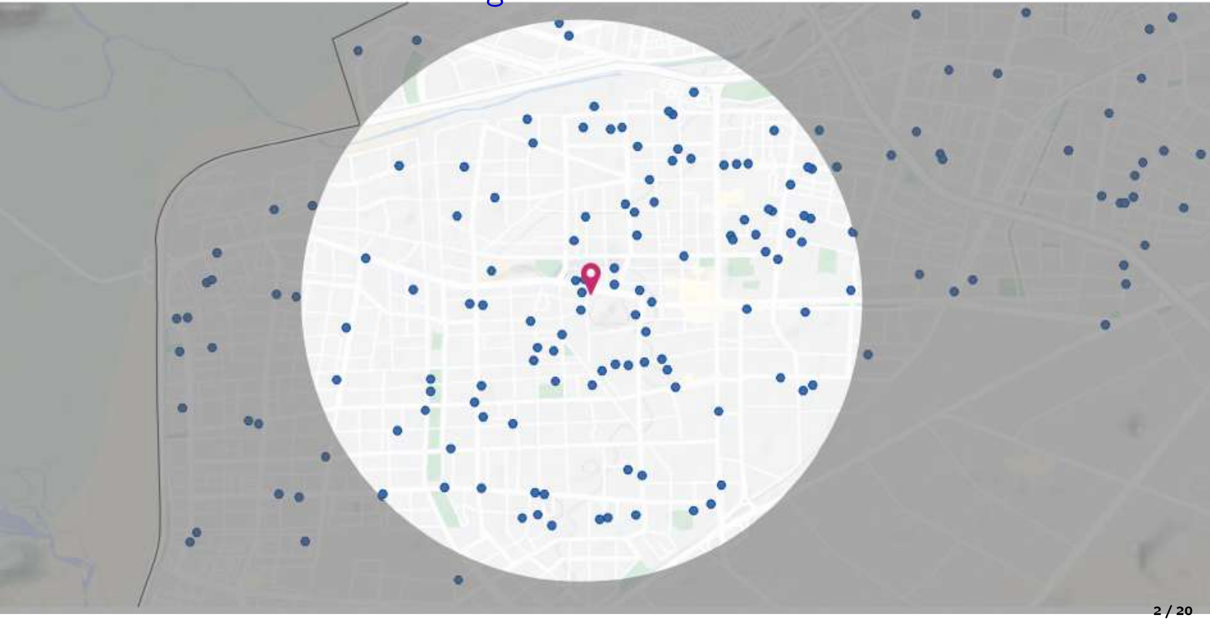
Setting: Education Markets in Chile



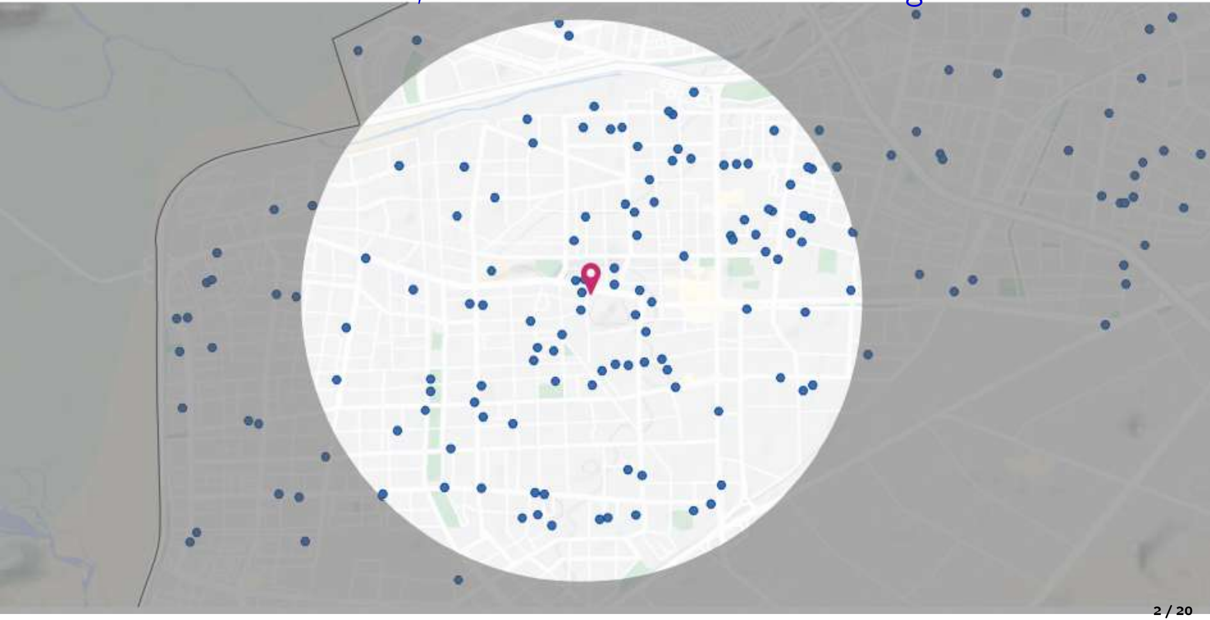
Zoom into a 2km Radius Area



Almost 100 Schools Offering K



Placement is uncertain; must submit ranked list to assignment mech.



This paper

- Nationwide school choice process: Collaboration w/ Government and EdTech NGO

This paper

- Nationwide school choice process: Collaboration w/ Government and EdTech NGO
- Methods: novel microdata + field experiments + model:

This paper

- Nationwide school choice process: Collaboration w/ Government and EdTech NGO
- Methods: novel microdata + field experiments + model:

1. We model demand for schools w/ heterogeneous preferences + search costs:

This paper

- Nationwide school choice process: Collaboration w/ Government and EdTech NGO
- Methods: novel microdata + field experiments + model:

1. We model demand for schools w/ heterogeneous preferences + search costs:
    - Imperfect awareness of schools

This paper

- Nationwide school choice process: Collaboration w/ Government and EdTech NGO
- Methods: novel microdata + field experiments + model:

1. We model demand for schools w/ heterogeneous preferences + search costs:
    - Imperfect awareness of schools
    - Endogenous (sequential) search decisions

This paper

- Nationwide school choice process: Collaboration w/ Government and EdTech NGO
- Methods: novel microdata + field experiments + model:

1. We model demand for schools w/ heterogeneous preferences + search costs:
    - Imperfect awareness of schools
    - Endogenous (sequential) search decisions
    - Misperceptions of known schools' price, quality score, match quality, admission chance.

This paper

- Nationwide school choice process: Collaboration w/ Government and EdTech NGO
- Methods: novel microdata + field experiments + model:

1. We model demand for schools w/ heterogeneous preferences + search costs:
   - Imperfect awareness of schools
   - Endogenous (sequential) search decisions
   - Misperceptions of known schools' price, quality score, match quality, admission chance.
   - Biased beliefs about these four objects over unknown schools.

This paper

- Nationwide school choice process: Collaboration w/ Government and EdTech NGO
- Methods: novel microdata + field experiments + model:

1. We model demand for schools w/ heterogeneous preferences + search costs:
   - Imperfect awareness of schools
   - Endogenous (sequential) search decisions
   - Misperceptions of known schools' price, quality score, match quality, admission chance.
   - Biased beliefs about these four objects over unknown schools.
2. Novel search data and surveys:
   - Set up "school explorer" platform w/ personalized information, track search activity
   - 3 survey waves; measure beliefs, perceptions, awareness, preferences pre/post
   - Link survey and clicks to admin applications and enrollment

This paper

- Nationwide school choice process: Collaboration w/ Government and EdTech NGO
- Methods: novel microdata + field experiments + model:

1. We model demand for schools w/ heterogeneous preferences + search costs:
    - Imperfect awareness of schools
    - Endogenous (sequential) search decisions
    - Misperceptions of known schools' price, quality score, match quality, admission chance.
    - Biased beliefs about these four objects over unknown schools.
2. Novel search data and surveys:
    - Set up "school explorer" platform w/ personalized information, track search activity
    - 3 survey waves; measure beliefs, perceptions, awareness, preferences pre/post
    - Link survey and clicks to admin applications and enrollment
3. RCTs:
    1. "Search Aid" via explorer: dist'n of school characteristics + salience of good schools
    2. "Feedback" using app data: info about "known" schools, targeting misperceptions

Today's Talk

- Many frictions exist. But the most important are:
  1. imperfectly-revealing search technology
  2. misperceptions about observables of known schools
  - Fix (1) -> welfare gains + $\approx$ 50% more search
  - Fix (1) + (2) -> more welfare; quality gains $\approx$ full info; close SES school quality gap.

Today's Talk

- Many frictions exist. But the most important are:
  1. imperfectly-revealing search technology
  2. misperceptions about observables of known schools
  - Fix (1) -> welfare gains + $\approx$ 50% more search
  - Fix (1) + (2) -> more welfare; quality gains $\approx$ full info; close SES school quality gap.

- Rest of talk:
  1. example
  2. data + descriptives
  3. experiments
  4. model
  5. results

Motivating Example

- Household knows outside option (sure payoff 0), one "inside" school.
  - Payoff $u_1 = 1$ if assigned to this school.
  - Rejected with probability $r_1 \in [0, 1]$.

Motivating Example

- Household knows outside option (sure payoff 0), one "inside" school.
  - Payoff $u_1 = 1$ if assigned to this school.
  - Rejected with probability $r_1 \in [0,1]$.
- Can pay to draw one more school $(u_2, r_2) \sim f(\cdot)$ before submitting ranking.

Motivating Example

- Household knows outside option (sure payoff 0), one "inside" school.
  - Payoff $u_1 = 1$ if assigned to this school.
  - Rejected with probability $r_1 \in [0, 1]$.
- Can pay to draw one more school $(u_2, r_2) \sim f(\cdot)$ before submitting ranking.
- Chile uses student-proposing deferred acceptance with independent lotteries:

Motivating Example

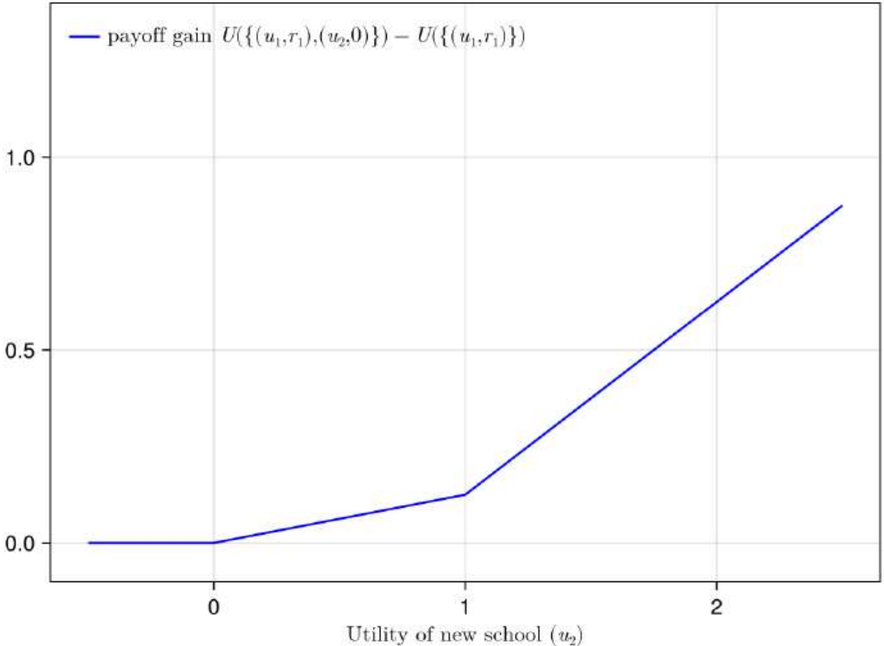
- Household knows outside option (sure payoff 0), one "inside" school.
  - Payoff $u_1 = 1$ if assigned to this school.
  - Rejected with probability $r_1 \in [0,1]$.
- Can pay to draw one more school $(u_2, r_2) \sim f(\cdot)$ before submitting ranking.
- Chile uses student-proposing deferred acceptance with independent lotteries:

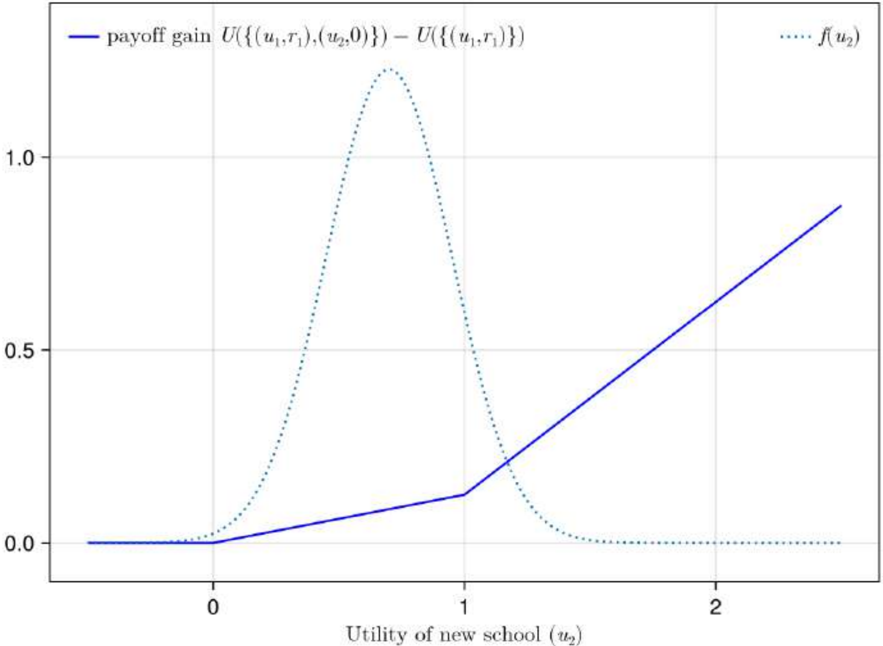
  $\Rightarrow$ optimal to rank truthfully.

Motivating Example

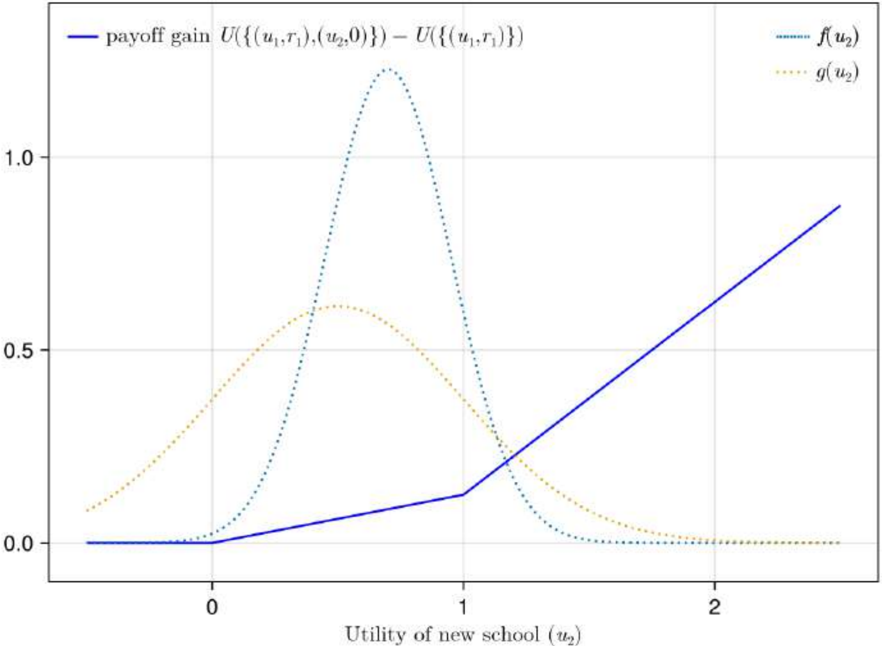
- Household knows outside option (sure payoff 0), one "inside" school.
  - Payoff $u_1 = 1$ if assigned to this school.
  - Rejected with probability $r_1 \in [0, 1]$.
- Can pay to draw one more school $(u_2, r_2) \sim f(\cdot)$ before submitting ranking.
- Chile uses student-proposing deferred acceptance with independent lotteries:

  $\Rightarrow$ optimal to rank truthfully.
- Expected payoff of optimal app: $\begin{cases} (1 - r_j)u_j + r_j(1 - r_k)u_k & \text{if } u_j > u_k > 0 \\ (1 - r_1)u_1 & \text{if } u_2 < 0 \text{ or don't search.} \end{cases}$

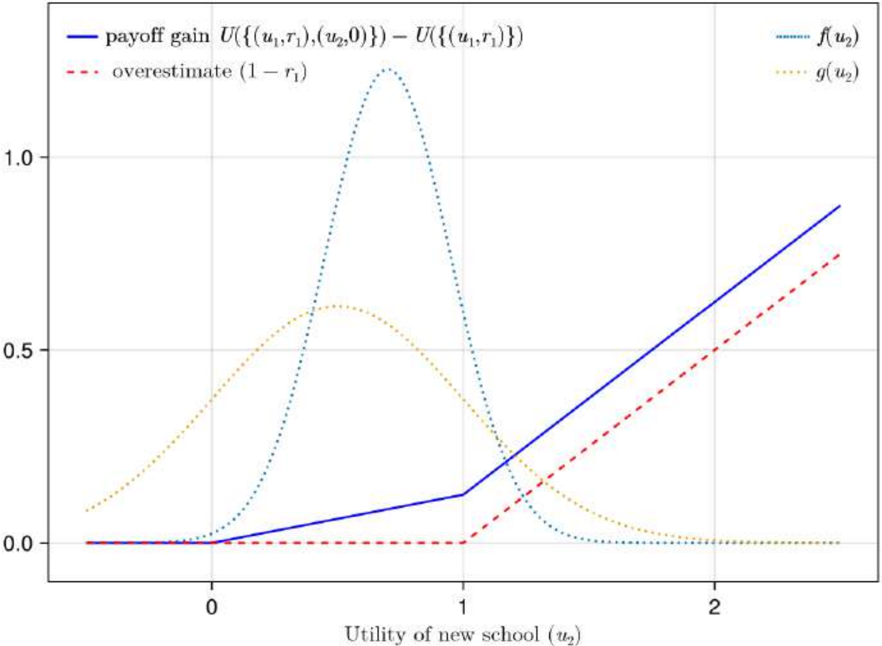

payoff gain $U(\{(u_1,r_1),(u_2,0)\}) - U(\{(u_1,r_1)\})$
1.0
0.5
0.0
0
1
2
Utility of new school (u_2)



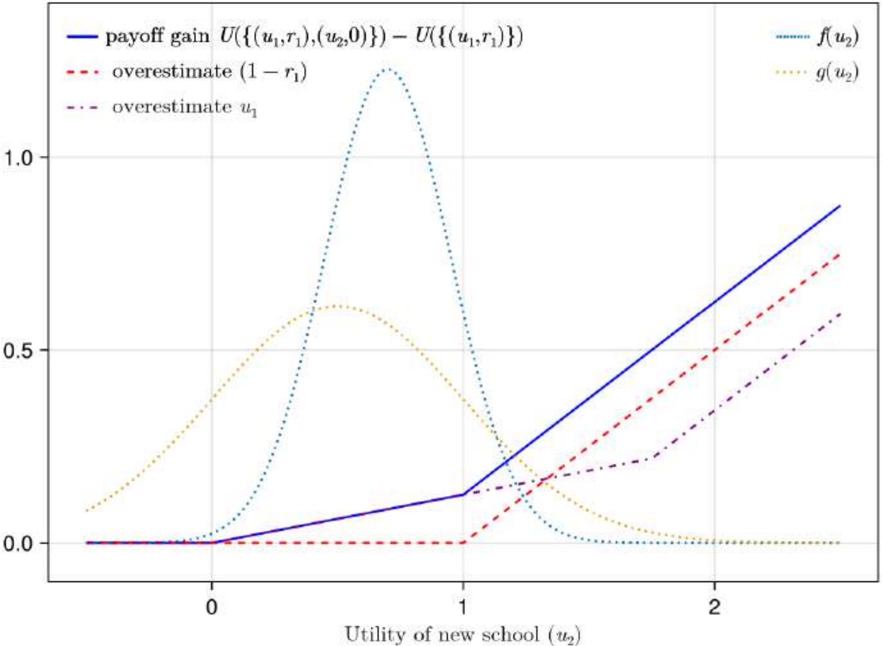
payoff gain $U(\{(u_1,r_1),(u_2,0)\}) - U(\{(u_1,r_1)\})$
$f(u_2)$
1.0
0.5
0.0
0
1
2
Utility of new school (u_2)



payoff gain $U(\{(u_1,r_1),(u_2,0)\}) - U(\{(u_1,r_1)\})$
$f(u_2)$
$g(u_2)$
1.0
0.5
0.0
0
1
2
Utility of new school (u_2)



payoff gain $U(\{(u_1,r_1),(u_2,0)\}) - U(\{(u_1,r_1)\})$
overestimate $(1 - r_1)$
$f(u_2)$
$g(u_2)$
1.0
0.5
0.0
0
1
2
Utility of new school (u_2)



payoff gain $U(\{(u_1,r_1),(u_2,0)\}) - U(\{(u_1,r_1)\})$
overestimate $(1 - r_1)$
overestimate u_1
$f(u_2)$
$g(u_2)$
1.0
0.5
0.0
0
1
2
Utility of new school (u_2)

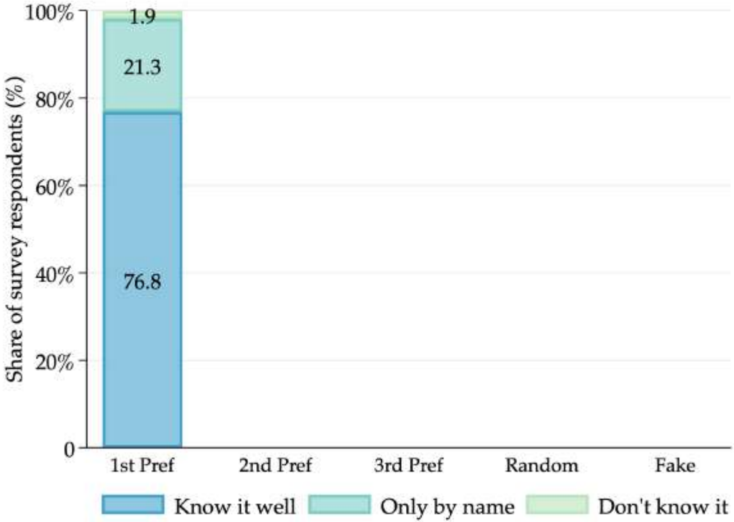
Data + Experiments Outline

- Surveys measure awareness, perceptions (known schools), beliefs (unknown schools)
  1. Baseline: also has subj. ROL, perceived returns to search
  2. Midline: gives repeated perception measures
  3. Endline: awareness and perceptions only
- Sample: recruited from preschools; restrict to first-time preK / K / 1st applicants
  - In study if and only if complete baseline survey
- Timing:
  - 1 Baseline survey. Search RCT. Explorer made available. Almost all on-platform search.
  - 2 Feedback RCT: uses submitted apps, 1 week before deadline.
  - 2-3 Midline survey occurs slightly before/after (timing varied).
  - 3 Final apps due; endline survey post-deadline.
- Admin data: demographics and repeated measures of rank-order lists:
  - Three snapshots: baseline (survey), "just before feedback" and final (admin).

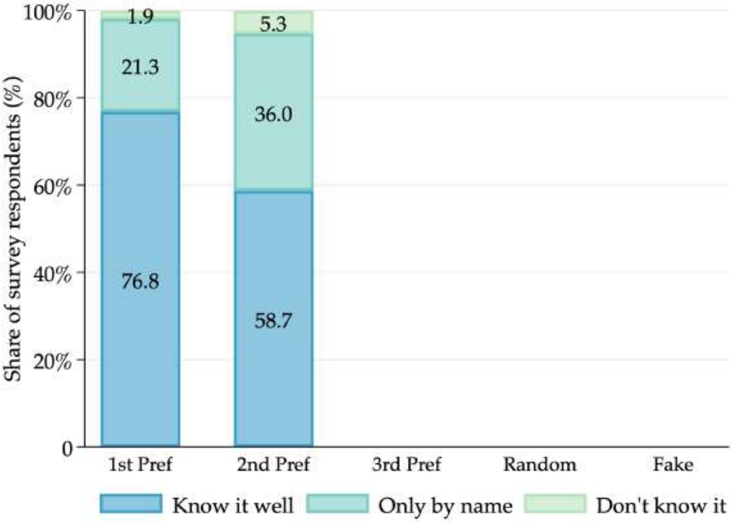
Descriptive Analysis

- **Awareness:**
  - Families don't know all the schools at baseline **[TODAY]**
  - Closer/higher quality schools more likely to be known ▸ Figure
- **Perceptions of known schools' characteristics:**
  - Key vertical characteristic: official gov't quality index $\in \{1, \ldots, 4\}$.
  - Households overestimate quality of schools they like, especially 1st choice **[TODAY]**
  - Households also mispredict price (too high), admissions chances (compression) ▸ Figure
- **Beliefs about unknown schools:**
  - Households overestimate quality and price of unknown schools ▸ Figure
- **Preferences:**
  - ROL explained by subjective perceptions of quality and price, not truth ▸ Figure

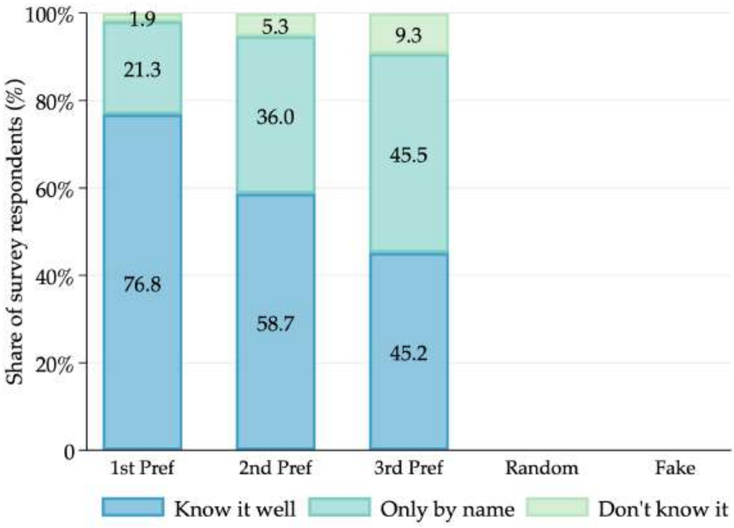
Families don't know all schools at baseline



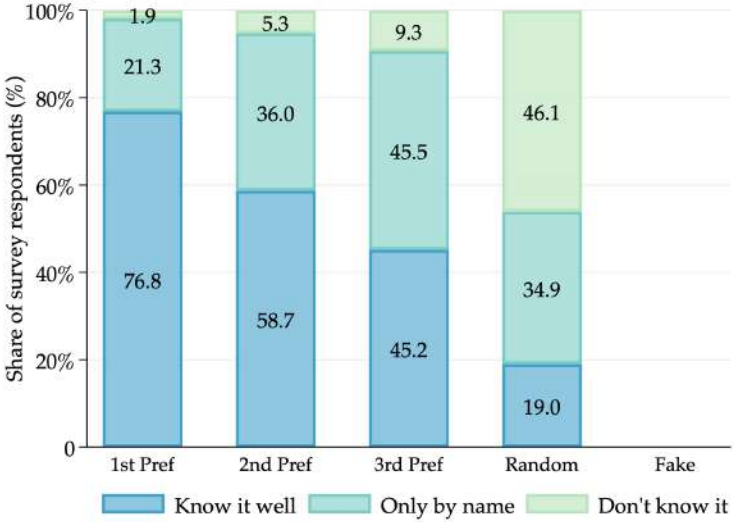
Families don't know all schools at baseline



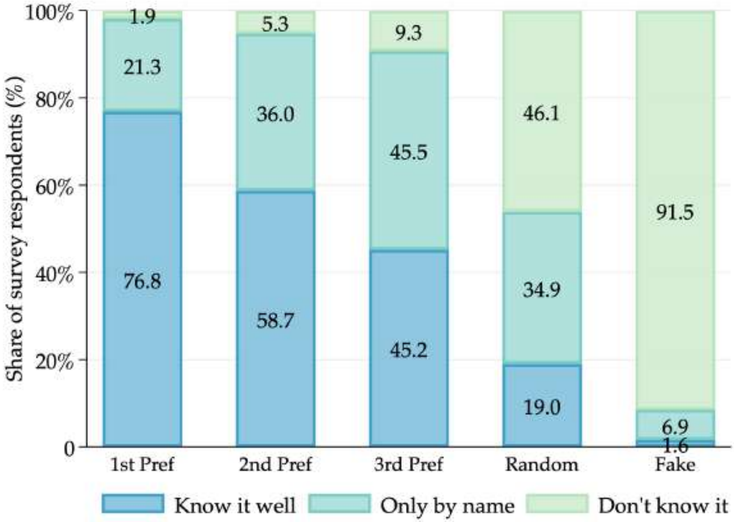
Families don't know all schools at baseline



Families don't know all schools at baseline

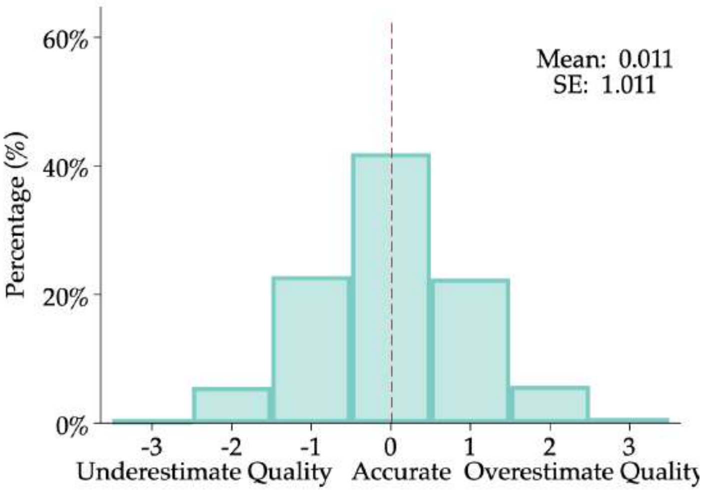


Families don't know all schools at baseline



Families overestimate quality of known, liked schools

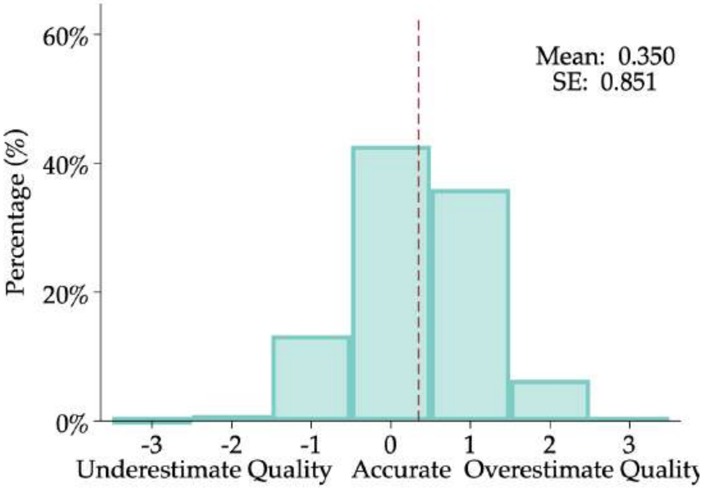
Winner's curse?



Random not-in-app

Families overestimate quality of known, liked schools

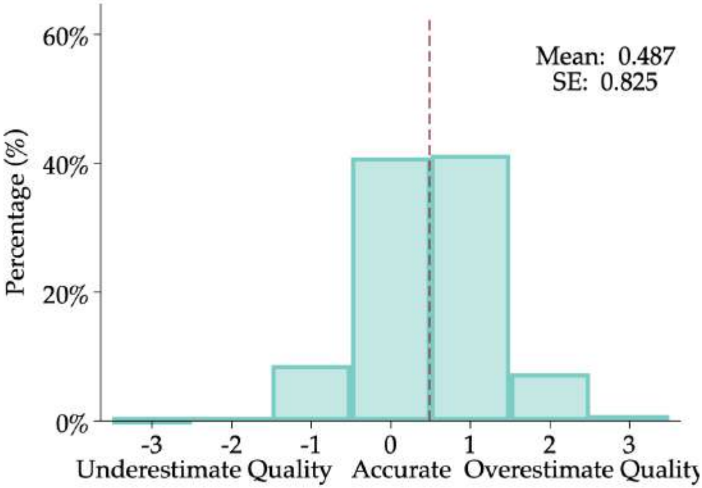
Winner's curse?



Random in-app

Families overestimate quality of known, liked schools

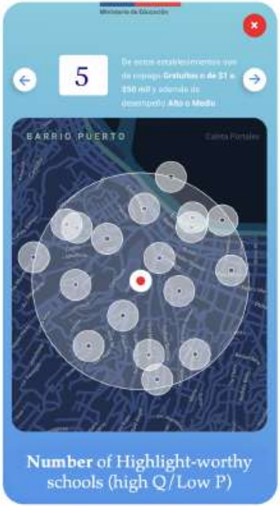
Winner's curse?



1st Preference

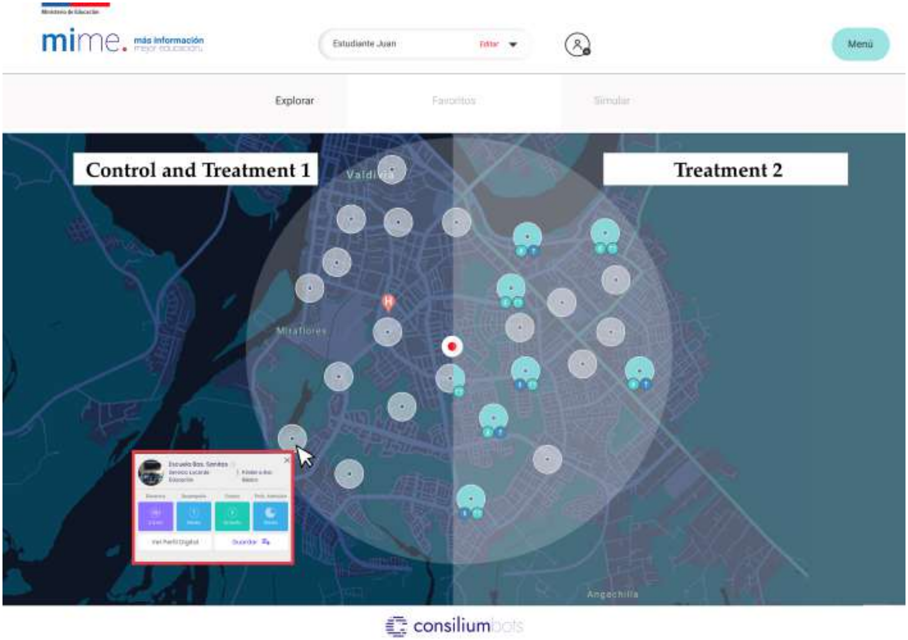
Search Experiment:

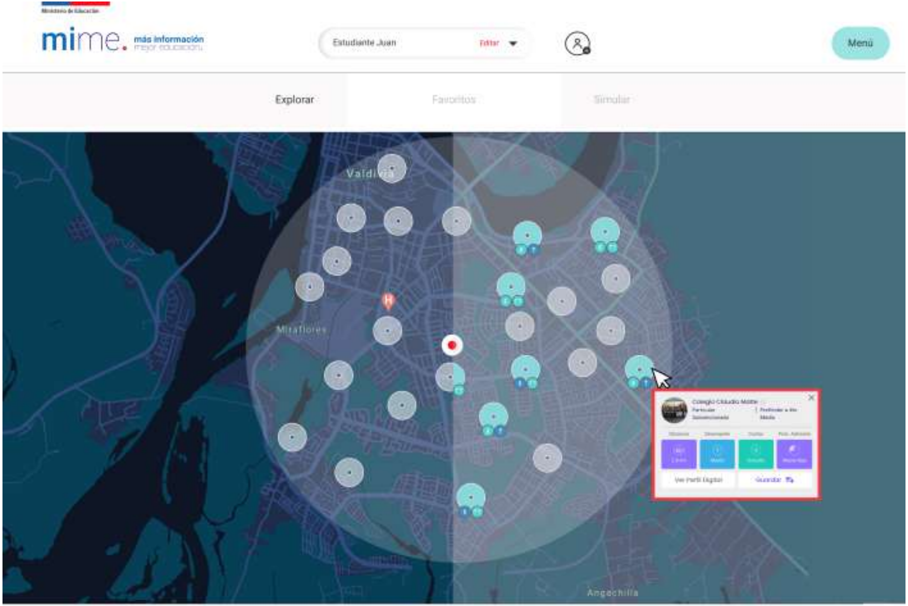
Treatment 1 and 2



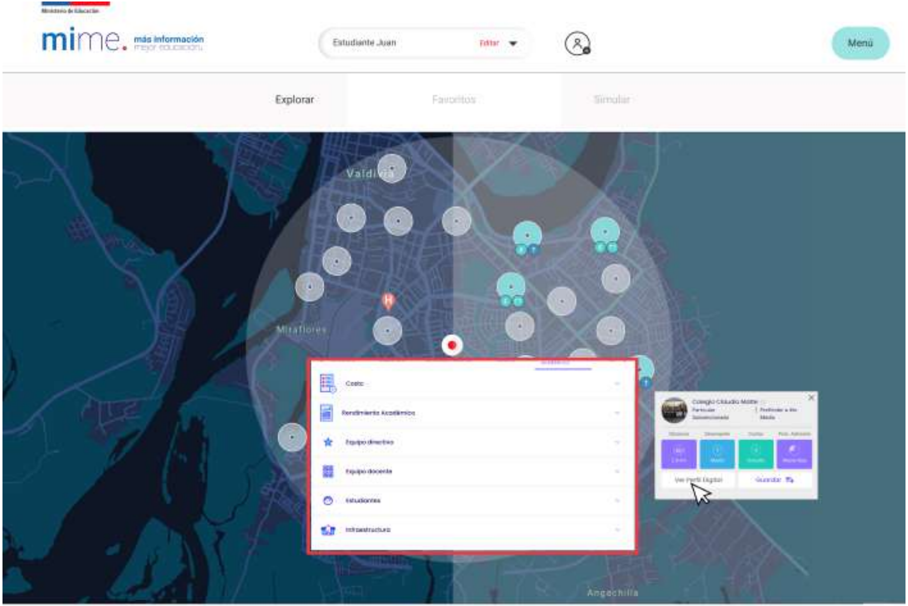
Treatment 1

Treatment 2

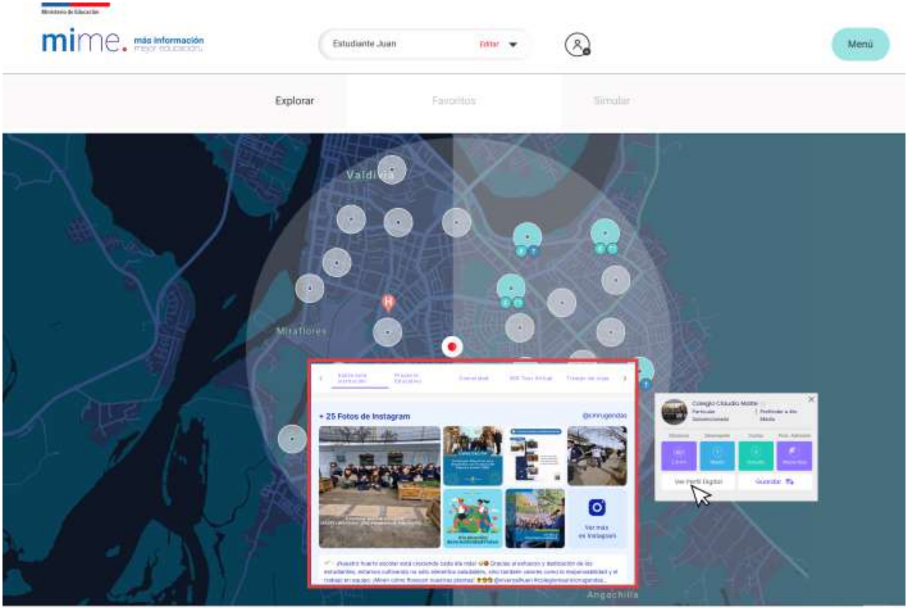

mime. más información
Estudiante Juan
Editar
Menú
Explorar
Favoritos
Simular
Control and Treatment 1
Treatment 2
Valdivia
Miraflores
Angachilla
consiliumbots

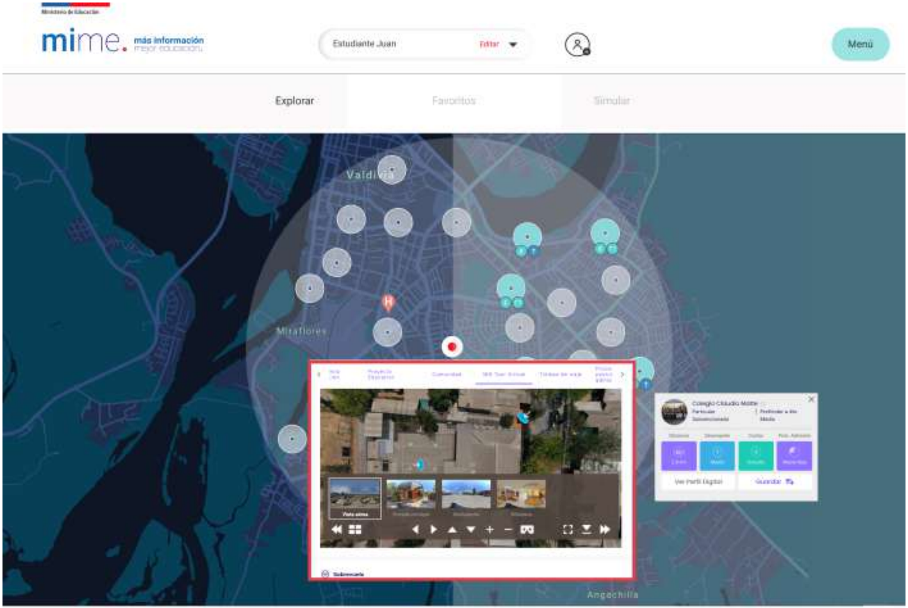


consiliumbots



consiliumbots


mime. más información
Estudiante Juan
Editar
Menú
Explorar
Favoritos
Simular
Valdivia
Miraflores
+ 25 Fotos de Instagram
Angachilla



consiliumbots

Experimental Results in Paper

Search Interventions: ▸ Table

- T1: Inform $X's$ distr. → Effects on beliefs (L/H), knowledge (H), search (H)
  → No effects on application outcomes
- T2: Inform $X's$ distr. → Effects on beliefs (L/H), knowledge (H), applications (H)
  + search tech. → No effects on "number of searches"

Feedback Intervention: ▸ Table

- T: Inform $X's$ + recs → Effects on perceptions (Q for H/L, P for L)
  → Effects on applications (L) and assignment (L)

Search intervention effects concentrated among H (high-SES), Feedback intervention among L (low-SES)

Model Overview

- Will present (and estimate) model in two steps:

1. Preferences, awareness, perceptions of known schools' characteristics:
    - Suffices for demand under counterfactuals with given info assignment (e.g. full info)
2. Admissions chances, beliefs about unknowns, search costs and technology:
    - Needed for endogenous info acquisition, e.g. change in info before search decisions

Step 1 Details

- Information:

$$\pi^*_{ijt} = z_{ij}\alpha^z + w_{ijt}\alpha^w + w^{rc}_{ijt}\alpha^{rc}_i + \eta_j + \nu_{ijt} \tag{1}$$

$$\pi_{ijt} = 1(\pi^*_{ijt} > 0) + 1(\pi^*_{ijt} > 1). \tag{2}$$

Step 1 Details

- Information:

$$\pi^*_{ijt} = z_{ij}\alpha^z + w_{ijt}\alpha^w + w^{rc}_{ijt}\alpha^{rc}_i + \eta_j + \nu_{ijt} \tag{1}$$

$$\pi_{ijt} = 1(\pi^*_{ijt} > 0) + 1(\pi^*_{ijt} > 1). \tag{2}$$

- Payoffs and Perceptions:

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

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

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

Step 1 Details

- Information:

$$\pi^*_{ijt} = z_{ij}\alpha^z + w_{ijt}\alpha^w + w^{rc}_{ijt}\alpha^{rc}_i + \eta_j + \nu_{ijt} \tag{1}$$

$$\pi_{ijt} = 1(\pi^*_{ijt} > 0) + 1(\pi^*_{ijt} > 1). \tag{2}$$

- Payoffs and Perceptions:

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

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

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

- π^*_{ijt}: index for awareness (can't apply if $\pi^*_{ijt} < 0$), perceptions (more accurate if $\pi^*_{ijt} > 1$.)

Step 1 Details

- Information:

$$\pi^*_{ijt} = z_{ij}\alpha^z + w_{ijt}\alpha^w + w^{rc}_{ijt}\alpha^{rc}_i + \eta_j + \nu_{ijt} \tag{1}$$
$$\pi_{ijt} = 1(\pi^*_{ijt} > 0) + 1(\pi^*_{ijt} > 1). \tag{2}$$

- Payoffs and Perceptions:

$$\hat{u}^{(2)}_{ij} = z_{ij}\beta^z + \hat{x}^{2,rc}_{ij}\beta^x_i + \delta_j + \hat{x}^{(2)}_{ij}\gamma + \varepsilon_{ij} \tag{3}$$
$$\hat{u}^{(1)}_{ij} = z_{ij}\beta^z + \hat{x}^{1,rc}_{ij}\beta^x_i + \delta_j + \hat{x}^{(1)}_{ij}\gamma + \hat{\varepsilon}^{(1)}_{ij} \tag{4}$$
$$\hat{x}^{(1)}_{ij} \sim \Gamma(\cdot|x_j), \quad \hat{x}^{(2)}_{ij} = (x_j \text{ w.p. } p^h, \text{ otherwise } \hat{x}^{(1)}_{ij}) \tag{5}$$

- π^*_{ijt}: index for awareness (can't apply if $\pi^*_{ijt} < 0$), perceptions (more accurate if $\pi^*_{ijt} > 1$.)
  - z: distance

Step 1 Details

- Information:

$$\pi^*_{ijt} = z_{ij}\alpha^z + w_{ijt}\alpha^w + w^{rc}_{ijt}\alpha^{rc}_i + \eta_j + \nu_{ijt} \tag{1}$$

$$\pi_{ijt} = 1(\pi^*_{ijt} > 0) + 1(\pi^*_{ijt} > 1). \tag{2}$$

- Payoffs and Perceptions:

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

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

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

- π^*_{ijt}: index for awareness (can't apply if $\pi^*_{ijt} < 0$), perceptions (more accurate if $\pi^*_{ijt} > 1$.)
  - Excluded info shifters w: treatment indicators, search (pin clicks, profile views)

Step 1 Details

- Information:

$$\pi^*_{ijt} = z_{ij}\alpha^z + w_{ijt}\alpha^w + w^{rc}_{ijt}\alpha^{rc}_i + \eta_j + \nu_{ijt} \tag{1}$$

$$\pi_{ijt} = 1(\pi^*_{ijt} > 0) + 1(\pi^*_{ijt} > 1). \tag{2}$$

- Payoffs and Perceptions:

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

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

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

- π^*_{ijt}: index for awareness (can't apply if $\pi^*_{ijt} < 0$), perceptions (more accurate if $\pi^*_{ijt} > 1$.)
  - Off-platform learning: shocks $\nu_{ij} \sim N(0, \Sigma^\nu)$, time indicators $w^{rc}_{ijt} \sim N(\mu^{rc,\pi}, \Sigma^{rc,\pi})$.

Step 1 Details

- Information:

$$\pi^*_{ijt} = z_{ij}\alpha^z + w_{ijt}\alpha^w + w^{rc}_{ijt}\alpha^{rc}_i + \eta_j + \nu_{ijt} \tag{1}$$
$$\pi_{ijt} = 1(\pi^*_{ijt} > 0) + 1(\pi^*_{ijt} > 1). \tag{2}$$

- Payoffs and Perceptions:

$$\hat{u}^{(2)}_{ij} = z_{ij}\beta^z + \hat{x}^{2,rc}_{ij}\beta^x_i + \delta_j + \hat{x}^{(2)}_{ij}\gamma + \varepsilon_{ij} \tag{3}$$
$$\hat{u}^{(1)}_{ij} = z_{ij}\beta^z + \hat{x}^{1,rc}_{ij}\beta^x_i + \delta_j + \hat{x}^{(1)}_{ij}\gamma + \hat{\varepsilon}^{(1)}_{ij} \tag{4}$$
$$\hat{x}^{(1)}_{ij} \sim \Gamma(\cdot|x_j), \quad \hat{x}^{(2)}_{ij} = (x_j \text{ w.p. } p^h, \text{ otherwise } \hat{x}^{(1)}_{ij}) \tag{5}$$

- $\hat{u}^{\pi_{ijt}}_{ij}$: subj. expected payoff of j given i's info at time t.
  - True price and quality category $x \in \{1, \ldots, 4\}^2$. Subj. perceptions: $\hat{x}$.

Step 1 Details

- Information:

$$\pi^*_{ijt} = z_{ij}\alpha^z + w_{ijt}\alpha^w + w^{rc}_{ijt}\alpha^{rc}_i + \eta_j + \nu_{ijt} \tag{1}$$

$$\pi_{ijt} = 1(\pi^*_{ijt} > 0) + 1(\pi^*_{ijt} > 1). \tag{2}$$

- Payoffs and Perceptions:

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

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

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

- $\hat{u}^{\pi_{ijt}}_{ij}$: subj. expected payoff of j given i's info at time t.
  - Shock $\varepsilon_{ij} \sim N(0, \sigma^2)$. Low-info $\hat{\varepsilon}^1_{ij}$: classical meas. error; shrink using misspecified model.

Step 1 Details

- Information:

$$\pi^*_{ijt} = z_{ij}\alpha^z + w_{ijt}\alpha^w + w^{rc}_{ijt}\alpha^{rc}_i + \eta_j + \nu_{ijt} \tag{1}$$

$$\pi_{ijt} = 1(\pi^*_{ijt} > 0) + 1(\pi^*_{ijt} > 1). \tag{2}$$

- Payoffs and Perceptions:

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

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

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

- $\hat{u}^{\pi_{ijt}}_{ij}$: subj. expected payoff of j given i's info at time t.
  - RC's: $\hat{x}, 1, z$.

Step 1 Details

- Information:

$$\pi^*_{ijt} = z_{ij}\alpha^z + w_{ijt}\alpha^w + w^{rc}_{ijt}\alpha^{rc}_i + \eta_j + \nu_{ijt} \quad (1)$$

$$\pi_{ijt} = 1(\pi^*_{ijt} > 0) + 1(\pi^*_{ijt} > 1). \quad (2)$$

- Payoffs and Perceptions:

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

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

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

- $\hat{u}^{\pi_{ijt}}_{ij}$: subj. expected payoff of j given i's info at time t.
    - Mean utility and "discoverability" (δ, η): correlated RE; means depend on (true) x.

Model and Estimation

- Step 1 Estimation: Gibbs Sampler

Model and Estimation

- Step 1 Estimation: Gibbs Sampler
- Exploit repeated within-person measurements of $\hat{x}, \pi$, rankings.
  - Many objects (awareness, perceptions) observed.
  - $\hat{\varepsilon}$: are "known by name" schools excessively penalized / overdispersed?

Model and Estimation

- Step 1 Estimation: Gibbs Sampler
- Exploit repeated within-person measurements of $\hat{x}, \pi$, rankings.
  - Many objects (awareness, perceptions) observed.
  - $\hat{\varepsilon}$: are "known by name" schools excessively penalized / overdispersed?
- Allow meas. error on every survey variable.

Model and Estimation

- Step 1 Estimation: Gibbs Sampler
- Exploit repeated within-person measurements of $\hat{x}, \pi$, rankings.
  - Many objects (awareness, perceptions) observed.
  - $\hat{\varepsilon}$: are "known by name" schools excessively penalized / overdispersed?
- Allow meas. error on every survey variable.
- ID: info shifters are (random, exogenous) treatments, (endog.) clicks.
  - "Change of variables, then DiD".

Model and Estimation

- Step 1 Estimation: Gibbs Sampler
- Exploit repeated within-person measurements of $\hat{x}, \pi$, rankings.
  - Many objects (awareness, perceptions) observed.
  - $\hat{\varepsilon}$: are "known by name" schools excessively penalized / overdispersed?
- Allow meas. error on every survey variable.
- ID: info shifters are (random, exogenous) treatments, (endog.) clicks.
  - "Change of variables, then DiD".

- Step 2: model admissions optimism and compression; subj. beliefs $\hat{F}(x)$; subj. dist'n of $(\hat{\varepsilon}_1, \varepsilon)$; click probs if search; search costs.

Model and Estimation

- Step 1 Estimation: Gibbs Sampler
- Exploit repeated within-person measurements of $\hat{x}, \pi$, rankings.
  - Many objects (awareness, perceptions) observed.
  - $\hat{\varepsilon}$: are "known by name" schools excessively penalized / overdispersed?
- Allow meas. error on every survey variable.
- ID: info shifters are (random, exogenous) treatments, (endog.) clicks.
  - "Change of variables, then DiD".

- Step 2: model admissions optimism and compression; subj. beliefs $\hat{F}(x)$; subj. dist'n of $(\hat{\varepsilon}_1, \varepsilon)$; click probs if search; search costs.
  - Search is sequential.

Model and Estimation

- Step 1 Estimation: Gibbs Sampler
- Exploit repeated within-person measurements of $\hat{x}, \pi$, rankings.
  - Many objects (awareness, perceptions) observed.
  - $\hat{\varepsilon}$: are "known by name" schools excessively penalized / overdispersed?
- Allow meas. error on every survey variable.
- ID: info shifters are (random, exogenous) treatments, (endog.) clicks.
  - "Change of variables, then DiD".

- Step 2: model admissions optimism and compression; subj. beliefs $\hat{F}(x)$; subj. dist'n of $(\hat{\varepsilon}_1, \varepsilon)$; click probs if search; search costs.
  - Search is sequential.
  - conditional on "one more pin click", discovered school is stochastic.

Model and Estimation

- Step 1 Estimation: Gibbs Sampler
- Exploit repeated within-person measurements of $\hat{x}, \pi$, rankings.
  - Many objects (awareness, perceptions) observed.
  - $\hat{\varepsilon}$: are "known by name" schools excessively penalized / overdispersed?
- Allow meas. error on every survey variable.
- ID: info shifters are (random, exogenous) treatments, (endog.) clicks.
  - "Change of variables, then DiD".

- Step 2: model admissions optimism and compression; subj. beliefs $\hat{F}(x)$; subj. dist'n of $(\hat{\varepsilon}_1, \varepsilon)$; click probs if search; search costs.
  - Search is sequential.
  - conditional on "one more pin click", discovered school is stochastic.
  - "one-step lookahead" heuristic: search if subj. gain $E(\hat{U}_i(\pi')|\pi_{it}) - \hat{U}(\pi_{it})$ exceeds cost.

Model and Estimation

- Step 1 Estimation: Gibbs Sampler
- Exploit repeated within-person measurements of $\hat{x}, \pi$, rankings.
  - Many objects (awareness, perceptions) observed.
  - $\hat{\varepsilon}$: are "known by name" schools excessively penalized / overdispersed?
- Allow meas. error on every survey variable.
- ID: info shifters are (random, exogenous) treatments, (endog.) clicks.
  - "Change of variables, then DiD".

- Step 2: model admissions optimism and compression; subj. beliefs $\hat{F}(x)$; subj. dist'n of $(\hat{\varepsilon}_1, \varepsilon)$; click probs if search; search costs.
  - Search is sequential.
  - conditional on "one more pin click", discovered school is stochastic.
  - "one-step lookahead" heuristic: search if subj. gain $E(\hat{U}_i(\pi')|\pi_{it}) - \hat{U}(\pi_{it})$ exceeds cost.
  - Beliefs over unknowns: latent heterogeneity; nests Bayesian updating

Model and Estimation

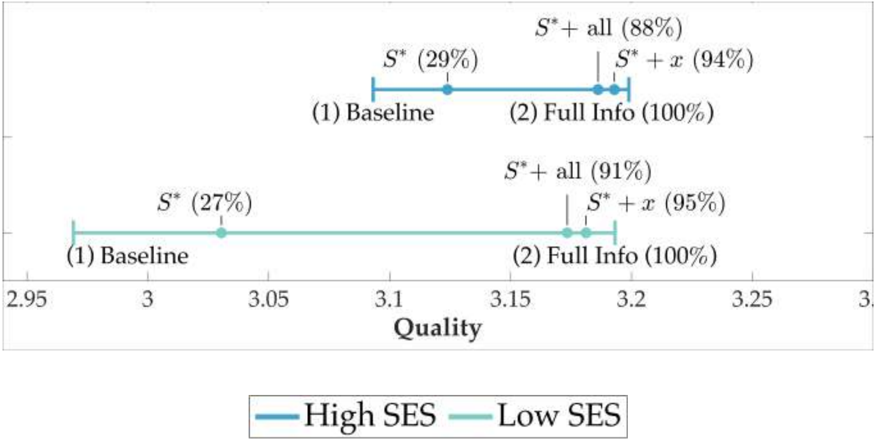
- Step 1 Estimation: Gibbs Sampler
- Exploit repeated within-person measurements of $\hat{x}, \pi$, rankings.
  - Many objects (awareness, perceptions) observed.
  - $\hat{\varepsilon}$: are "known by name" schools excessively penalized / overdispersed?
- Allow meas. error on every survey variable.
- ID: info shifters are (random, exogenous) treatments, (endog.) clicks.
  - "Change of variables, then DiD".

- Step 2: model admissions optimism and compression; subj. beliefs $\hat{F}(x)$; subj. dist'n of $(\hat{\varepsilon}_1, \varepsilon)$; click probs if search; search costs.
  - Search is sequential.
  - conditional on "one more pin click", discovered school is stochastic.
  - "one-step lookahead" heuristic: search if subj. gain $E(\hat{U}_i(\pi')|\pi_{it}) - \hat{U}(\pi_{it})$ exceeds cost.
  - Beliefs over unknowns: latent heterogeneity; nests Bayesian updating
- Estimation: MLE/SMLE.

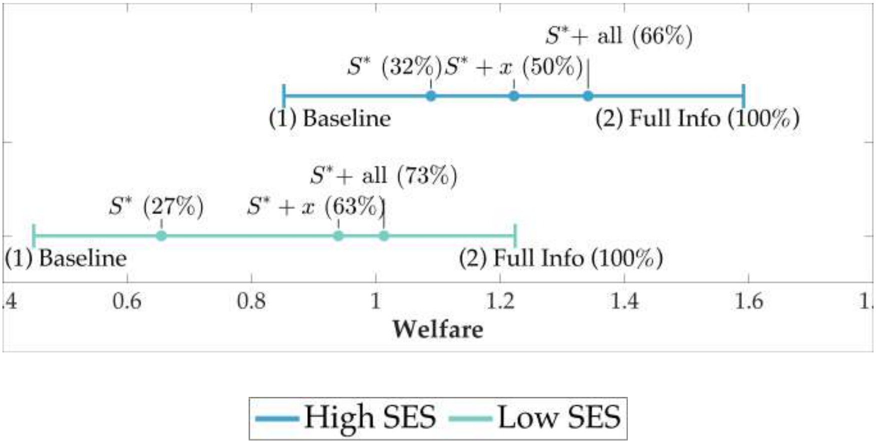
Counterfactuals Overview

1. **Gains from full information + decomposition:**
   - Base simulation: remove treatments
   - Full information: full information about all schools + we correct all misperceptions
   - Better search (S^*): improves and simplifies the search technology
   - Better search (S^*) + correct biases, misperceptions, and imperfect information
2. **Search Activity and Search Costs:**
   - Does individual level search change with counterfactuals?
   - Gradual Reduction in Search Costs
3. **Misspecified models:**
   - Estimate the model dropping data on $\hat{x}$ and ignoring misperceptions
   - Effects on countrerfactual analysis

Gains of Full Information + Decomposition



Gains of Full Information + Decomposition



Other Results in Paper:

- Improving search technology induces 47% more search.
- $S^* \to S^* + X$: zero avg. effect on search, but large absolute changes; some search more, others less.
- Gradual reduction in search costs: need to almost eliminate to beat better info.
- Misspecified models: assume $\hat{x} = x$, get *wrong sign* of quality impacts of info provision.

Conclusions

- **Results:** Households' inaccurate perceptions distort search and applications
  - Households value quality — but respond to perception, not truth.
  - Systematically overestimate quality of initial most-preferred "known" school.
- **Counterfactuals:** Perfect takeup of info intervention would close quality SES gap
  - Differences between groups: perceptions of x's, not admissions optimism or prefs.
  - But in practice, high-education households respond more to our search intervention.
- **Methods:** Crucial to model biases/imperfect awareness of "known" options.
- **Agenda:** This paper takes schools' quality, peers as given. Input for eqbm analysis.