School Choice under Endogenous Awareness Sets: High Search Costs or Biased Beliefs?

Christopher Neilson
(with Patrick Agte, Claudia Allende and Adam Kapor)

January 3, 2020
Coauthors

**Claudia Allende Santa Cruz**
Columbia University
PhD student, *On the Market 2019!*
Industrial Organization, Education and Labor Economics.
Coauthors

**Claudia Allende Santa Cruz**
Columbia University
PhD student, *On the Market 2019!*
Industrial Organization, Education and Labor Economics.

**Patrick Agte**
Princeton University
PhD student, *Coming on the market soon!*
Development Economics and Industrial Organization.
Motivation

- Families sometimes send their children to low performing schools when better options are available. Why?
Motivation

- Families sometimes send their children to low performing schools when better options are available. Why?

- Do families not care about school quality? Unlikely whole story.
Motivation

- Families sometimes send their children to low performing schools when better options are available. Why?
- Do families not care about school quality? Unlikely whole story.
- Do they not know that nearby schools are better?
Motivation

- Families sometimes send their children to low performing schools when better options are available. Why?
- Do families not care about school quality? Unlikely whole story.
- Do they not know that nearby schools are better? Maybe.
Motivation

- Families sometimes send their children to low performing schools when better options are available. Why?
- Do families not care about school quality? Unlikely whole story.
- Do they not know that nearby schools are better? Maybe. But why don’t parents search more?
Motivation

- Families sometimes send their children to low performing schools when better options are available. Why?

- Do families not care about school quality? Unlikely whole story.

- Do they not know that nearby schools are better? Maybe. But why don’t parents search more?
  - Acquiring information is costly?
Motivation

- Families sometimes send their children to low performing schools when better options are available. Why?

- Do families not care about school quality? Unlikely whole story.

- Do they not know that nearby schools are better? Maybe. But why don’t parents search more?
  - Acquiring information is costly? Search costs $\rightarrow$ unrealistically high to explain the full behavior.
Families sometimes send their children to low performing schools when better options are available. Why?

Do families not care about school quality? Unlikely whole story.

Do they not know that nearby schools are better? Maybe. But why don’t parents search more?

- Acquiring information is costly? Search costs → unrealistically high to explain the full behavior.
- **Alternative Idea**: Families have incorrect beliefs about the distribution of schools.
Santo Domingo, Dominican Republic
This Project

- Research Question:
  - Quantify to what extent households underinvest in search due to preferences, biased beliefs and search costs.
This Project

- Research Question:
  ⇒ Quantify to what extent households underinvest in search due to preferences, biased beliefs and search costs.

- Framework:
  ⇒ Construct an empirical model of costly sequential search that allows for biased beliefs about the joint distribution of schools’ quality and price.
This Project

- **Research Question:**
  ⇒ Quantify to what extent households underinvest in search due to preferences, biased beliefs and search costs.

- **Framework:**
  ⇒ Construct an empirical model of costly sequential search that allows for biased beliefs about the joint distribution of schools’ quality and price.

- **Data:**
  ⇒ Conduct household survey to inform the model: collect data on preferences, beliefs and awareness sets.
This Project

- Research Question:
  ⇒ Quantify to what extent households underinvest in search due to preferences, biased beliefs and search costs.

- Framework:
  ⇒ Construct an empirical model of costly sequential search that allows for biased beliefs about the joint distribution of schools’ quality and price.

- Data:
  ⇒ Conduct household survey to inform the model: collect data on preferences, beliefs and awareness sets.

- Identification:
  ⇒ Design a field experiment that creates variation to identify key parameters of the model.
This Project

- **Research Question:**
  ⇒ Quantify to what extent households underinvest in search due to preferences, biased beliefs and search costs.

- **Framework:**
  ⇒ Construct an empirical model of costly sequential search that allows for biased beliefs about the joint distribution of schools’ quality and price.

- **Data:**
  ⇒ Conduct household survey to inform the model: collect data on preferences, beliefs and awareness sets.

- **Identification:**
  ⇒ Design a field experiment that creates variation to identify key parameters of the model.

- **Status:**
  ⇒ Pilot study in the Dominican Republic, won an NSF grant in 2019 to implement the full study in spring 2020.
Framework: How do parents choose schools?

- Parents \( (i) \) have preferences over school \( (j) \) characteristics.
Parents \((i)\) have preferences over school \((j)\) characteristics.

- Observables \((x_j)\): price, quality, full day, location
Parents \((i)\) have preferences over school \((j)\) characteristics.

- Observables \((x_j : \text{price, quality, full day, location})\)
- Unobservables \((\xi_j)\)
Parents \((i)\) have preferences over school \((j)\) characteristics.

- Observables \((x_j: \text{price, quality, full day, location})\)
- Unobservables \((\xi_j)\)
- i.i.d. preference shocks \((\epsilon_{ij})\)
Framework: How do parents choose schools?

- Parents \((i)\) have preferences over school \((j)\) characteristics.
  - Observables \((x_j : \text{price, quality, full day, location})\)
  - Unobservables \((\xi_j)\)
  - i.i.d. preference shocks \((\epsilon_{ij})\)

- The utility of households \(i\) from choosing school \(j\) is given by:

\[
u_{ij} = x_j' \gamma + \xi_j + \epsilon_{ij}\]
Framework: How do parents choose schools?

Parents ($i$) have preferences over school ($j$) characteristics.

- Observables ($x_j$ : price, quality, full day, location)
- Unobservables ($\xi_j$)
- i.i.d. preference shocks ($\epsilon_{ij}$)

The utility of households $i$ from choosing school $j$ is given by:

$$u_{ij} = x_j' \gamma + \xi_j + \epsilon_{ij}$$

Parents choose the option with the highest utility among the schools that they know (the set $\Omega_i^T$):
Framework

Let $\Omega^t_i$ be the awareness set of family $i$ by time $t$. 
Let $\Omega^t_i$ be the **awareness set** of family $i$ by time $t$. At each $t$, parents observe all the characteristics of schools in $\Omega^t_i$. 
Let $\Omega^t_i$ be the awareness set of family $i$ by time $t$.

At each $t$, parents observe all the characteristics of schools in $\Omega^t_i$.

The school choice decision is made at $t = T$. 

Christopher Neilson

Search and Biased Beliefs

January 3, 2020 7 / 23
Let $\Omega^t_i$ be the **awareness set** of family $i$ by time $t$.

At each $t$, parents observe all the characteristics of schools in $\Omega^t_i$.

The school choice decision is made at $t = T$.

The set of schools in $\Omega^T_i$ is critical for the outcome of the school choice decision.
Let $\Omega^t_i$ be the \textbf{awareness set} of family $i$ by time $t$.

At each $t$, parents observe all the characteristics of schools in $\Omega^t_i$.

The school choice decision is made at $t = T$.

The set of schools in $\Omega^T_i$ is critical for the outcome of the school choice decision.

- How do parents form $\Omega^T_i$?
Let $\Omega_i^t$ be the **awareness set** of family $i$ by time $t$.

At each $t$, parents observe all the characteristics of schools in $\Omega_i^t$.

The school choice decision is made at $t = T$.

The set of schools in $\Omega_i^T$ is critical for the outcome of the school choice decision.

- How do parents form $\Omega_i^T$?
- What explains the sequence of search and the evolution of $\Omega_i^t$?
Model: How do parents search for schools and form?

Timing:

$t=0$ Parents inherit an initial information set that contains:

- The school awareness set at $t=0$: $(\Omega^0_i)$
- Beliefs over unknown schools' attributes: $(\tilde{F}^0_i(x))$
Model: How do parents search for schools and form?

Timing:

$t=0$ Parents inherit an initial information set that contains:

- The school awareness set at $t=0$: $(\Omega^0_i)$
- Beliefs over unknown schools' attributes: $(\tilde{F}^0_i(x))$

At each period, parents are able to enrich $\Omega^0_i$ and update $(\tilde{F}^0_i(x))$ to do so they must decide to engage in search:

- They can pay cost $\psi_i$ to draw a school from the true distribution: $(F(x))$
- Every searched school is added to the awareness set (perfect recall): $\Omega^{t+1}_i = \Omega^t_i \cup j$
Model: How do parents search for schools and form?

Timing:

\( t=0 \) Parents inherit an initial information set that contains:

- The school awareness set at \( t=0 \) : \( (\Omega_i^0) \)
- Beliefs over unknown schools’ attributes. \( (\tilde{F}_i^0(x)) \)
Model: How do parents search for schools and form?

Timing:

\( t=0 \) Parents inherit an initial information set that contains:

- The school awareness set at \( t=0 \) : \( (\Omega_i^0) \)
- Beliefs over unknown schools' attributes. \( (\tilde{F}_i^0(x)) \)

\( t=\{1,\ldots,T-1\} \) At each period, parents are able to enrich \( \Omega_i^0 \) and update \( \tilde{F}_i^0(x) \). To do so they must decide to engage in search:
Model: How do parents search for schools and form?

Timing:

$t=0$ Parents inherit an initial information set that contains:
- The school awareness set at $t=0$: $(\Omega^0_i)$
- Beliefs over unknown schools’ attributes: $(\tilde{F}^0_i(x))$

$t=\{1,\ldots,T-1\}$ At each period, parents are able to enrich $\Omega^0_i$ and update $\tilde{F}^0_i(x)$
To do so they must decide to engage in search:
- They can pay cost $\psi_i$ to draw a school from the true distribution: $(F(x))$
Model: How do parents search for schools and form?

Timing:

\( t=0 \) Parents inherit an initial information set that contains:

- The school awareness set at \( t=0 \): \((\Omega_i^0)\)
- Beliefs over unknown schools’ attributes: \((\tilde{F}_i^0(x))\)

\( t=\{1, \ldots, T-1\} \) At each period, parents are able to enrich \( \Omega_i^0 \) and update \( \tilde{F}_i^0(x) \).

To do so they must decide to engage in search:

- They can pay cost \( \psi_i \) to draw a school from the true distribution: \((F(x))\)
- Every searched school is added to the awareness set (perfect recall)

\[
\Omega_i^{t+1} = \Omega_i^t \cup j
\]
Optimal search is determined by a cutoff rule based on:

i. Search costs $\psi_i$

ii. The awareness set in $t$ ($\Omega_i^t$)

iii. Current beliefs. ($\tilde{F}^t_i (x)$)
Taking Model To Data

- Initial awareness set $\Omega_i^0$:
  - We assume it to be exogenous conditional on family characteristics (SES, location, social network)
  - We measure $\Omega_i^0$ and $\Omega_i^T$ directly in the baseline and endline survey of parents.
Taking Model To Data

- Initial awareness set $\Omega_i^0$:
  - We assume it to be exogenous conditional on family characteristics (SES, location, social network)
  - We measure $\Omega_i^0$ and $\Omega_i^T$ directly in the baseline and endline survey of parents.

- Beliefs over unknown schools’ attributes $\tilde{F}_i^t(x)$:
  - Beliefs $\tilde{F}_i^t(x)$ can change over time with search and random arrival of information.
  - We measure $\tilde{F}_i^0(x)$ and $\tilde{F}_i^T(x)$ directly in the surveys.
Taking Model To Data

- Initial awareness set $\Omega^0_i$:
  - We assume it to be exogenous conditional on family characteristics (SES, location, social network)
  - We measure $\Omega^0_i$ and $\Omega^T_i$ directly in the baseline and endline survey of parents.

- Beliefs over unknown schools’ attributes $\tilde{F}^t_i(x)$:
  - Beliefs $\tilde{F}^t_i(x)$ can change over time with search and random arrival of information.
  - We measure $\tilde{F}^0_i(x)$ and $\tilde{F}^T_i(x)$ directly in the surveys.

- Preference $\gamma_i, \xi_j$:
  - We measure preference rankings directly in the survey.
  - Observe choice in administrative data.
Setting

- Sample area: low-income neighborhoods in Santo Domingo.
- Decentralized school choice mechanism in which parents apply directly to primary schools.
- No fixed admission criteria.
- Many households receive a conditional cash transfer that requires that children aged 5+ years go to school.
Sample Area
Sample Area
Sample Area
Sample Area
What Do Parents Value?

**Figure: Rank of School Attributes**

Average Rank (3 = Most Important)

- Price: 0.5
- Quality: 1
- Full-Day: 1.5
- Location: 2.5

Christopher Neilson

Search and Biased Beliefs

January 3, 2020
How Do Parents Search?

- Share of Respondents
- Other Media
- Called School
- Talked to Friends
- Visited School

- Other Media: 0.01
- Called School: 0.03
- Talked to Friends: 0.42
- Visited School: 0.64
Belief Elicitation
Beliefs vs. True Distribution

**Figure**: True and Believed Quality Distribution

- **Private Schools**
  - Belief vs. Truth
  - Passing Rate vs. Quality Tercile

- **Public Schools**
  - Belief vs. Truth
  - Passing Rate vs. Quality Tercile

Christopher Neilson
Beliefs vs. True Distribution

Figure: True and Believed Joint Distribution (Private Schools)
There is significant price dispersion in the market.

Parents learn about new schools through visits and friends.

Parents’ beliefs are not aligned with the true distribution:

- Parents overestimate the quality of public schools.
- Parents overestimate the correlation between price and quality for private schools.
Sample: 1,000 households in Santo Domingo

Treatments

- C: Video + Brochure with general information on schooling.
- T1: Video + Brochure with information on the quality and price distribution of schools in their cluster.
- T2: Video + Brochure with information on the quality and price distribution of schools in their cluster + information on attributes of individual schools.
¿Cómo se relaciona el desempeño en Pruebas Nacionales con la mensualidad de las escuelas?

Para cada quintil de desempeño en Pruebas Nacionales se indica el porcentaje de escuelas en cada tramo de mensualidad.

<table>
<thead>
<tr>
<th>Grupo alto [66%-100%]</th>
<th>25</th>
<th>25</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grupo medio [33%-66%]</td>
<td>39</td>
<td>28</td>
<td>33</td>
</tr>
<tr>
<td>Grupo bajo [0%-33%]</td>
<td>63</td>
<td>32</td>
<td>5</td>
</tr>
</tbody>
</table>

Grado: 🔴 Gratis, 🔵 Entre $0 y $1500, 🔴 Más de $1500
## LAS ESCUELAS EN DETALLE

<table>
<thead>
<tr>
<th>Número</th>
<th>Nombre del Centro</th>
<th>Ofrece Nivel Inicial</th>
<th>Ofrece Media</th>
<th>Tanda</th>
<th>Estudiantes por Profesor</th>
<th>Cobro mensual</th>
<th>Aprobación en PN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Abejitas, Las</td>
<td>Sí</td>
<td>Sí</td>
<td>M</td>
<td>11</td>
<td>$2000</td>
<td>94</td>
</tr>
<tr>
<td>2</td>
<td>Angeles Sin Paraíso</td>
<td>No</td>
<td>No</td>
<td>V</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>C.E. Abc Santo Domingo Norte</td>
<td>Sí</td>
<td>Sí</td>
<td>M</td>
<td>17</td>
<td>$2500</td>
<td>96</td>
</tr>
<tr>
<td>4</td>
<td>C.E. Arena</td>
<td>Sí</td>
<td>No</td>
<td>M</td>
<td>12</td>
<td>$1800</td>
<td>99</td>
</tr>
<tr>
<td>5</td>
<td>C.E. Huellas Del Saber</td>
<td>No</td>
<td>No</td>
<td>M</td>
<td></td>
<td></td>
<td>75</td>
</tr>
<tr>
<td>6</td>
<td>C.E. Nuevo Mella</td>
<td>Sí</td>
<td>Sí</td>
<td>M</td>
<td>14</td>
<td>$1000</td>
<td>82</td>
</tr>
<tr>
<td>7</td>
<td>C.E. Vigotsky “Tia Hellen”</td>
<td>Sí</td>
<td>No</td>
<td>M</td>
<td>6</td>
<td>$2000</td>
<td>92</td>
</tr>
<tr>
<td>8</td>
<td>Centro Infantil Trincito De Colores</td>
<td>Sí</td>
<td>No</td>
<td>M</td>
<td>13</td>
<td>$800</td>
<td>96</td>
</tr>
<tr>
<td>9</td>
<td>Chaparral</td>
<td>Sí</td>
<td>No</td>
<td>V</td>
<td>51</td>
<td>$0</td>
<td>91</td>
</tr>
<tr>
<td>10</td>
<td>Cariaco Maria Fe Y Alejandro Marathon II</td>
<td>Sí</td>
<td>No</td>
<td>V</td>
<td>17</td>
<td>$0</td>
<td>95</td>
</tr>
<tr>
<td>11</td>
<td>Claribel</td>
<td>No</td>
<td>No</td>
<td>M</td>
<td></td>
<td></td>
<td>80</td>
</tr>
<tr>
<td>12</td>
<td>Colegio 2000</td>
<td>Sí</td>
<td>No</td>
<td>M</td>
<td>15</td>
<td>$600</td>
<td>74</td>
</tr>
<tr>
<td>13</td>
<td>Colegio Cristiano Hidekel</td>
<td>No</td>
<td>No</td>
<td>M</td>
<td></td>
<td></td>
<td>83</td>
</tr>
<tr>
<td>14</td>
<td>Colegio Santo Cristo</td>
<td>No</td>
<td>No</td>
<td>M</td>
<td></td>
<td></td>
<td>79</td>
</tr>
<tr>
<td>15</td>
<td>Colegio Vivire</td>
<td>Sí</td>
<td>No</td>
<td>M</td>
<td>10</td>
<td>$1000</td>
<td>100</td>
</tr>
</tbody>
</table>
## Pilot Results

### Treatment Effects on Knowledge

<table>
<thead>
<tr>
<th></th>
<th>Knows New Schools</th>
<th>Number of New Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>T1: Distributions</td>
<td>0.041</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>[0.061]</td>
<td>[0.061]</td>
</tr>
<tr>
<td>T2: Distributions + Info</td>
<td>0.056</td>
<td>0.077</td>
</tr>
<tr>
<td></td>
<td>[0.065]</td>
<td>[0.067]</td>
</tr>
<tr>
<td>Controls</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Mean of Control Group</td>
<td>0.327</td>
<td>0.327</td>
</tr>
<tr>
<td>Observations</td>
<td>350</td>
<td>350</td>
</tr>
</tbody>
</table>
Conclusion

- We propose a new mechanism for why parents make schooling choices with little information.

- Combination of search model and field experiment allows us to study how families search for schools endogenously.

- Use results to evaluate potential policies that could help reduce segregation and inequality in education systems.