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

PRINCETON UNIVERSITY

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

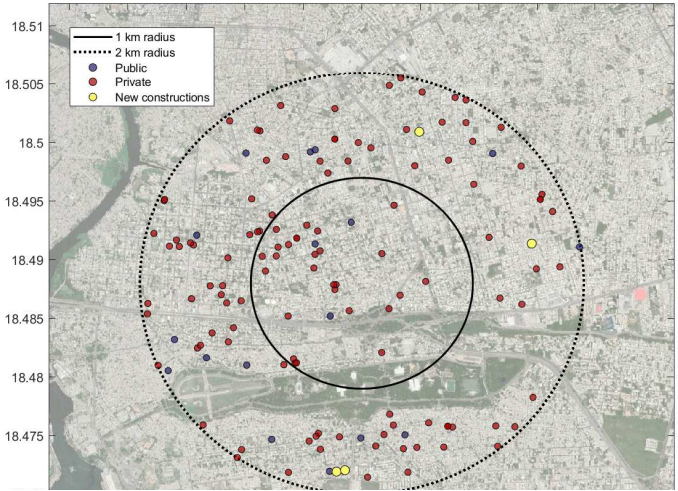
- Families sometimes send their children to low performing schools when better options are available. Why?
- Do families not care about school quality? Unlikely whole story.
- Do they not know that nearby schools are better? Maybe.
  But why don't parents search more?
  - Acquiring information is costly?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

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.
  - **Alternative Idea**: Families have incorrect beliefs about the distribution of schools.

Santo Domingo, Dominican Republic

Santo Domingo, Dominican Republic



This Project

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

This Project

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

- Framework:
  $\Rightarrow$ 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 :
  $\Rightarrow$ Quantify to what extent households underinvest in search due to **preferences**, **biased beliefs** and **search costs**.

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

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

This Project

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

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

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

- Identification:
  $\Rightarrow$ Design a field experiment that creates variation to identify key parameters of the model.

This Project

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

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

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

- Identification:
  $\Rightarrow$ Design a field experiment that creates variation to identify key parameters of the model.

- Status:
  $\Rightarrow$ 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.

Framework: How do parents choose schools?

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

Framework: How do parents choose schools?

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

Framework: How do parents choose schools?

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

Framework: How do parents choose schools?

- Parents (i) have preferences over school (j) characteristics.
  - Observables (x_j : price, quality, full day, location)
  - Unobservables (ξ_j)
  - i.i.d. preference shocks (ϵ_{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 (ξ_j)
  - i.i.d. preference shocks (ϵ_{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 Ω_i^T):

Framework

- Let Ω_i^t be the **awareness set** of family i by time t.

Framework

- Let Ω_i^t be the **awareness set** of family i by time t.
- At each t, parents observe all the characteristics of schools in Ω_i^t

Framework

- Let Ω_i^t be the **awareness set** of family i by time t.
- At each t, parents observe all the characteristics of schools in Ω_i^t
- The school choice decision is made at $t = T$

Framework

- Let Ω_i^t be the **awareness set** of family i by time t.
- At each t, parents observe all the characteristics of schools in Ω_i^t
- The school choice decision is made at $t = T$
- The set of schools in Ω_i^T is critical for the outcome of the school choice decision

Framework

- Let Ω_i^t be the **awareness set** of family i by time t.
- At each t, parents observe all the characteristics of schools in Ω_i^t
- The school choice decision is made at $t = T$
- The set of schools in Ω_i^T is critical for the outcome of the school choice decision
  - How do parents form Ω_i^T?

Framework

- Let Ω_i^t be the **awareness set** of family i by time t.
- At each t, parents observe all the characteristics of schools in Ω_i^t
- The school choice decision is made at $t = T$
- The set of schools in Ω_i^T is critical for the outcome of the school choice decision
  - How do parents form Ω_i^T?
  - What explains the sequence of search and the evolution of Ω_i^t?

Model: How do parents search for schools and form ?

Timing:

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

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 : (Ω_i^0)

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 : (Ω_i^0)
- Beliefs over unknown schools' attributes. $(\widetilde{F}_i^0(x))$

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 : (Ω_i^0)
- Beliefs over unknown schools' attributes. $(\widetilde{F}_i^0(x))$

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

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 : (Ω_i^0)
- Beliefs over unknown schools' attributes. $(\widetilde{F}_i^0(x))$

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

To do so they must decide to engage in search:

- They can pay cost ψ_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 : (Ω_i^0)
- Beliefs over unknown schools' attributes. $(\widetilde{F}_i^0(x))$

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

To do so they must decide to engage in search:

- They can pay cost ψ_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 ψ_i

ii. The awareness set in t (Ω_i^t)

iii. Current beliefs. $(\widetilde{F}_i^{t-1}(x))$

Taking Model To Data

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

Taking Model To Data

- Initial awareness set Ω_i^0:
  - We assume it to be exogenous conditional on family characteristics (SES, location, social network)
  - We **measure Ω_i^0 and Ω_i^T directly** in the baseline and endline survey of parents.
- Beliefs over unknown schools' attributes $\widetilde{F}_i^t(x)$:
  - Beliefs $\widetilde{F}_i^t(x)$ can change over time with search and random arrival of information.
  - We **measure $\widetilde{F}_i^0(x)$ and $\widetilde{F}_i^T(x)$ directly** in the surveys.

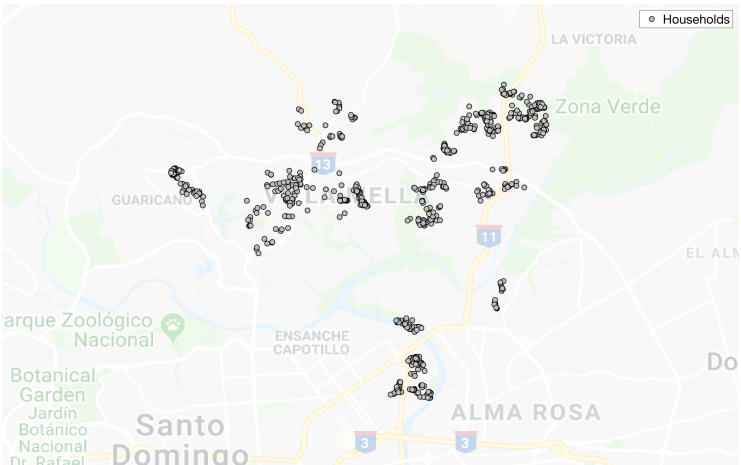
Taking Model To Data

- Initial awareness set Ω_i^0:
  - We assume it to be exogenous conditional on family characteristics (SES, location, social network)
  - We **measure Ω_i^0 and Ω_i^T directly** in the baseline and endline survey of parents.
- Beliefs over unknown schools' attributes $\widetilde{F}_i^t(x)$:
  - Beliefs $\widetilde{F}_i^t(x)$ can change over time with search and random arrival of information.
  - We **measure $\widetilde{F}_i^0(x)$ and $\widetilde{F}_i^T(x)$ directly** in the surveys.
- Preference γ_i, ξ_j:
  - We **measure preference rankings directly** in the survey.
  - Observe choice in **administrative data**.

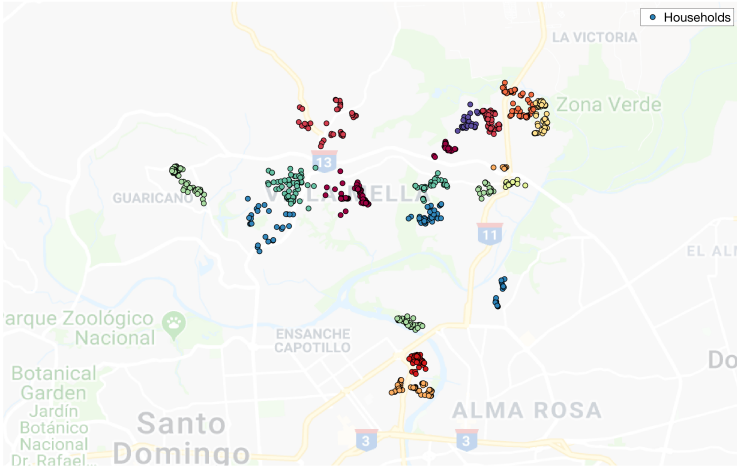
Setting

- 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.

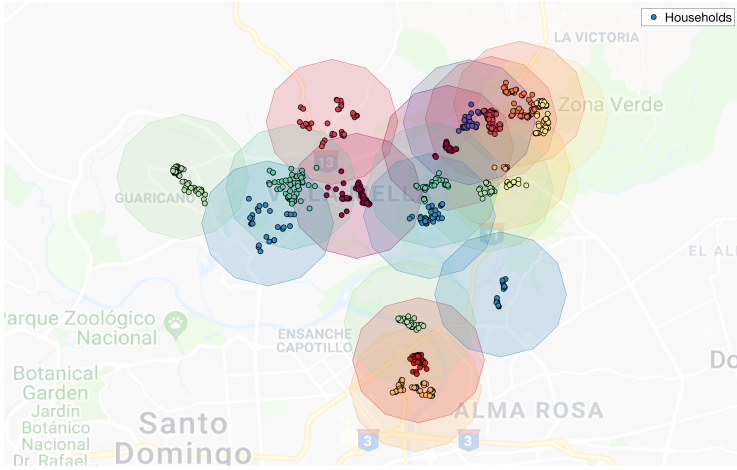
Sample Area


Households
LA VICTORIA
Zona Verde
GUARICANO
18
11
EL ALM
arque Zoológico
Nacional
ENSANCHE
CAPOTILLO
Botanical
Garden
Jardín
Botánico
Nacional
Dr. Rafael...
Santo
Domingo
ALMA ROSA
3
3

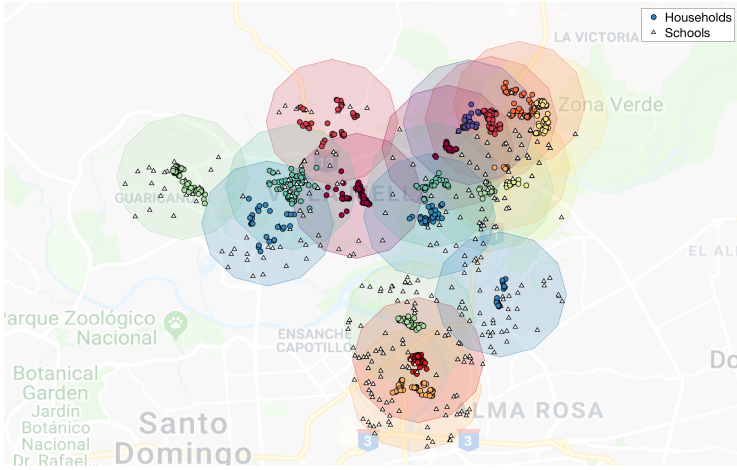
Sample Area


Households
LA VICTORIA
Zona Verde
GUARICANO
ENSANCHE
CAPOTILLO
ALMA ROSA
Botanical
Garden
Jardín
Botánico
Nacional
Santo
Domingo
3
11

Sample Area

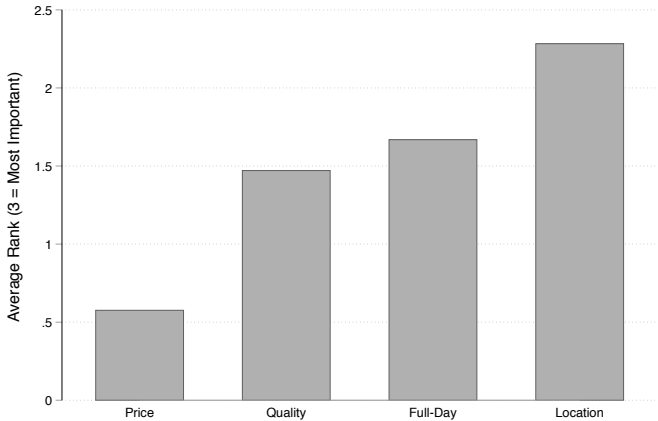

Households
LA VICTORIA
Zona Verde
GUARICANO
EL ALM
ENSANCHE
CAPOTILLO
Botanical
Garden
Jardín
Botánico
Nacional
Dr. Rafael...
Santo
Domingo
ALMA ROSA
3
3

Sample Area

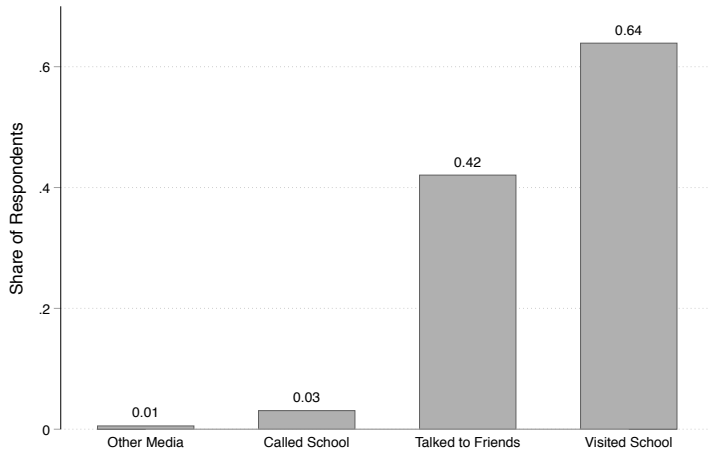

Households
Schools
LA VICTORIA
Zona Verde
EL ALM
Parque Zoológico Nacional
Botanical Garden
Jardín Botánico Nacional Dr. Rafael...
ENSANCHE CAPOTILLO
ALMA ROSA
Santo Domingo
3
3

What Do Parents Value?

Figure: Rank of School Attributes



How Do Parents Search?

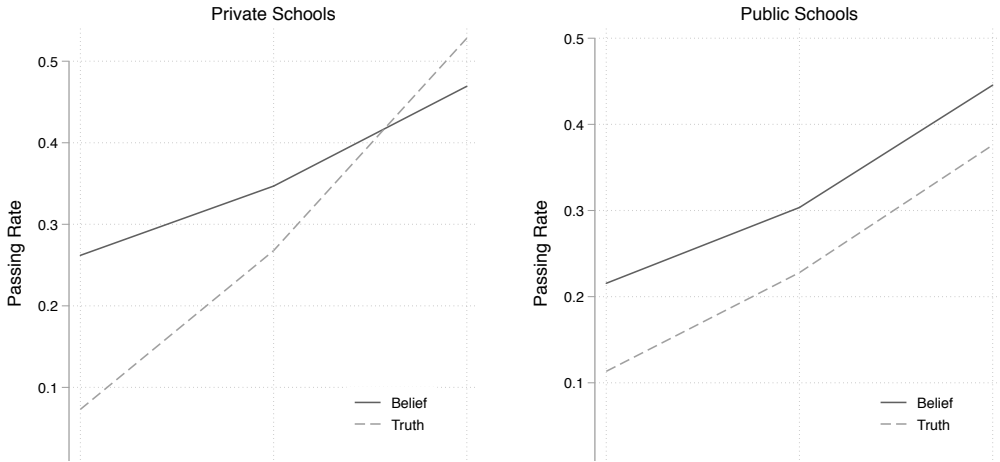


Belief Elicitation

3

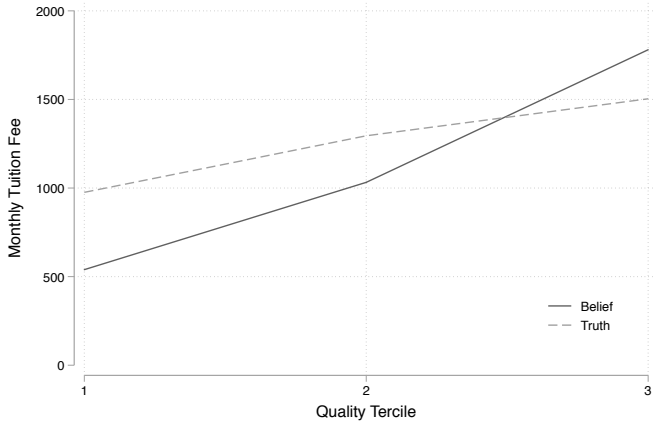
Beliefs vs. True Distribution

Figure: True and Believed Quality Distribution



Beliefs vs. True Distribution

Figure: True and Believed Joint Distribution (Private Schools)



Pilot Study - Stylized Facts

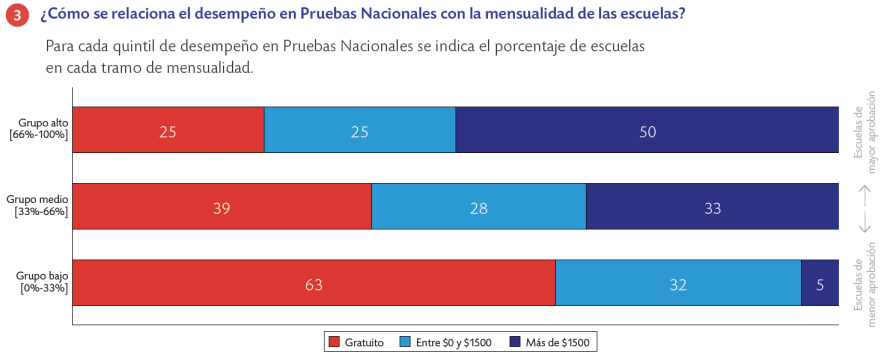
- There is significant price dispersion in the market.
- Parents learn about new schools through visits and friends.
- Parents' beliefs are not aligned with the true distribution:
  - Parents overestimate of the quality of public schools.
  - Parents overestimate the correlation between price and quality for private schools.

Field Experiment

- 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.

Treatment 1

Joint Distribution of Price and Quality



Treatment 2

LAS ESCUELAS EN DETALLE

	Nombre Centro	Ofrece Nivel Inicial	Ofrece Media	Tanda	Estudiantes por Profesor	Cobro mensual	Aprobacion en PN
1	Abejitas, Las	Sí	Sí	M	11	$2000	94
2	Angeles Sin Paraiso	No	No	V			.
3	C.E. Abc Santo Domingo Norte	Sí	Sí	M	17	$2500	96
4	C.E. Amela	Sí	No	M	12	$1800	99
5	C.E. Huellas Del Saber	No	No	M			75
6	C.E. Nuevo Mella	Sí	Sí	M	14	$1000	82
7	C.E. Vigotsky "Tia Hellen"	Sí	No	M	6	$2000	92
8	Centro Infantil Trencito De Colores	Sí	No	M	13	$800	96
9	Chaparral	Sí	No	V	51	$0	91
10	Ciriaco Maria Fe Y Alegria Marañon Ii	Sí	No	V	17	$0	95
11	Claribel	No	No	M			80
12	Colegio 2000	Sí	No	M	15	$600	74
13	Colegio Cristiano Hidekel	No	No	M			83
14	Colegio Santo Cristo	No	No	M			79

Pilot Results

Treatment Effects on Knowledge

	Knows New Schools		Number of New Schools	
	(1)	(2)	(3)	(4)
T1: Distributions	0.041	0.044	0.225**	0.230**
	[0.061]	[0.061]	[0.096]	[0.098]
T2: Distributions + Info	0.056	0.077	0.139	0.165*
	[0.065]	[0.067]	[0.089]	[0.091]
Controls	No	Yes	No	Yes
Mean of Control Group	0.327	0.327	0.354	0.354
Observations	350	350	350	350

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.