Education Markets II

Christopher Campos

Economics of Education, Spring 2026

Outline for Today

1. Rise & Theory of Coordinated Assignment
   - → Neilson (2024): the global landscape of coordinated systems
   - → Abdulkadiroğlu & Sönmez (2003): the mechanism design framework
   - → Abdulkadiroğlu, Agarwal & Pathak (2017): welfare effects of coordination

2. Design Details Matter
   - → Kapor, Neilson & Zimmerman (2020): heterogeneous beliefs
   - → Terrier, Pathak, and Ren (2026): natural experiment transitioning from IA to DA
   - → Arteaga, Kapor, Neilson & Zimmerman (2022): search and information

3. Self-Selection
   - → The selection problem
   - → Characterizing more general classes of treatment effects

4. System Design: Campos, Bruhn, Chyn & Vazquez (2026)
   - → National catalog, LAUSD case study, structural model, counterfactuals

Coordinated Assignment

From School Choice to Market Design

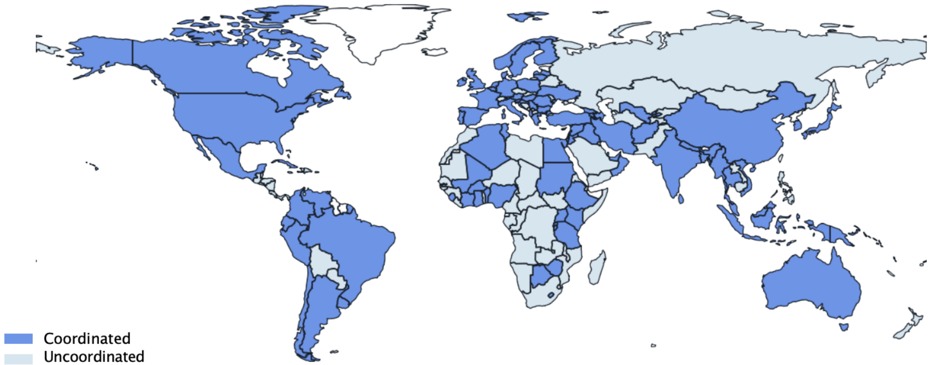
- Last week: school choice initiatives can produce supply-side effects, but results seem to be fragile and market-dependent
  - → Abdulkadiroğlu et al. (2020): parents value school quality, but demand is imperfectly aligned with effectiveness
  - → Campos & Kearns (2024): introducing choice raised achievement, but effects depended on the competitive environment and information available to families
- This week: The assignment mechanism, the information environment, and the participation rule all shape who benefits
  - → How do different design choices translate into welfare and outcomes? How should we design markets? And how have they actually been designed in practice?
  - → We begin with the global picture — a massive expansion of coordinated choice and assignment systems (CCAS) over the past two decades

The Global Rise of Coordinated Systems

- In education markets, equilibrium prices rarely determine allocation — societies instead develop rules and regulations to allocate access
- A coordinated choice and assignment system (CCAS) requires three conditions:
  - → Applicants submit a ranked list of preferences
  - → An external agency coordinates enrollment across institutions
  - → The system makes a single offer based on predefined rules and criteria
- Neilson (2024) reviews 149 countries across primary, secondary, and higher education
- 60 percent of countries now use some form of CCAS to determine access to education
- Adoption is positively correlated with income, urbanization, and government transparency — but 57 low- and middle-income countries also use coordinated mechanisms

CCAS Adoption Around the World

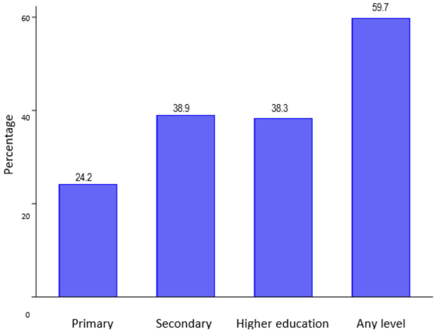
Map 1 Countries with at least one coordinated system



Source: Original calculations for the *World Development Report 2024.*

CCAS Adoption Around the World

Figure 1 Percentage of countries with coordinated systems by education level

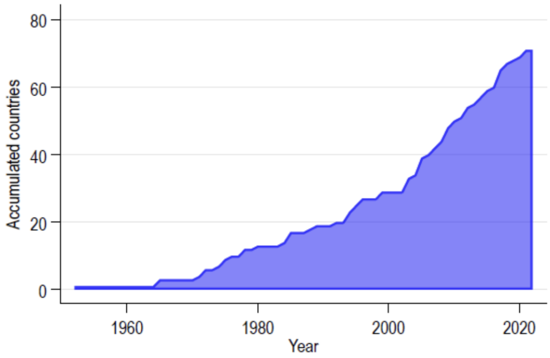


Source: Original calculations for the *World Development Report 2024.*

- CCAS more common at secondary (38.9%) and higher education (38.3%) than primary (24.2%)
- Standardized tests at higher levels facilitate priority-based assignment
- Primary systems require alternative criteria (e.g., distance, siblings)

Rapid and Accelerating Adoption

Figure 2 Number of accumulated countries with a CCAS



- Higher education was the early adopter — coordinated university admissions have existed for decades
- Primary and secondary adoption has accelerated sharply since 2000

Significant Heterogeneity “In the Wild”

- There is vast heterogeneity in policy choices “in the wild” and often do not align with theoretical best practices
- Key design dimensions that vary across systems:
  - → Assignment mechanism: Immediate Acceptance, Deferred Acceptance, Serial Dictatorship, Top Trading Cycles, or other
  - → Preference list length: some systems cap the number of ranked options
  - → Priority criteria: academic performance, distance, socioeconomic equity, siblings, gender, special needs
  - → Information provision: what families know about schools, cutoffs, and admissions probabilities
  - → Off-platform options: aftermarket frictions, penalties for rejecting offers
- Many of these design choices have been studied individually but most real systems combine features in configurations that have not been thoroughly studied
- The more modern theory poses assignment of students to schools (in choice environments) as an engineering problem

The School Choice Problem

- A set of students $I = \{i_1, \ldots, i_n\}$ and a set of schools $S = \{s_1, \ldots, s_m\}$
- Each school s has a capacity q_s (maximum enrollment)
- Each student has strict preferences over schools (a complete ranking)
- Each school has a strict priority ordering over students
  - → Priorities are *not* preferences — they are set by law or policy
  - → Typical priority categories: sibling at school $>$ walk zone $>$ lottery number
- A matching assigns each student to one school; no school exceeds capacity
- Important definitions:
  - → A mechanism is *strategy-proof* if no student can ever benefit from misreporting her preferences
  - → A matching is *Pareto efficient* if no student can be made better off without making another worse off
  - → A matching is *stable* if no student–school pair (i, s) exists such that i prefers s to her assignment and s either has empty seats or ranks i above some admitted student

The Boston Mechanism (Immediate Acceptance)

- The mechanism used in Boston (since 1999), Columbus, Minneapolis, Seattle, and many other districts
- Round 1: Each student applies to her *first* choice
  - → For each school: accept applicants one at a time by priority until capacity is filled
  - → Accepted students are permanently assigned. Seats are final
  - → Remaining applicants are rejected
- Round 2: Rejected students apply to their *second* choice
  - → Schools with remaining seats accept by priority; reject the rest
- Round k: Continue with kth choices at schools with remaining seats
- Algorithm terminates when all students are assigned or no seats remain
- Key design feature: Some families may have incentives to “misreport” their preferences

Boston Mechanism: The Strategic Problem

- A student with high priority at school s loses that priority unless she lists s first
- Consider a concrete example:
  - → Family prefers School A (popular, oversubscribed) but has high priority at School B (sibling attends)
  - → If they list A first and get rejected in Round 1, they apply to B in Round 2
  - → But B may have already filled its seats with Round 1 applicants
  - → The family's priority advantage at B is *wasted*
- Rational response: list B first even though they prefer A
- The mechanism forces families to play a complicated admissions game:
  - → Which schools are oversubscribed? What are my odds?
  - → Should I "waste" my first choice on a reach school?
- Sophisticated families with better information may gain an advantage
- Pathak & Sönmez (2008): removing the strategic burden "levels the playing field" — disadvantaged families benefit most from strategy-proof mechanisms

Boston Mechanism: Formally Not Strategy-Proof

- The Boston mechanism is not strategy-proof: there exist preference profiles under which a student strictly benefits from misreporting
- The mechanism gives "very strong incentives to students and their parents to misrepresent their preferences by improving ranks of those schools for which they have high priority"
- Two consequences:
  - → Unfairness: families who misread the strategic environment are punished — they "waste" their top choice and lose priority elsewhere
  - → Inefficiency: since families misrepresent, the mechanism operates on distorted inputs. The resulting matching is unlikely to be Pareto efficient

Deferred Acceptance (Student-Proposing Gale-Shapley)

- Interpret school priorities as preferences $\rightarrow$ apply the Gale-Shapley (1962) college admissions algorithm
- Step 1: Each student proposes to her first choice
  - $\rightarrow$ Each school tentatively holds the highest-priority applicants up to capacity
  - $\rightarrow$ Remaining applicants are rejected
  - $\rightarrow$ Key: holds are tentative, not final!
- Step k: Each rejected student proposes to her next choice
  - $\rightarrow$ Each school considers *new proposers together with currently held students*
  - $\rightarrow$ Keeps the top applicants by priority up to capacity; rejects the rest
- Algorithm terminates when no student is rejected (no new proposals)
- Each student is assigned her final tentative hold
- The difference from Boston: a student held at step k can still be displaced at step $k + 1$ by a higher-priority applicant $\rightarrow$ priorities are not "wasted"

Deferred Acceptance: Properties

- Student-proposing DA Pareto-dominates any other mechanism that eliminates justified envy (Gale and Shapley 1962)
  - → It is student-optimal among all stable matchings
- DA is strategy-proof (Roth 1982)
  - → Truthful preference revelation is a *weakly dominant strategy*
  - → Families never benefit from misreporting
- DA eliminates justified envy (stability):
  - → No unmatched pair (i, s) where student i prefers s to her assignment and has higher priority at s than some student assigned there
- Limitation: DA is not "problem free" because stability can conflict with Pareto efficiency
  - → There exist preference profiles where the DA outcome is Pareto-dominated

Top Trading Cycles (TTC)

- An alternative that prioritizes efficiency over stability
- Step 1: Each student points to her favorite school. Each school points to its highest-priority student. Since the graph is finite, at least one cycle forms.
  - → Execute all cycles: assign each student in a cycle to the school she points to
  - → Remove assigned students; reduce school counters by one
- Step k: Repeat among remaining students and schools with available seats
- TTC is Pareto efficient
  - → Any student who leaves at step k gets her top choice among remaining schools — no one can be made better off without hurting someone who left earlier
- TTC is strategy-proof
  - → Intuition: at each step, cycles that form before you leave are unaffected by your report. Truthfully pointing to your best remaining option weakly dominates.
- Not stable: TTC does not eliminate all justified envy — a student may envy another's assignment even when she has higher priority

Three Mechanisms: IA, DA, and TTC

- Immediate Acceptance (IA / Boston): Process rounds by student ranks; rejected students lose seats permanently
- Deferred Acceptance (DA / Gale-Shapley): Rejections are tentative — students can displace others in later rounds
- Top Trading Cycles (TTC): Students “trade” priority claims; cycles of mutual improvement are executed

Property	IA	DA	TTC
Strategy-proof	×	✓	✓
Pareto efficient	×	×	✓
Eliminates justified envy (stable)	×	✓	×

- IA punishes naïve families who rank truthfully — sophisticated families strategically avoid listing competitive schools
- DA vs. TTC depends on how policy makers interpret priorities:
  - → Priorities as hard constraints (fairness/legal) → DA
  - → Priorities as soft signals (tradeable opportunity) → TTC
- Does all of this matter in practice?

Abdulkadiroğlu, Agarwal & Pathak (2017): Welfare Effects of Coordinated Assignment

- NYC high school assignment had a multi-offer, uncoordinated system until 2003
- Pre-2003 *uncoordinated* mechanism:
  - → Students rank at most 5 programs; schools see the full rank-order list
  - → Schools independently admit, waitlist, or reject; students may receive multiple offers
  - → 3 rounds of offers/acceptances — insufficient to clear the market
  - → Some schools only consider students who ranked them first
- Result: 37% of students administratively placed at schools near home after the main round
- Only half of students placed in the main round; high-achieving and Manhattan students overrepresented in early rounds
- In fall 2003, NYC adopted a coordinated mechanism based on student-proposing DA: rank up to 12 programs, schools cannot see rankings, single lottery breaks ties

AAP (2017): Descriptive Evidence on the Reform

- Under DA, 82% of students assigned in the main round (vs. 48% under old system)
- Administrative assignments fall from 37% to 11%
- Administratively assigned students experience severe mismatch:
  - → Ranked schools avg 5.1 miles from home; assigned to schools 1.6 miles away
  - → Math achievement gap of 4.4 pp; ranked schools have 400+ fewer 9th graders
- Average distance to assigned school *increases* by 0.69 miles under the new mechanism
- But exit rates fall (8.5% → 6.4%) and enrollment at non-assigned schools falls (18.6% → 11.4%)
- The new mechanism distributes access more evenly across boroughs, races, and achievement levels
- How does this all translate to welfare?

AAP (2017): Demand Model

- DA is strategy-proof $\Rightarrow$ treat submitted rankings as truthful preferences
- Utility of student i from program j:

$$u_{ij} = \delta_j + \sum_l \alpha^l z_i^l x_j^l + \sum_k \gamma_i^k x_j^k - d_{ij} + \varepsilon_{ij}$$

  with $\delta_j = \mathbf{x}_j \beta + \xi_j, \;\; \gamma_i \sim \mathcal{N}(0, \Sigma_\gamma), \;\; \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$
- Distance d_{ij} enters with coefficient -1 (numeraire): welfare measured in miles of willingness to travel
- Key identifying assumption: $(\gamma_i, \varepsilon_{ij}) \perp d_{ij} \mid \mathbf{z}_i, \mathbf{x}_j, \xi_j$
- Estimation via Gibbs sampling using full rank-order data
- Findings: high-achieving families students value achievement more; minority students less attracted to high-%-white schools; substantial unobserved heterogeneity via random coefficients
- With an estimated demand system, AAP can then calculate welfare for numerous counterfactual environments

AAP (2017): Mechanism Design Tradeoffs

- Welfare measured as $\bar{W}(\mu) = \frac{1}{|\mathcal{I}|} \sum_i E[u_{i\mu(i)} \mid \mathbf{r}_i]$ in distance units
  - → Why not $\bar{W}(\mu) = \frac{1}{|\mathcal{I}|} \sum_i u_{i\mu(i)}$?
- Two benchmarks define the welfare range:
  - → *Neighborhood assignment* (closest school subject to capacity): lower bound
  - → *Utilitarian optimum* ($\max \sum u_{ij} a_{ij}$ s.t. capacity): upper bound
  - → Range: 18.96 miles
- Where do different mechanisms fall?
  - → DA (coordinated mechanism): -3.73 miles from optimum ⇒ 80% of range
  - → Student-optimal stable (Erdil & Ergin 2008): additional 0.11 miles (0.6%)
  - → Pareto efficient (TTC): additional 0.62 miles total (3.3%)
- Simply having coordinated choice captures the vast majority of welfare; algorithmic refinements are second-order

AAP (2017): Coordinated vs. Uncoordinated Mechanism

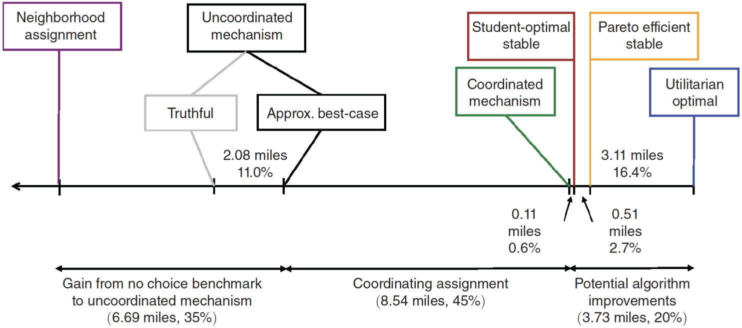


Figure 5. Coordinating Assignments versus Algorithm Improvements

- Coordination alone represents $\sim$ 45% of the total welfare range (8.54 miles)
- This is more than double the possible gain from all algorithm improvements (20% of range)
- Gains positive for every subgroup; new mechanism assigns students further from home but to much more preferred schools

AAP (2017): Who Benefits?

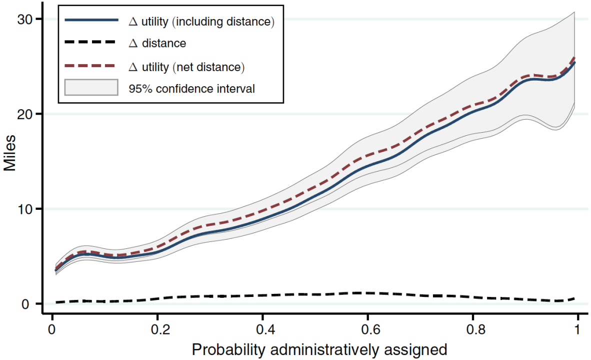
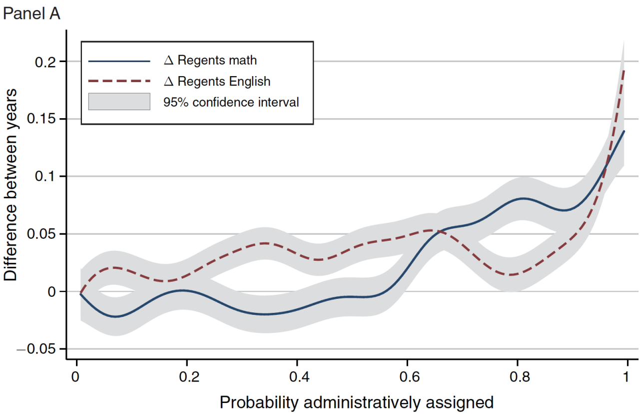


FIGURE 6. CHANGE IN STUDENT WELFARE BY PROPENSITY TO BE ADMINISTRATIVELY ASSIGNED IN THE UNCOORDINATED MECHANISM

- Construct admin-assignment propensity score from student demographics and census tract FEs
- The welfare impacts across this propensity score are increasing: those most likely to be administratively assigned had largest welfare gains

AAP (2017): Who Benefits?



- Similar exercise like before but now consider student achievement instead of welfare
- Cross-cohort comparisons of future outcomes, holding constant risk of administrative assignment
- Suggestive evidence that those who faced largest admin. assignment risk, also had the largest achievement gains

AAP (2017): Takeaways

- Uncoordinated assignment generated severe congestion: 37% of students administratively placed at undesirable schools
- The coordinated DA mechanism achieves 80% of the welfare range from neighborhood assignment to the utilitarian max
- Coordinating admissions ($\sim$45% of range) dwarfs potential gains from algorithm refinements ($\sim$3.3% of range)
- The largest beneficiaries are students most likely to have been administratively assigned
- Allocative improvements are mirrored by suggestive gains in test scores and graduation for the most affected students
- The theory assumes families know their preferences, know their admissions chances (when relevant), know all of their options, etc. What happens if some of this is not true?

Design Details Matter

Kapor, Neilson, and Zimmerman (2020): Heterogeneous Beliefs and School Choice Mechanisms

- Setting: New Haven, CT — a low-income, majority-minority district
  - → Centralized school choice since at least 1997
  - → Mechanism closely resembles the Boston / Immediate Acceptance mechanism
  - → Rising 9th graders choose among 12 high schools; can list up to 4
- Theoretical debate: which mechanism is better?
  - → Boston mechanism lets families express cardinal preference intensity through strategic ranking
  - → DA is strategy-proof but loses this expressiveness channel
  - → Under rational expectations, theory is ambiguous — an *empirical* question
- The key innovation: original household survey of 417 rising 9th graders
  - → In-person interviews at parents' homes (2015 and 2017)
  - → Elicit *preferences*, *beliefs about admissions chances*, and *strategic sophistication*
  - → Link survey data to administrative records of the choice process
- First paper to collect belief data from school choice participants

Families Don't Understand the Mechanism

- Belief errors are enormous:
  - → Mean absolute error between subjective beliefs and rational-expectations admissions probabilities: 37 percentage points
  - → Respondents are 8 pp more optimistic on average, with wide spread
- Misunderstanding is systematic:
  - → Only 3.4% correctly identified both the priority group ordering *and* the role of ranking in determining admissions
  - → Families are 42 pp more optimistic about second-ranked options than first-ranked — they don't understand how much *ranking position* matters
- Strategic behavior is common but often mistaken:
  - → 32% of respondents play strategically (list a non-favorite school first)
  - → But 48% of strategic players are "mistaken strategic" — they would have been better off listing their true favorite first
  - → These families attempt to game the system but do so based on wrong beliefs
- Subjective beliefs predict behavior; rational expectations do not

Incorporating Real Beliefs Reverses the Policy Conclusion

- Whether DA beats Boston depends *entirely* on what you assume about beliefs

	Baseline	DA	DA − Baseline	No survey DA − Baseline
Mean welfare	14.42	18.35	+3.93 (27% of mean)	−1.80 (reversal)

Welfare in miles-traveled equivalents. "No survey DA" imposes rational expectations.

- With real belief data: switching Boston → DA increases welfare by 3.9 miles
- With rational expectations assumption: switching decreases welfare by 1.8 miles
- The difference is 5.7 miles — equal to 40% of mean welfare at baseline
- Incorporating real beliefs flips the policy recommendation

Takeaways

- The theoretical literature frames Boston vs. DA as a tradeoff:
  - → Boston: rewards strategic play, allows expression of cardinal preferences
  - → DA: strategy-proof, but sacrifices the expressiveness channel
- KNZ's contribution: this tradeoff depends on *how informed families actually are*
  - → If families have rational expectations $\Rightarrow$ Boston does better (expressiveness $>$ mistakes)
  - → If families have heterogeneous, error-prone beliefs $\Rightarrow$ DA performs better (mistakes $>$ expressiveness)
- KNZ provide a nice welfare analysis but like all papers has some limitations
  - → Beliefs are from a selected sample; representative of the typical NH family? Are they understating or overstating the importance of biased beliefs?
  - → Model-based results lack policy variation. Can we find a natural experiment?

Terrier, Pathak, and Ren (2026): From Immediate Acceptance to Deferred Acceptance

- England assigns students to secondary schools (age 11) through 152 local authorities (LAs)
- Before 2008, LAs could use either Immediate Acceptance (IA) or Deferred Acceptance (DA)
  - → 16 “pure IA” LAs: all schools use First Preference First (FPF) criterion
  - → 49 “pure DA” LAs: all schools use Equal Preference (EP) criterion
- The 2007 School Admissions Code banned IA/FPF nationwide, effective September 2008
  - → Motivated by equity concerns: “those who get it wrong or don’t understand, lose out”
- Key institutional features:
  - → Parents rank 3–6 schools; all IA LAs limited lists to 3
  - → Selective schools (grammar schools) admit via test scores — parents did not know results before submitting preferences
  - → Distance-based priorities at non-selective schools $\Rightarrow$ residential segregation maps into priority segregation

Conceptual Framework: Two Competing Effects

- Under IA, sophisticated (high-SES) parents strategically avoid ranking oversubscribed schools where admission is uncertain
- **Competition-for-top-schools effect:** Sophisticated parents' strategic avoidance reduces competition for popular schools, benefiting sincere (low-SES) applicants who rank truthfully
- **Trickle-down effect:** Sincere parents who are rejected from their first choice lose priority at second/third choices under IA (because other students ranked those schools first)
- The net effect for low-SES students depends on:
  - → Uncertainty in admission chances (amplifies strategic mistakes)
  - → Level of competition for top schools (amplifies competition effect)
  - → Oversubscription at lower-ranked schools (amplifies trickle-down)
  - → Availability of outside options / private schools (dampens strategic incentives)

Conceptual Framework: Two Competing Effects

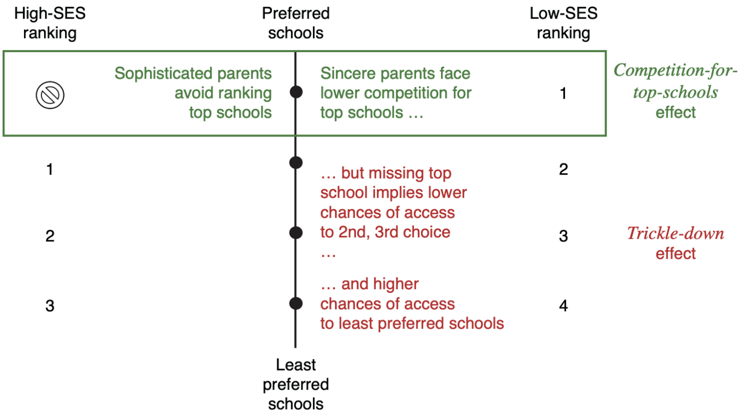


FIGURE 1. ILLUSTRATION OF THE *COMPETITION-FOR-TOP-SCHOOLS* AND *TRICKLE-DOWN* EFFECTS

Research Design and Data

- Difference-in-differences: Compare outcomes in 16 IA LAs (treated) vs. 49 DA LAs (control), before and after the 2008 ban
- Triple-difference specification to estimate heterogeneous effects by SES:

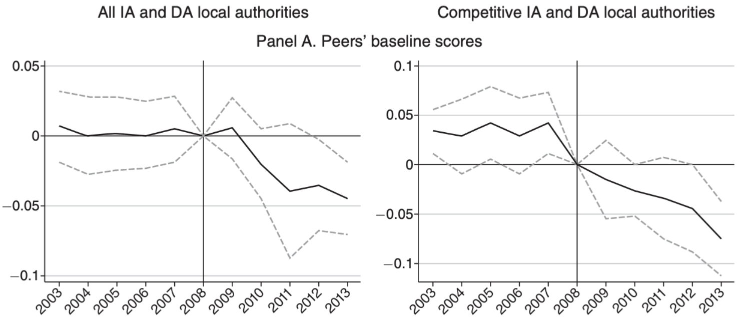
$$Y_{ilt} = \mu + \alpha \cdot IA_l + \beta \cdot Post_t + \phi \cdot LowSES_i + \gamma \cdot IA_l \cdot Post_t + \eta \cdot IA_l \cdot Post_t \cdot LowSES_i + \delta X_{lt} + \varepsilon_{ilt}$$

  → γ: effect on high-SES students in IA LAs relative to DA
  → η: differential effect for low-SES vs. high-SES (key parameter)

- Data: National Pupil Database, 2002–2014: universe of English students
  → KS2 scores (age 11, pre-assignment), KS3 (age 14), GCSE (age 16)
  → School VA computed from GCSE outcomes with EB shrinkage

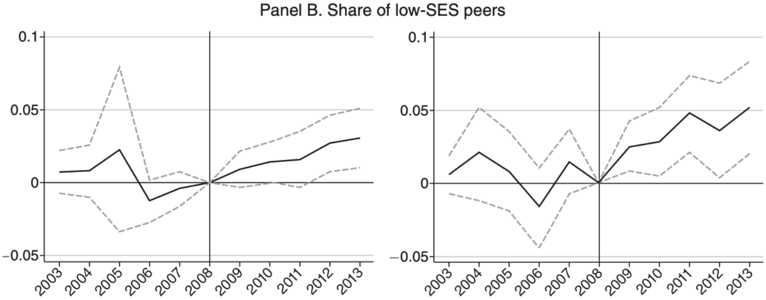
The IA-to-DA transition **widened the SES gap** in school quality measures

Enroll in schools with lower achieving students



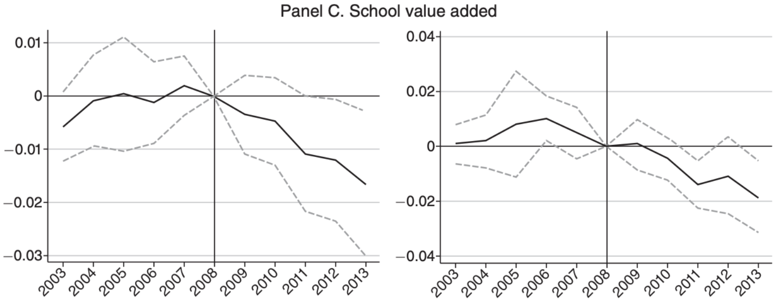
The IA-to-DA transition **widened the SES gap** in school quality measures

Enroll in schools with higher concentration of low-SES students



The IA-to-DA transition **widened the SES gap** in school quality measures

Enroll in lower quality schools



Takeaways

- Although low-SES students enroll in worse schools, no achievement impacts are found
- Strategy-proofness is not automatically equity-enhancing
  - → Under IA, sophisticated parents' strategic mistakes inadvertently benefited low-SES students
  - → DA removes this "accidental equity" channel by encouraging truthful reporting from all
- The competitive environment matters: effects are concentrated where schools are vertically differentiated and oversubscribed
- Not in this paper, but Akbarpour et al. (2022) find that the presence of outside options further distorts who benefits from IA-to-DA transitions.
  - → Families with better outside options are more likely to report truthfully
- Policy implication: transitioning to DA may require complementary policies (de-screening, reserves for disadvantaged groups) if one wants to preserve or improve equity
- What other mistakes or inefficiencies can arise in these environments?

Artega et al. (2022): Smart Matching Platforms

- The canonical school choice model assumes applicants know which schools are available to them
- In reality, learning about schools is costly sequential search:
  - → Internet research, school visits, phone calls, word of mouth
- If a family is overoptimistic about placement at listed schools, the perceived value of searching for backups falls → she stops too early
- Result: short application portfolios, high nonplacement risk, worse matches
- Strategy-proofness solves the problem *inside* the mechanism but not the search problem that *precedes* it
- Can "smart matching platforms" alleviate this concern and increase efficiency in centralized assignment systems?

A Model for Searching Schools with Biased Beliefs

Preliminaries

- Applicants start with consideration set $\mathcal{C}_0 = \{1, 2, 3, \cdots, N_0\} \subset \mathcal{J}$ where $\mathcal{J}$ is the set of all schools
- Applicants receive utility u_j from placement at j
- Assume $u_1 > u_2 > \cdots > u_{N0} > 0$ measured relative to an outside option with utility 0
- Applicants know the u_j for each $j \in \mathcal{C}_0$ and have subjective beliefs about their admission chances p_j
- Applicants ignore correlation of admissions probabilities across schools

Searching

- Applicants may choose to pay a cost κ to add a school to $\mathcal{C}_0$
- If they choose to do so, the subjective probability and utility are drawn from $F_{p,u}(p, u)$
- Applicants therefore learn about a new school's u_s and have a subjective belief about their admission chances p_s

The Value of Learning About a School

Let $R_j = 1 - p_j$ be the subjective risk of non-placement at j. The value of the optimal portfolio given $\mathcal{C}_0$ in the absence of searching is

$$V(\mathcal{C}_0) = p_1 u_1 + p_2 u_2 R_1 + \cdots + p_{N_0} u_{N_0} \prod_{j<N_0} R_j$$

The value of the optimal portfolio given a consideration set with an augmented school s, $\mathcal{C} = \mathcal{C}_0 \cup \{s\}$ is

$$V(\mathcal{C}) = \underbrace{\sum_{j=1}^{r-1} p_j u_j \prod_{j'<j} R_{j'}}_{\text{Schools preferred to } s} + \underbrace{p_s u_s \prod_{j'<r} R_{j'}}_{\text{New school } s} + \underbrace{\sum_{j=r}^{N_0} p_j u_j R_s \prod_{j'<j} R_{j'}}_{\text{Schools less preferred than } s}$$

The value of searching is therefore

$$V(\mathcal{C}) - V(\mathcal{C}_0) = p_s (u_s - \Gamma_r) \prod_{j<r} R_j$$

where

$$\Gamma_r = \sum_{j=r}^{N_0} p_j u_j \prod_{j'=r}^{j-1} R_{j'}$$

is the expected value of schools with rank $r = \min\{j \in \mathcal{C}_0 : u_j < u_s\}$ or greater

How Optimism Influences Value of Adding Schools

- Assume belief errors are multiplicative, $R_j = (1-a)R_j^*$, where R_j^* is the true risk. Similarly, let $p_j^* = 1 - R_j^*$ denote the true admission chances at j.
- As a gets larger, people assume their admissions chances at schools are higher than they really are
- Taking logs of the value of search we have

$$\log(V(\mathcal{C}) - V(\mathcal{C}_0)) = \log(1-(1-a)R_s^*) + \log(u_s - \Gamma_r) + \sum_{j<r}(\log(1-a) + \log(R_j^*))$$

- The derivative with respect to a:

$$\frac{d\log(V(C)-V(C_0))}{da} = \underbrace{\frac{1-r}{1-a}}_{\text{Value of adding } s;<0} + \underbrace{\frac{R_s^*}{1-R_s^*(1-a)}}_{\text{Additional expected utility due to optimism; }>0} + \underbrace{\frac{d\Gamma_r}{da}\frac{1}{\Gamma_r - u_s}}_{\text{Effect on expected value of being ranked below } s;<0}$$

Intervention Design

Setting

- Chile adopts a centralized assignment system for their K-12 system in 2016
- DA assignment system without any list length limits
- Sibling, school employee, and alumni priorities + random tiebreaker used to ration access at oversubscribed schools
- Applicants apply during August each year; platform open for roughly one month

Policy Motivation and Intervention

- Nonplacement risk was a major concern for education policy makers in Chile (21 % unplaced in the first round of admissions)
- Worked with policymakers to design an intervention to “warn” applicants of their nonplacement risk (using pop-ups or reminders) and encourage adding more schools to their application

Intervention Design

Pop-ups

- Compute a predicted risk for each applicant
- Applications with nonplacement risk exceeding 30% flagged as “risky” and receive a pop-up
- Flagged applicants receive a warning message encouraging them to add more schools

Reminders

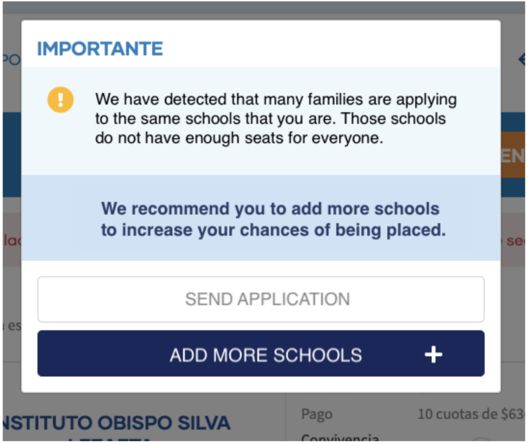
- In 2018, the Chilean Ministry of Education sent an SMS message to all risky students four days before the application deadline
- In 2019, no reminders were sent
- In 2020, three reminders are randomized to risky applicants

Survey

- Mineduc contacts applicants in 2020; 13% response rate
- Modules about preferences, beliefs, and search designed to provide context for the interventions

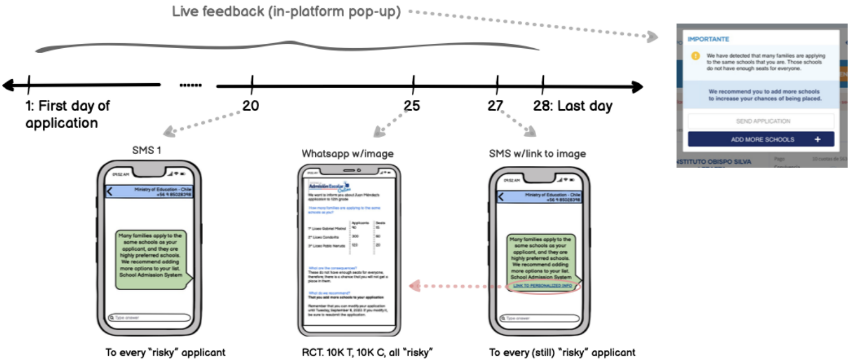
Pop-up Example

Figure B.I
Platform Pop-Up Intervention– 2018 and 2019



Reminder Timeline

Figure B.II
Timeline of Feedback Interventions– 2020



Survey Evidence

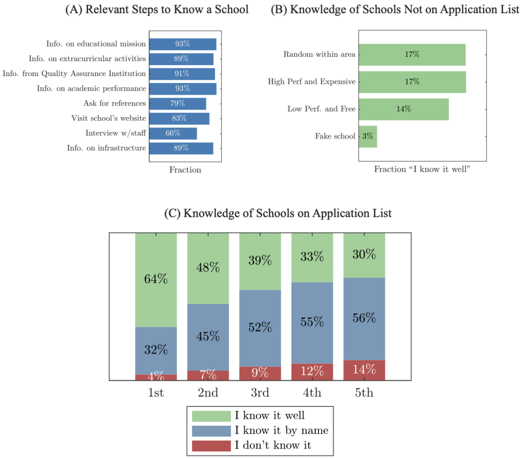
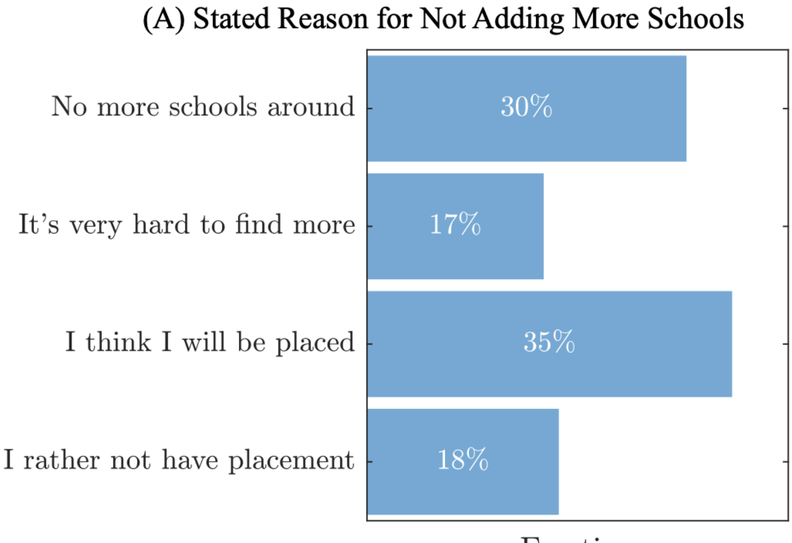


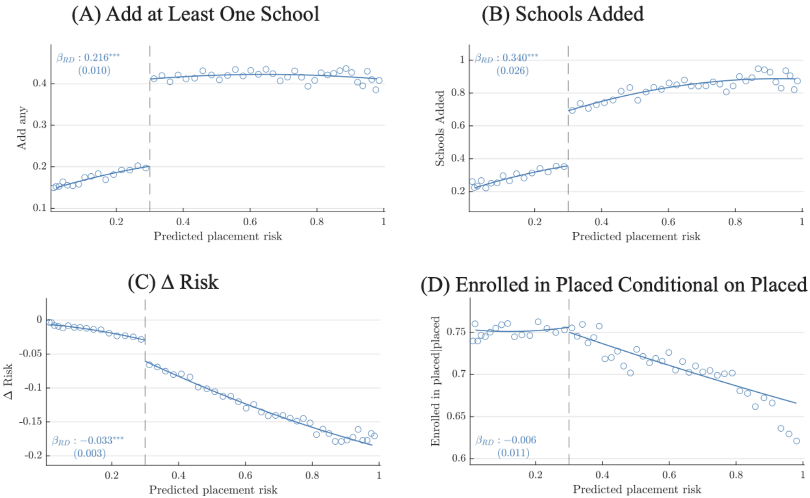
FIGURE II

Knowledge of and Search for Schools

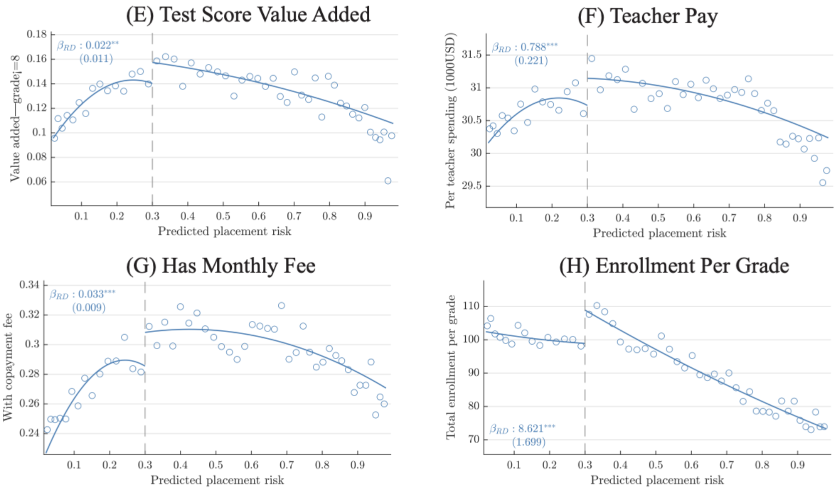
Survey Evidence



RD Evidence


(A) Add at Least One School
β_{RD} : 0.216***
(0.010)
Add any
Predicted placement risk
(B) Schools Added
β_{RD} : 0.340***
(0.026)
Schools Added
Predicted placement risk
(C) Δ Risk
β_{RD} : −0.033***
(0.003)
Δ Risk
Predicted placement risk
(D) Enrolled in Placed Conditional on Placed
β_{RD} : −0.006
(0.011)
Enrolled in placed|placed
Predicted placement risk

RD Evidence


(E) Test Score Value Added
β_{RD} : 0.022**
(0.011)
Value added—grade=8
Predicted placement risk
(F) Teacher Pay
β_{RD} : 0.788***
(0.221)
Per teacher spending (1000USD)
Predicted placement risk
(G) Has Monthly Fee
β_{RD} : 0.033***
(0.009)
With copayment fee
Predicted placement risk
(H) Enrollment Per Grade
β_{RD} : 8.621***
(1.699)
Total enrollment per grade
Predicted placement risk

Effects on Beliefs

TABLE V
RD ESTIMATES OF PLATFORM POP-UP EFFECTS ON SUBJECTIVE BELIEFS

	2020 (1)	$E[Y \mid X = 0.3^{-}]$ (2)
Panel A: Survey takeup and completion		
Survey take-up	– 0.020	0.173
	(0.010)	
Answered expectation questions	– 0.013	0.150
	(0.010)	
Panel B: Application behavior in survey sample		
Add any	0.196	0.265
	(0.033)	
Δ Risk	– 0.016	– 0.027
	(0.008)	
Panel C: Subjective beliefs		
Subjective *P*(not assigned to any)	0.036	0.165
	(0.017)	
Subjective *P*(assigned to 1st)	– 0.049	0.754
	(0.021)	
Panel D: Stated preferences		
Satisfaction if assigned to 1st choice (1–7)	– 0.017	6.855
	(0.047)	
N to the left of the cutoff	1,381	
N to the right of the cutoff	1,500	

Testing the Impacts of Personalization

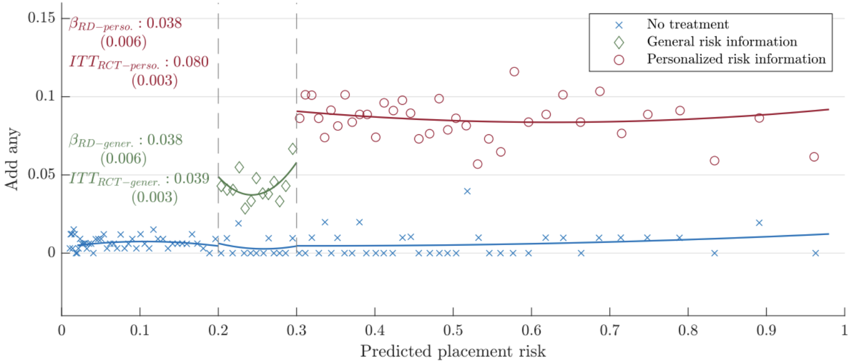


FIGURE VII

Arteaga et al. (2022) Takeaways

- The mechanism design literature focuses on what happens *after* families submit their lists
- But the hardest part of school choice may be what comes *before*: finding schools, learning about them, and building a portfolio
- Strategy-proofness $\rightarrow$ no need to misrepresent preferences
- But strategy-proofness does *not* solve:
  - $\rightarrow$ The search problem: which schools exist and are worth listing?
  - $\rightarrow$ The beliefs problem: how likely am I to be placed?
  - $\rightarrow$ The information problem: what are these schools actually like?

Transition: From Welfare to Outcomes

- We have shown that:
  - → Mechanism choice matters for welfare (Abdulkadiroğlu, Agarwal & Pathak 2017)
  - → Beliefs matter for behavior and welfare rankings (Kapor, Neilson & Zimmerman 2020)
  - → Changing mechanisms affects allocation of students to schools, potentially hurting low-SES students but the details matter (Terrier, Pathak, and Ren 2026)
  - → Search frictions matter (Arteaga et al. 2022)
- These papers show that the allocation of students to schools is changes and as a consequence we should expect changes in outcomes
- But we have limited evidence on how system design choices map to student achievement outcomes
- Campos & Kearns (2024) is one exception but natural experiments are hard to find
- To formally assess how market-level design features shape outcomes — who participates, who benefits, and by how much — we need to combine quasi-experimental variation with structural modeling that characterizes the distribution of treatment effects
- This motivates our next detour: selection models

Self-Selection

Self-Selection

- Last lecture we took a discrete choice detour: tools to measure what families want
- Now we need a self-selection detour: tools that help us characterize who benefits
- A core problem: in the most compelling designs (lotteries), we are limited to the set of families that apply to oversubscribed schools
  - → Families self-select into applying
  - → It is likely their benefits differ from the typical student
  - → Observational comparisons between applicants and non-applicants are unlikely to isolate causal effects
  - → We are therefore unable to say much about what benefits would look like for non-participants
- This is a class of selection problem (Heckman 1974, 1979)
- Why this is important in general?
  - → A useful balance of external vs internal validity
  - → Relevant anywhere there's self-selection with heterogeneous returns (job training, insurance take-up, migration, college major)
  - → Useful to answer richer counterfactual and system/market-level questions

The Selection Problem: Setup

- Potential outcomes (e.g., test scores with/without a choice school):

$$Y_i(1) = \alpha_1 + X_i'\beta_1 + \epsilon_{i1}$$
$$Y_i(0) = \alpha_0 + X_i'\beta_0 + \epsilon_{i0}$$

  where $E[\epsilon_{it} \mid X_i] = 0$

- Selection rule (e.g., applies to a choice school):

$$D_i = \mathbf{1}\{X_i'\gamma + Z_i'\pi - \eta_i > 0\}$$

  where Z_i shifts participation but is excludable from potential outcomes

- We observe $Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0)$, so each potential outcome is observed on a selected sample

- The bias: If Cov$(\epsilon_{it}, \eta_i) \neq 0$, comparing mean outcomes across $D_i = 1$ and $D_i = 0$ confounds the causal effect with the selection effect:

$$E[Y_i(1) \mid X_i, D_i = 1] - E[Y_i(0) \mid X_i, D_i = 0] \neq ATE$$

Correcting for Selection

- One approach is to assume $E[\epsilon_{it} \mid X_i, \eta_i] = \delta_t \eta_i$ (linearity of each potential outcome error in the selection error)
- Under normality of η_i, the CEFs of observed outcomes are:

$$E[Y_i \mid X_i, Z_i, D_i = 1] = \alpha_1 + X_i'\beta_1 + \rho_1\sigma_1 \cdot \lambda^-(X_i'\gamma + Z_i'\pi)$$
$$E[Y_i \mid X_i, Z_i, D_i = 0] = \alpha_0 + X_i'\beta_0 + \rho_0\sigma_0 \cdot \lambda^+(X_i'\gamma + Z_i'\pi)$$

- The selection correction terms are

$$\lambda^-(v) = \frac{\phi(v)}{\Phi(v)}, \quad \lambda^+(v) = \frac{\phi(v)}{1 - \Phi(v)}$$

- The two-step procedure is *semiparametric*—it requires linearity of $E[\epsilon_{it} \mid \eta_i]$, not full joint normality or homoskedasticity
- Joint normality buys us a known functional form for $\lambda(\cdot)$, but the bias-correction logic doesn't depend on it

Two-Sided Heckit

$$E[Y_i \mid X_i, Z_i, D_i = 1] = \alpha_1 + X_i'\beta_1 + \rho_1\sigma_1\lambda^-(v_i)$$
$$E[Y_i \mid X_i, Z_i, D_i = 0] = \alpha_0 + X_i'\beta_0 + \rho_0\sigma_0\lambda^+(v_i)$$

- We can estimate this model in two steps: first estimate a probit for D_i, then selection-correct *both* the treated and untreated samples
- Under valid exclusion restrictions and the maintained functional-form assumptions, this delivers mean potential-outcome objects such as ATE, TOT, and TUT
- But what *identifies* the selection correction separately from the causal parameters?

What Identifies the Selection Correction?

- To see the issue clearly, consider one side with no covariates:

$$Y_i = \alpha_1 + \rho_1\sigma_1 \cdot \lambda^-(\hat{\pi}) + u_i = \psi + u_i$$

  The constant identifies ψ, but α_1 and $\rho_1\sigma_1$ are not separately identified

- With covariates, suppose outcome and selection are both saturated in X_i:

$$Y_i = \sum_k \beta_k \mathbf{1}\{X_i = k\} + \rho_1\sigma_1 \cdot \sum_k \lambda^-(\hat{\gamma}_k)\mathbf{1}\{X_i = k\} + u_i$$

$$= \sum_k \underbrace{\left[\beta_k + \rho_1\sigma_1 \cdot \lambda^-(\hat{\gamma}_k)\right]}_{\text{not separately identified}} \mathbf{1}\{X_i = k\} + u_i$$

- X_i main effects and the Mills ratio are collinear when both are unrestricted functions of X_i

Functional Form Is Not Enough

- With a linear (non-saturated) specification in X_i:

$$Y_i = X_i'\beta_1 + \rho_1\sigma_1 \cdot \lambda^-(X_i'\hat{\gamma}) + u_i$$

  $X_i'\beta_1$ is linear and $\lambda^-(X_i'\hat{\gamma})$ is nonlinear, so they are technically separable

- A selection model with the same variables in both steps is identified *only by functional form restrictions*

- The same logic applies to both sides: without additional variation, α_t and $\rho_t\sigma_t$ are not separately identified on either

Exclusion Restrictions Solve the Problem

- Include variables Z_i in the selection equation that are excluded from potential outcomes:

$$D_i = \mathbf{1}\{X_i'\gamma + Z_i'\pi - \eta_i > 0\}, \qquad E[\epsilon_{it} \mid X_i, Z_i] = 0$$

- The corrected outcome equation (treated side) becomes:

$$E[Y_i \mid X_i, Z_i, D_i = 1] = \alpha_1 + X_i'\beta_1 + \rho_1\sigma_1 \cdot \lambda^-(X_i'\gamma + Z_i'\pi)$$

- If $\pi \neq 0$, variation in Z_i separately identifies the selection term, even if X_i is saturated

- The requirements for Z_i are the same as for a good instrument:
  - → *Relevance:* Z_i shifts the probability of participation ($\pi \neq 0$)
  - → *Exclusion:* Z_i doesn't belong in the potential outcome equations

- This should be familiar: control function and IV are methods for solving the same problem

Other Functional Forms

- Normality is not the only distributional assumption one can make
- Consider the one-sided model

$$Y_i = X_i'\beta + \epsilon_i$$
$$D_i = \mathbf{1}\{X_i'\gamma + Z_i'\pi - \eta_i > 0\}$$

- We can assume

$$E[\epsilon_i \mid X_i, Z_i, \eta_i] = \delta\eta_i \qquad \text{and} \qquad \eta_i \mid X_i, Z_i \sim U(0,1)$$

- Also assume interior support:

$$0 \leq X_i'\gamma + Z_i'\pi \leq 1$$

so the implied selection probability stays inside the unit interval

LPM Selection

- The first stage is now a linear probability model:

$$P(D_i = 1 \mid X_i, Z_i) = X_i'\gamma + Z_i'\pi$$

- The selected CEF for the LPM selection model is

$$\begin{aligned} E[Y_i \mid X_i, Z_i, D_i = 1] &= X_i'\beta + \delta\, E[\eta_i \mid X_i, Z_i,\ \eta_i < X_i'\gamma + Z_i'\pi] \\ &= X_i'\beta + \frac{\delta}{2}(X_i'\gamma + Z_i'\pi) \end{aligned}$$

- With no instrument Z_i, the selection correction is a second linear function of X_i—exactly the collinearity problem from earlier
- Exclusion restrictions are needed for identification regardless of the distributional assumption

General Approach

- More generally, assume $\eta_i \mid X_i, Z_i$ has some arbitrary distribution function $F_\eta(\cdot)$
- We can write the selected outcome CEF:

$$\begin{aligned}
E[Y_i \mid X_i, Z_i, D_i = 1] &= X_i'\beta + E[\epsilon_i \mid X_i, Z_i, D_i = 1] \\
&= X_i'\beta + E[\epsilon_i \mid X_i, Z_i,\ \eta_i < X_i'\gamma + Z_i'\pi] \\
&= X_i'\beta + E[\epsilon_i \mid X_i, Z_i,\ F_\eta(\eta_i) < F_\eta(X_i'\gamma + Z_i'\pi)] \\
&= X_i'\beta + g\Big(F_\eta(X_i'\gamma + Z_i'\pi)\Big) \\
&= X_i'\beta + g(\Pr[D_i = 1 \mid X_i, Z_i])
\end{aligned}$$

- The selection bias is some unknown function of the probability of selection

General Approach

$$E[Y_i \mid X_i, Z_i, D_i = 1] = X_i'\beta + g(\Pr[D_i = 1 \mid X_i, Z_i])$$

- Linear selection model uses $g(p) = \delta\, p$
- Heckit model uses $g(p) = \delta\, \lambda\left(\Phi^{-1}(p)\right)$
- Semiparametric approach: flexible model for $g(\cdot)$
- In all cases, identification of β separately from $g(\cdot)$ requires that Z_i generates variation in $\Pr[D_i = 1 \mid X_i, Z_i]$ beyond what X_i alone provides

What Does the Selection Correction Tell Us?

- We've established that the selection bias is some function of the propensity score:

$$E[Y_i \mid X_i, Z_i, D_i = 1] = X_i'\beta + g(p(X_i, Z_i))$$

- So far we've treated $g(\cdot)$ as a *nuisance*—something to control for
- But $g(\cdot)$ is substantively informative: it encodes how unobserved determinants of outcomes covary with the propensity to participate
- In particular, consider the local effect of expanding participation by marginally raising p:

$$\frac{\partial}{\partial p} E[Y_i \mid X_i, D_i = 1, P_i = p] \; = \; g'(p)$$

- The derivative $g'(p)$ tells us how outcomes change as we move along the population in terms of their proclivity to self-select into treatment
- This is the key idea behind the marginal treatment effect and has close connection to the selection framework we have outlined

Moving beyond mean effects

- Suppose a policymaker wants to *expand* a school choice program. The relevant question is not the average effect for current participants: it is what happens to the *next* student brought in?
- A two-sided selection model targets several mean objects:
  - → ATE $= E[Y_i(1) - Y_i(0)]$ – average for the entire population
  - → TOT $= E[Y_i(1) - Y_i(0) \mid D_i = 1]$ – average for participants
  - → TUT $= E[Y_i(1) - Y_i(0) \mid D_i = 0]$ – average for non-participants
- Positive selection on gains (TOT $>$ ATE $>$ TUT): those who benefit most are more likely to participate
  - → Expanding access brings in students who benefit *less* $\Rightarrow$ average gains fall
- Negative selection on gains (TUT $>$ ATE $>$ TOT): those who benefit most are less likely to participate
  - → Expanding access brings in students who benefit *more* $\Rightarrow$ average gains rise
- But this is still coarse: TOT and TUT average over heterogeneous groups. We want to know what happens at the margin and as we vary that margin

The Marginal Treatment Effect

- Recall from the general approach: $g(\cdot)$ operates on $F_\eta(X_i'\gamma + Z_i'\pi)$
- Define $U_i = F_\eta(\eta_i) \sim U[0,1]$
- Rewrite the selection rule:

$$D_i = \mathbf{1}\{P_i \geq U_i\}, \qquad P_i = \Pr[D_i = 1 \mid X_i, Z_i]$$

  → Low U_i: eager participants (low unobserved resistance)
  → High U_i: reluctant participants (high unobserved resistance)

- For a given propensity score $P_i = p$, the marginal student is the one with $U_i = p$
- If policy raises p by changing Z_i, the newly induced students are exactly those at $U_i = p$
- So the key object for evaluating a policy expansion is the marginal treatment effect:

$$E[Y_i(1) - Y_i(0) \mid X_i = x,\, U_i = p]$$

Marginal Treatment Effects under the Normal Model

- The marginal treatment effect is

$$MTE(x,u) = E[Y_i(1) - Y_i(0) \mid X_i = x, U_i = u]$$

- Under the normal control-function model,

$$MTE(x,u) = \underbrace{x'(\beta_1 - \beta_0) + (\alpha_1 - \alpha_0)}_{\text{ATE}(x)} + (\rho_1\sigma_1 - \rho_0\sigma_0)\Phi^{-1}(u)$$

- All standard parameters are weighted integrals of MTE:

$$ATE(x) = \int_0^1 MTE(x,u)\,du = x'(\beta_1 - \beta_0) + (\alpha_1 - \alpha_0)$$

$$LATE(x,\, p \to p') = \frac{1}{p' - p}\int_p^{p'} MTE(x,u)\,du = ATE(x) + (\rho_1\sigma_1 - \rho_0\sigma_0)\cdot\frac{\phi\big(\Phi^{-1}(p)\big) - \phi\big(\Phi^{-1}(p')\big)}{p' - p}$$

- Positive selection on gains means $(\rho_1\sigma_1 - \rho_0\sigma_0) < 0$, so $MTE(x,u)$ declines in u
  - → ATE averages over the full unit interval, so $\Phi^{-1}(u)$ integrates to zero
  - → $LATE$ averages over $[p, p']$—it captures effects for *compliers* induced by a specific shift in the instrument
- Given a selection model and its assumptions, we can derive the model-implied MTE and recover any treatment-effect parameter as some weighted average (see Heckman, Urzua, and Vytlacil 2006)

System Design and Outcomes

The U.S. Landscape: Decentralized and Poorly Documented

- Education in the U.S. is highly decentralized: $\sim$13,000 districts, no federal mandate for how school districts are run or organized
- Some districts adopted coordinated DA following Abdulkadiroğlu & Sönmez (2003):
  - → NYC redesigned its high school match in 2003–04
  - → Boston abandoned the Boston Mechanism in 2005
- But what about the other thousands of districts?
- Until very recently, there was no systematic documentation of how districts have actually implemented choice systems
- The examples we know — NYC, Boston, New Haven, Charlotte — come from a handful of districts. They are not representative
- We lack even the most basic facts: what fraction of districts use algorithms? Require participation? Provide clear information?

Filling the Gap

- We need two things:

  1. A systematic catalog of how large U.S. districts have designed their choice systems — moving beyond the handful of well-studied cases
  2. A structural framework that links system design to student achievement — allowing us to evaluate counterfactual designs, not just the status quo

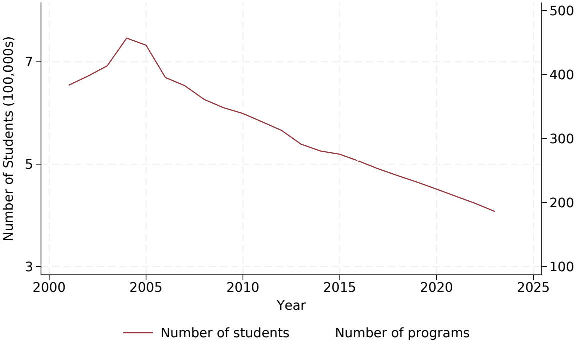
- Campos, Bruhn, Chyn & Vazquez (2026) fill both gaps:
  - → Original data collection from the 150 largest U.S. districts
  - → A deep-dive case study in LAUSD — the largest opt-in system in the country
  - → A generalized Roy model linking applications, enrollment, and achievement
  - → Counterfactual simulations of alternative system designs

Who Chooses and Who Benefits?

The Design of Public School Choice Systems

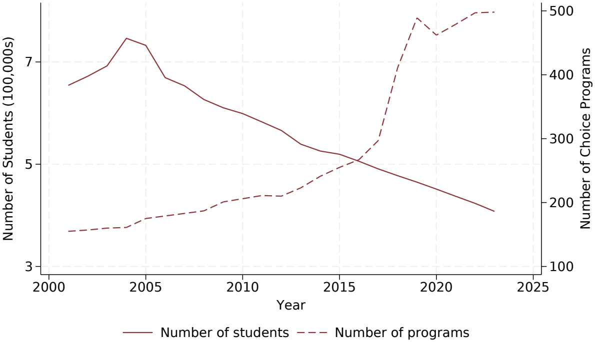
Campos, Bruhn, Chyn & Vazquez (2026)

Enrollment Decline in Urban U.S. School Districts



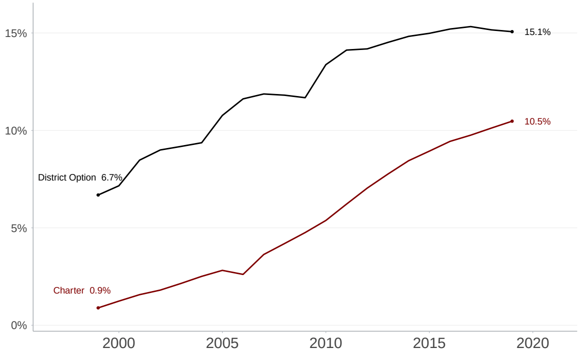
- There has been pronounced enrollment decline in urban U.S. school districts
- In Los Angeles, roughly a 50% drop in enrollment from the peak in 2004
- Primarily driven by demographic shifts and charter expansion
- Patterns are not unique to LA — similar trends across large urban districts

Districts Are Responding by Expanding Public School Choice



- Districts have responded by expanding public school choice offerings — magnets, themed schools, dual-language, affiliated charters
- In LAUSD, the number of choice programs more than doubled in two decades
- Patterns are not unique to Los Angeles
- Motivating question: How have school districts organized public education markets as they expand public school choice?

District-Run Choice Has More Than Doubled



- Enrollment in district-run choice schools has more than doubled in two decades
- District choice enrollment (∼15%) now exceeds charter enrollment (∼10.5%) by over 40%
- Districts have responded competitively to the charter sector by expanding their own choice portfolios
- Introducing more options into a system with neighborhood based assignment poses an access and market design problem

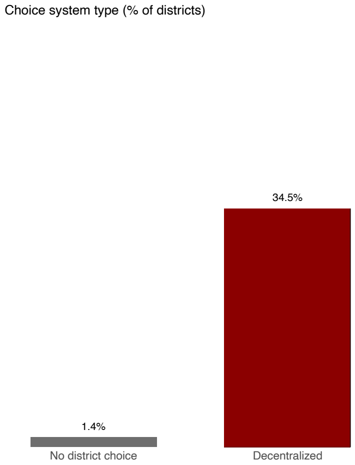
Original Data Collection: 150 Largest Districts

- First systematic catalog of how large U.S. districts design their school choice systems
- Sample: 150 largest school districts, covering 27% of all U.S. public school students
- Key design dimensions documented:
  - → Does the district offer intra-district choice at all?
  - → Is there a centralized algorithm for assignment?
  - → Is participation mandatory or opt-in?
  - → How difficult is the system to navigate?
- Goal: uncover unknown facts about the variataion in design across districts

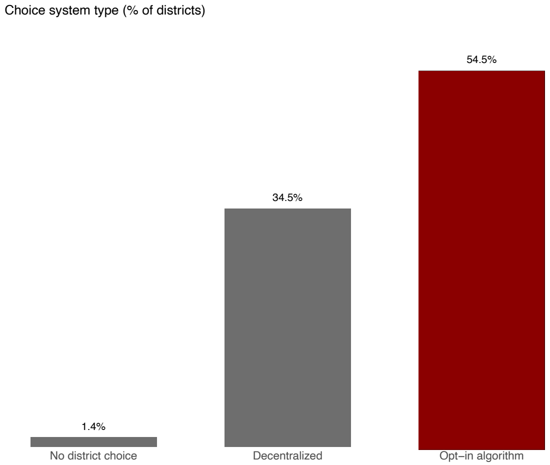
We survey the 150 largest school districts (25% of public enrollment nationwide)...

Nearly all offer some kind of intra-district choice

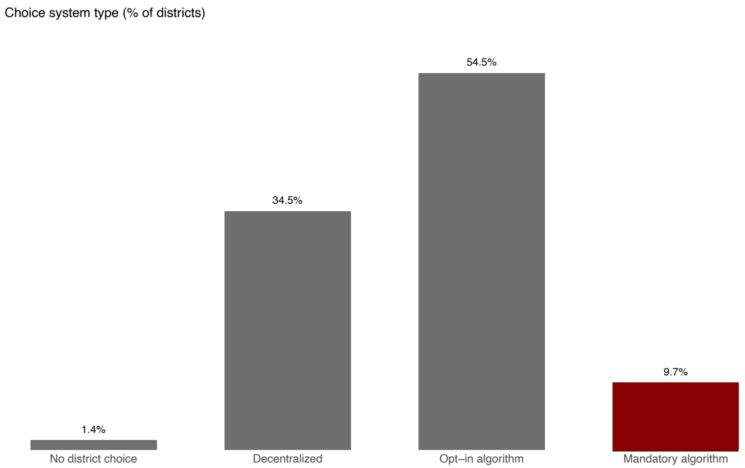
One third are decentralized i.e., have not adopted a centralized assignment system



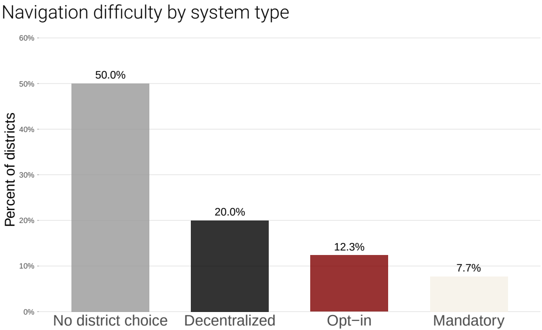
A majority of districts with centralized clearinghouses have an “opt-in” design



De facto mandatory systems—commonly studied in the literature—are rare

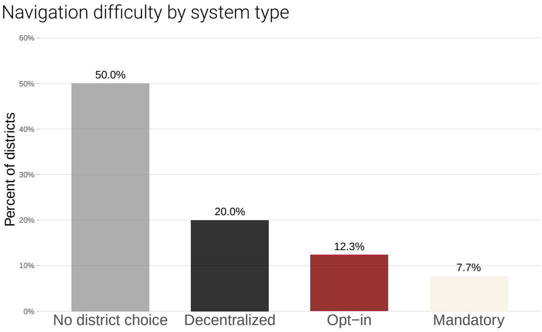


Centralization Reduces Barriers and Broadens Access

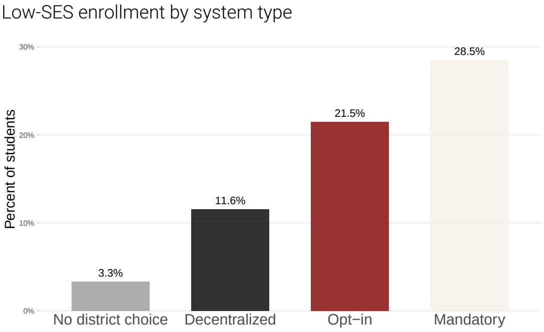


- 50% of no-choice districts are difficult to navigate → only 7.7% of mandatory systems
- More centralized → easier to find and use

Centralization Reduces Barriers and Broadens Access



- 50% of no-choice districts are difficult to navigate → only 7.7% of mandatory systems
- More centralized → easier to find and use



- Low-SES enrollment: 3.3% (no choice) → 28.5% (mandatory)
- Navigation difficulty and low-SES enrollment exhibit an inverse pattern

Two Primary Questions

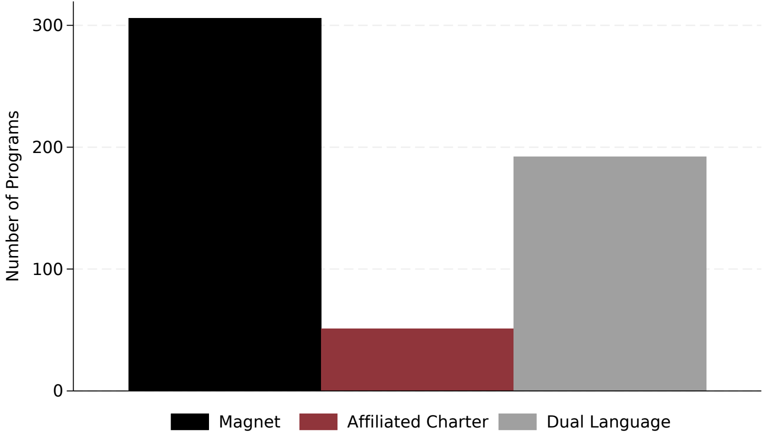
Two primary questions:

1. What are the implications of opt-in choice systems for aggregate achievement and inequality?
   - → **Who** opts in to LAUSD's voluntary choice system?
   - → **The What:** Are LAUSD's choice options vertically differentiated, high-return programs?
   - → **The How:** How do application costs, travel, and preferences shape participation and who benefits?
2. How do alternative market designs or policy interventions change who benefits from public school choice?
   - → Holding constant the participation architecture, how far can policies that expand access go?
   - → What are the effects of decentralized markets and de facto participation mandates (e.g., NYC-style deferred acceptance)?

The Largest Opt-In System in the US

- LAUSD would be the 35th largest school district in the country if its choice sector were its own district
- In 2024, there were roughly 500 programs/schools students could apply to
- Participation rates have increased:
  - → In 2000, roughly 10% of families participated
  - → In 2024, roughly 25% of families participated
- Between 2000–2013, applicants could rank at most one option; in more recent years an immediate acceptance mechanism is used to allocate students to schools
- If a student does not participate, they are defaulted to their neighborhood school
- Oversubscribed programs $\Rightarrow$ randomized lotteries by priority groups (geographic, sibling, waiting list)

School Choice Offerings in Los Angeles

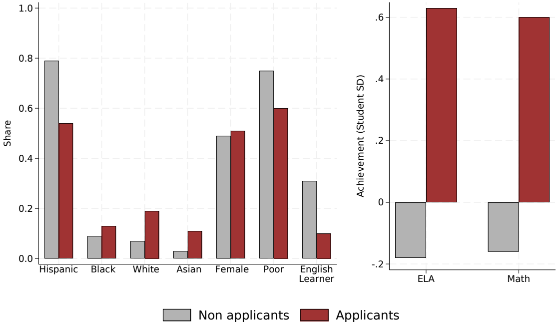


Type of choice offerings have evolved over the past twenty years — magnets, dual-language immersion, STEM-focused, affiliated charters

Data

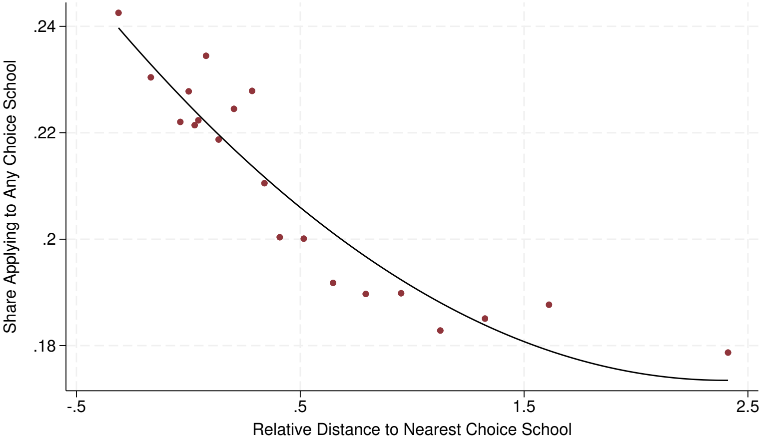
- Administrative student records (2000–2017):
  - → Demographics, addresses, and test scores
- Choice program records (2004–2017):
  - → Applications, program capacities, admission records
- Three analysis samples:
  - → Baseline sample (2004–2017): all fifth grade students
  - → Lottery sample (2004–2017): restricts to students who apply to oversubscribed choice middle schools
  - → Structural sample (2004–2013): restricts to students with addresses who enroll in LAUSD middle schools and for whom at least one test score outcome is observed

Applicants Are Positively Selected on Baseline Achievement



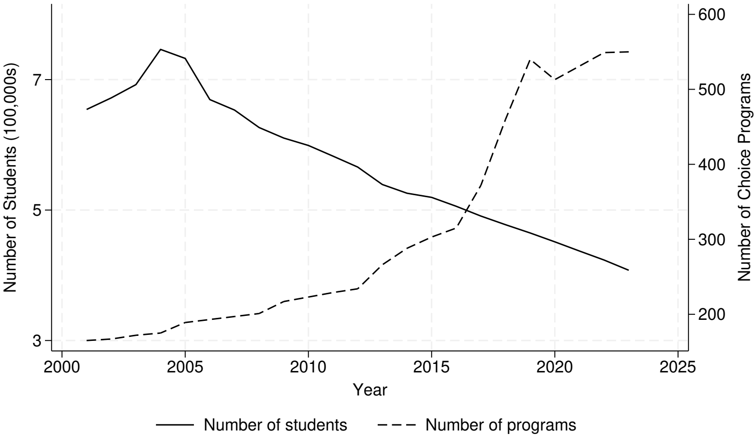
- Only 20% of 5th graders submit an application
- Applicants are 0.8σ higher in baseline achievement
- Applicants are 14 pp less likely to be low-income
- Applicants are 21 pp less likely to be English Learners
- Applicants are more racially diverse (more Black, White, Asian; fewer Hispanic)
- Next: How does proximity and access shift participation over time?

Distance Predicts Participation

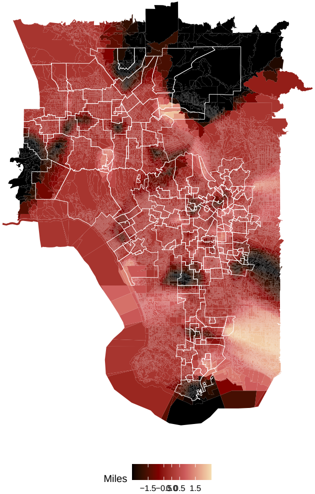


Distance to the nearest choice school is a strong predictor of whether families participate — a cost shifter that is useful for identification later

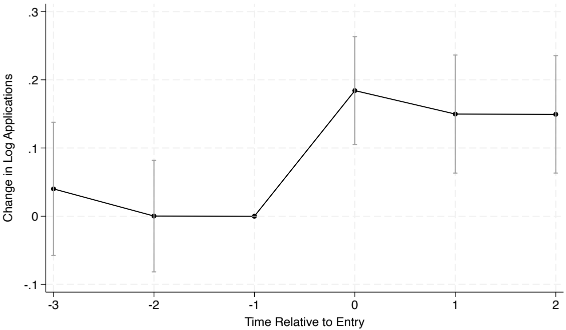
Access Has Evolved Over Time



Changes in Access Across Space



Choice School Entry Induces More Applications

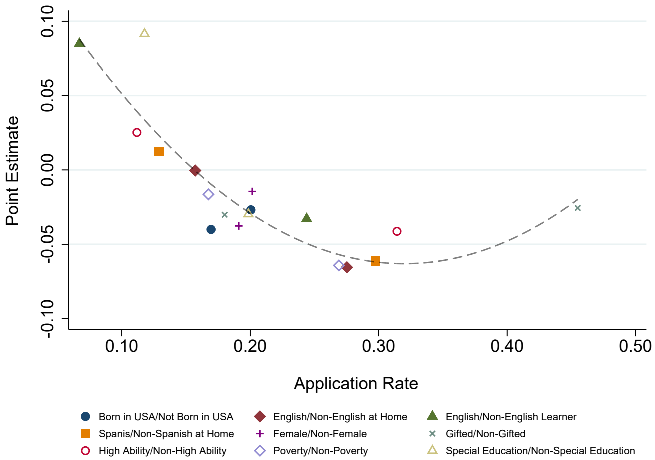


- LAUSD's choice portfolio more than doubled in the 2000s — creating quasi-experimental variation in geographic access
- Stacked event study: "treated" neighborhoods experience a reduction in distance to nearest choice program; "control" neighborhoods do not
- Result: a ∼20% jump in neighborhood applications precisely timed with choice school entry
- *Changes* in access/distance are cost shifters for participation
- If the distance margin shifts who applies, does it also predict who benefits when an offer arrives?

Reduced Form Exercises

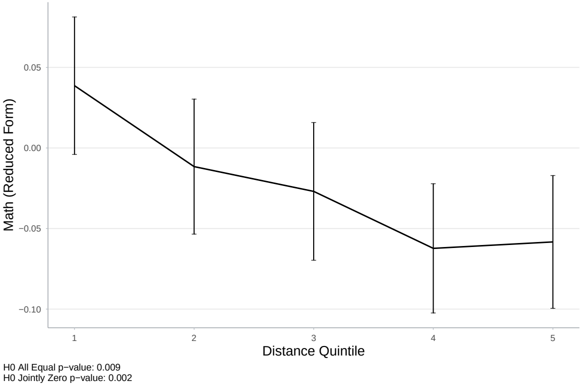
- **Key Question: Does demand predict treatment effect heterogeneity?**
- Use oversubscribed admission lotteries to estimate reduced form offer effects
- Two reduced form exercises to assess this:
  1. Lottery Effects Against Application Rates by Group
  2. Lottery Effect Heterogeneity by Distance

Lottery Effects and Application Rates by Group



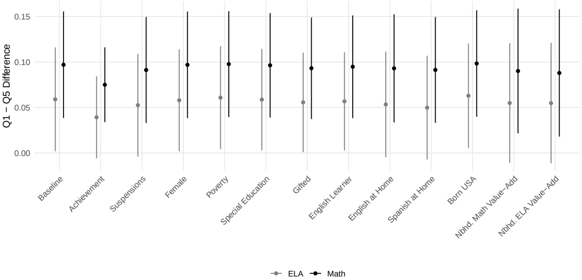
Groups with lower application rates tend to experience larger lottery effects — a first indication of negative selection on gains

Distance Strongly Predicts Treatment Effects



- Nearest quintile (Q1): $+0.05\sigma$ in math
- Farthest quintile (Q5): -0.05σ in math
- Difference: $\sim$0.10σ
- First-stage offer effects are comparable across quintiles (0.38–0.45) — heterogeneity is not driven by differential take-up

Distance Effects Not Explained by Other Observables



- Could the distance pattern be driven by treatment effect heterogeneity with other observables correlated with distance?
- Augment model with interactions between X_i and Z_i and report the Q1–Q5 distance difference
- The distance gradient persists after controlling for baseline achievement, poverty, race, English Learner status, and neighborhood school quality

Understanding treatment effect heterogeneity

Two pieces of evidence *suggest* latent preferences are associated with causal effects:

- Can we further isolate the relationship between preference heterogeneity and causal effects?

Understanding treatment effect heterogeneity

Two pieces of evidence *suggest* latent preferences are associated with causal effects:

- Can we further isolate the relationship between preference heterogeneity and causal effects?

Beyond Lottery Effects

- Our approach: Model preferences, application decisions, enrollment decisions, and outcomes which allows us to:
  - → Characterize selection into application *enrollment*
  - → Estimate and validate reduced form causal effects, e.g., LATE
  - → Extrapolate causal effects beyond the lottery sample
- Identification will require instruments that affect application and enrollment decisions but do not directly influence outcomes
  - → Offers to choice programs
  - → Policy-induced *changes* in distance

Understanding treatment effect heterogeneity

Two pieces of evidence *suggest* latent preferences are associated with causal effects:

- Can we further isolate the relationship between preference heterogeneity and causal effects?

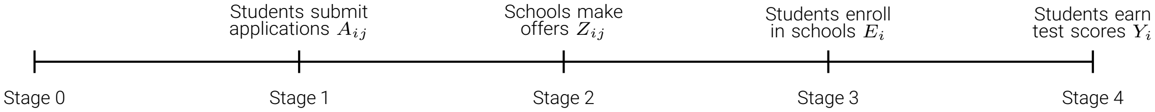
Beyond Lottery Effects

- Our approach: Model preferences, application decisions, enrollment decisions, and outcomes which allows us to:
  - → Characterize selection into application *enrollment*
  - → Estimate and validate reduced form causal effects, e.g., LATE
  - → Extrapolate causal effects beyond the lottery sample
- Identification will require instruments that affect application and enrollment decisions but do not directly influence outcomes
  - → Offers to choice programs
  - → Policy-induced *changes* in distance

This approach allows us to better understand "The What" and "The How" and evaluate counterfactual policies of interest

Model Overview: Timeline

- Goal: recover the distribution of treatment effects for the entire population — not just lottery participants
- Generalized Roy model linking four stages:



- Framework allows us to model selection into application and enrollment and selection-correct
- Identification from two sources of exogenous variation:
  - → Lottery offers at oversubscribed programs (randomized within strata)
  - → Quasi-experimental distance variation from the expansion of LAUSD's choice school portfolio

Model Overview: Application and Enrollment Stages

Application Stage:

- Decision to apply depends on school utility $U_{ij}(X_i, \theta_i, D_{ij})$ and application cost $c(A_i, X_i, \eta_i)$
- School utility depends on: school popularity δ_j, student observables X_i, relative distance D_{ij}, idiosyncratic choice school taste $\theta_i \sim F(\theta)$, and preference shock ϵ_{ij}
- θ_i is drawn from a finite mixture of normals with K types — some families love choice schools, others strongly prefer their neighborhood school
- Application costs c depend on observables X_i and cost shock η_i

Enrollment Stage:

- If applied and received an offer, student chooses between choice school and neighborhood school based on school utilities $v_j(X_i, \theta_i, D_{ij})$ plus post-lottery shock ξ_{ij}
- Students without offers (or who did not apply) default to neighborhood school

Potential Outcomes

We assume the following restrictions on mean potential outcomes (Y_{ij}):

$$E[Y_{ij} \mid X_i, D_i, \tau_i, E_i = j] = \qquad j = 1, \ldots, J$$

Potential Outcomes

We assume the following restrictions on mean potential outcomes (Y_{ij}):

$$E[Y_{ij} \mid X_i, D_i, \tau_i, E_i = j] = \gamma_x' X_i \qquad j = 1, \ldots, J$$

- X_i is a mean-zero vector of student attributes

Potential Outcomes

We assume the following restrictions on mean potential outcomes (Y_{ij}):

$$E[Y_{ij} \mid X_i, D_i, \tau_i, E_i = j] = \alpha_j + \gamma_x' X_i \qquad j = 1, \ldots, J$$

- X_i is a mean-zero vector of student attributes
- α_j is the average achievement at school j for the average student in the district

Potential Outcomes

We assume the following restrictions on mean potential outcomes (Y_{ij}):

$$E[Y_{ij} \mid X_i, D_i, \tau_i, E_i = j] = \alpha_j + \gamma_x' X_i + \gamma_{cx}' X_i \times \mathbf{1}\{j > 0\} \qquad j = 1, \ldots, J$$

- X_i is a mean-zero vector of student attributes
- α_j is the average achievement at school j for the average student in the district
- $\gamma_{cx}' X_i \times \mathbf{1}\{j > 0\}$ allows for heterogeneity in gains with respect to X_i at choice schools

Potential Outcomes

We assume the following restrictions on mean potential outcomes (Y_{ij}):

$$E[Y_{ij} \mid X_i, D_i, \tau_i, E_i = j] = \alpha_j + \gamma_x' X_i + \gamma_{cx}' X_i \times \mathbf{1}\{j > 0\} + h_j(\tau_i) \qquad j = 1, \ldots, J$$

- X_i is a mean-zero vector of student attributes
- α_j is the average achievement at school j for the average student in the district
- $\gamma_{cx}' X_i \times \mathbf{1}\{j > 0\}$ allows for heterogeneity in gains with respect to X_i at choice schools
- $h_j(\tau_i)$ models selection which depends on a student's latent type $\tau_i \in \{1, \ldots, K\}$

Potential Outcomes

We assume the following restrictions on mean potential outcomes (Y_{ij}):

$$E[Y_{ij} \mid X_i, D_i, \tau_i, E_i = j] = \alpha_j + \gamma_x' X_i + \gamma_{cx}' X_i \times \mathbf{1}\{j > 0\} + h_j(\tau_i) \qquad j = 1, \ldots, J$$

- X_i is a mean-zero vector of student attributes
- α_j is the average achievement at school j for the average student in the district
- $\gamma_{cx}' X_i \times \mathbf{1}\{j > 0\}$ allows for heterogeneity in gains with respect to X_i at choice schools
- $h_j(\tau_i)$ models selection which depends on a student's latent type $\tau_i \in \{1, \ldots, K\}$

We model selection in the following way:

$$h_j(\tau_i) = \sum_{k \neq 1} \gamma_k T_{ik} + \sum_{k \neq 1} \gamma_{ck} T_{ik} \times \mathbf{1}\{j > 0\}$$

- T_{ik} are indicators for belonging to type $k \in \{1, \ldots, K\}$
- γ_k allow for k-specific differences in achievement levels
- γ_{ck} allow for k-specific differences in achievement gains at choice schools ($j > 0$)

Empirical Outcome Model

- Equipped with model estimates, we can estimate the posterior individual-specific type probabilities ▸ Details

$$p^*_{ik} = E[T_{ik} \mid X_i, D_i, A_i, Z_i, E_i]$$

- p^*_{ik} serves two roles:
  - → p^*_{ik} serves as a control function to account for selection
  - → We allow for flexible heterogeneity by group membership

Empirical Outcome Model

- Equipped with model estimates, we can estimate the posterior individual-specific type probabilities ▶ Details

$$p^*_{ik} = E[T_{ik} \mid X_i, D_i, A_i, Z_i, E_i]$$

- p^*_{ik} serves two roles:
  - → p^*_{ik} serves as a control function to account for selection
  - → We allow for flexible heterogeneity by group membership

- For expositional purposes, let C_i be an indicator for choice school attendance and let $E[X_i] = 0$. Our empirical outcome model is:

Empirical Outcome Model

- Equipped with model estimates, we can estimate the posterior individual-specific type probabilities ▸ Details

$$p^*_{ik} = E[T_{ik} \mid X_i, D_i, A_i, Z_i, E_i]$$

- p^*_{ik} serves two roles:
  - → p^*_{ik} serves as a control function to account for selection
  - → We allow for flexible heterogeneity by group membership

- For expositional purposes, let C_i be an indicator for choice school attendance and let $E[X_i] = 0$. Our empirical outcome model is:

$$Y_i = \underbrace{X_i'\gamma_x + \sum_{k \neq 1} \gamma_k p^*_{ik}}_{\text{Ability}}$$

Empirical Outcome Model

- Equipped with model estimates, we can estimate the posterior individual-specific type probabilities ▸ Details

$$p^*_{ik} = E[T_{ik} \mid X_i, D_i, A_i, Z_i, E_i]$$

- p^*_{ik} serves two roles:
  - → p^*_{ik} serves as a control function to account for selection
  - → We allow for flexible heterogeneity by group membership

- For expositional purposes, let C_i be an indicator for choice school attendance and let $E[X_i] = 0$. Our empirical outcome model is:

$$Y_i = \underbrace{X_i'\gamma_x + \sum_{k\neq 1} \gamma_k p^*_{ik}}_{\text{Ability}} + \underbrace{\beta C_i}_{\beta = ATE}$$

Empirical Outcome Model

- Equipped with model estimates, we can estimate the posterior individual-specific type probabilities ▶ Details

$$p^*_{ik} = E[T_{ik} \mid X_i, D_i, A_i, Z_i, E_i]$$

- p^*_{ik} serves two roles:
  - → p^*_{ik} serves as a control function to account for selection
  - → We allow for flexible heterogeneity by group membership

- For expositional purposes, let C_i be an indicator for choice school attendance and let $E[X_i] = 0$. Our empirical outcome model is:

$$Y_i = \underbrace{X_i'\gamma_x + \sum_{k \neq 1} \gamma_k p^*_{ik}}_{\text{Ability}} + \underbrace{\beta C_i}_{\beta = ATE} + \underbrace{C_i \times \left[X_i'\delta_x + \sum_{k \neq 1} \gamma_{ck} p^*_{ik}\right]}_{X_i \text{ and } \theta_i \text{ TE heterogeneity}} + e_i$$

Empirical Outcome Model

- Equipped with model estimates, we can estimate the posterior individual-specific type probabilities ▸ Details

$$p^*_{ik} = E[T_{ik} \mid X_i, \mathbf{D_i}, A_i, \mathbf{Z_i}, E_i]$$

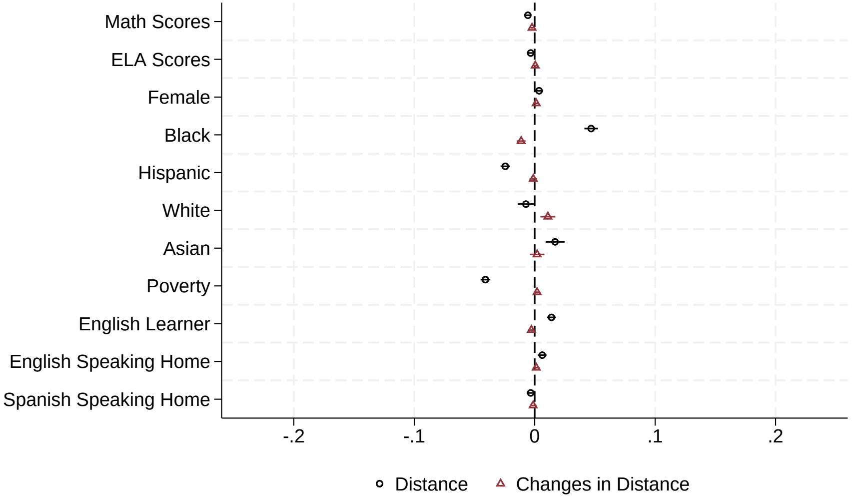
- p^*_{ik} serves two roles:
  - → p^*_{ik} serves as a control function to account for selection
  - → We allow for flexible heterogeneity by group membership

- For expositional purposes, let C_i be an indicator for choice school attendance and let $E[X_i] = 0$. Our empirical outcome model is:

$$Y_i = \underbrace{\mu_{n(i)} + \mu_{t(i)}}_{\text{Block and Time Effects}} + \underbrace{X_i'\gamma_x + \sum_{k \neq 1} \gamma_k p^*_{ik}}_{\text{Ability}} + \underbrace{\beta C_i}_{\beta = ATE} + \underbrace{C_i \times \left[X_i'\delta_x + \sum_{k \neq 1} \gamma_{ck} p^*_{ik} \right]}_{X_i \text{ and } \theta_i \text{ TE heterogeneity}} + e_i$$

Distance Balance

Cross-sectional distance is not balanced, but policy-induced changes are mostly balanced



Summary of Demand Estimates

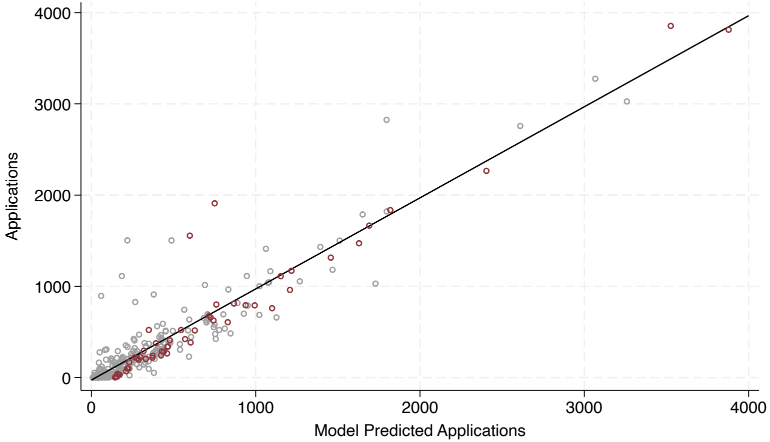
- **Model Fit:** The model produces forecast unbiased school-by-group application and enrollment rates in a holdout sample ▸ Evidence
- **Type Heterogeneity:** The data suggest a model with $K = 3$ types ▸ Evidence

$$\mu = \begin{pmatrix} -0.88 \\ -0.43 \\ 2.75 \end{pmatrix}, \qquad \sigma = \begin{pmatrix} 0.19 \\ 0.13 \\ 0.54 \end{pmatrix}, \qquad p = \begin{pmatrix} 0.44 \\ 0.49 \\ 0.07 \end{pmatrix}$$

- **Dominant Friction:** Application costs are substantially higher than travel costs ▸ Evidence

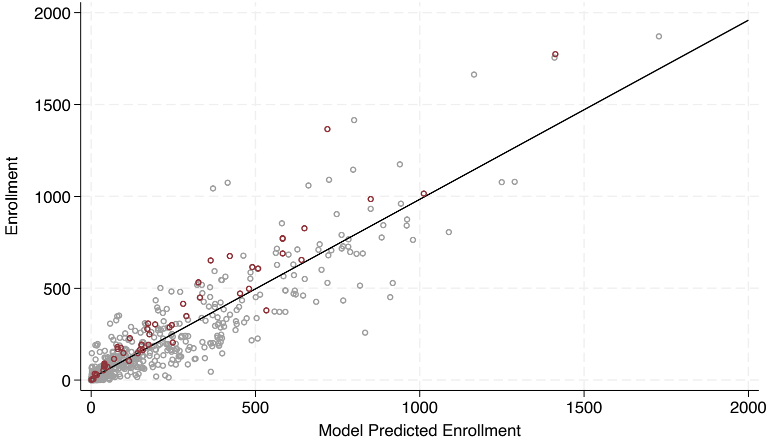
Model Fit

Out-of-sample predicted application decisions are forecast unbiased

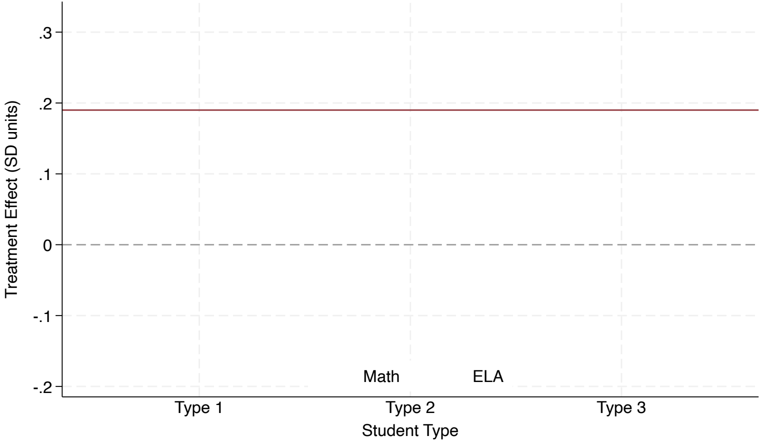


Model Fit

Out-of-sample predicted enrollment decisions are forecast unbiased

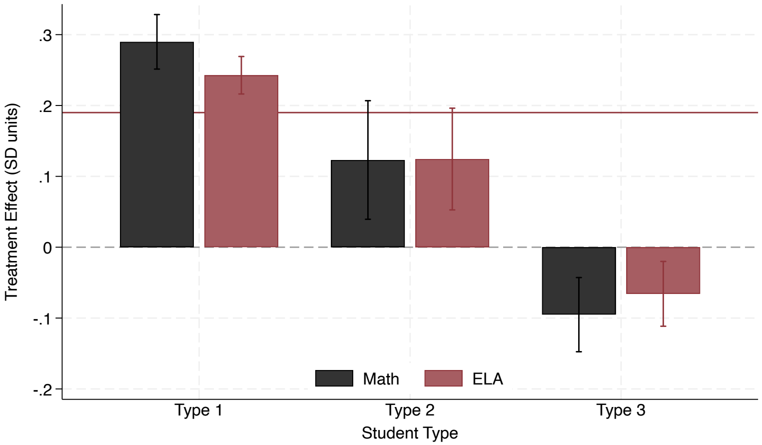


Choice schools are effective for the average student



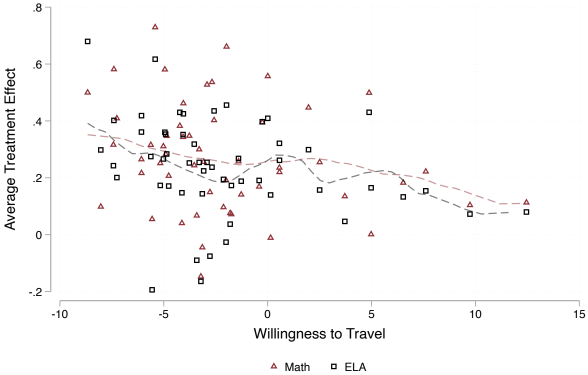
Validation Exercise

Evidence of negative selection on gains



▸ Validation Exercise

Demand for School Effectiveness Is Imperfectly Aligned



- Negative association between school popularity and causal effects
- Families' preferences reflect a mix of academic, social, and locational factors — the demand-based allocation does not coincide with the achievement-maximizing one

Three Types of Students

- Type 1: Would gain the most but do not participate
  - → Face largest barriers—lack of information, less parental motivation
- Type 2: Would benefit but many are screened out
  - → Relatively high potential gains from attending a choice school
  - → Rarely apply ⇒ never treated ⇒ gains are left on the table
- Type 3: Participate but benefit little
  - → Low treatment effects, but strong preferences for choice schools
  - → Occupy seats that could generate larger gains if allocated to Type 1 or Type 2 students

Policy can change the participation architecture and as a consequence change who benefits and by how much

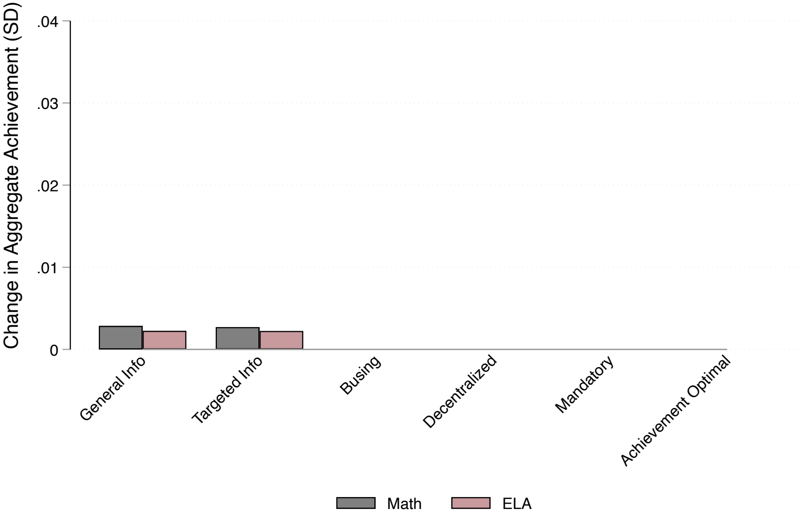
What Would Happen Under Alternative System Designs?

Counterfactual Policies of Interest

- **Policies that expand access:**
  - → Information nudge: 50% of students receive a boost to θ_i at the application and enrollment stages
  - → Generous busing: effectively eliminates travel costs
- **Alternative market designs:**
  - → Decentralized markets: families can apply to many schools and receive many offers ($\approx 35\%$ of districts)
  - → Mandatory participation with deferred acceptance and list-length limits: $c(a \mid X_i, \eta_i) = 0$, but families can still rank their neighborhood school first (NYC/Boston/Denver-style)
- **What we hold fixed vs. what varies:**
  - → Fixed across policies: menu of programs, capacities, school effectiveness
  - → Varying across policies: application costs, travel costs, participation rule, assignment mechanism

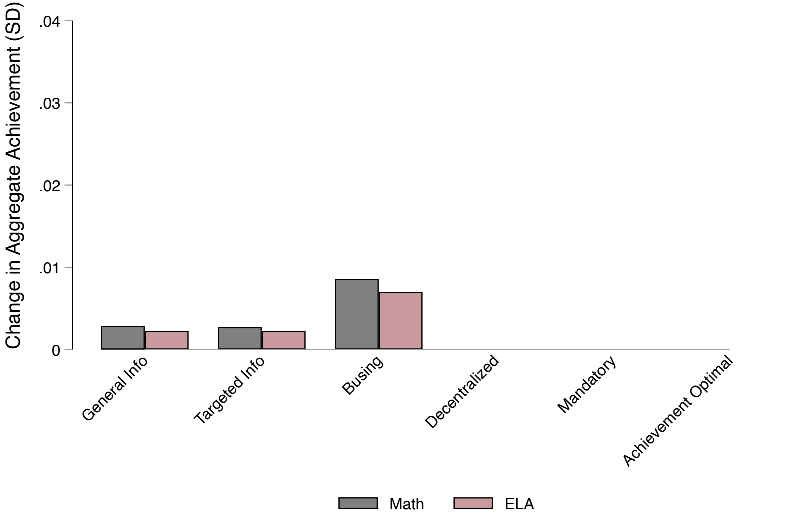
Changes in District-level Average Achievement

Information nudges moderately improve aggregate achievement



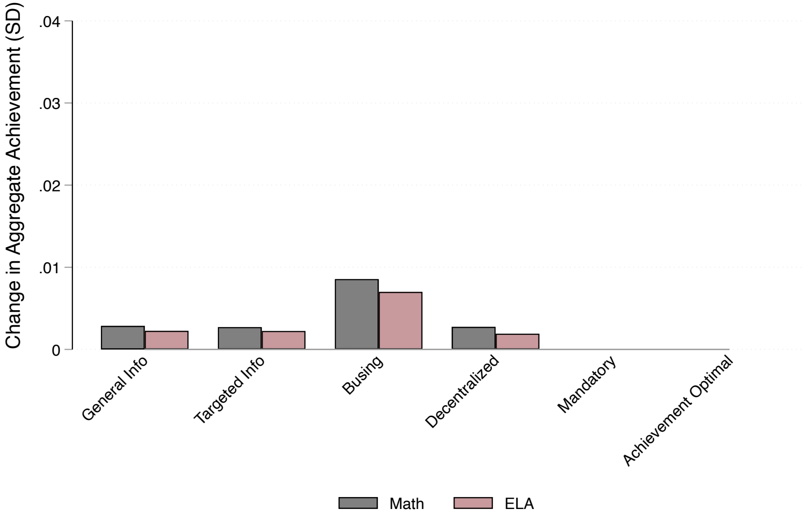
Changes in District-level Average Achievement

Fiscally unrealistic busing policies do better



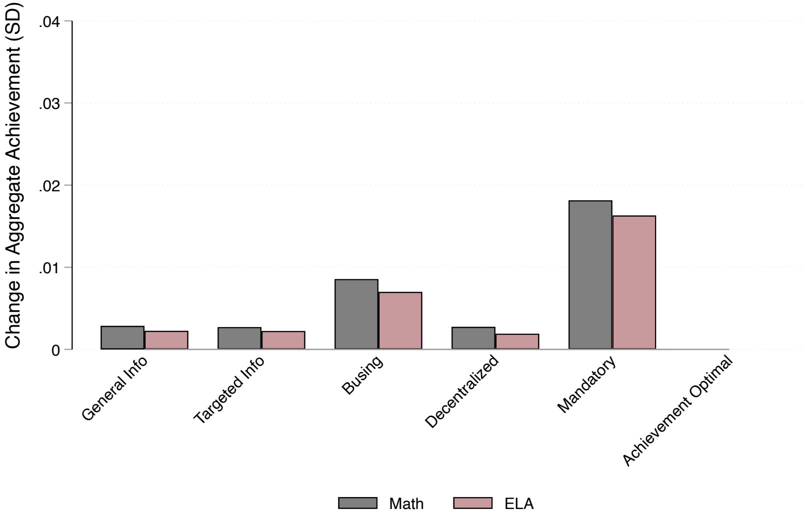
Changes in District-level Average Achievement

Decentralized markets do moderately better



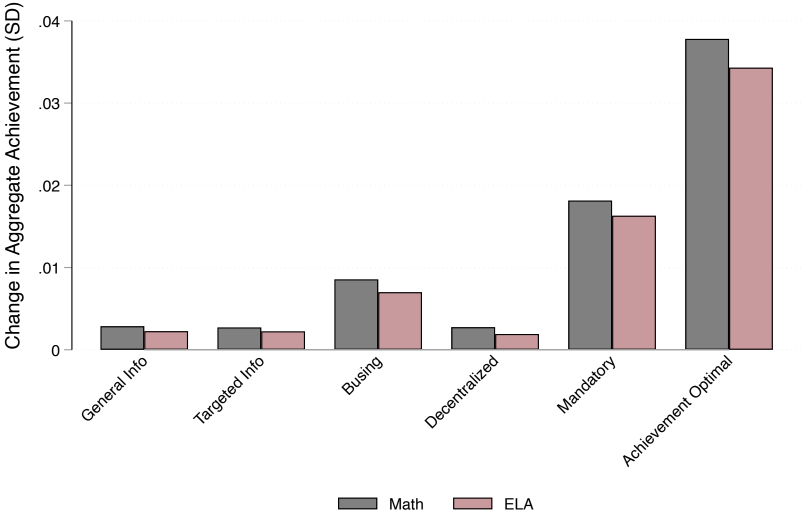
Changes in District-level Average Achievement

Mandatory, single-offer coordinated systems does substantially better

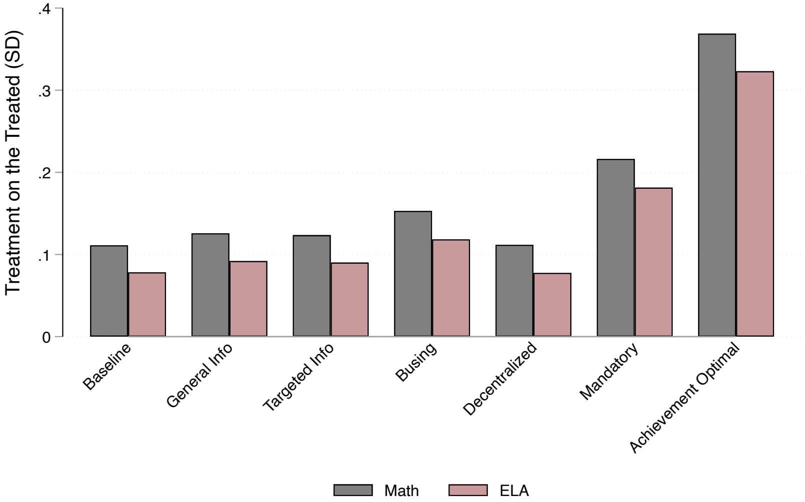


Changes in District-level Average Achievement

Even the best design captures only ∼55% — demand for quality does not fully align with effectiveness

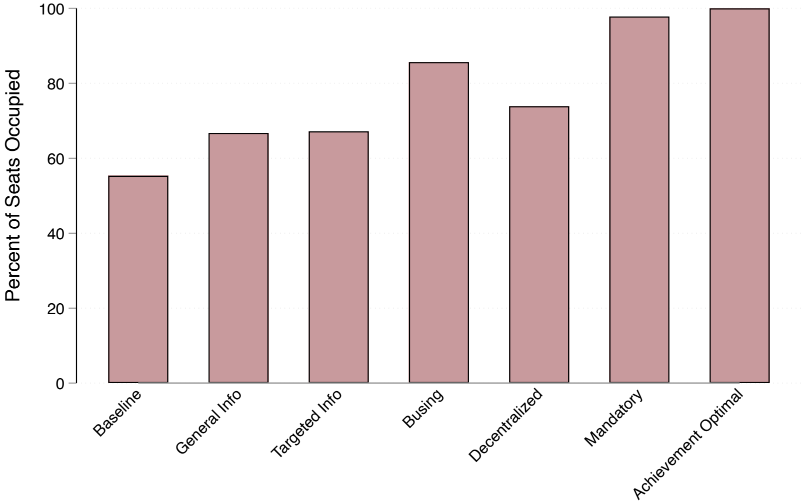


System Design Affects the TOT



Many seats are unfilled with opt-in design

The "first-mile problem" — getting students to apply — is a first-order barrier to the effectiveness of public school choice



The Punchline

- New fact: Opt-in is the most common way school districts have organized their choice offerings
- Who: Opt-in systems tend to segment the public education system based on achievement
- What: The average student has a sizable treatment effect ($\approx 0.19\sigma$)
- How: Negative selection on gains interacts with opt-in design to produce an allocative inefficiency in achievement
- Takeaways and implications for policy:
  - → Mandatory participation with DA raises district-level achievement by ≈ 0.016–0.019σ; TOT doubles
  - → Information and busing help at the margins, but the largest gains come from changing the participation architecture
  - → If opt-in is the political constraint, the task is to design systems that deliver gains despite it

System design — not school effectiveness alone —

shapes who benefits from public school choice

Lecture Takeaways

- Market design has real welfare effects (Abdulkadiroğlu, Agarwal & Pathak 2017)
- The participation rule (opt-in vs. mandatory) is a first-order design dimension that the theoretical literature has largely overlooked
- Negative selection on gains: opt-in systems screen out the students who would benefit most $\Rightarrow$ centralizing participation is the most effective lever but politically challenging option
- Alignment of demand tends to be a limiting factor: as long as families' preferences don't fully align with school effectiveness, there are limits to what can materialize from choice initiatives