Boundaries in Motion: Effectiveness and Efficiency of School Redistricting

Tomás Monarrez
Philadelphia Fed[1]

David Schönholzer
USC

Yifei Wang
USC

March 2026

[1]The views presented in this study are those of the authors alone and do not represent the position of the Federal Reserve Bank of Philadelphia, the Federal Reserve Board, or the Federal Reserve System.

School Boundaries Matter

Large literature on **school choice** (algorithms, charter schools)

Abdulkadiroğlu & Sönmez 2003, Pathak 2011, Abdulkadiroğlu, Angrist, Dynarski, Kane & Pathak 2011, Walters 2018, Campos & Kearns 2024

- But in most districts, neighborhood assignment still dominates
- Opt-in choice concentrated among advantaged families
  Campos, Bruhn, Chyn & Vazquez 2025

Districts must **draw attendance boundaries**, facing tradeoffs Monarrez 2023

- Capacity, proximity, diversity, and choice
- Boundaries adapt to openings, closures, and demographic shifts

→ **Redistricting** is the key policy lever — yet **no large-scale causal evidence**

This Paper: What We Do

Step 1: New **parcel-level data** on school attendance boundaries

- Annual SAB polygons for thousands of districts, 2011–2021 Precisely
- Linked to parcel-level house prices CoreLogic and school/district outcomes CCD, EdFacts
- Observe which parcels are assigned to which schools, and how this changes

Step 2: Evaluate district-wide impact of **redistricting events**

- Districts that significantly redraw their boundary arrangement
- Including changes in overlap between school zones (choice vs. zoning)
- Compare outcomes to matched control districts over time

Step 3 (in progress)**: Distributional** impacts of redistricting

- Which parcels are reassigned to better or worse schools?
- Do effects differ by house price rank and owner race/ethnicity?

What We Find

1. Redistricting reshapes districts along four dimensions

1. Proximity: changes in distance between parcels and assigned schools
2. Diversity: shifts in racial/ethnic composition of school catchment areas
3. Capacity: enrollment responds elastically, but composition is sticky
4. Choice: changes in the set of schools available to households

2. Impacts on outcomes

1. Effectiveness: student achievement → Tentative effects on average test scores
2. Efficiency: house prices → Little evidence of aggregate house price effect
3. Redistribution: [TBD]

Contributions to the Literature

1. School assignment & choice

Hoxby 2000; Abdulkadiroğlu et al. 2011; Campos & Kearns 2024

We study the extensive margin of choice — most districts blend neighborhood assignment with overlapping boundaries

2. Residential sorting and school composition

Billings, Deming & Rockoff 2014; Lutz 2011; Monarrez 2023

We estimate how redistricting events reshape diversity and enrollment district-wide

3. Capitalization & student outcomes

Black 1999; Bayer, Ferreira & McMillan 2007; Ding et al. 2024

First credible estimates of district-wide effects of boundary changes

Roadmap

1. Data
2. Measurement
3. Conceptual Framework
4. Empirical Strategy
5. Results
6. Summary and Next Steps

Roadmap

1. Data
2. Measurement
3. Conceptual Framework
4. Empirical Strategy
5. Results
6. Summary and Next Steps

Data Sources

Source	Contents	Coverage
Precisely	School attendance boundary polygons	2011–2021
CoreLogic	Residential parcel, locations, chars, prices	1999–2024
CCD	School enrollment by race/grade, status	2009–2021
Census / ACS	Block-level demographics	2010, 2020
CRDC / EdFacts	Test scores, school characteristics	2009–2021
SEDA	District-level achievement	2009–2018

→ **Key empirical innovation**: spatial join of millions of residential parcels to school attendance zones, observed annually for 11 years

School Attendance Boundaries (Precisely)

- **What**: Geographic polygons defining the catchment zone for each school
- **Coverage**: 11 annual snapshots (2011–2021), $\sim$786,000 school-level-year polygon observations
  - Near-national coverage, growing over time
- Three vendor eras:
  - Maponics (2011–2016)
  - Pitney Bowes (2017–2019)
  - Precisely (2020–2021)
- Substantial data cleaning: school ID deduplication, grade-level classification, spotty data interpolation

Residental Parcels (CoreLogic)

CoreLogic is a U.S. property data provider covering 99% of U.S. parcels

Property Basic Table

- Property characteristics (beds, baths, sq ft, year built, parcel coordinates)
- Land use and zoning codes
- Assessed & market values
- Geographic identifiers (APN, FIPS, lat/lon)

Owner Transfer Table

- Arms-length & non-arms-length sales
- Transaction price & date
- Buyer/seller names and type
- Mortgage amount & lender info

Parcel-to-School Assignment (CoreLogic)

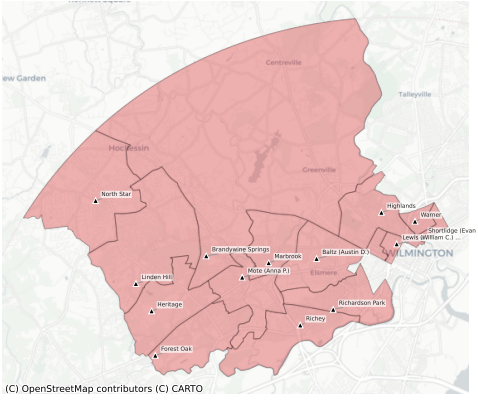
- **CoreLogic**: universe of residential property parcels with precise geocoded location
- **Spatial join**: every parcel → the SAB polygon it falls within → the school it is assigned to
- Schools serve grade levels: E (Elementary), M (Middle), H (High School)
- Parcels can enter data during study period (new residential development)
- Hundreds of millions of parcel × school × grade level × year assignments
- This crosswalk gives us the **assignment matrix** $A(t)$ — fundamental data object

→ Unlike prior work using district and block-group data aggregates, we observe the parcel-level assignment structure and how it changes year to year

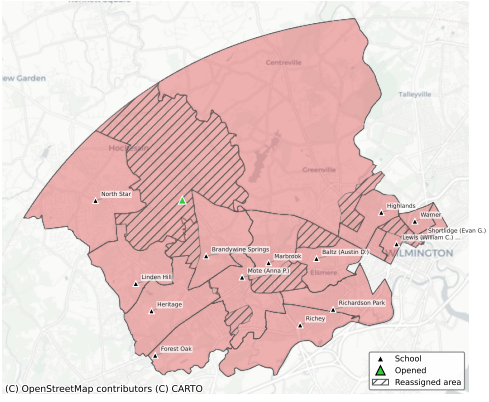
Example: Red Clay DE — Elementary School Redistricting

Red Clay Consolidated School District (DE) | Elementary | Event 2016

Pre (2015) (14 zones)



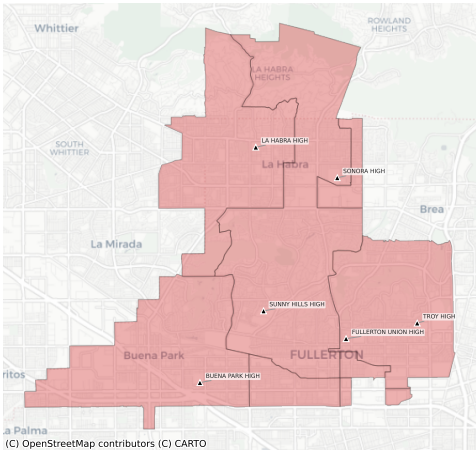
Post (2016) (15 zones)



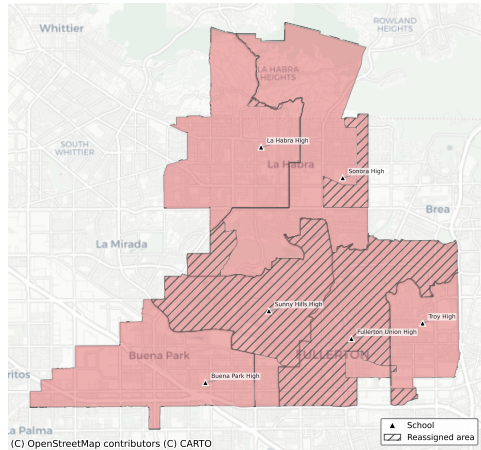
Example: Fullerton CA — High School Redistricting

Fullerton Joint Union High (CA) | High School | Event 2015

Pre (2014) (6 zones)



Post (2015) (6 zones)



School and Outcome Data

Common Core of Data (CCD):

- Enrollment by race and grade for every public school, 2009–2021
  - Socioeconomic status not measured reliably over time
- School directory: location, grade span, charter status

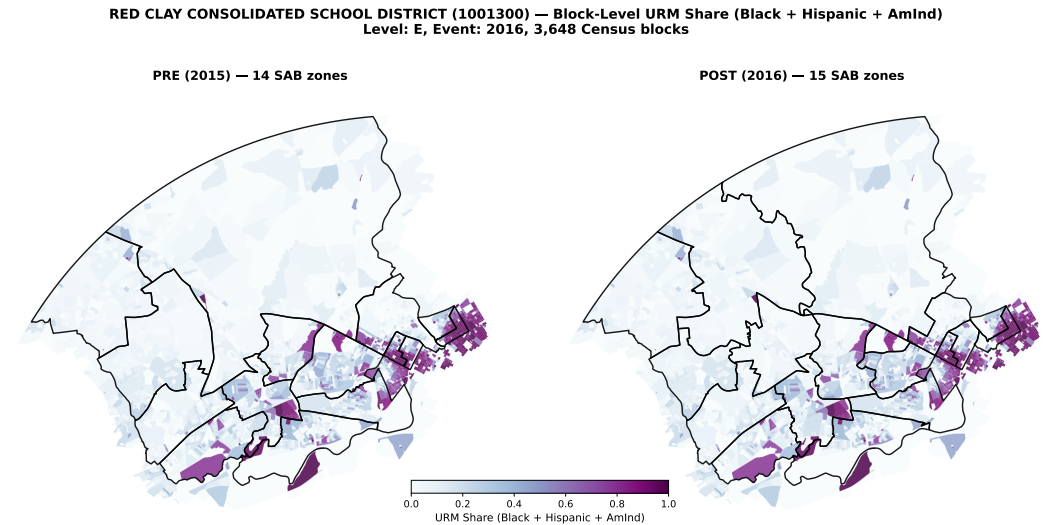
Outcomes:

- Test scores: EdFacts proficiency rates (math, reading); SEDA achievement
- Housing: CoreLogic transaction prices $\rightarrow \log(\text{sale amount})$

Residential Demographics (Census)

- 2010 and 2020 Decennial Census: block-level population by race
  - TIGER Shapefiles
  - NHGIS demographic tables
- 2015 and 2021 ACS: block-group-level socioeconomic characteristics
- **Area-weighted crosswalk**: Census blocks → SAB zones
  - Zone demographics change over time even when Census data is fixed — because boundaries move different blocks into different zones
- We focus on 2010 demographic patterns, to shut off possible endogeneity of residential makeup to school redistricting (more on this later!)

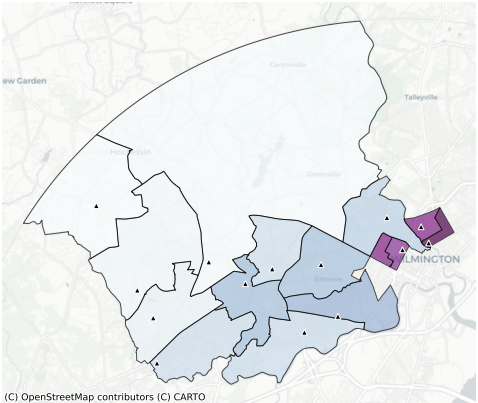
Example: Red Clay DE — 2010 Census Blocks



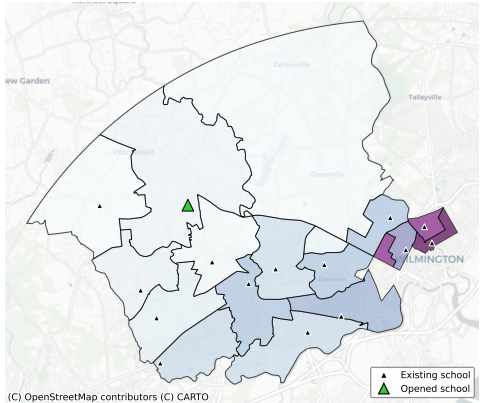
Example: Red Clay DE — Segregation in Assignments

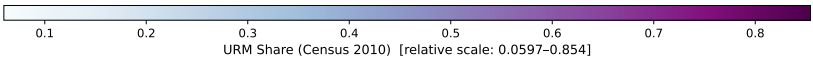
Red Clay Consolidated (DE) | Elementary | Event 2016 | fixed
URM Share (Census 2010)

Pre (2015) (14 zones)

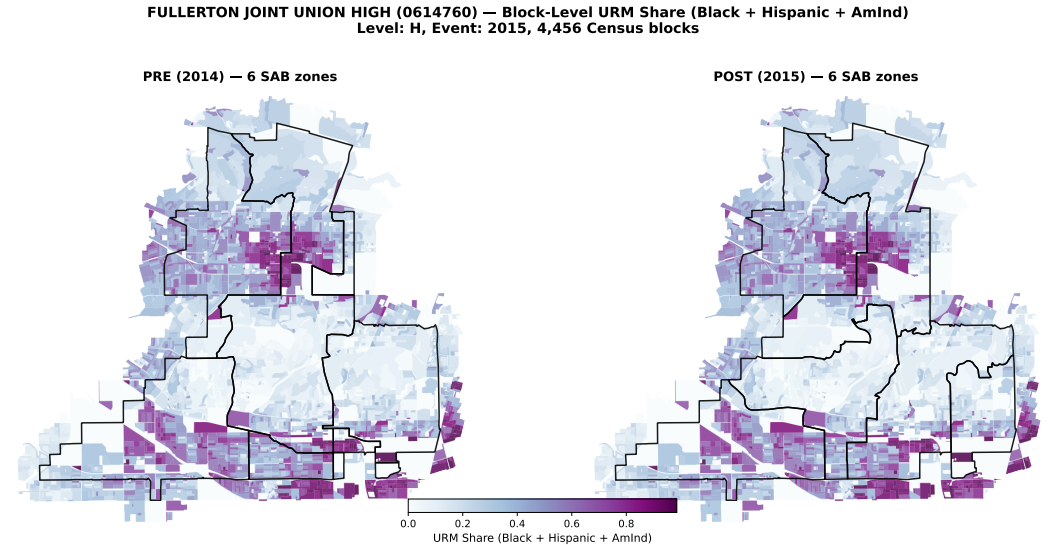


Post (2016) (15 zones)

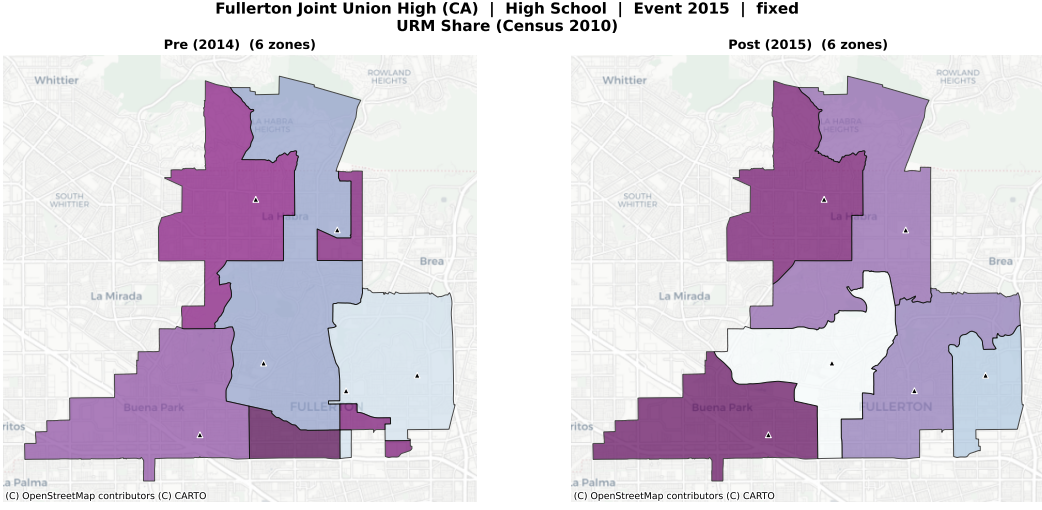




Example: Fullerton CA — 2010 Census Blocks



Example: Fullerton CA — Segregation in Assignments

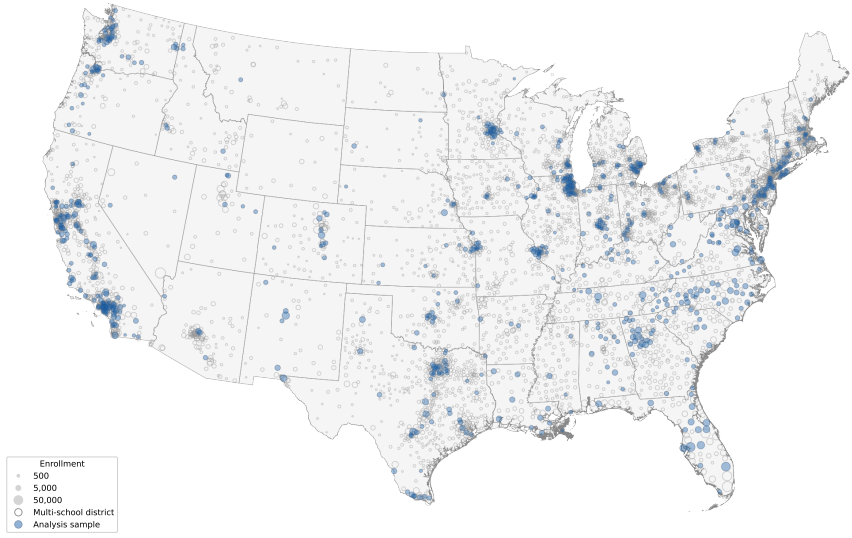


Sample Construction

Stage	Count
All public school districts	13,030
Multi-school	5,604
Precisely data coverage ($\geqslant$7 years)	1,978

Sample Coverage

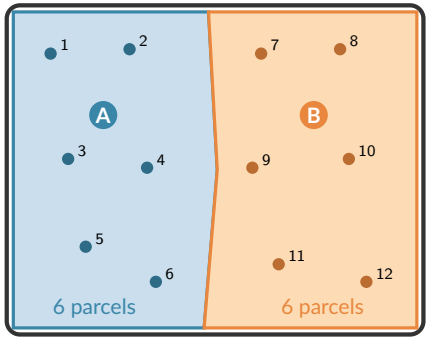
Multi-School Districts: All (5,558) vs. Analysis Sample (770)



Roadmap

1. Data
2. Measurement
3. Conceptual Framework
4. Empirical Strategy
5. Results
6. Summary and Next Steps

Parcels Within Boundaries → Assignment Matrix



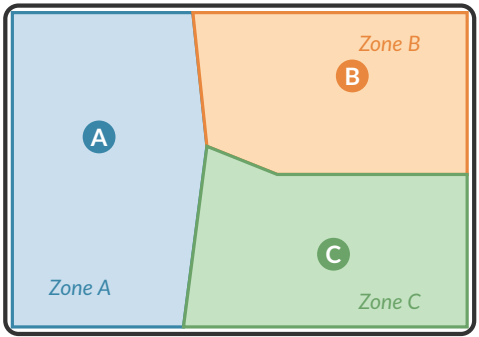
spatial join

$A(t)$

	A	B	
p_1	1	0	Zone A
p_2	1	0	
p_3	1	0	
p_4	1	0	
p_5	1	0	
p_6	1	0	
p_7	0	1	Zone B
p_8	0	1	
p_9	0	1	
p_{10}	0	1	
p_{11}	0	1	
p_{12}	0	1	

Each parcel is geocoded and spatially joined to the SAB polygon it falls within → one row per parcel in $A(t)$.

Case 1: Perfect Partition (Zoned District)

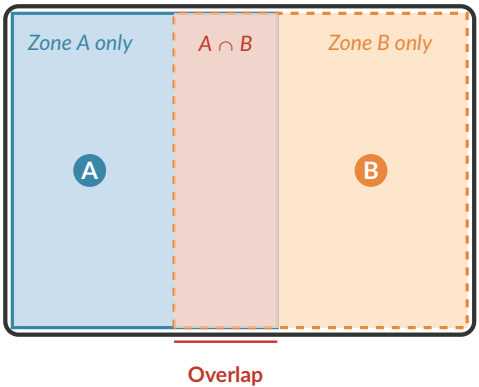


Properties:

- Zones are mutually exclusive: no parcel belongs to two zones
- Zones are exhaustive: every parcel in the district is assigned
- Each parcel → exactly one school
- $A(t)$ has exactly one 1 per row

Most US school districts approximate this structure at each grade level (elementary, middle, high).

Case 2: Overlapping Boundaries (Mixed District)



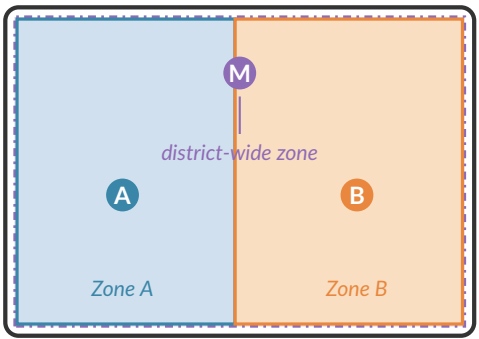
Properties:

- Parcels in overlap region assigned to both schools
- Overlap $\Rightarrow$ rows of $A(t)$ with multiple 1s
- Arises from magnet zones, transfer zones, dual-assignment areas

Implications:

- School choice index > 1 in overlap
- Coverage density $> 1/J$
- Redistricting can create or dissolve overlap zones

Case 3: Open Enrollment (OE) School



Open enrollment (OE):

- Magnet M's zone $=$ entire district
- Every parcel can attend M
- Column M in $A(t)$ is all 1s
- Coverage$_M$ $=$ N parcels

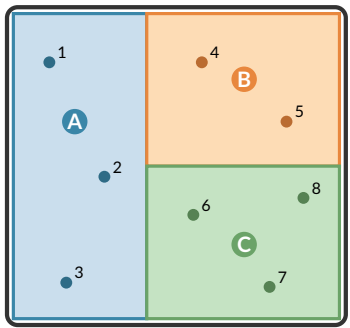
In the data:

- OE $\iff$ school covers $\geqslant 95\%$ of district parcels
- OE schools excluded from boundary change measures
- Common for magnets, charters

Redistricting: Boundary Change

Assume district initial condition:

8 parcels, 3 schools



Assignment matrix $A(t)$

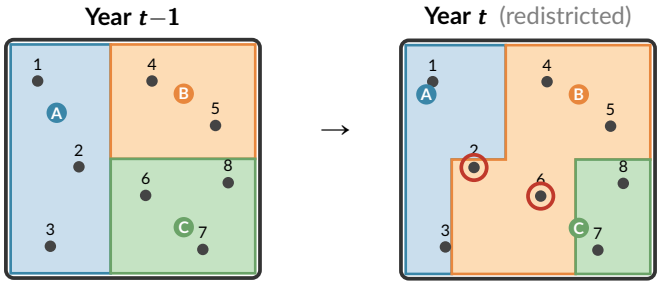
$$
\begin{array}{c} \\ p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \\ p_7 \\ p_8 \end{array}
\begin{array}{c}
\begin{array}{ccc} A & B & C \end{array} \\
\left[\begin{array}{ccc}
1 & 0 & 0 \\
1 & 0 & 0 \\
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & 0 & 1 \\
0 & 0 & 1
\end{array}\right]
\end{array}
$$

Row sums (access): all $= 1$ (perfect partition)

Col sums (coverage): $(3,\ 2,\ 3)$

Density: $8/24 = 1/3$

Redistricting: Boundary Change $\rightarrow \Delta A(t)$



Change matrix $\Delta A(t)$

$$\begin{array}{c} \\ p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \\ p_7 \\ p_8 \end{array} \begin{bmatrix} 0 & 0 & 0 \\ -1 & +1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & +1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- p_2: School A $\rightarrow$ B
- p_6: School C $\rightarrow$ B
- 6 parcels unchanged

Change rate $= 4/24 \approx 17\%$

The Assignment Change Matrix $\Delta A(t)$

Consider a district with N parcels and J schools observed in year t

Define the $N \times J$ **assignment matrix**:

$$A_{ij}(t) = \mathbf{1}\big[\text{parcel } i \text{ is assigned to school } j \text{ at time } t\big]$$

School redistricting is captured by changes in the matrix

$$\Delta A(t) \;=\; A(t) - A(t-1)$$

Entries take values in $\{-1,\ 0,\ +1\}$:

- $+1$: parcel newly assigned to school
- -1: parcel no longer assigned to school
- 0: no change

Treatment: Stable-Territory Change Rate

The raw change rate conflates boundary changes with mechanical changes:

- New parcels entering (developed during study period)
- Occasional spottiness in both CoreLogic and Precisely

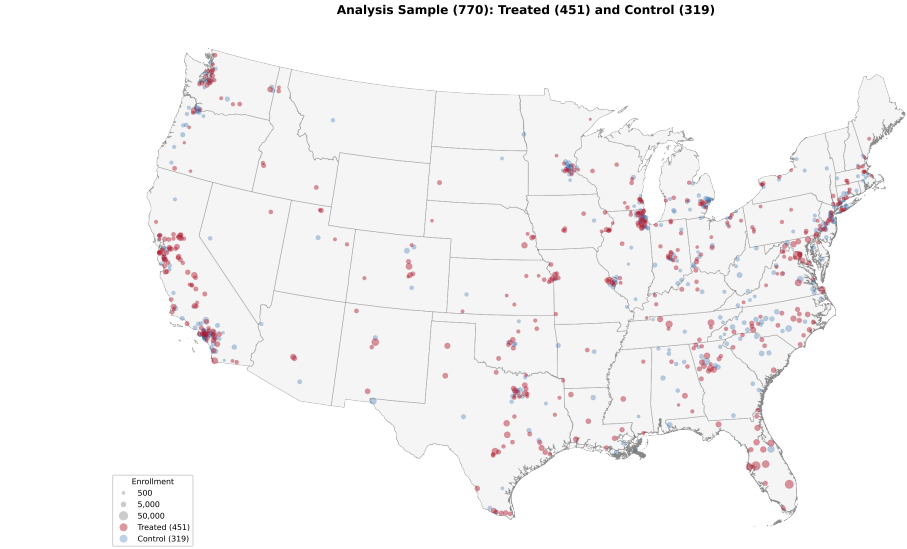
Adjustment: restrict $\Delta A(t)$ to parcels and schools active in *both* $t-1$ and t:

$$D(t) = \frac{1}{N_{\text{stable}} \cdot J_{\text{stable}}} \sum_{\substack{i \in \mathcal{N}_{t-1} \cap \mathcal{N}_t \\ j \in \mathcal{J}_{t-1} \cap \mathcal{J}_t}} \left| \Delta A_{ij}(t) \right|$$

Treatment definition:

- **Treated**: exactly one year with $D(t) > 0.05$
- **Control**: $D(t) < 0.01$ in all years

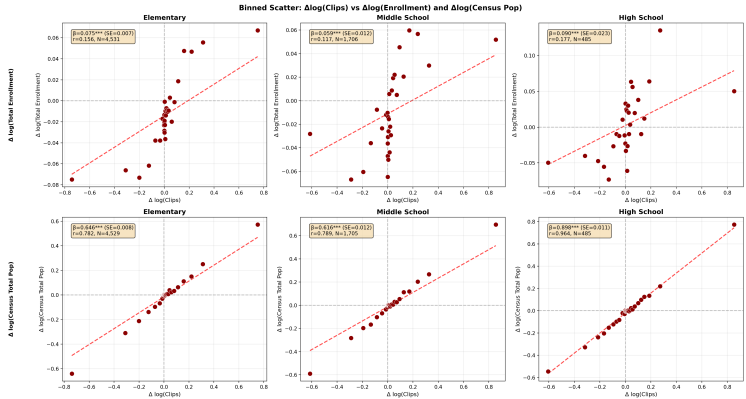
Distribution of Redistricting Events



506 treated districts, 668 district×level matched pairs across 2012–2021 (353

Assignment Changes and Enrollment Dynamics

- **Elasticity** ≈ 0.10: a 10% increase in assigned parcels $\rightarrow$ $\sim$10% enrollment increase
- Consistent across racial groups
- School enrollment counts and composition are sticky



Roadmap

1. Data
2. Measurement
3. Conceptual Framework
4. Empirical Strategy
5. Results
6. Summary and Next Steps

Districts' Boundary Drawing Problem

Districts face competing objectives when setting SABs

Commonly, debate over school redistricting revolves around 4 key objectives

- **Proximity** (d): Minimizing daily commuting cost
- **Capacity** (h): Balancing school sizes (enrollment/staffing)
- **Diversity** (s): Avoiding isolation of demographic/socioeconomic groups
- **Choice** (c): Providing flexibility in assignments

District k chooses boundaries that maximize objective

$$A_k^* = \arg\max_{A \in \mathbf{A}} \; U_k(d, h, s, c)$$

Proximity Objective: Average Parcel-to-School Distance

Average distance: mean distance from each parcel to its assigned school(s)

$$d_k = \frac{1}{\sum_{i,j} A_{ij}(t)} \sum_{i=1}^{N} \sum_{j=1}^{J} A_{ij}(t) \cdot d(i,j)$$

where $d(i,j)$ is the distance from parcel i to school j

Redistricting channel:

- School openings tend to reduce average distance (new school closer to some parcels)
- School closures tend to increase distance
- Pure boundary redraws: ambiguous

Capacity Objective: School Imbalance

School Capacity Imbalance index measures how evenly parcels are distributed across schools

Define

$$\text{Coverage}_j(t) = \sum_{i=1}^{N} A_{ij}(t)$$

Index derived from column sums of $A(t)$:

$$h_k(t) = 1 - \text{HHI}_k = 1 - \sum_{j=1}^{J} \left(\frac{\text{Coverage}_j}{\sum_k \text{Coverage}_k} \right)^2$$

- $h_k(t) = 1 - 1/J$: perfectly equal territory distribution
- $h_k(t) \approx 0$: one school dominates all territory

Redistricting channel: Boundary changes can equalize or concentrate school sizes, affecting resource allocation and crowding

Diversity Objective: District-wide Segregation

School j with subgroup population n_j^a, total n_j, and school-level share $s_j^a = n_j^a / n_j$

Variance ratio — excess within-zone exposure to group a relative to cross-zone exposure:

$$V = \sum_{j=1}^{J} \left[\frac{n_j^a}{N^a} s_j^a - \frac{n_j^b}{N^b} s_j^a \right] \quad \text{where } n_j^b = n_j - n_j^a,\ N^b = N - N^a$$

Dissimilarity index — share of group a that would need to move across zones for even distribution:

$$D = \frac{1}{2\, N\, \bar{s}^a\, (1 - \bar{s}^a)} \sum_{j=1}^{J} n_j \left| s_j^a - \bar{s}^a \right|$$

Two versions, same formulas:

- Census-based segregation changes driven purely by boundary movement, holding residential demographics fixed
- Enrollment-based actual attendance patterns including behavioral responses

Choice Objective: School Choice Index

School choice index measures the extent of overlapping assignment zones in a district.

Each parcel i is assigned to $\sum_j A_{ij}(t)$ schools. In a perfect partition, every parcel maps to exactly one school, so total assignments $= N$. With overlapping zones, total assignments $> N$.

$$c_k(t) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{J} A_{ij}(t)}{N \cdot J}$$

- $c_k = 1/J$: perfect partition (each parcel $\rightarrow$ exactly one school)
- $c_k > 1/J$: some parcels assigned to multiple schools (choice)
- $c_k = 1$: full open enrollment (every parcel can attend every school)

Conceptual Framework Summary

From assignment matrix to econometrics

Object		Yields
$A(t)$	$\rightarrow$	School choice, capacity imbalance
$A(t) \times \mathbf{d}$	$\rightarrow$	Mean assignment distance
$A(t) \times$ Census	$\rightarrow$	Segregation in assignments
$\Delta A(t)$	$\rightarrow$	Redistricting (treatment variable)

$\rightarrow$ The assignment matrix provides a **unified framework**: treatment, mechanisms, and some outcomes all derive from the same object

Roadmap

1. Data
2. Measurement
3. Conceptual Framework
4. Empirical Strategy
5. Results
6. Summary and Next Steps

Matching Strategy

Goal: Pair each treated district with a comparable control observed at the same baseline

Mahalanobis Distance

For covariate vectors $\mathbf{x}_i, \mathbf{x}_j$ measured at $t = -1$:

$$d(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)' \hat{\Sigma}^{-1} (\mathbf{x}_i - \mathbf{x}_j)}$$

- Scale-invariant: normalizes by the covariance structure of $\mathbf{X}$
- Accounts for correlation across covariates
- 1:1 nearest-neighbor, without replacement
- Exact match on school level (E / M / H)

Matching Covariates ($t = -1$)

1. Log total enrollment
2. Variance Ratio index (URM)
3. School choice index
4. Log avg. parcel–school distance
5. Capacity Imbalance index

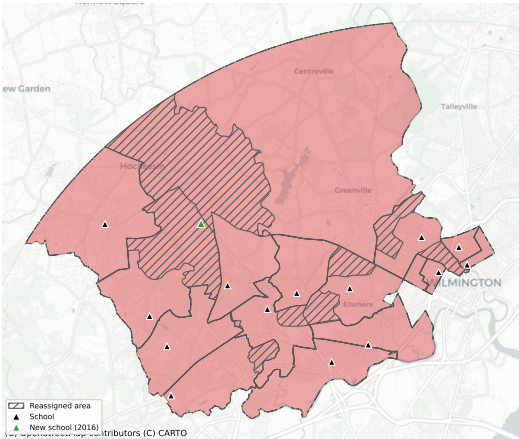
→ **668 matched pairs** (906 unique districts)

353 Elementary + 215 Middle + 100 High School

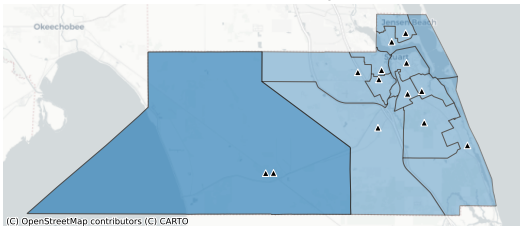
Example: Red Clay DE — Matched Pair

SAB Boundaries — Elementary

Treated: Red Clay, DE
14 zones



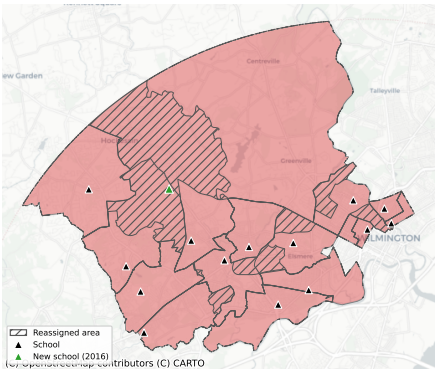
Control: Martin, FL
13 zones, 1.6x overlap



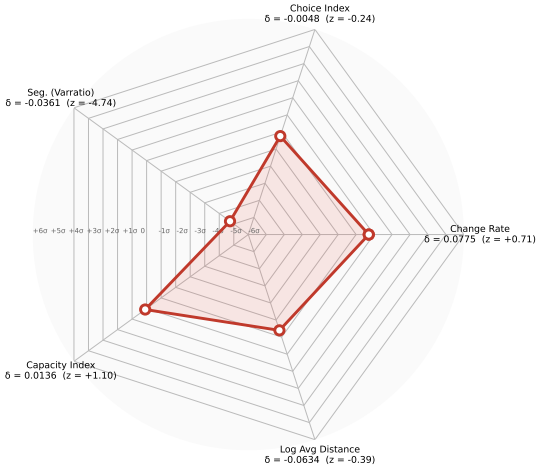
Example: Red Clay DE — Matched Pair

Red Clay, DE vs Martin, FL
Elementary | Event year 2016

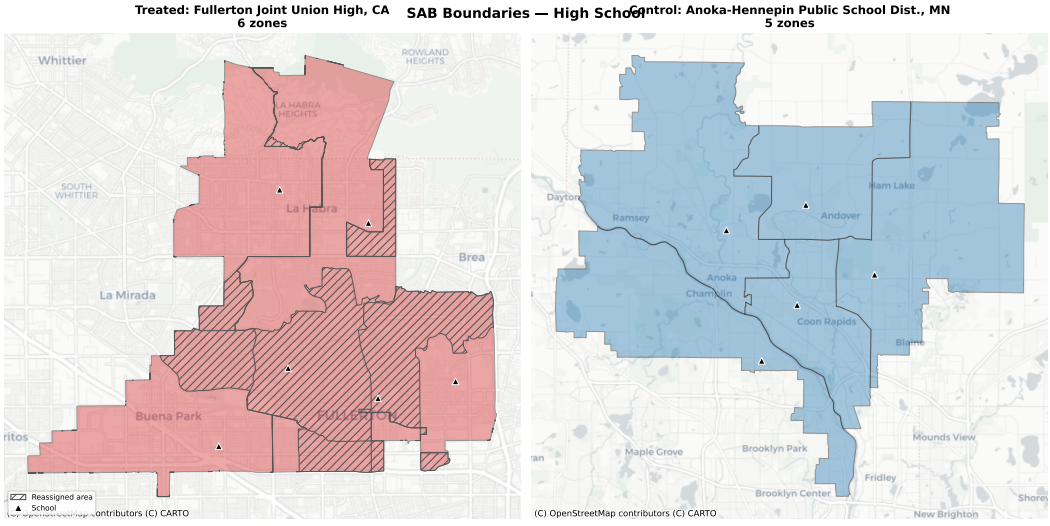
Treated: Red Clay, DE
14 zones



Red Clay, DE vs Martin, FL
Elementary | Event 2016 | z-scored vs 353 E-level pairs



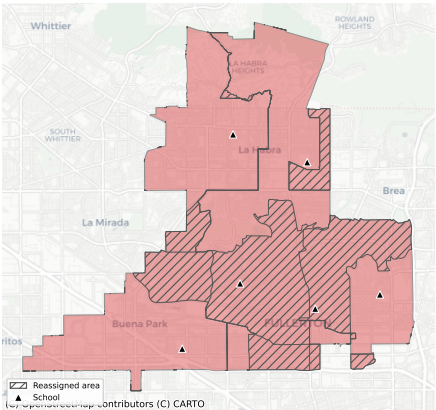
Example: Fullerton CA — Matched Pair



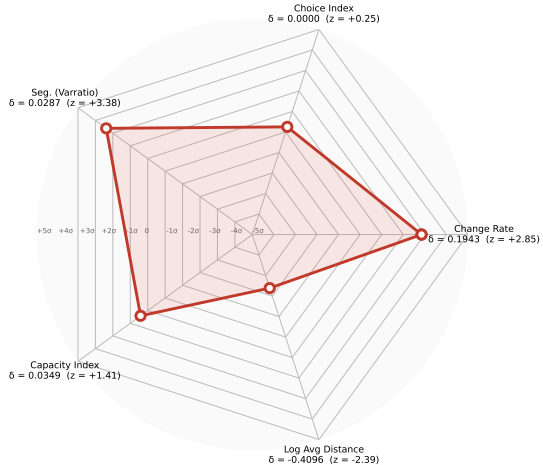
Example: Fullerton CA — Matched Pair

Fullerton Joint Union High, CA vs Anoka-Hennepin Public School Dist., MN
High School | Event year 2015

Treated: Fullerton Joint Union High, CA
6 zones



Fullerton Joint Union High, CA vs Anoka-Hennepin Public School Dist., MN
High School | Event 2015 | z-scored vs 100 H-level pairs



Difference-in-Differences Specification

Event-study model for matched pair p, district i, event time k:

$$y_{ipk} = \sum_{\substack{k=-3 \\ k\neq -1}}^{3} \beta_k \cdot \text{Treated}_i \cdot \mathbf{1}[t = k] + \alpha_{pk} + \gamma_i + \varepsilon_{ipk}$$

- α_{pk}: pair $\times$ event-time fixed effects; γ_i: district fixed effects
- Reference period: $k = -1$ (year before redistricting); Event window: ± 3 years
- Standard errors clustered at the district level

→ **Identification**: β_k captures the causal effect of redistricting at horizon k, under parallel trends assumption (no anticipation and SUTVA)

→ **Robustness to new DD lit.**: Matched pairs, untreated controls, pair-event time effects, ensure valid comparisons (positive weighting)

Dimensions of Heterogeneity

1. **Pooled model**: Mean effect of all redistricting events
   - Difficult to interpret

2. **Redistricting Trigger**: School Opening, Closure, Fixed Redistricting
   - Fixed redistricting is more common in our sample
   - Existing literature for openings and closures Engberg et al. 2012; Carlson & Lavertu 2016

3. **Grade Level**: Elementary, High
   - Evidence of heterogeneity in capitalization effects Schonholzer & Lafortune 2022

4. **Redistricting Objectives**
   - Proximity, Capacity, Diversity, Choice
   - Requires data-driven assessment of revealed redistricting objective

Clustering Redistricting Objectives ("First Stage")

Goal: Redistricting events are heterogeneous — districts redraw boundaries for different reasons. We classify events by their revealed objectives

Setup. For each matched pair p, compute:

$$\hat{\delta}_p \;=\; \left(\bar{Y}^T_{p,\text{post}} - \bar{Y}^T_{p,\text{pre}}\right) - \left(\bar{Y}^C_{p,\text{post}} - \bar{Y}^C_{p,\text{pre}}\right)$$

yielding a vector $\hat{\delta}_p = (\hat{\delta}_{p1}, \ldots, \hat{\delta}_{pJ})' \in \mathbb{R}^J$ for each pair across outcomes j

Clustering variables ($J = 4$):

1. Dissimilarity index — segregation
2. School choice index — choice
3. Log average distance — proximity
4. Adjusted capacity index — capacity

All school-level pairs are pooled into a single clustering ($N = 668$).

K-Means Procedure and Cluster Selection

Standardize. Let $\tilde{\delta}_p = \text{diag}(\hat{\sigma})^{-1}(\hat{\delta}_p - \bar{\delta})$ so each variable contributes equally.

K-means. Partition pairs into K clusters $\mathcal{C}_1, \ldots, \mathcal{C}_K$ by solving:

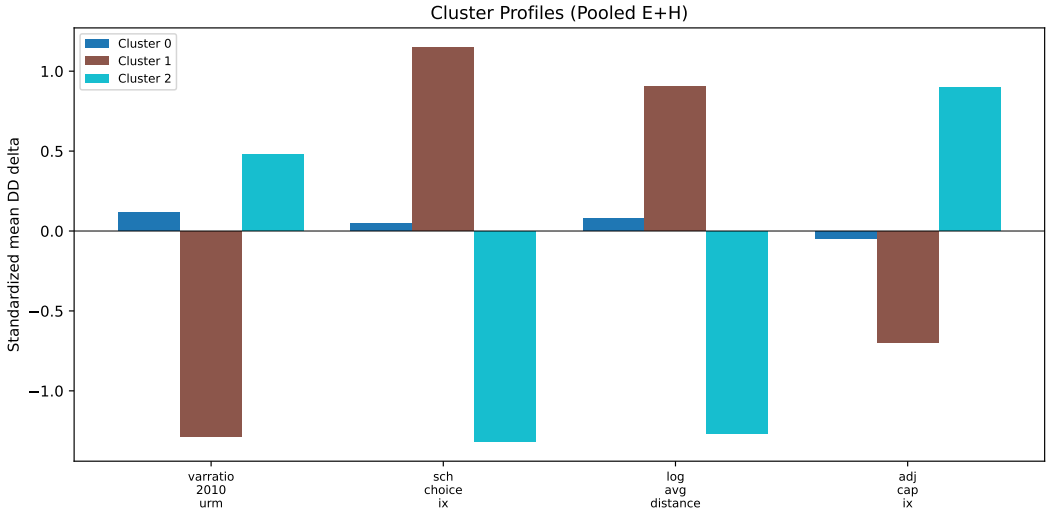
$$\min_{\{\mathcal{C}_k\},\{\boldsymbol{\mu}_k\}} \sum_{k=1}^{K} \sum_{p \in \mathcal{C}_k} \left\| \tilde{\delta}_p - \boldsymbol{\mu}_k \right\|^2$$

where $\boldsymbol{\mu}_k = |\mathcal{C}_k|^{-1} \sum_{p \in \mathcal{C}_k} \tilde{\delta}_p$ is the cluster centroid.

Choosing K. Goodness of fit scores over $K \in \{2, \ldots, 8\} \rightarrow$ select $K = 3$

Cluster 0 ($n = 504$)	**Cluster 1** ($n = 78$)	**Cluster 2** ($n = 86$)
Low-impact	Choice-expanding	Distance-Reducing
Minimal shifts on all dimensions	↑ choice, ↑ distance, ↓ capacity, ↓ segregation	↓ choice, ↓ distance, ↑ capacity, ↑ segregation

K-Means Clustering Profiles



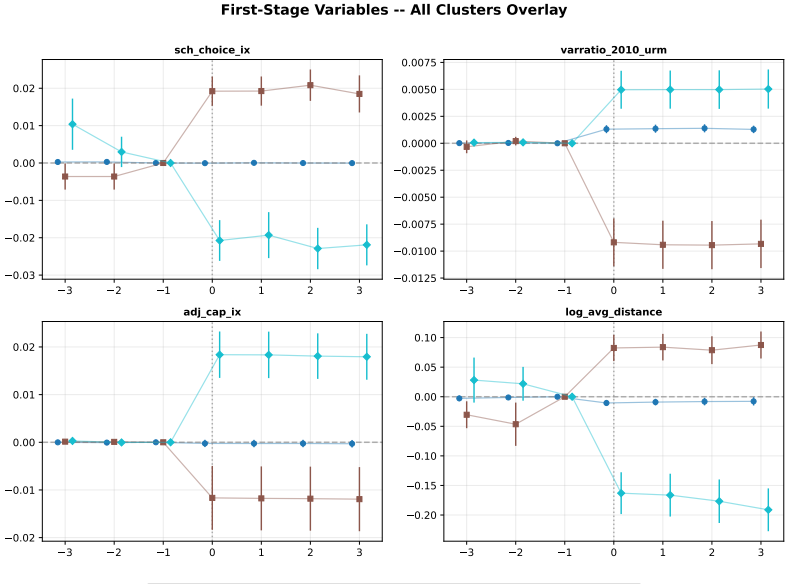
Roadmap

1. Data
2. Measurement
3. Conceptual Framework
4. Empirical Strategy
5. Results
6. Summary and Next Steps

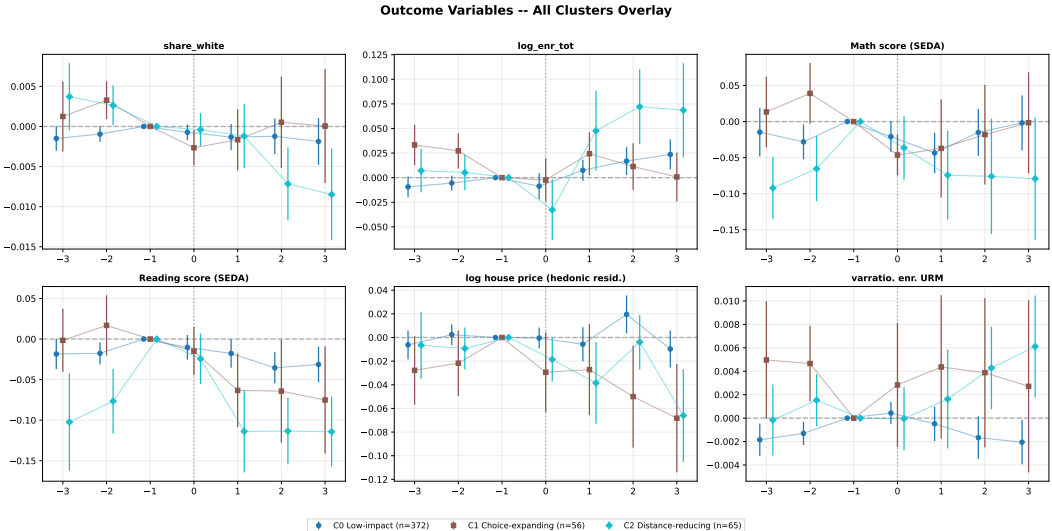
Summary Statistics — District Level

	All	Multi-sch.	Sample	Treated	Control	Matched Ctrl
Panel A: Demographics						
Enrollment (1,000s)	3.49	4.86	8.42	15.52	6.54	9.61
Number of schools	6.83	8.89	14.58	24.39	12.05	16.38
Share URM (enrollment)	0.32	0.34	0.45	0.52	0.43	0.49
Share URM (population)	0.18	0.19	0.25	0.28	0.24	0.27
Panel B: Socioeconomic						
Median HH income ($1,000s)	41.53	44.45	58.09	64.97	56.29	62.02
Share college educated	0.26	0.27	0.31	0.35	0.30	0.33
Poverty rate	0.12	0.12	0.12	0.11	0.12	0.11
House price ($1,000s)	228.04	230.03	292.39	324.70	281.63	306.47
Panel C: Segregation						
Variance ratio URM (Census)	0.03	0.03	0.03	0.03	0.03	0.03
Variance ratio URM (enrollment)	0.03	0.03	0.05	0.06	0.05	0.05
Panel D: SAB Indices						
School choice index		0.76	0.35	0.28	0.36	0.30
Adjusted capacity index		0.78	0.77	0.81	0.76	0.80
N (districts)	13030	8033	2056	512	1632	424

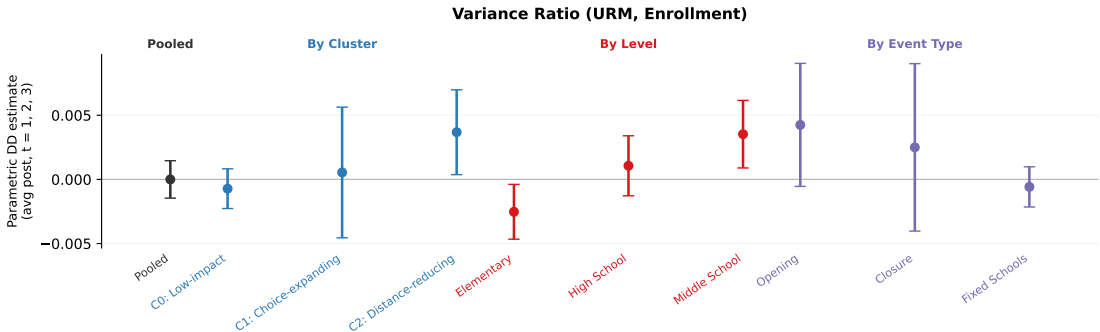
First Stage Effects



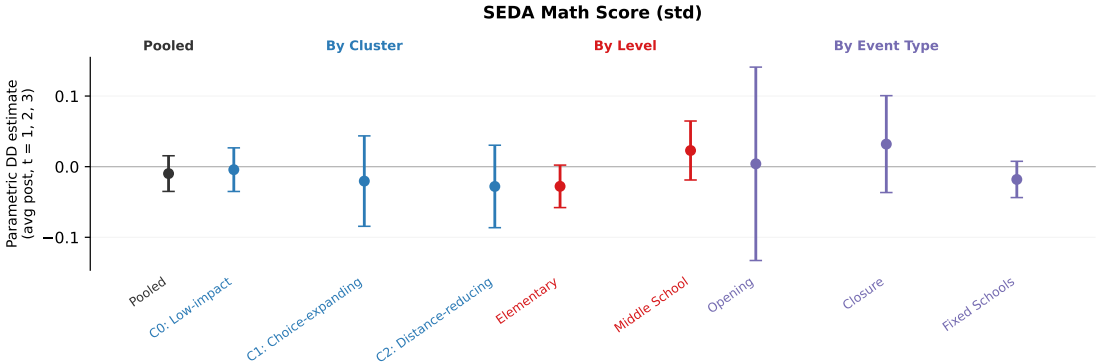
Main Outcomes



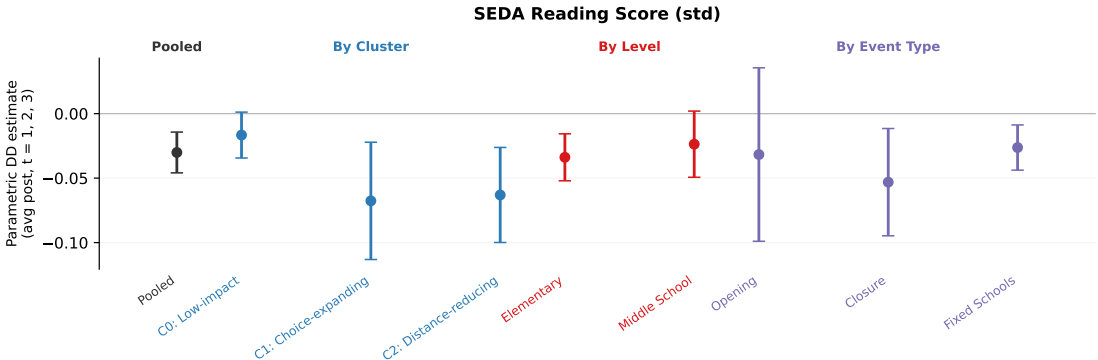
Parametric DD Estimates: Segregation of Enrollment



Parametric DD Estimates: Math Scores (SEDA)



Parametric DD Estimates: Reading Scores (SEDA)



Housing Price Residuals: Hedonic Models

House prices reflect both structural characteristics and location value. We isolate the location component by estimating hedonic regressions within each state, then using residuals as outcomes.

Dependent variable: $\ln(\text{sale_amount}_{it})$ for parcel i in year t.

Hedonic covariates (from CoreLogic property records):

- Log acres, log living sqft, log building sqft
- Year built (indicators, winsorized $\geqslant 1875$)
- Bedrooms (indicators, capped at 30)
- Bathrooms (indicators, capped at 15)
- Number of buildings (indicators, capped at 10)
- Stories (indicators, capped at 20)
- Pool (binary), Fireplace (binary)
- Missing-value dummies for each covariate

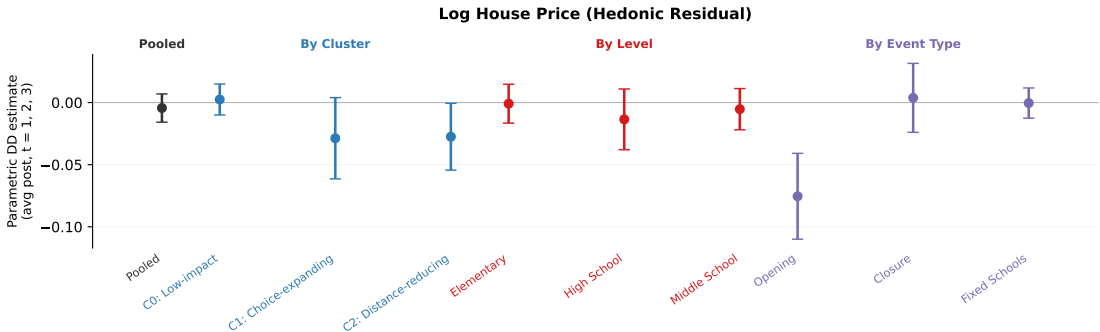
Housing Price Residuals: Hedonic Models

Three nested models (residuals r_1–r_3):

	Specification	Hedonics	Prop. type ×	Year FE
r_1	Hedonics only	✓		
r_2	+ Property type interactions	✓	✓	
r_3	+ Year FE	✓	✓	✓

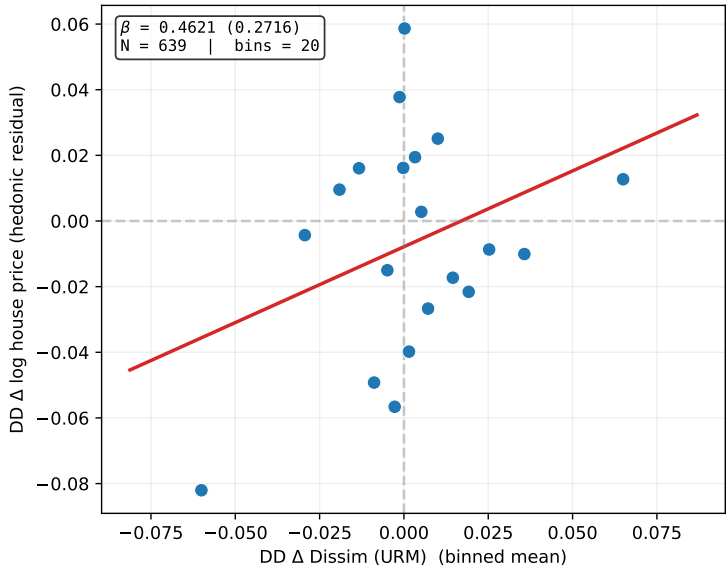
→ r_3 is the preferred residual: absorbs structural heterogeneity and property type, but preserves cross-district variation needed for DID.

Parametric DD Estimates: House Price (Hedonic Residual)



Linearization and Dosage

- While district objectives are multidimenstional, they correlate direclty via $A(t)$
- We can linearize with respect to one of the FS variables to gain stat. power and examine dosage effects
- Results suggest segregating redistricting events are associated with house price increases



Roadmap

1. Data
2. Measurement
3. Conceptual Framework
4. Empirical Strategy
5. Results
6. Summary and Next Steps

Summary

1. **Data**: First large-scale parcel-level panel linking school attendance boundaries to outcomes (11 years, 1,978 districts)
2. **Measurement**: The assignment matrix framework provides a unified approach to measuring boundary structure, treatment, and mechanisms
3. **Theory**: Districts optimize over 4 competings objetives: proximity, diversity, capacity, and choice
4. **Identification**: Matched DID on 668 treated-control pairs (906 unique districts)
5. **Results**: District revealed objectives: proximity and choice (at the cost of segregation)
   Limited evidence of aggregate house price effect
   Tentative effects on student achievement

Limitations and Next Steps

Limitations:

- Data coverage gaps (Precisely vendor transitions, heterogeneous data quality)
- Short post-treatment window ($\leqslant 3$ years)

Next steps for aggregate analysis:

- Block-based measurement robustness to structuring build around census blocks (as opposed to parcels)
- Improve power for achievement outcomes (New SEDA data release)

Distributional effects: Parcel-level analysis of who gains and loses, classified by pre-treatment price rank and owner race/ethnicity (BISG)

Next Steps: Distributional Effects

Motivation: Aggregate effects mask within-district redistribution.
Who gains and who loses from redistricting?

Parcel-level classification (pre-treatment):

1. **Price rank:** Residualize sale price on year FE + hedonics; above/below district median
2. **Owner race/ethnicity:** BISG (Bayesian Improved Surname Geocoding) applied to CoreLogic buyer names cf. Diamond et al. 2019

For each reassigned parcel, measure:

- School quality change: $\Delta q_i = \bar{q}_i^{\text{post}} - \bar{q}_i^{\text{pre}}$ (pre-treatment test scores of assigned school)
- House price change: $\Delta \log p_i$ relative to control parcels

Two counterfactuals:

(a) Non-reassigned parcels in same district (outside changed zones)

(b) Matched-characteristic parcels in matched control district

Thank you

For comments/questions: tomas.monarrez@phil.frb.org

Appendix

Change Rates

From $\Delta A(t)$, define district-level rates (share of NJ cells):

Change rate (total churn):

$$\frac{1}{NJ}\sum_{i=1}^{N}\sum_{j=1}^{J}\left|\Delta A_{ij}(t)\right|$$

Gross expansion rate (new assignments):

$$\frac{1}{NJ}\sum_{i,j}\mathbf{1}\left[\Delta A_{ij}(t)=+1\right]$$

Gross contraction rate (lost assignments):

$$\frac{1}{NJ}\sum_{i,j}\mathbf{1}\left[\Delta A_{ij}(t)=-1\right]$$

$\rightarrow$ Change rate $=$ Expansion $+$ Contraction

Event Quality Flag Definitions

Flag	Type	Threshold	Description
Raw rate > 1	Critical	> 1.0	Impossible change rate
Degenerate geometry	Critical	—	Invalid SAB polygons
Zone coverage change	Critical	> 0.20	Large school count change
Low Precisely coverage	Critical	> 0.20	High CCD-only school fraction
Clip drop	Critical	> 0.30	Large parcel count decline
Closure-dominated	Warning	> 0.50	Closures explain most change
High exit fraction	Warning	> 0.33	Many parcels leave district
Low territorial stability	Warning	< 0.95	Unstable school territories
Data gap	Warning	—	Missing years in window
Reverting	Warning	—	Change reverses next year
Gradual	Warning	—	Change spread over multiple years

Treated Districts by Event Type

	All Treated	Opening	Closure	Fixed
Panel A: Demographics				
Enrollment (1,000s)	15.52	27.23	21.59	14.56
Number of schools	24.39	41.13	35.71	22.83
Share URM (enrollment)	0.52	0.56	0.54	0.51
Share URM (population)	0.28	0.31	0.31	0.28
Panel B: Socioeconomic				
Median HH income ($1,000s)	64.97	71.05	60.57	65.08
Share college educated	0.35	0.38	0.32	0.35
Poverty rate	0.11	0.11	0.12	0.11
House price ($1,000s)	324.70	410.23	280.34	319.00
Panel C: Segregation				
Variance ratio URM (Census)	0.03	0.06	0.04	0.03
Variance ratio URM (enrollment)	0.06	0.10	0.06	0.06
Panel D: SAB Indices				
School choice index	0.28	0.19	0.25	0.29
Adjusted capacity index	0.81	0.85	0.84	0.81
N (districts)	512	47	59	427

Matching Sensitivity Analysis

Three matching specifications tested:

Specification	Covariates
Baseline	Share white, log enrollment
FS only	Variance ratio, school choice, distance, capacity
Combined	Log enrollment + 4 FS variables

Robustness: Alternative Matching Approaches

Propensity Score Matching

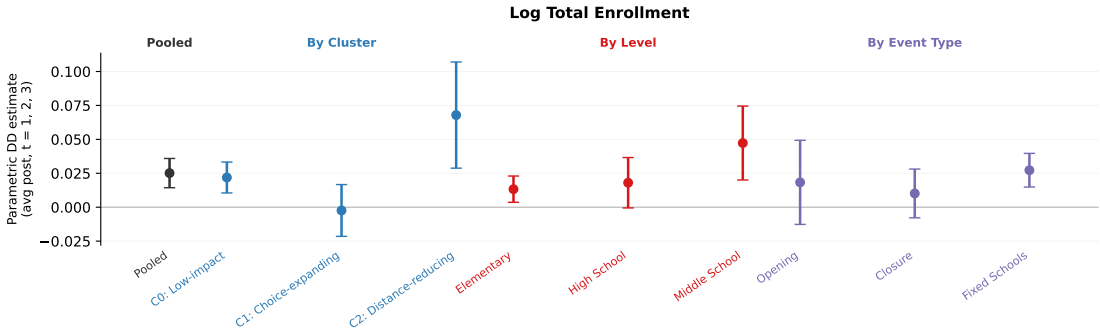
- Estimate $\hat{e}(\mathbf{x}) = \Pr(\text{Treated} = 1 \mid \mathbf{x})$ via logistic regression
- Match on $|\hat{e}(\mathbf{x}_i) - \hat{e}(\mathbf{x}_j)|$
- Reduces $\mathbf{x}$ to a scalar $\rightarrow$ avoids curse of dimensionality
- Implemented: same covariates, nearest-neighbor without replacement

LASSO Variable Selection

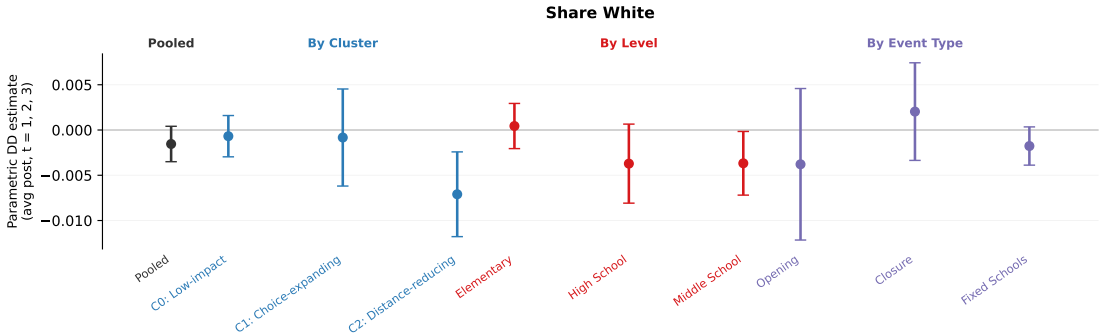
- Penalized logistic regression selects covariates from a large candidate set:
  - SAB geometry, Census segregation, CCD enrollment, district characteristics
- Data-driven covariate choice $\rightarrow$ reduces researcher degrees of freedom
- Selected variables fed into Mahalanobis or PS matching

$\rightarrow$ **Sensitivity analysis**: Three specifications tested (baseline, FS-only, combined) — balance is robust across all specs

Parametric DD Estimates: Total District Enrollment



Parametric DD Estimates: White Share of Enrollment



Event Study Plots: Test Scores

