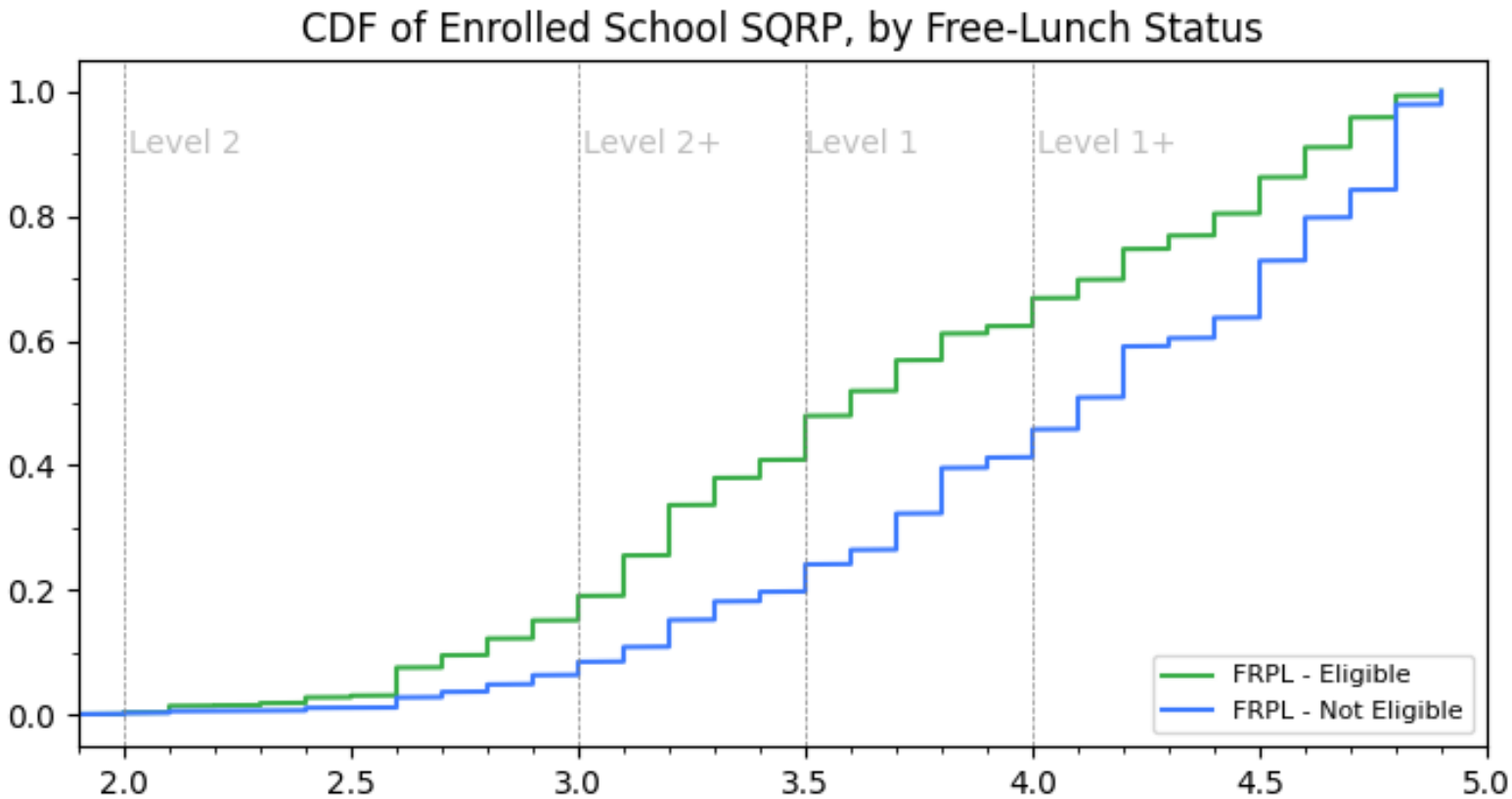
# Transportation as a Barrier to Education Access: Evidence from Chicago Public Schools

Cecilia Moreira (Stanford)
Steve Puller (Texas A&M)
Ini Umosen (Berkeley)

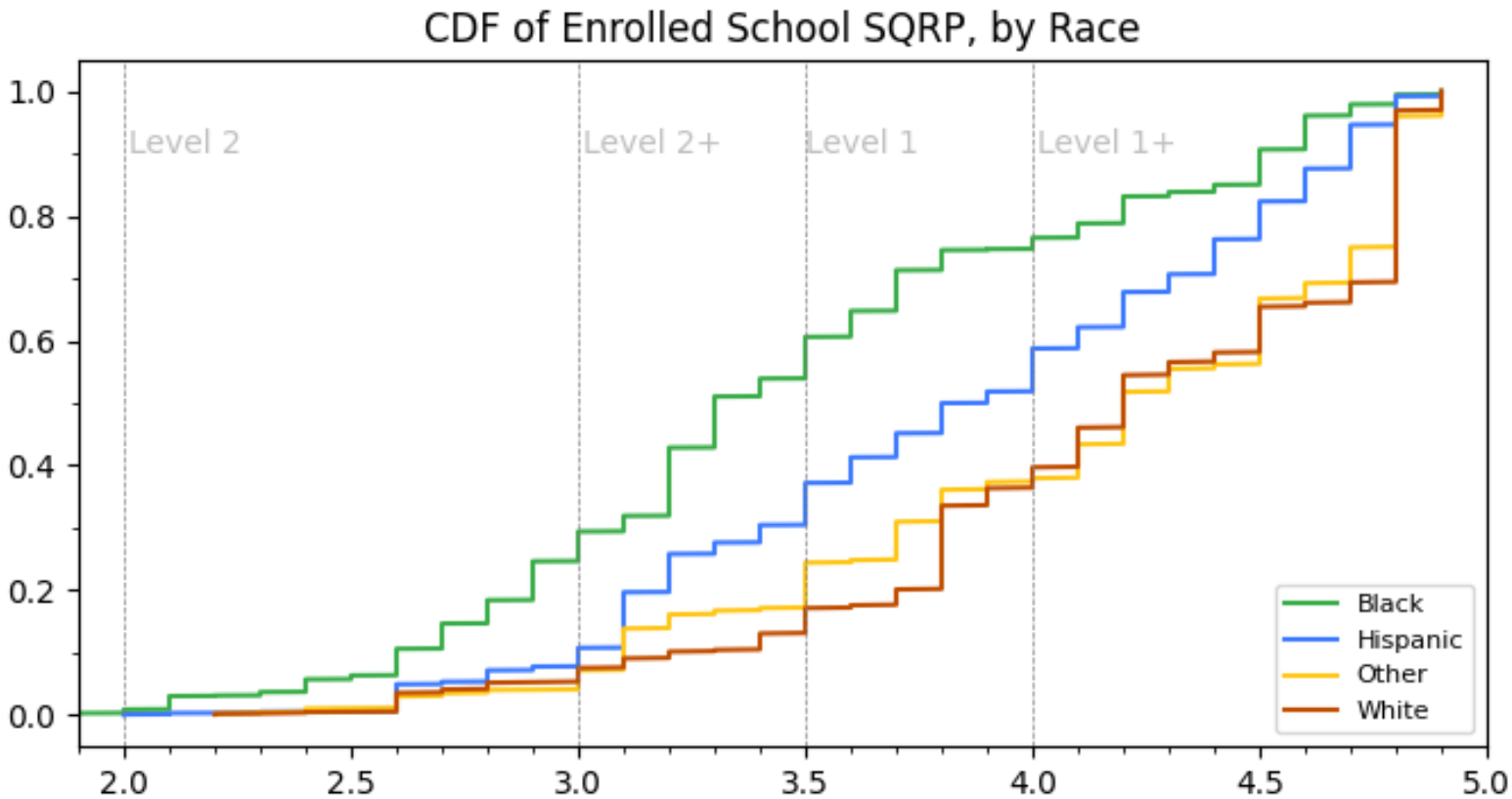
# School Choice and Access to Education

- Public education systems have embraced 'school choice' as policy to improve access to high-quality education for underserved populations
- 56% of 100+ large US school districts allowed school choice in 2015-16, up from 29% in 2000-01
- Include:
  - Vouchers
  - Charter schools
  - **'Open enrollment' via centralized assignment system**
- All of these policies involve **transportation** to a school that may not be as close as nearest school

# Low-Income Students Attend Lower Rated Schools



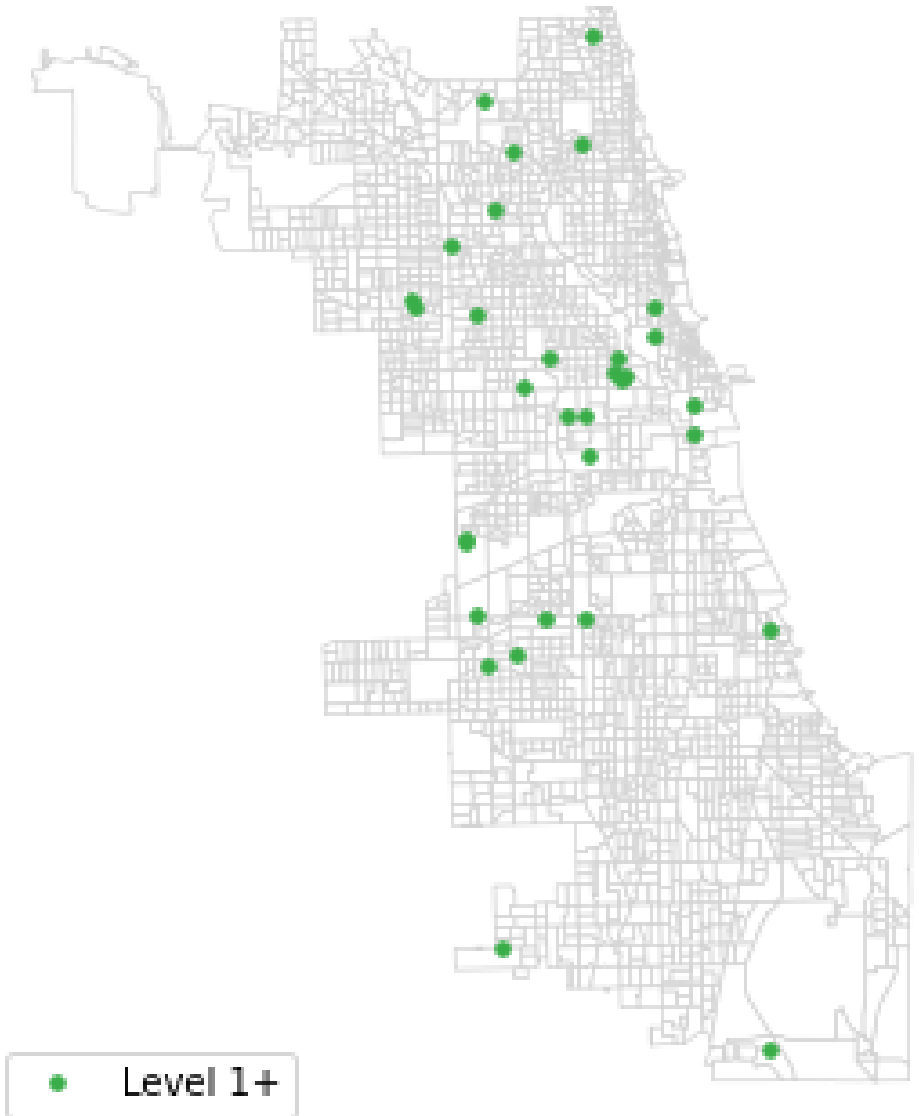
# Minority Students Attend Lower Rated Schools



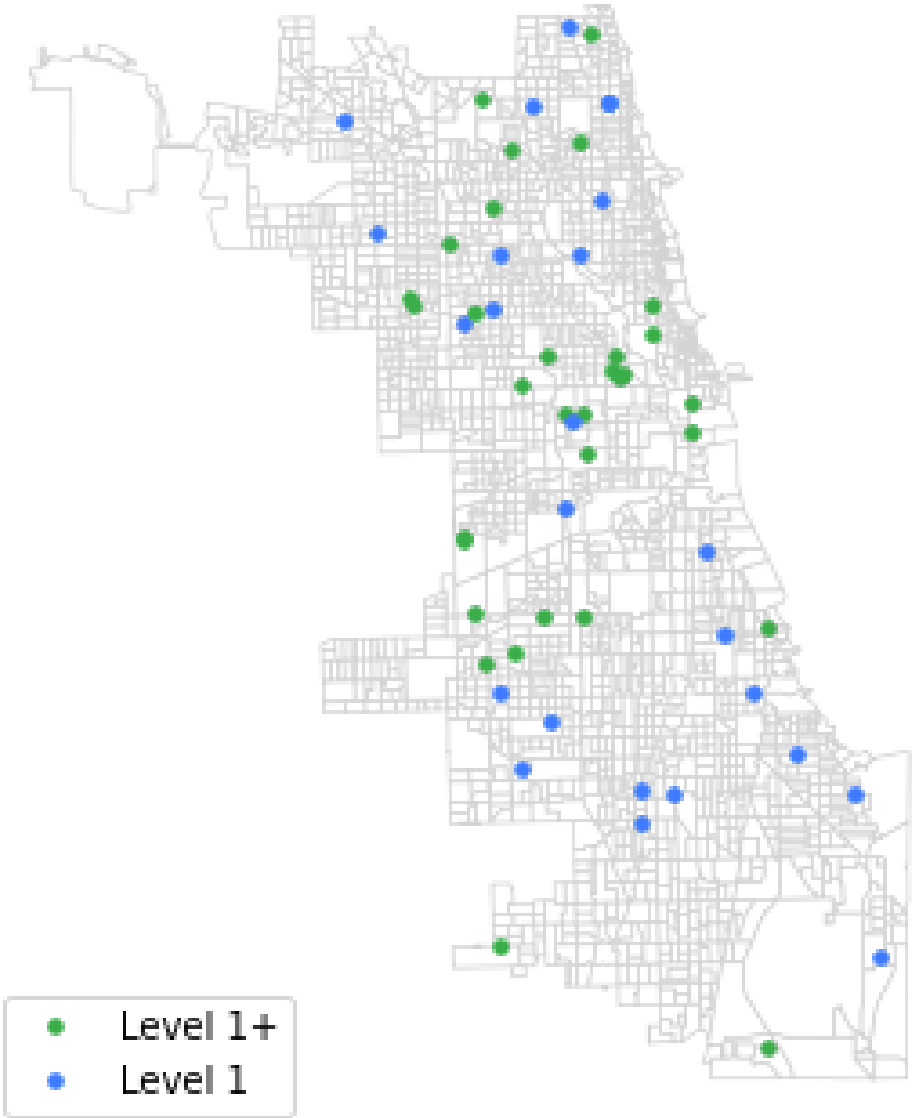
# School Choice and Access to Education

- Success of open enrollment will hinge on parent willingness to send children to schools other than neighborhood school
- Distaste for distance has been shown (Hastings et al. [2009], Burgess et al. [2015], Stein et al. [2021], Laverde [2022])
- Transportation time could drive decision, especially systems that rely on public transit
- **Does transportation create a barrier to access?**

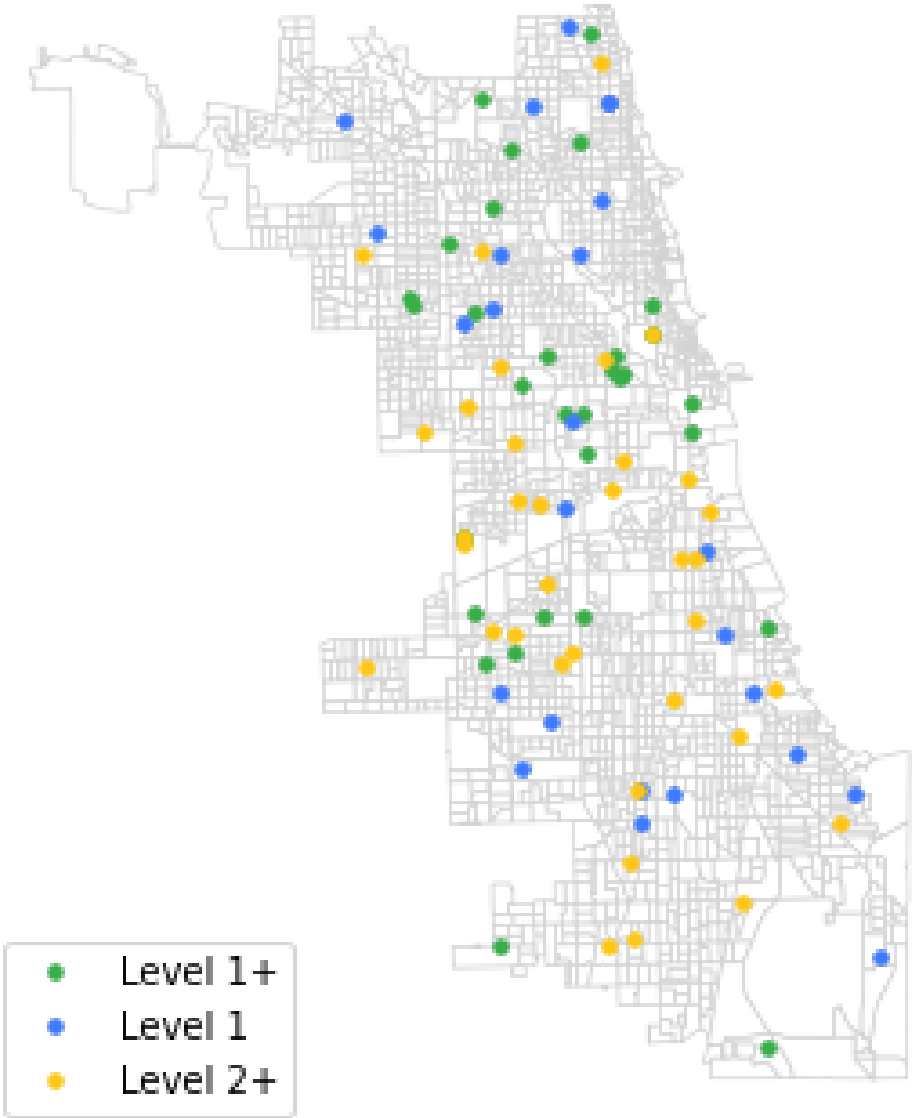
# Geographic Distribution of High Schools



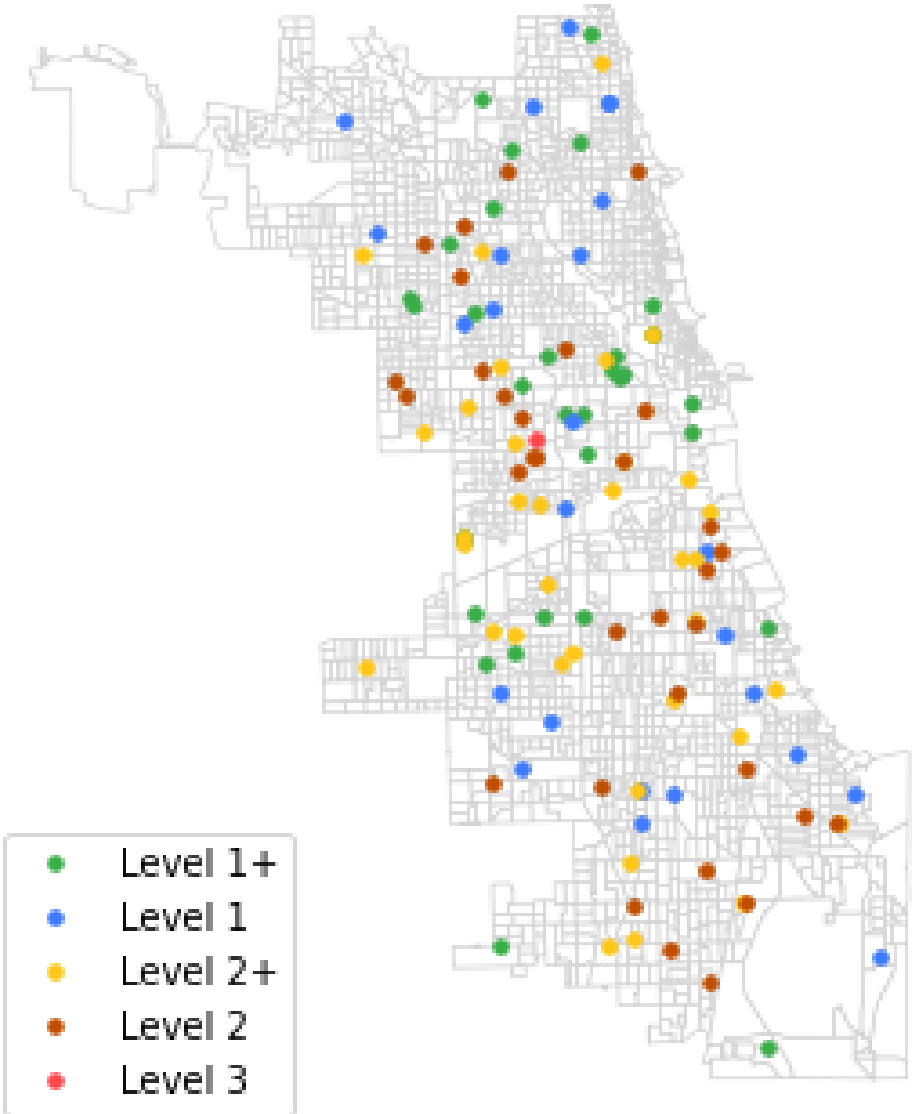
# Geographic Distribution of High Schools



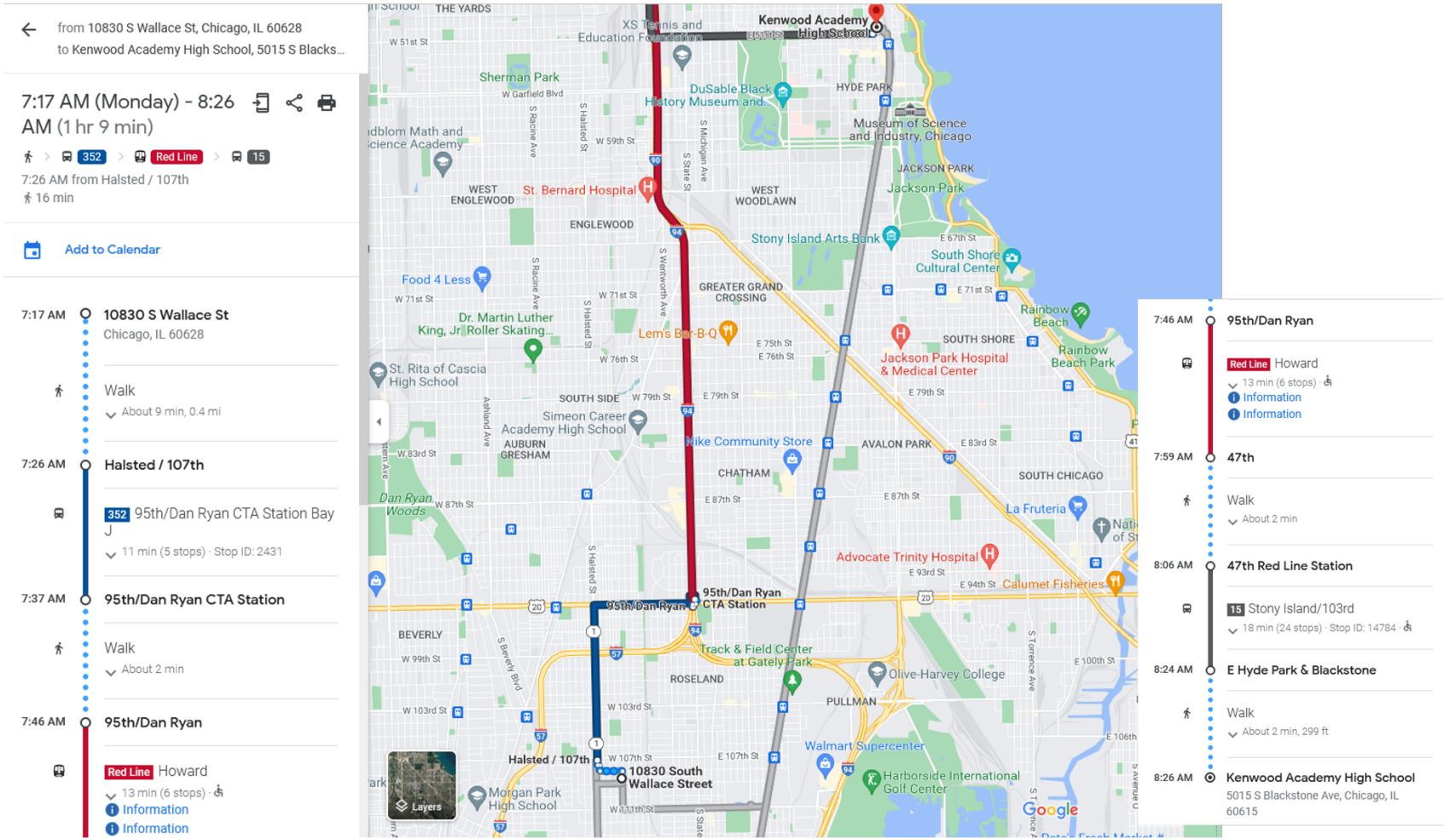
# Geographic Distribution of High Schools



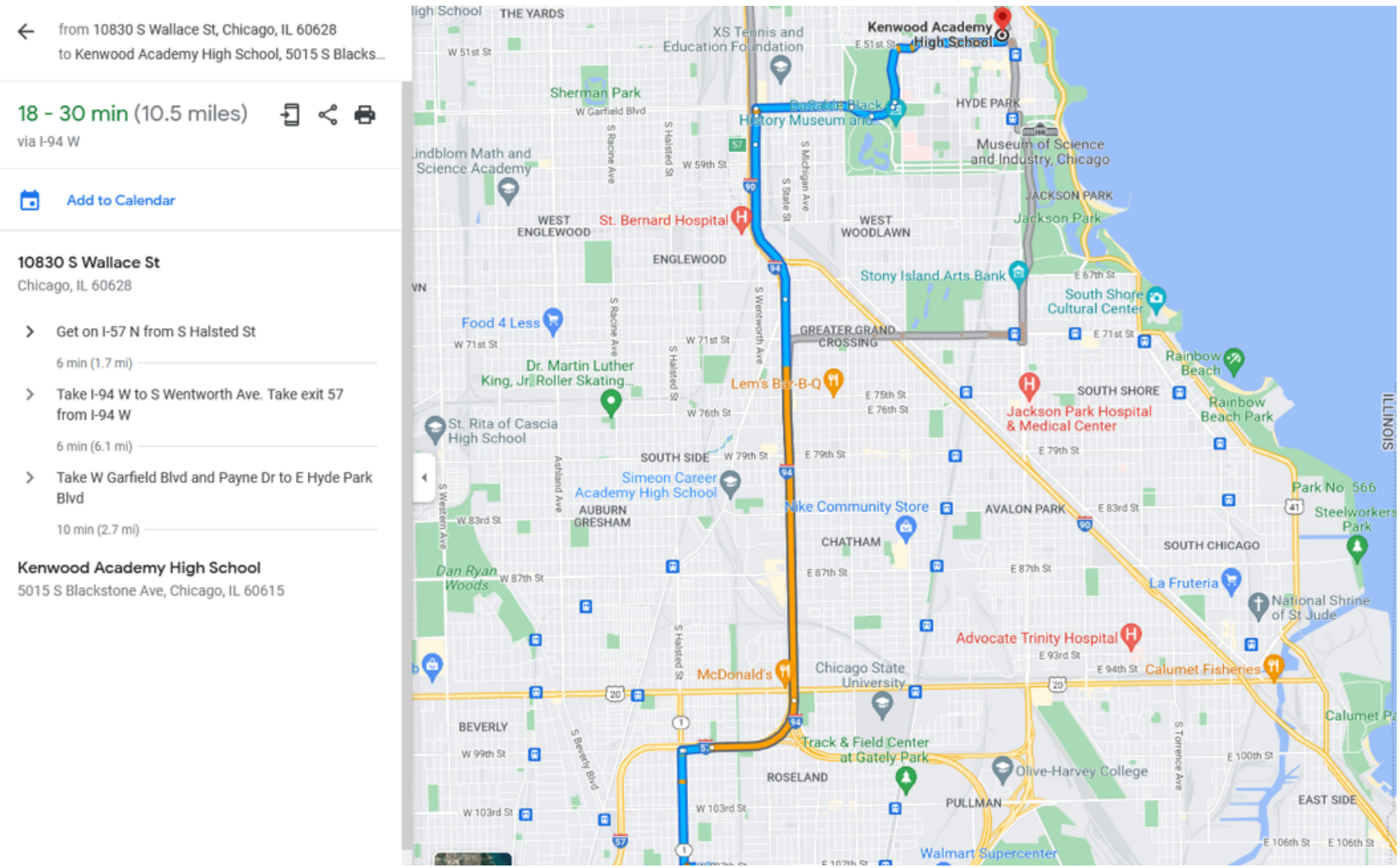
# Geographic Distribution of High Schools



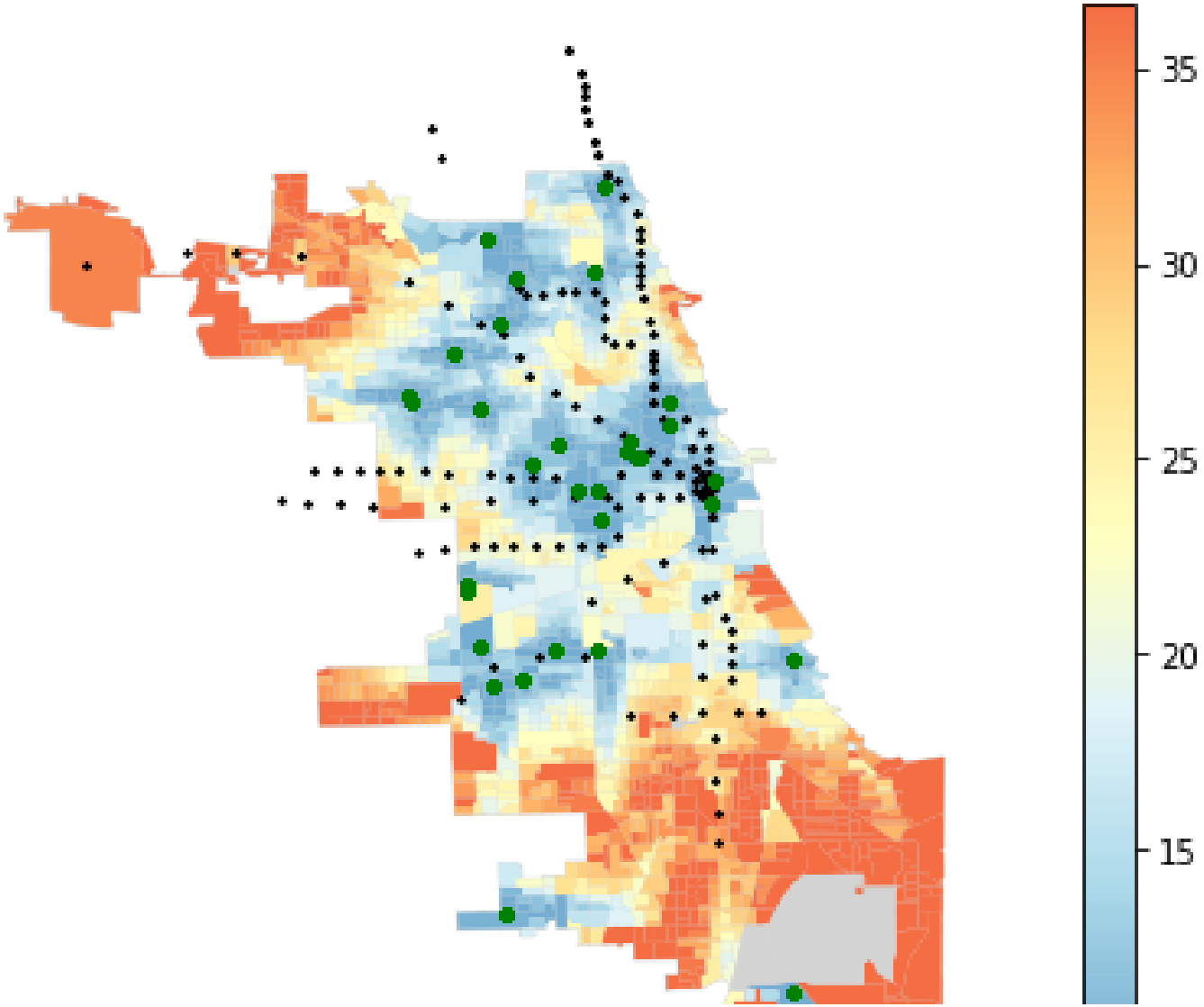
# Example of Public Transit vs Driving to Top-Rated High School



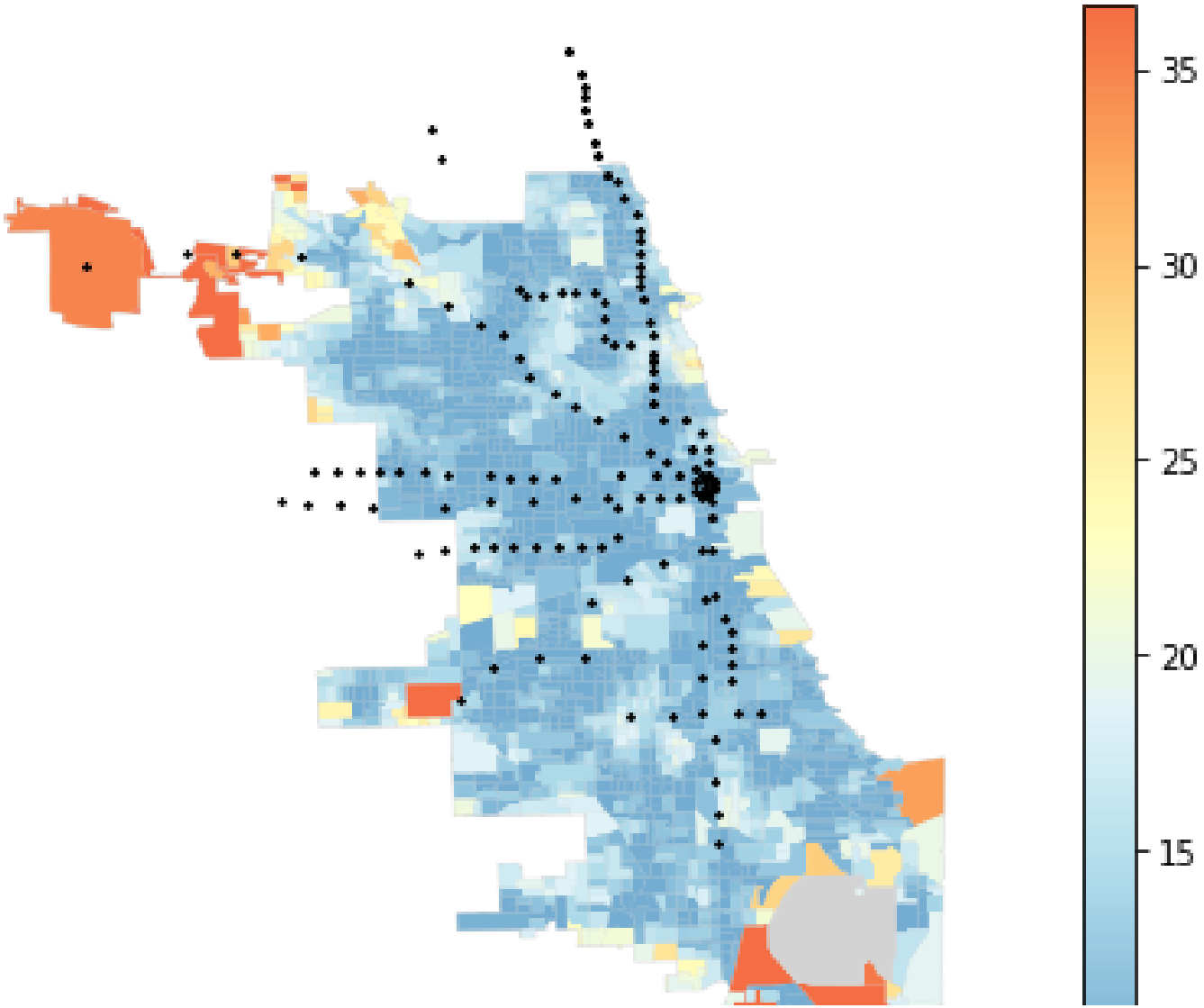
# Example of Public Transit vs Driving to Top-Rated High School



# Some Neighborhoods Face Long Transit Time to Top-Rated High School...



# ...But the Nearest High School is a Much Shorter Transit Ride



# Our Project

- Research Question: Are there socioeconomic differences in travel time that inhibit access to high-achieving high schools? Can new school transportation policies mitigate disparities?
- Empirical Strategy: Descriptive evidence and discrete choice logit model of demand to understand how parents tradeoff school quality and travel time
- Data
  - Student-level ranked school choices (CPS)
  - Travel times (GoogleMaps)
- Policy counterfactuals
  - Identify new specialized "express bus" routes from underserved neighborhoods to top-rated schools that induce families to prefer higher-quality schools

# Preview of Results

- Descriptive: Low-income and minority parents are less likely to rank top-rated schools first when choosing schools, and they face longer travel time from home to top-rated schools.
- Model: Parents face a sharp tradeoff of school quality vs. travel time
  - Despite placing significantly more value on higher-rated schools, parents choose lower-rated schools due to travel time
  - Estimated Preferences: An increase in one hour of travel time will induce parents to choose a lower-rated school (Level 2/2+ school rather than a Level 1+ school)
- Counterfactuals: Targeted express bus routes that reduce travel time from underserved neighborhoods to top-rated schools can induce parents to prefer higher quality schools

▸ Literature

# Chicago Public Schools (CPS)

- District
  - 3rd largest in country – 355,156 students in 2019-20
  - 49% Hispanic, 35% African-American, 10% White
  - Low income: 76% qualify for free/reduced price lunch
- Open Enrollment for High School
  - Growth in choice set from 78 schools in 2002 to 138 in 2016 (Barrow and Sartain [2017])
  - Researching and applying to schools was difficult until 2017 because there was no centralized system
  - In 2017, created centralized application platform: GoCPS. A 'one-stop-shop' for households to learn about school options
- No yellow school bus service for high school

# Chicago Public Schools (CPS)

- Application in fall prior to 9th grade: opens mid-October and closes mid-December
- Students submit preferred schools via online platform
- Can rank up to 20 "choice" programs and 6 "selective-enrollment" programs
  - We focus on "choice" programs
- Always have option of neighborhood school
- Deferred Acceptance algorithm assigns student an offer to highest-ranked school with seats available

# School Choice Platform: GoCPS

- Information available on platform:
  - School characteristics: average test scores, graduation rates, college enrollment rates, programs offered, any special application requirements
  - Demographics of students at school: race, income, English proficiency, measures of academic achievement

▸ View Platform

# Our Sample

- Entry into high school in first year of online platform to select in 2017
- Students who apply to at least one choice school
  - Represents 95% of applicants
  - Disproportionately minority and low-income students

# Data

- Student-level rankings of schools, student addresses linked to a Census block group, demographic variables including race and free/reduced lunch status
  - Provided by the Chicago Public School District
  - 22,902 incoming 9th graders
- Travel time (by both transit and by driving) from the centroid of every census block group in Chicago to every Chicago high school to arrive on a weekday morning
  - Google Maps Directions API

# Student Characteristics

| | All | Hispanic | Black | White | Asian/other | Missing race | FRPL | Not FRPL |
|---|---|---|---|---|---|---|---|---|
| Number of schools ranked | 6.05 | 5.89 | 7.52 | 3.02 | 4.86 | 3.97 | 6.53 | 4.13 |
| Share whose 1st choice is Level 1+ | 0.56 | 0.62 | 0.38 | 0.62 | 0.73 | 0.77 | 0.52 | 0.70 |
| Share whose enrolled school is Level 1+ | 0.43 | 0.47 | 0.27 | 0.61 | 0.63 | 0.73 | 0.38 | 0.66 |
| Transit time to neighborhood school | 0.29 | 0.25 | 0.33 | 0.31 | 0.31 | 0.28 | 0.29 | 0.30 |
| Drive time to neighborhood school | 0.11 | 0.10 | 0.12 | 0.11 | 0.12 | 0.11 | 0.11 | 0.11 |
| Transit time to 1st choice | 0.54 | 0.48 | 0.62 | 0.51 | 0.52 | 0.56 | 0.53 | 0.56 |
| Drive time to 1st choice | 0.22 | 0.20 | 0.24 | 0.20 | 0.22 | 0.24 | 0.22 | 0.23 |
| Transit time to enrolled school | 0.49 | 0.45 | 0.55 | 0.49 | 0.50 | 0.55 | 0.48 | 0.55 |
| Drive time to enrolled school | 0.20 | 0.19 | 0.21 | 0.20 | 0.20 | 0.23 | 0.19 | 0.22 |
| Observations | 22902 | 11038 | 7630 | 1847 | 1191 | 1196 | 18333 | 4569 |

Note: Transportation times in hours

▸ School Characteristics

# Model of School Choice (Overview)

Starting point: Ranked choices provide revealed preference data

Parents tradeoff distance (e.g. travel time), school quality (e.g. test scores, graduation rates), and other school characteristics

- Parents from different demographic groups may tradeoff these characteristics differently

**Ranked choices reflect the ordinal ranking of "utility" parents believe children would receive from attending each school.**

$U_{ij}$: "Utility" of student $i$ attending school $j$

- Consider the schools ranked #1-#5 for student $i$ among all $J$ schools the student is eligibile to attend
  - $U_{i1} \geq U_{ik}$ for $k = 2, 3, 4, 5, ....J$
  - $U_{i2} \geq U_{il}$ for $l = 3, 4, 5, ....J$
  - $U_{i3} \geq U_{im}$ for $m = 4, 5, ....J$

# Model of School Choice (Overview)

Key Assumption: Parents truthfully rank preferences

- Deferred Acceptance (DA) algorithm is strategy-proof (Abdulkadiroglu and Sonmez [2003])
- Rules out parents skipping over more preferred schools because they do not think will be admitted

Simplifying assumptions:

- Parents rank schools rather than programs
- No effects of siblings

▸ Details

# Model of School Choice

- Each school provides a bundle of characteristics that provides different value to each student.
- The average utility that school $j's$ entire bundle of characteristics yields to student in group $k$ is $\delta_{jk}$.
- $U_{ij} = \delta_{jk} - \lambda_k d_{ij} + \epsilon_{ij}$: The total utility that school $j$ yields to student $i$ is that mean utility $\delta_{jk}$ net of the time costs of transportation $\lambda_k d_{ij}$ plus a idiosyncratic student-specific value for the school $\epsilon_{ij}$.
  - $d_{ij}$: distance in travel time from student $i's$ home to school $j$ (varies by mode)
  - $\epsilon_{ij}$: characteristics of school $j$ valued by household $i$ that are observed by parent but not researcher
- Estimate $\lambda_k$ and $\delta_{jk}$ via Simulated Maximum Likelihood (mixed logit) with student-level choice data

▸ Estimation Details

# Threat to Identification: Residential Sorting

- If parents with (unobserved) tastes for a school move closer to the school, then we will obtain biased estimate of $\lambda$.
- Empirical challenge in literature on school choice (and urban generally).
- Future: will assess extent of sorting using techniques similar to literature (e.g. Davis, Dingel, Monras and Morales [2019])

# Results: Correlates with Estimated School Mean Utility

## All Students

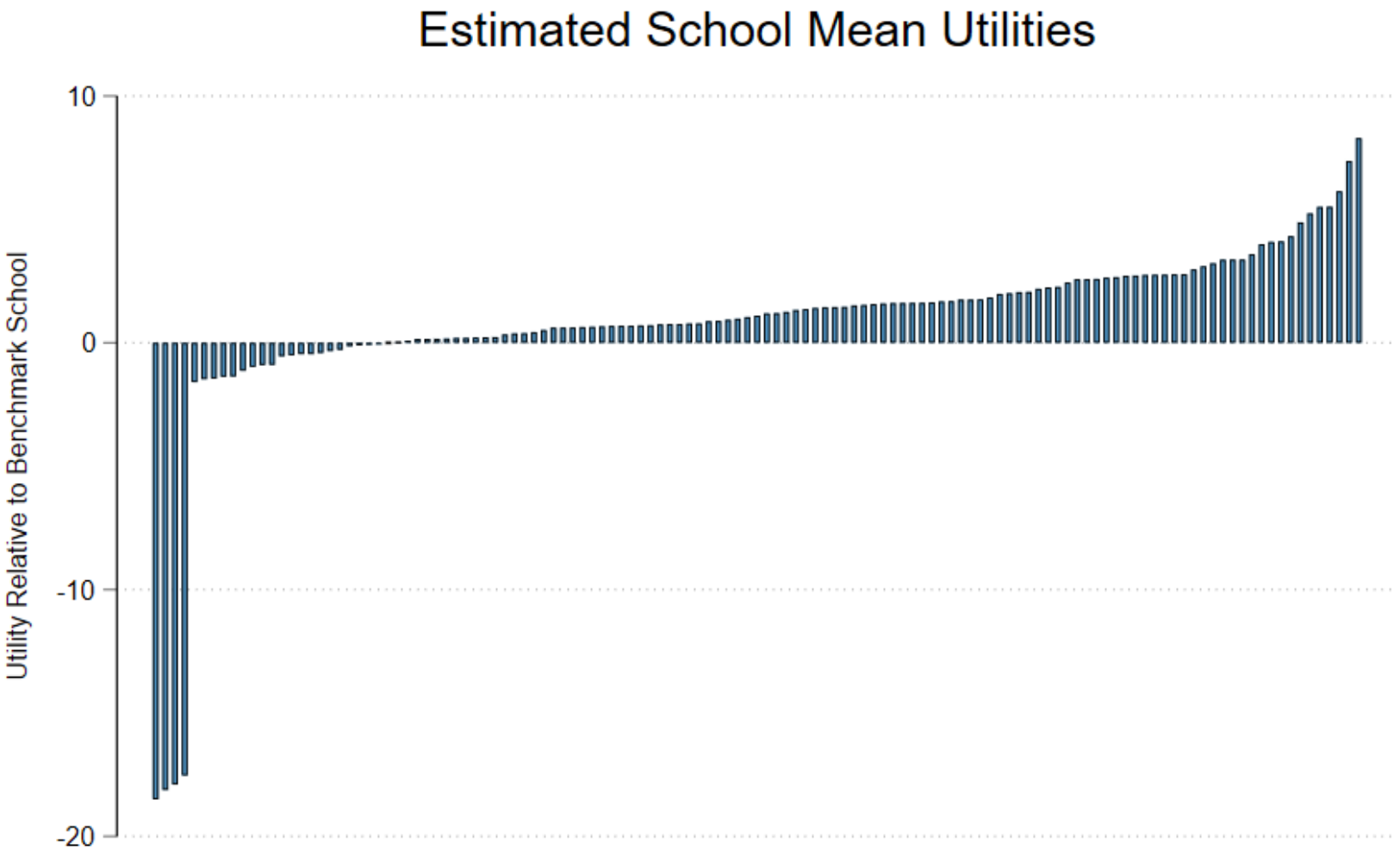
| | $\hat{\delta}_j$ | $\hat{\delta}_j$ | $\hat{\delta}_j$ | $\hat{\delta}_j$ | $\hat{\delta}_j$ |
|---|---|---|---|---|---|
| Pct FRPL | -0.0703 | | | | -0.0327 |
| | (0.0113) | | | | (0.0121) |
| Avg SAT Score | | 0.0119 | | | 0.00431 |
| | | (0.00133) | | | (0.00229) |
| 4-year Cohort Grad Rate | | | 0.0812 | | 0.0556 |
| | | | (0.00898) | | (0.0122) |
| College Enrollment Rate | | | | 0.0513 | -0.00367 |
| | | | | (0.00920) | (0.0106) |
| Constant | 6.594 | -10.40 | -5.608 | -2.712 | -4.481 |
| | (1.002) | (1.217) | (0.675) | (0.581) | (2.344) |
| Observations | 126 | 122 | 117 | 109 | 108 |

# Results: Preference Parameters

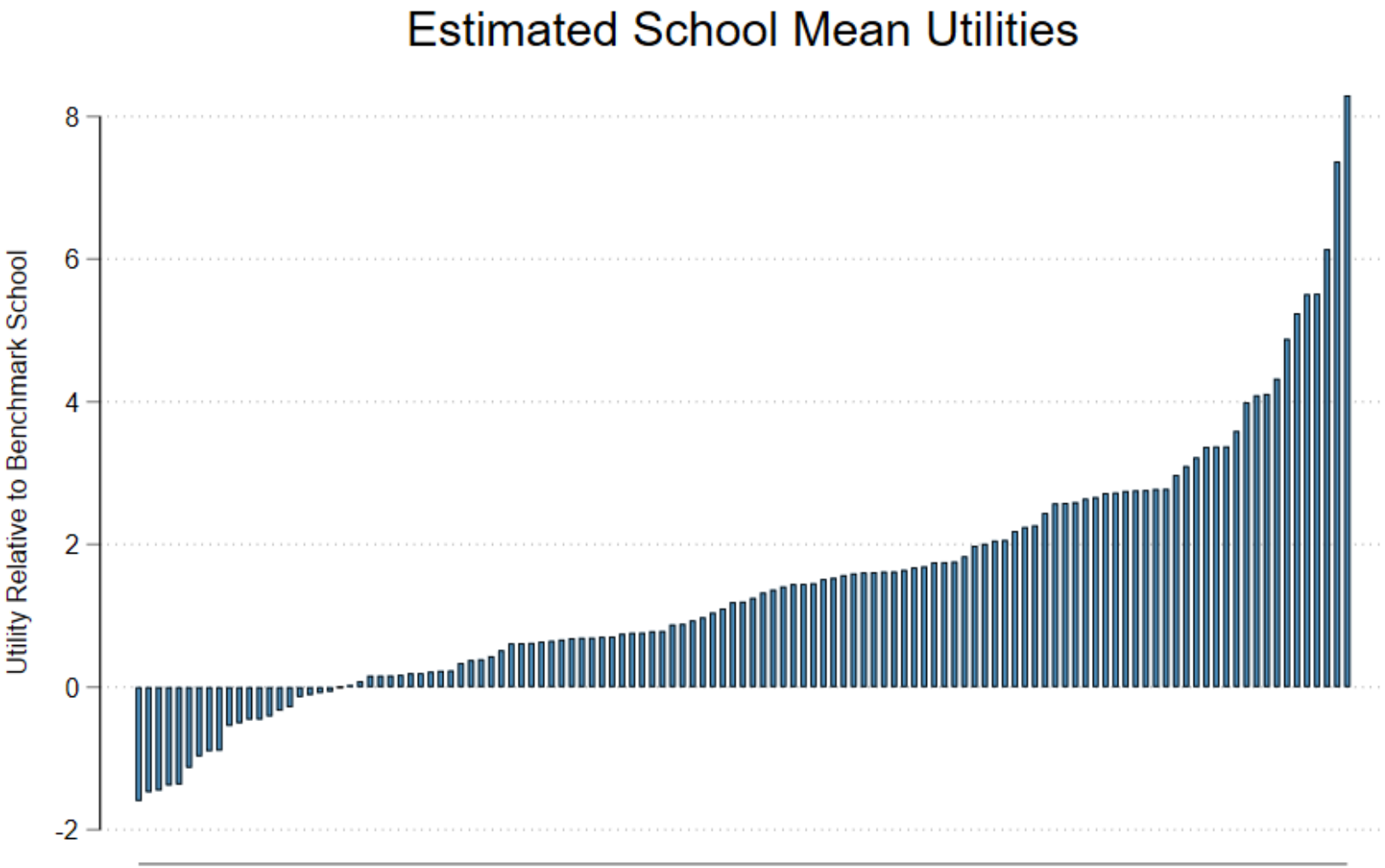
- $\hat{\lambda} = -4.47$ (utils/hour)
- **One hour** of additional travel time equivalent to differences in estimated mean utilities of:
  - Lincoln Park (Level 1+) - Roosevelt (Level 2)
  - Jones (Level 1+) - Noble Johnson (Level 2+)
  - Kenwood (Level 1+) - Orr (Level 2)

▸ SQRP ▸ By Race

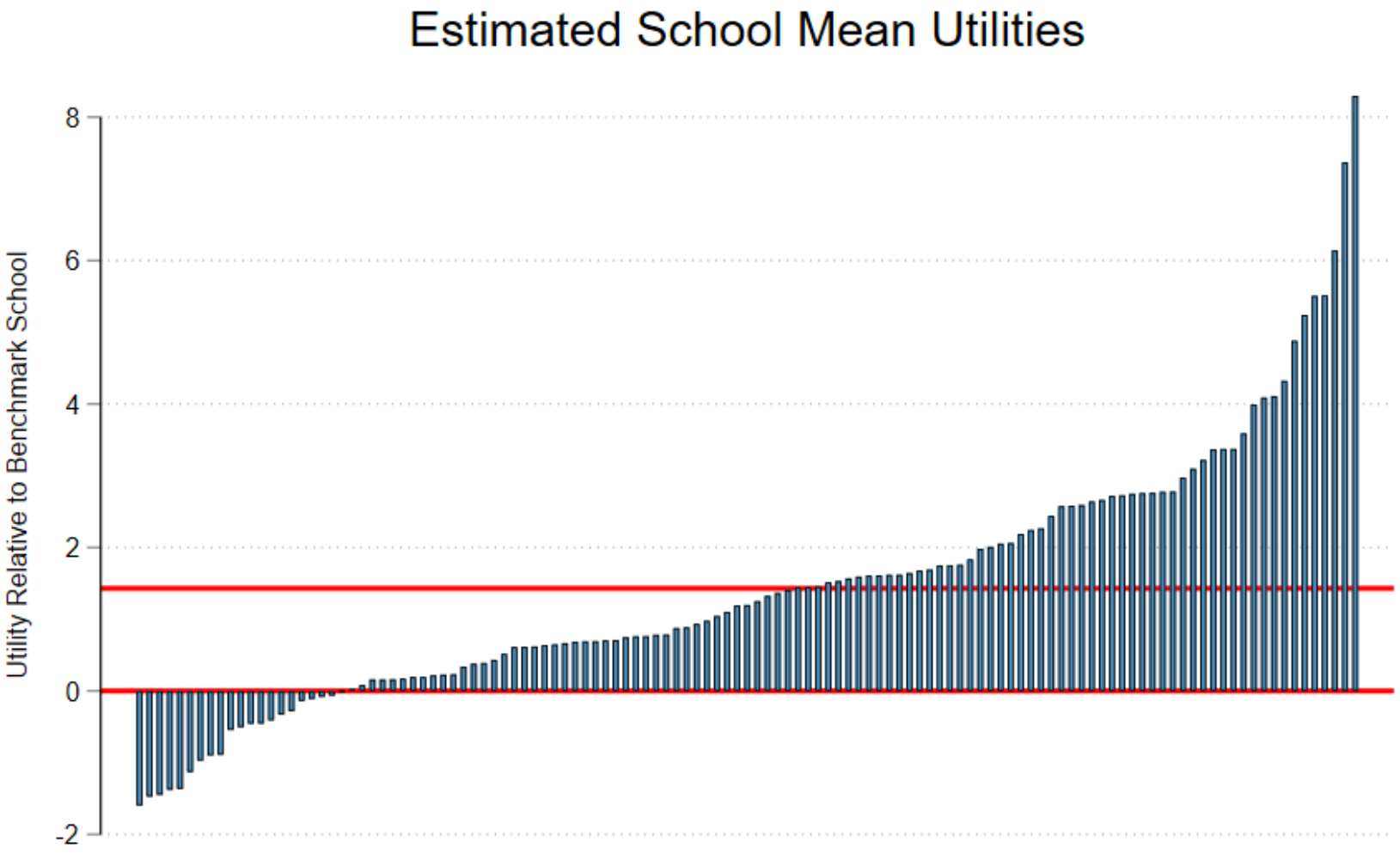
# Counterfactual: Potential to Induce Families to Choose Better Quality Schools



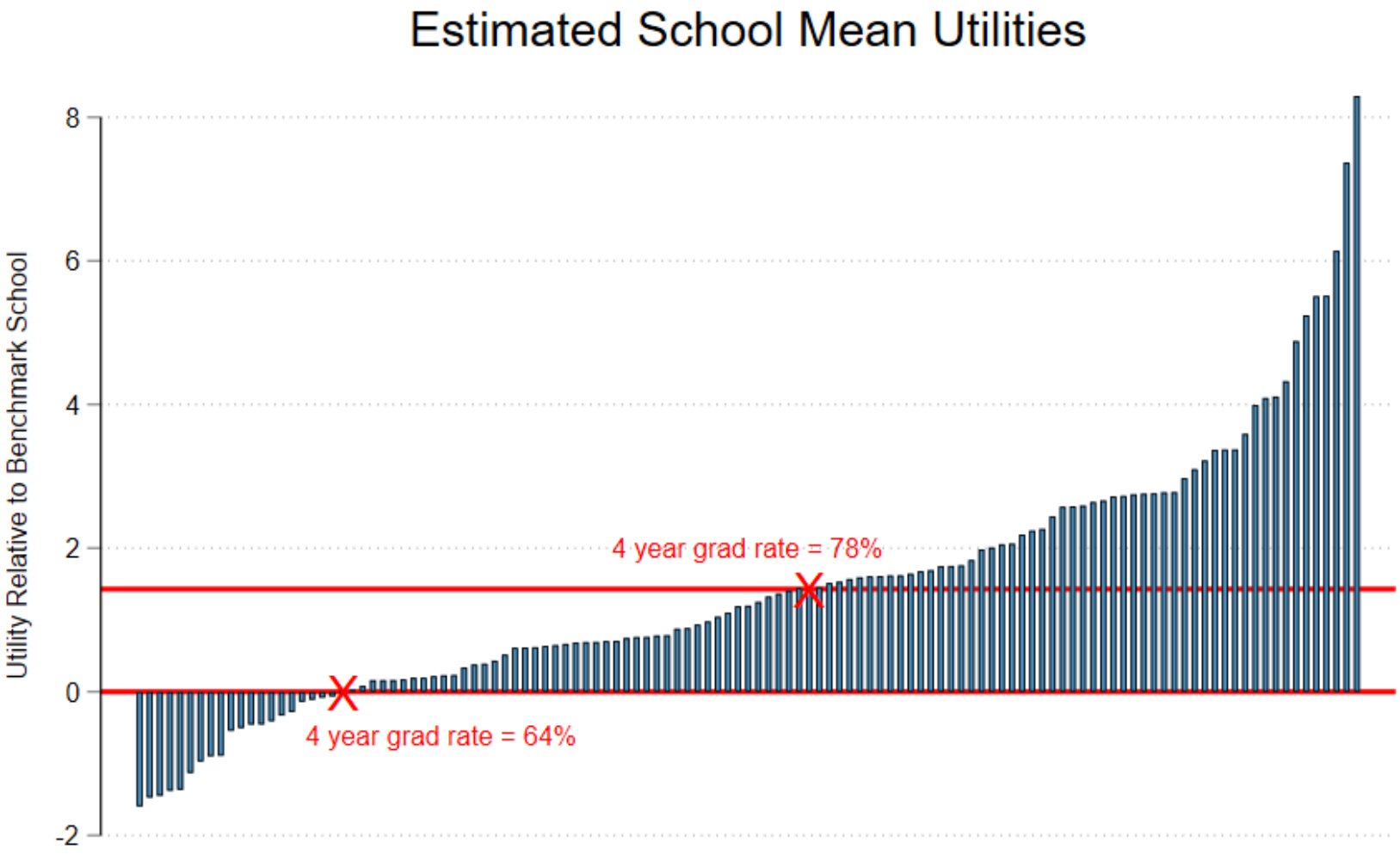
# Counterfactual: Potential to Induce Families to Choose Better Quality Schools



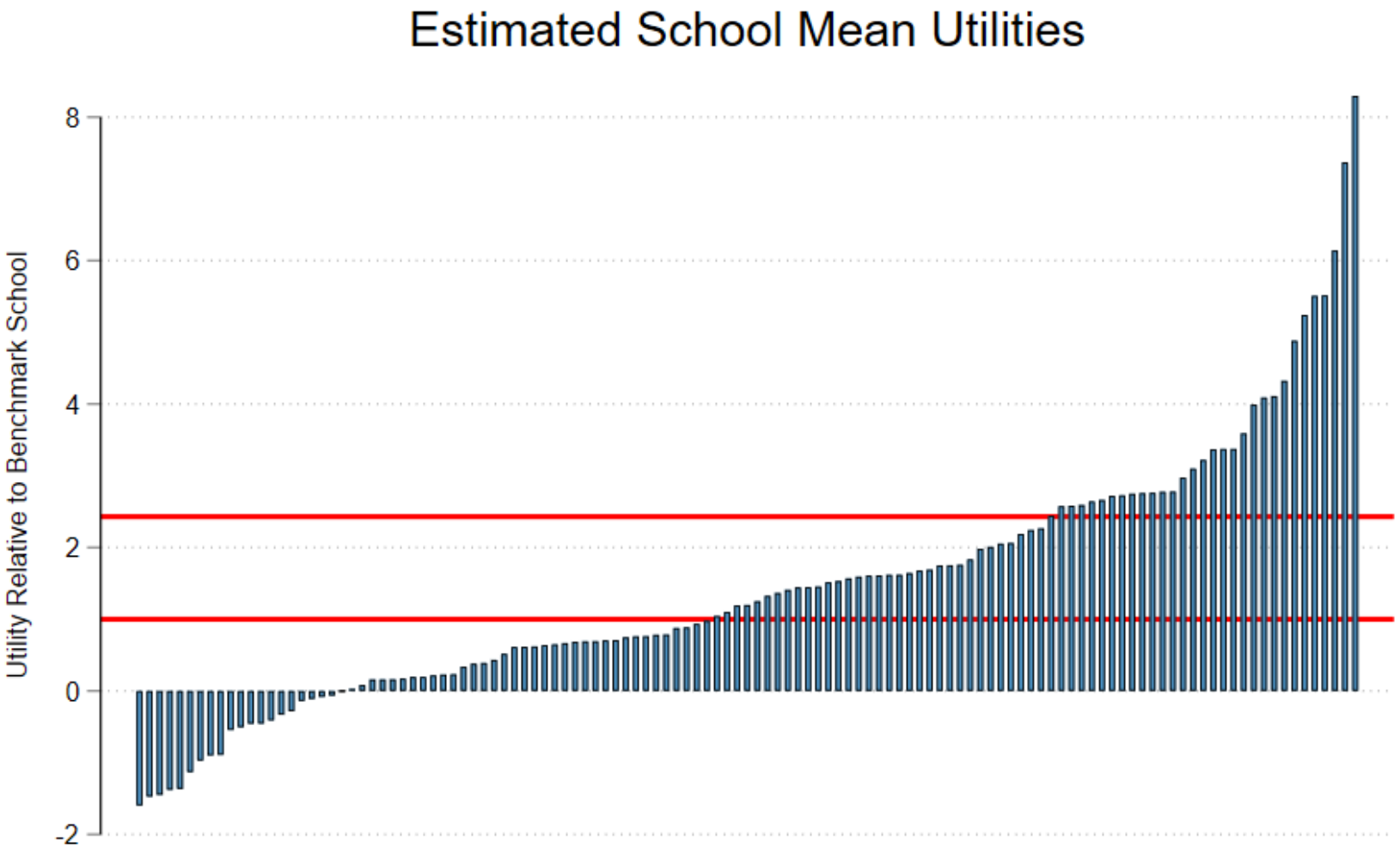
# Counterfactual: Potential to Induce Families to Choose Better Quality Schools



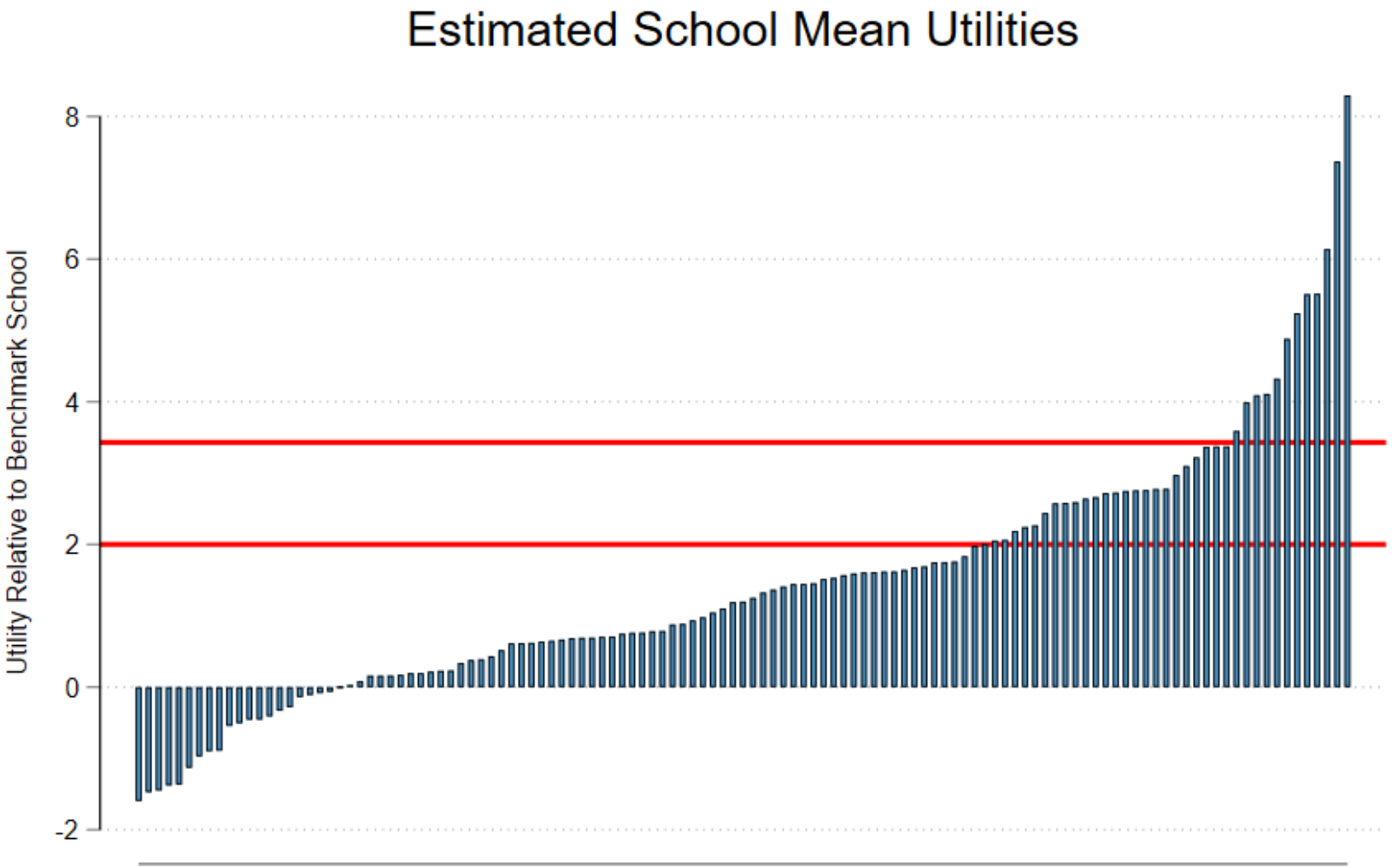
# Counterfactual: Potential to Induce Families to Choose Better Quality Schools



# Counterfactual: Potential to Induce Families to Choose Better Quality Schools



# Counterfactual: Potential to Induce Families to Choose Better Quality Schools



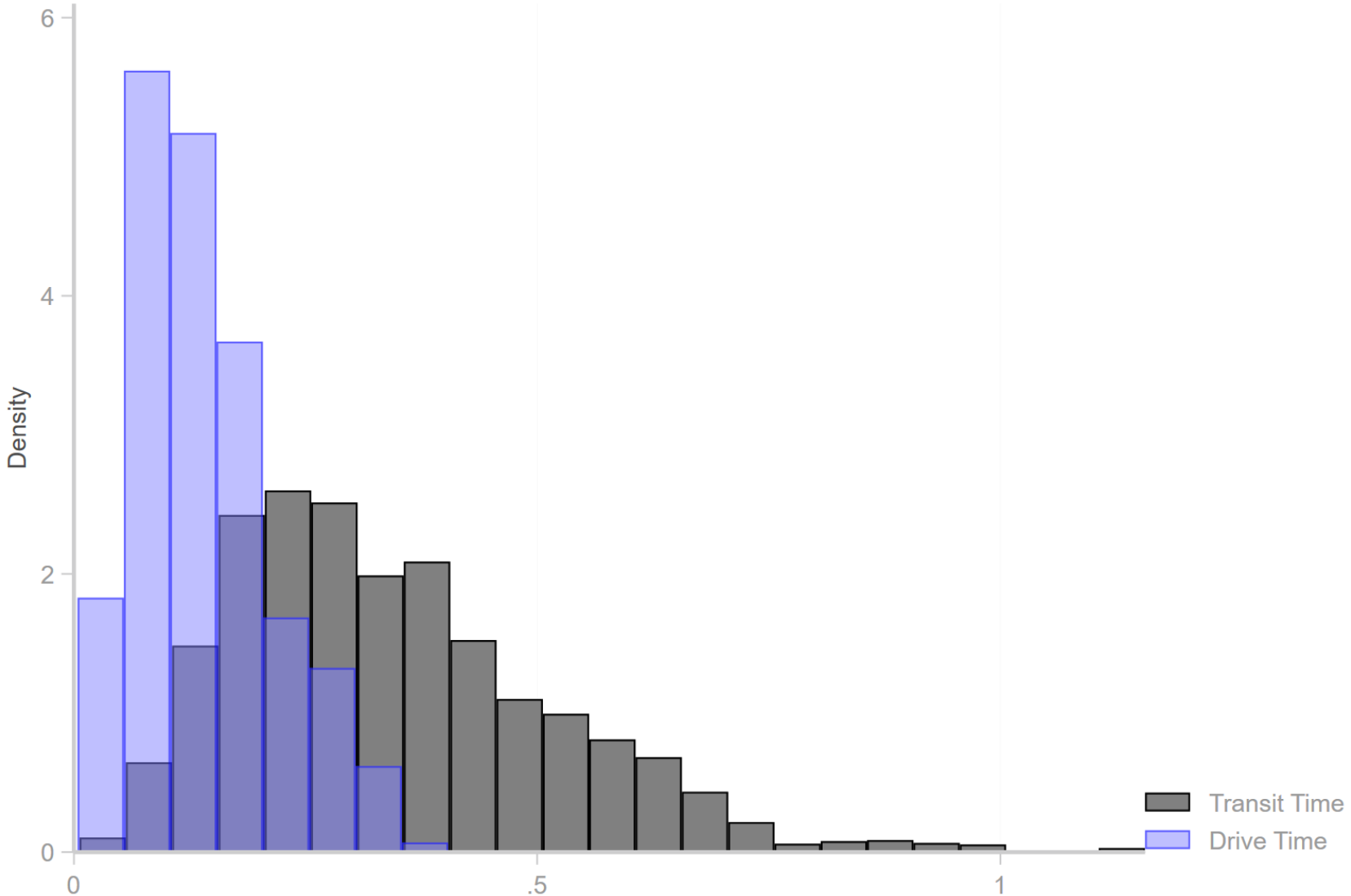
# Counterfactuals

- Can lowering transportation cost induce parents to apply for top-rated school?
- Express bus route that picks up in one neighborhood and delivers to one specific Level 1+ school
- Reduces 'distance' from *transit* time to *drive* time
  - Increase in student $i$'s probability of selecting target Level 1+ school $j$ as first choice estimated by:

$$\hat{P}_{ij}(d_{ij}^{drive}, \hat{\delta}_{jk}) - \hat{P}_{ij}(d_{ij}^{transit}, \hat{\delta}_{jk})$$

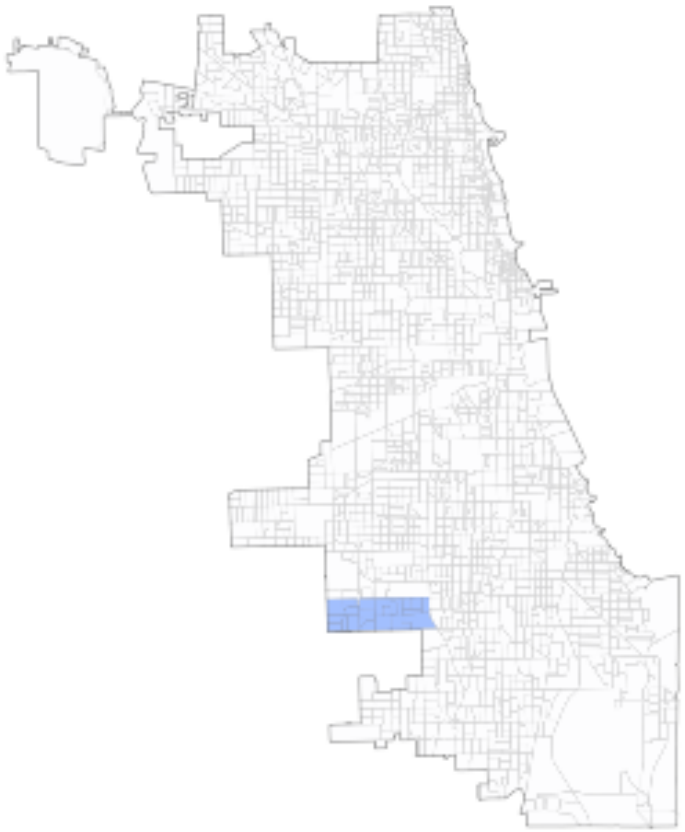
- Mean utilities held constant, so assumes students who would switch schools not large enough to change composition in way that affects mean utility
- Likely underestimates impact because non-time characteristics also change (e.g. safety)

# Counterfactual: Buses to Reduce Transportation Time to Nearest Level 1+



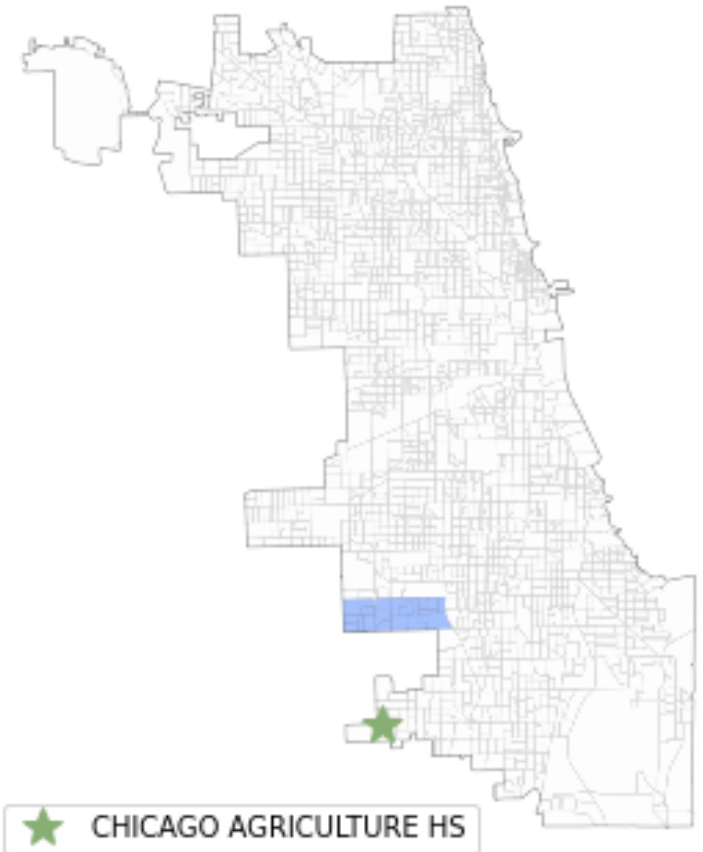
# Counterfactual 1: Bus Route To Level 1+ School

- 380 applicants
- 53% Hispanic, 39% Black
- 47% rank a Level 1+ school first (vs 56% across all neighborhoods)
- Nearest Level 1+ is 30 minutes away by transit



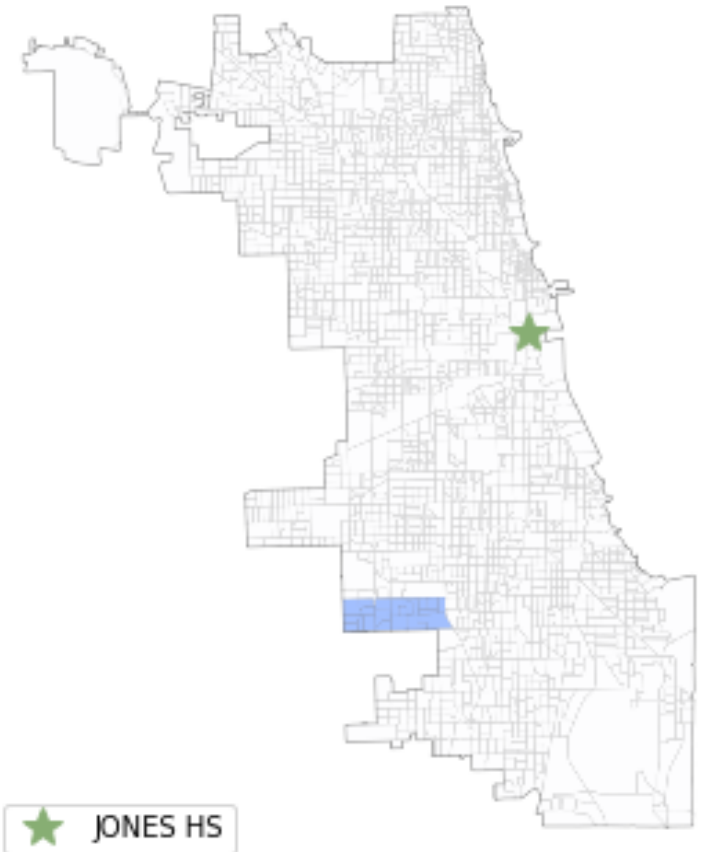
# Counterfactual 1: Bus Route To Level 1+ School

- **25%** of students would switch to this school if travel time reduced from transit time to drive time



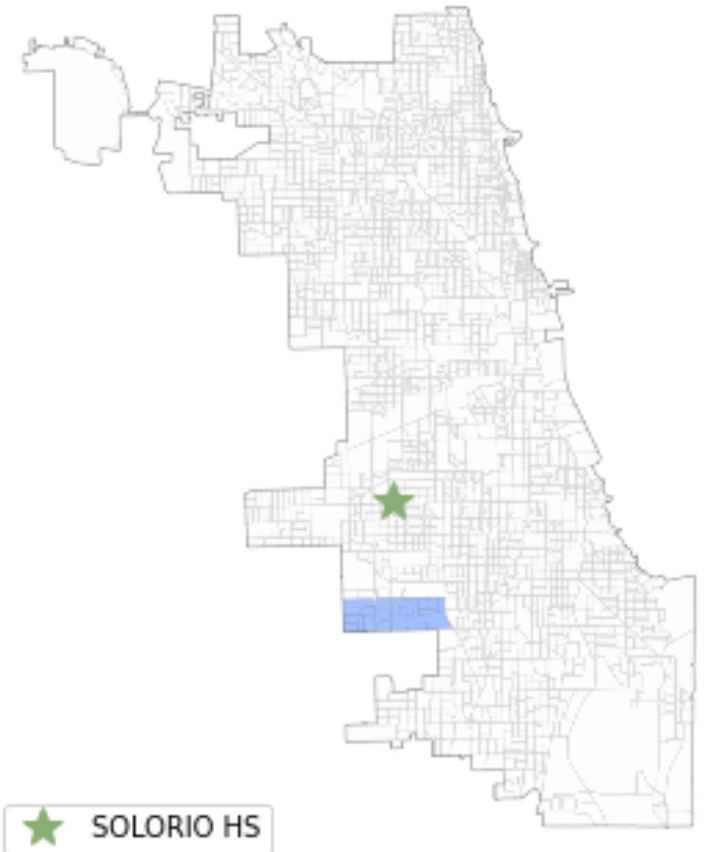
# Counterfactual 1: Bus Route To Level 1+ School

- **23%** of students would switch to this school if travel time reduced from transit time to drive time



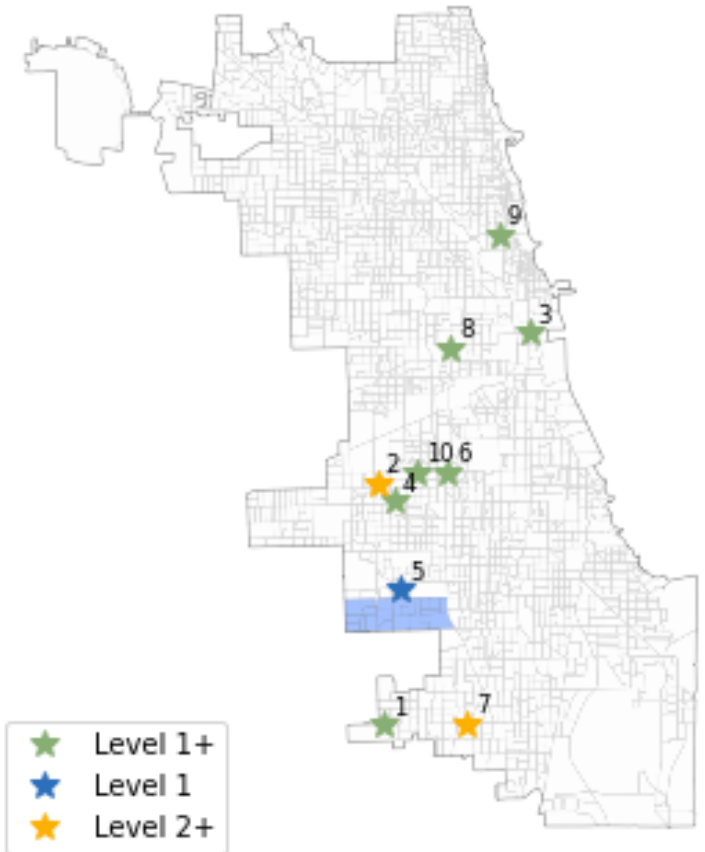
# Counterfactual 1: Bus Route To Level 1+ School

- **23%** of students would switch to this school if travel time reduced from transit time to drive time



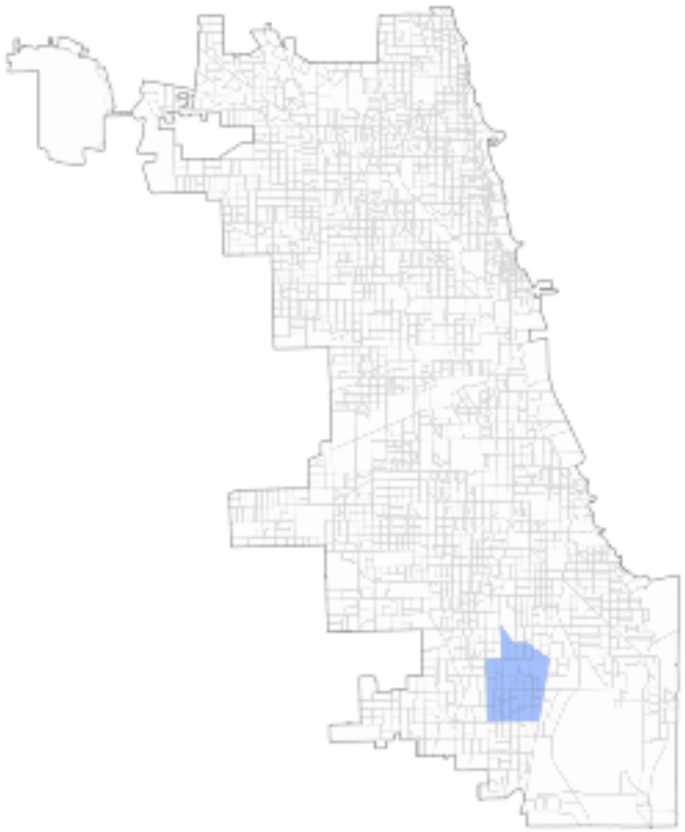
# Counterfactual 1: Bus Route To Top Schools

- Top 10 schools that a dedicated bus route would induce switches to (individually)



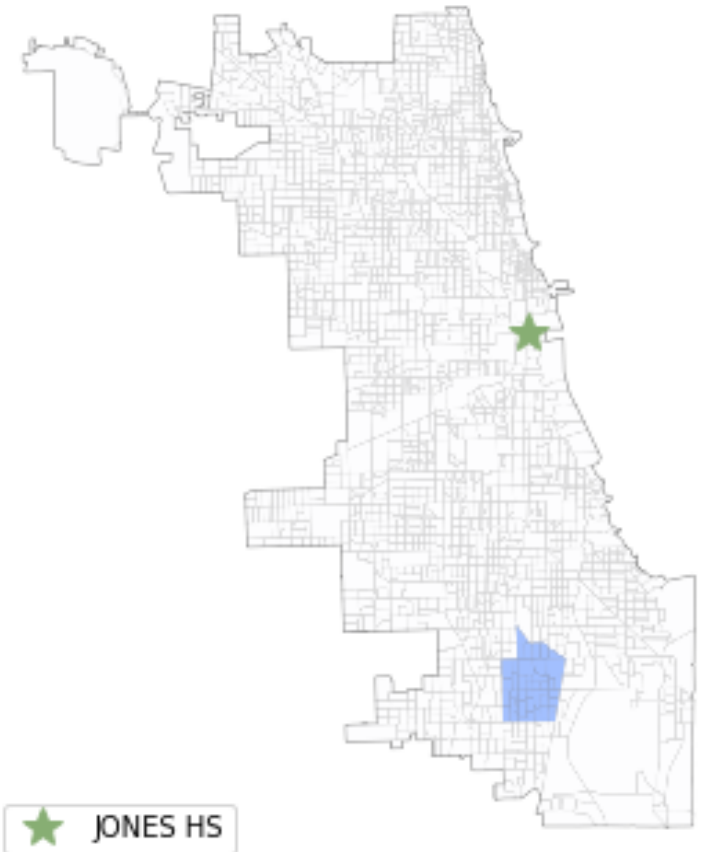
# Counterfactual 2: Bus Route To Level 1+ School

- 348 applicants
- 95% Black
- Only 27% rank a Level 1+ school first (vs 56% across all neighborhoods)
- Nearest Level 1+ is 35 minutes away by transit
- $\hat{\lambda} = -3.34$ (utils/hour)



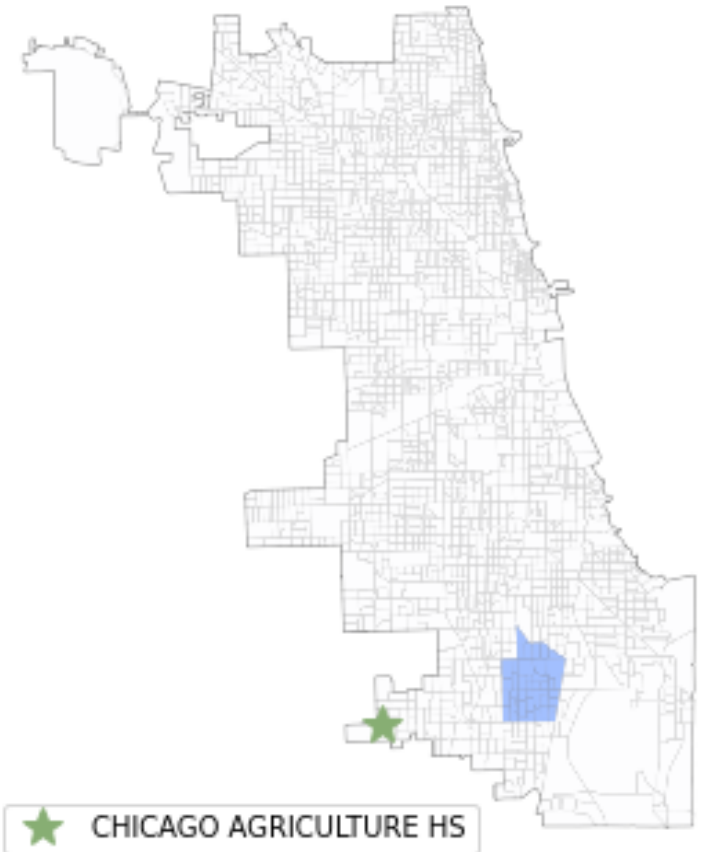
# Counterfactual 2: Bus Route To Level 1+ School

- **15%** of students would switch to this school if travel time reduced from transit time to drive time



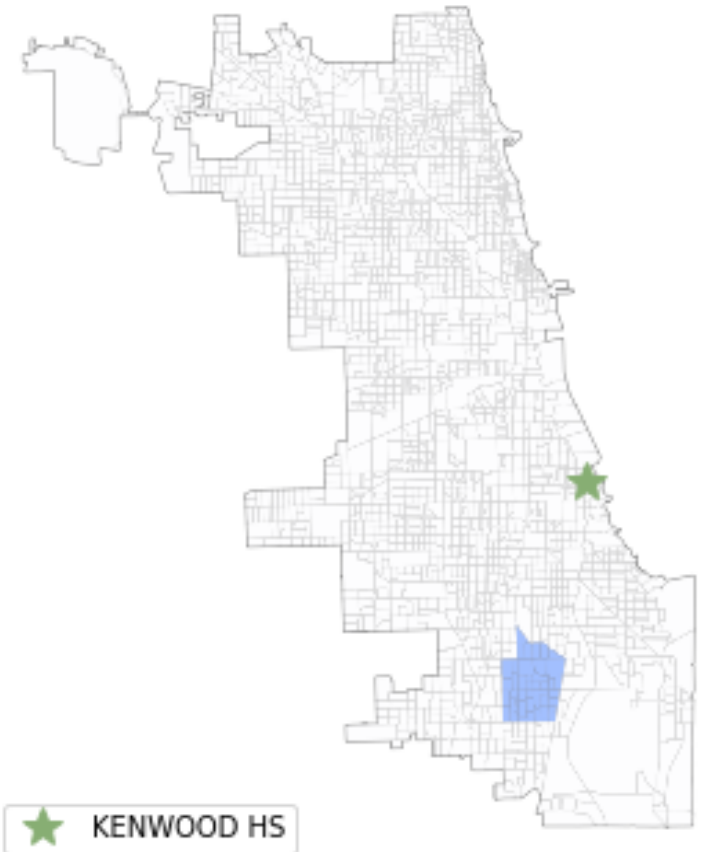
# Counterfactual 2: Bus Route To Level 1+ School

- **14%** of students would switch to this school if travel time reduced from transit time to drive time



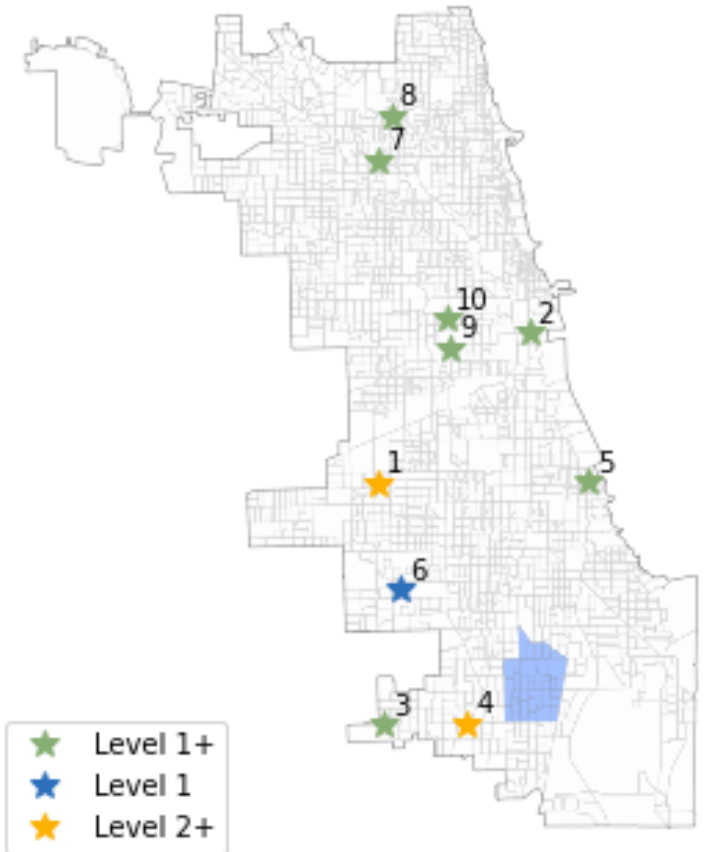
# Counterfactual 2: Bus Route To Level 1+ School

- **11%** of students would switch to this school if travel time reduced from transit time to drive time



# Counterfactual 2: Bus Route To Top Schools

- Top 10 schools that a dedicated bus route would induce switches to (individually)

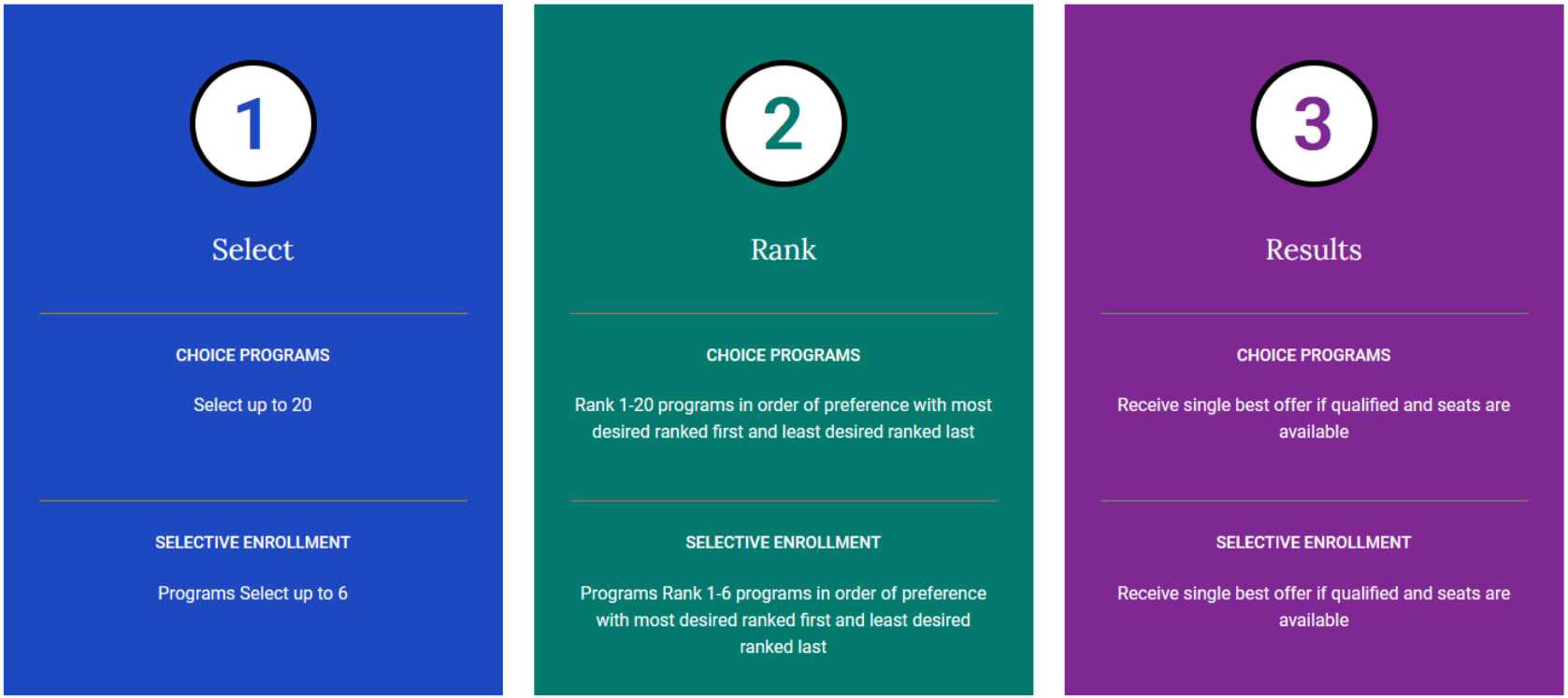


# Conclusions

- Key Policy Messages:
  1. Transportation time is a major driver of school choice decision and may play significant role in why low-income minority students disproportionately attend lower rated schools
  2. Well-targeted special express bus routes can induce parents to choose higher quality schools
- Future Work:
  - Assess magnitude of any residential sorting
  - Use model to identify *joint* bus routes with potential to maximize switchers from non-Level 1+ to Level 1+ schools
  - Identify routes to switch students from poor performing to better performing undersubscribed schools
  - Estimate number of underserved students who would attend an existing school that improved quality

# Thank you!

# School Choice Platform: GoCPS




**How to Rank Programs**

Rank the programs that are selected in the true order of preference. You will receive a single best offer for the highest-ranked program that is matched. Ranking a school higher does not increase your chances of getting an offer. Make sure to rank programs in the order that you would like to attend them.

**Changing School & Program Choices**

Students can change the schools and programs that have been selected up until the application deadline. Changes can be made directly in a student's GoCPS account or by calling the Office of Access and Enrollment at 773-553-2060. Make sure to resubmit all applications if any changes have been made.

**Re-Ranking**

Students are able to log into the GoCPS account and re-rank their choices up until the application deadline. No re-ranking will be possible once the application deadline has passed.

GUIDE TO RANKING AND CUT SCORES

# School Choice Platform: GoCPS




Back

# Model of School Choice

Utility of child $i$ attending school $j$ is a function of both household and school characteristics:

$$U_{ij} = \beta_k q_j - \lambda_k d_{ij} + \eta_k x_j + \epsilon_{ij}$$

- $q_j$: school quality
  - Composite metric presented to parents capturing growth in test scores, graduation rates, college enrollment rates, attendance rates
  - (future) Value-added
- $d_{ij}$: distance from student $i's$ home to school $j$
  - Commute time on public transit
  - (future) safety characteristics of routes
- $x_j$: vector of school characteristics
- $\epsilon_{ij}$: characteristics of school $j$ valued by household $i$ that are observed by parent
  - unobserved to us so modeled as i.i.d. Type I extreme value

# Model of School Choice

Utility of child $i$ attending school $j$ is a function of both household and school characteristics:

$$U_{ij} = \beta_k q_j - \lambda_k d_{ij} + \eta_k x_j + \epsilon_{ij}$$

- $\beta_k$: taste for quality by demographic group $k$
- $\lambda_k$: distaste for distance by demographic group $k$
- $\eta_k$: tastes for school characteristics by demographic group $k$

What we learn from these preference parameters:

1. Metric of how parents tradeoff distance and quality.
   - E.g. $\frac{\beta_k}{\lambda_k}$ = household values one unit increase in quality at one more unit of travel time
2. Use for counterfactuals
   - Reduce $i's$ travel time to any set of schools $j \rightarrow$ calculate $i's$ utility for each school $U_{ij}$ and estimate chosen school

Back

# Estimation

Enriching the "basic" discrete choice logit model

1. Mixed Logit: Allows for (random) heterogeneity in preferences and unrestricted substitution patterns (IIA)
2. Exploded Logit: Exploit information in 2nd, 3rd... ranked choices
   - $U_{i1} \geq U_{ik}$ for $k = 2, 3, 4, 5, ....J$
   - $U_{i2} \geq U_{il}$ for $l = 3, 4, 5, ....J$
   - $U_{i3} \geq U_{im}$ for $m = 4, 5, ....J$

For now, we estimate Mixed Logit but only utilize first choice data

# Estimation (with Mean Utility $\delta_{kj}$)

Observe first-ranked school $j$ for each student $i$

$$P_{ij} \equiv Prob(U_{ij} \geq U_{ik} \forall k \neq j, k \in J_i)$$

$$= \int \frac{exp(-\lambda_k d_{ij} + \delta_{jk})}{\sum_{l=1}^{J_i} exp(-\lambda_k d_{il} + \delta_{lk})} f(\gamma|\mu, \theta) d\gamma$$

where $\gamma \equiv (\lambda\, \delta)'$ are the preference parameters that come from pdf $f$ with mean$=\mu$ and covariance$=\theta$

Log-likelihood of first-ranked schools by all $N$ students:

$$LL(\mu, \theta) = \sum_{i=1}^{N} \sum_{j=1}^{J_i} d_{ij} ln(P_{ij})$$

where $d_{ij} = \mathbb{1}$(student $i$ ranked school $j$ first)

Estimate $\mu, \theta$ via Maximum Simulated Likelihood

◂ Back

# Contribution to Literature

- School choice and student achievement
  - e.g. Angrist, Pathak and Walters (2013), Campos and Kearns (2020), Barrow, Sartain and de la Torre (2020), Hastings and Weinstein (2008)
- Discrete choice models applied to school application data
  - Hastings, Kane and Staiger (2009), Burgess (2015), Laverde (2022), Kapor, Neilson and Zimmerman (2020)
  - Find first-order impact of distance on choices, trade off between academic achievement and proximity
- Emerging literature on role of transit access and school choice (Dustan and Ngo (2018))

**This is first paper (to our knowledge) to model how alternative transportation infrastructure can expand access to highly-rated schools.**

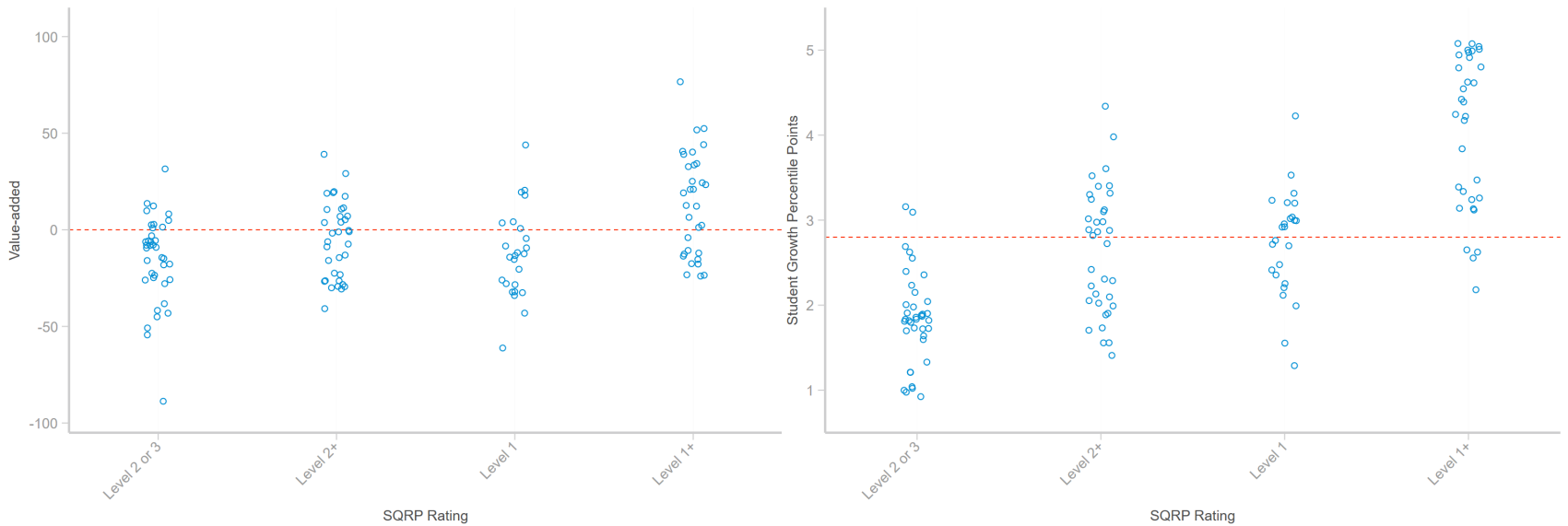
◂ Back

# School Characteristics by SQRP Rating

| | Level 1+ | Level 1 | Level 2+ | Level 2 |
|---|---|---|---|---|
| Black | 21.19 | 39.38 | 57.26 | 77.52 |
| Hispanic | 64.72 | 49.92 | 38.41 | 19.40 |
| White | 8.065 | 6.112 | 2.021 | 1.475 |
| Free/Reduced Lunch | 79.68 | 87.38 | 92.29 | 92.55 |
| Avg. SAT Score | 1022.9 | 912.2 | 879.5 | 833.7 |
| Observations | 31 | 24 | 34 | 32 |

Back

# SQRP Reflects School Quality




◂ Back

# Results: Correlates with Estimated School Mean Utility

## Black Students

| | $\delta_j$ | $\delta_j$ | $\delta_j$ | $\delta_j$ | $\delta_j$ |
|---|---|---|---|---|---|
| Pct FRPL | -0.0918 | | | | -0.0455 |
| | (0.0501) | | | | (0.0587) |
| Avg SAT Score | | 0.00928 | | | -0.0165 |
| | | (0.00648) | | | (0.0111) |
| 4-year Cohort Grad Rate | | | 0.0688 | | 0.0246 |
| | | | (0.0428) | | (0.0593) |
| College Enrollment Rate | | | | 0.124 | 0.148 |
| | | | | (0.0366) | (0.0517) |
| Constant | 8.633 | -7.745 | -4.246 | -6.828 | 8.971 |
| | (4.452) | (5.918) | (3.218) | (2.313) | (11.38) |
| Observations | 126 | 122 | 117 | 109 | 108 |