GeoGraphs for spatially embedded networks

Sergio Rey, Center for Open Geographical Sciences, SDSU
Elijah Knaap, Center for Open Geographical Sciences, SDSU
Zachary Neal, MSU
Clio Andris, GIT

School Boundaries vs Neighborhood Boundaries

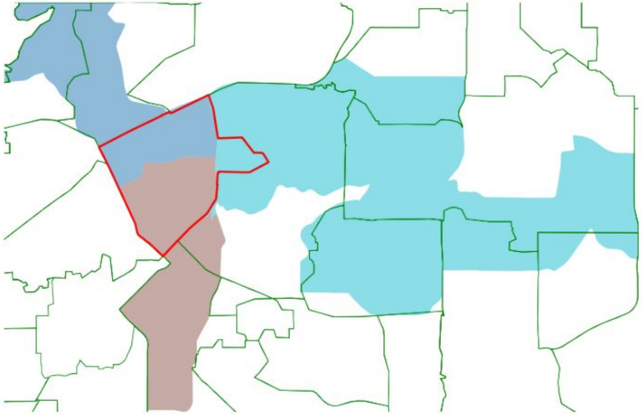


Figure 1: Lemon Avenue Elementary School Catchment (red) intersecting three neighborhoods (filled polygons). Other catchments shown with green boundaries.

School Boundaries vs Neighborhood Boundaries

How do neighborhoods influence schools?

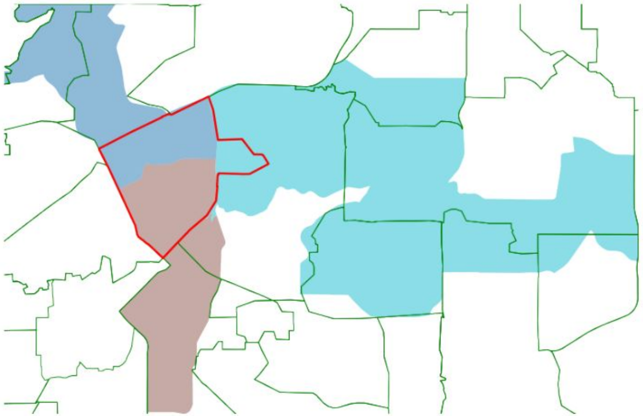


Figure 1: Lemon Avenue Elementary School Catchment (red) intersecting three neighborhoods (filled polygons). Other catchments shown with green boundaries.

School Boundaries vs Neighborhood Boundaries

How do schools influence neighborhoods?

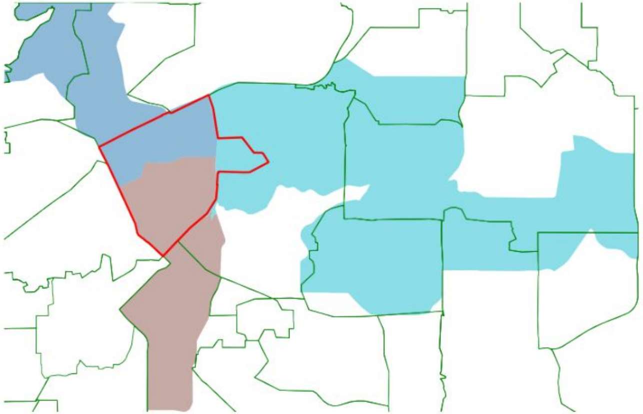
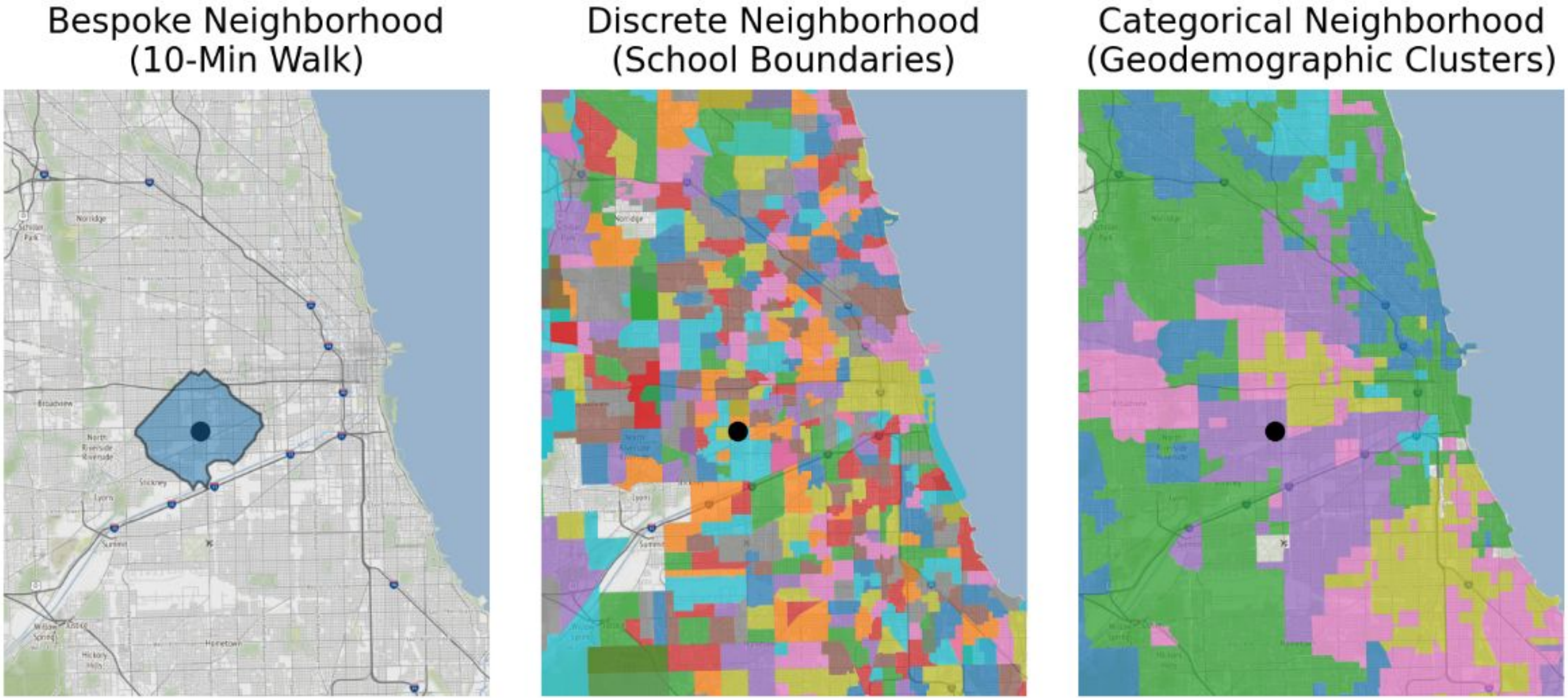


Figure 1: Lemon Avenue Elementary School Catchment (red) intersecting three neighborhoods (filled polygons). Other catchments shown with green boundaries.



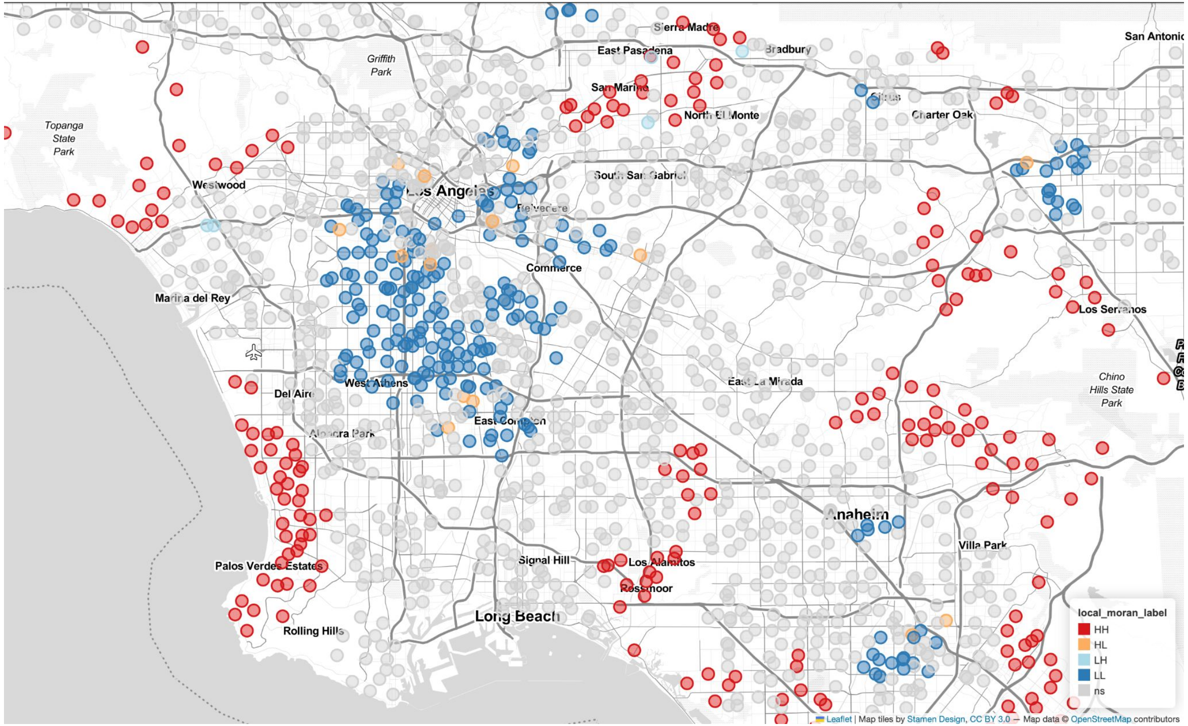
Neighborhoods

School performance

Student achievement outcomes are highly clustered in space

- **Blue**: coldspots of low achievement
- **Red**: hotspots of high achievement

Thus, a good model of student outcomes must account for spatial relationships. But at what scale and what unit? Space is complex.



Data: Stanford Educational Data Archive

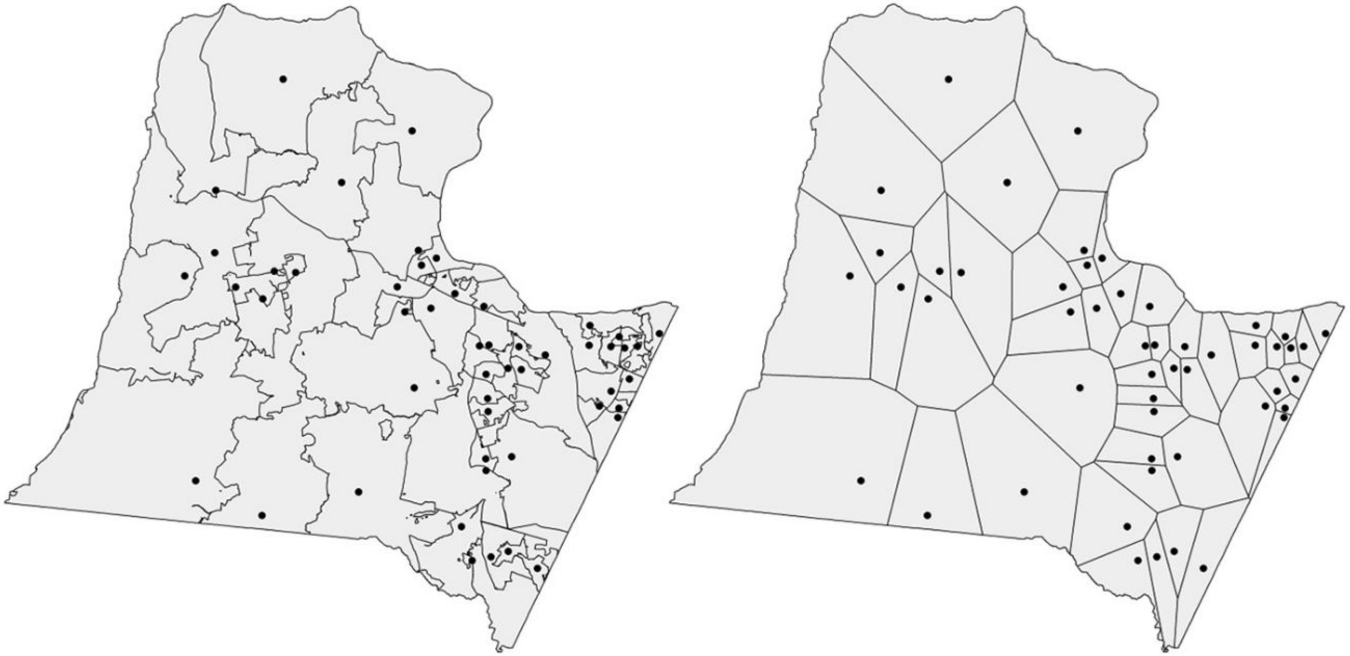
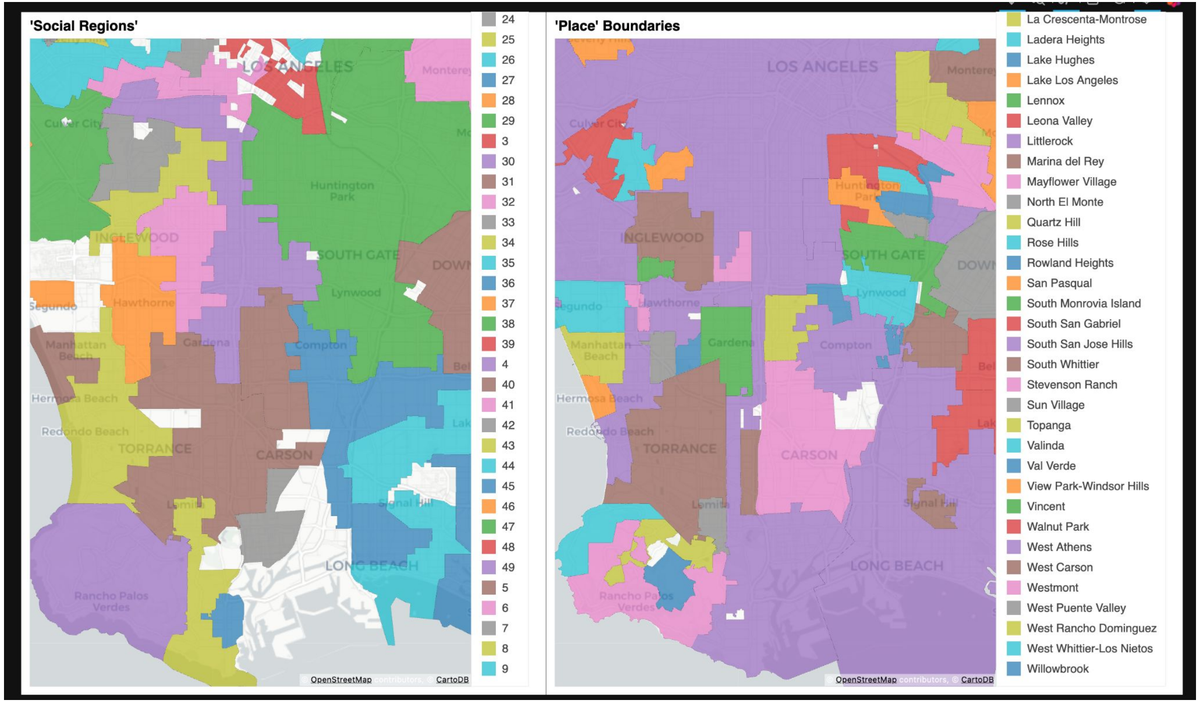


Figure 2. **School attendance zones. Left panel shows actual school attendance zones for schools in Loudoun County Public Schools, a suburban district outside Washington, D.C. Right panel shows Voronoi zones constructed around each school. School locations are marked with black dots.**

Social Regions vs Municipal Boundaries

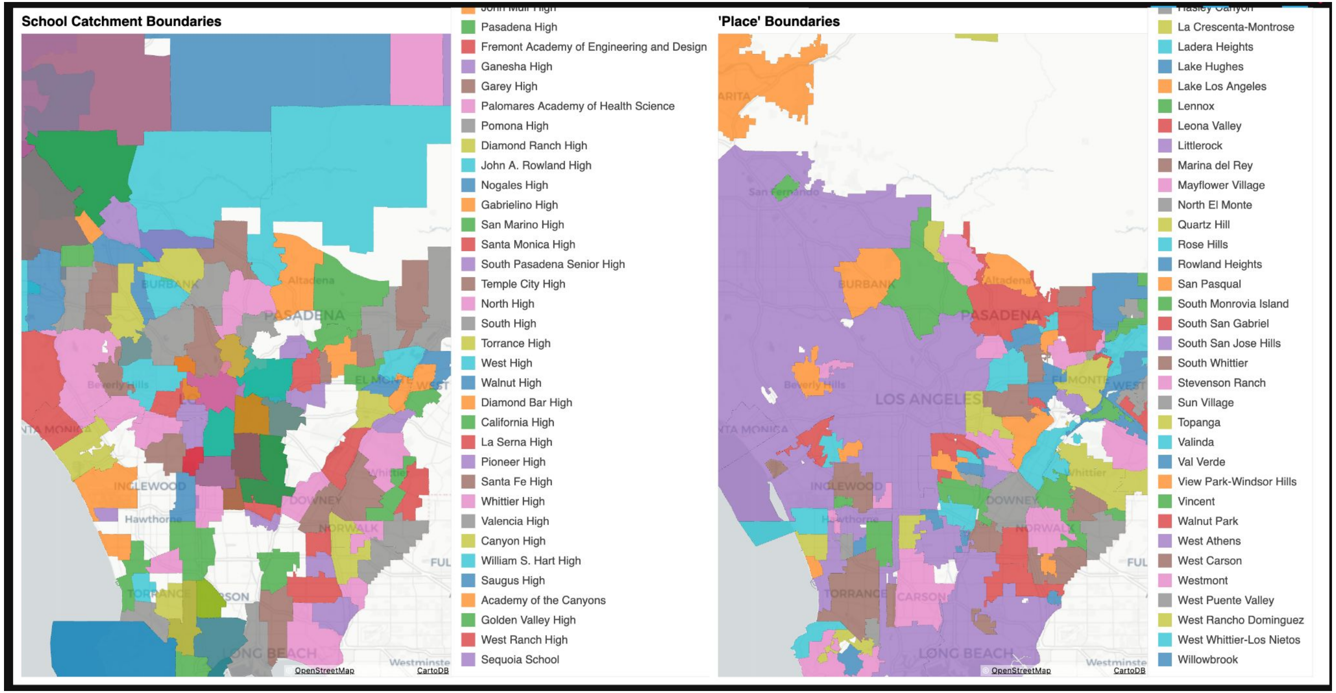
“Layers of Place” can also interact with each other, e.g. neighborhood←→city



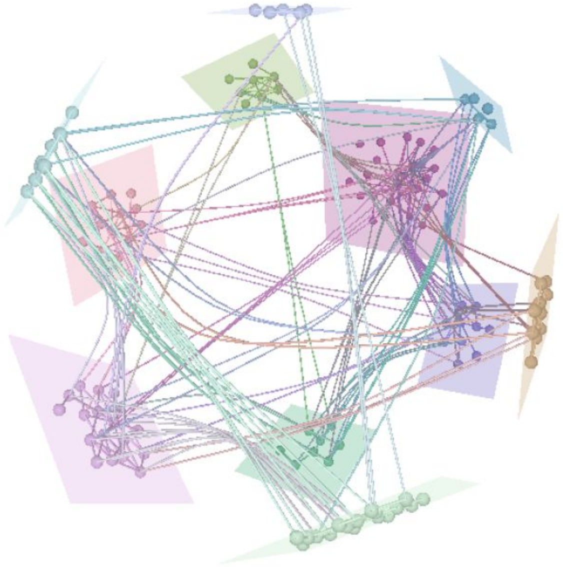
School Boundaries vs Municipal Boundaries

And school←→neighborhood,

etc.

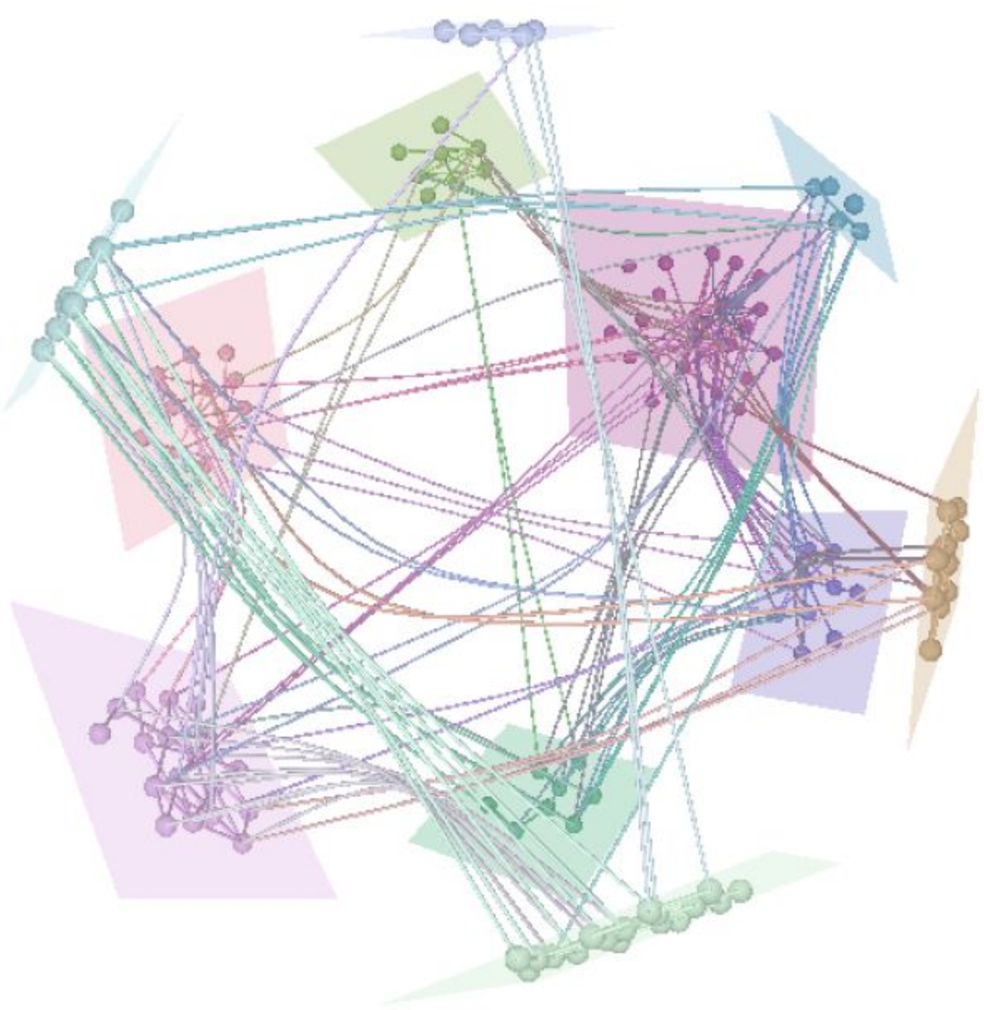


Understanding spatial outcomes requires parsing relationships among and between layers



A student is influenced simultaneously by her neighbors, her classmates, and her hometown—and each of those layers is influenced by other entities nearby (classmates of classmates, neighbors of neighbors)

Thus, a disturbance, innovation, or capital injection at one level can propagate through the system to affect other students. The critical transmitter of this transfer is **space,** which provides the mechanism for spillovers across layers



Conceptualizing The GeoGraph

To help model, analyze, and interrogate critical outcomes such as school achievement, we propose the GeoGraph as a first-class data structure that facilitates decomposition of these multiple feedback structures

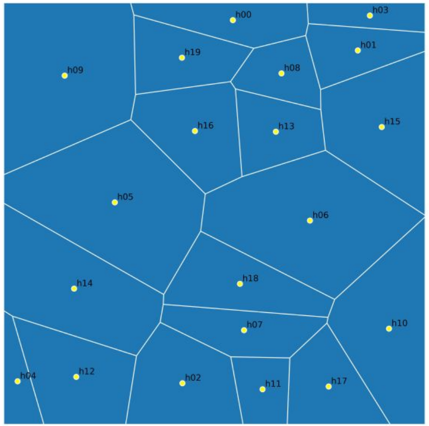


Figure 1. Neighborhoods for a community.

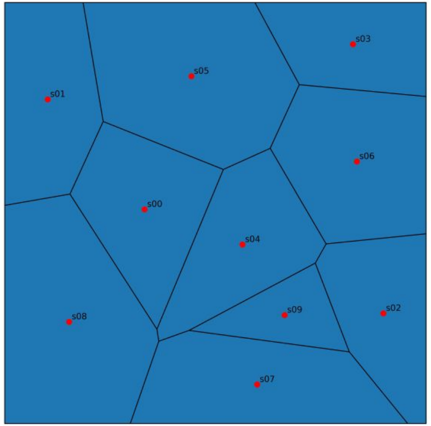


Figure 2. Catchments for a community.

Let C_s represent the polygon for catchment zone s, and N_i represent the polygon for neighborhood i. For each catchment zone C_s, we consider its spatial relations with the neighborhoods it draws children from. More formally,

$$\Omega_s = \frac{|C_s \cap G_s|}{|C_s \cup G_s|} \tag{1}$$

where G_s is defined as:

$$G_s = \bigcup_{i \in \Gamma_s} N_i, \{i \in \Gamma_s \mid N_i \cap C_s \neq \emptyset\} \tag{2}$$

Neighborhood to School Congruence. We can also measure the congruence from the perspective of a neighborhood unit.

$$\Omega_i = \frac{|N_i \cap G_i|}{|N_i \cup G_i|} \tag{5}$$

where G_i is defined as:

$$G_i = \bigcup_{s \in \Gamma_i} C_s, \{s \in \Gamma_i \mid C_s \cap N_i \neq \emptyset\} \tag{6}$$

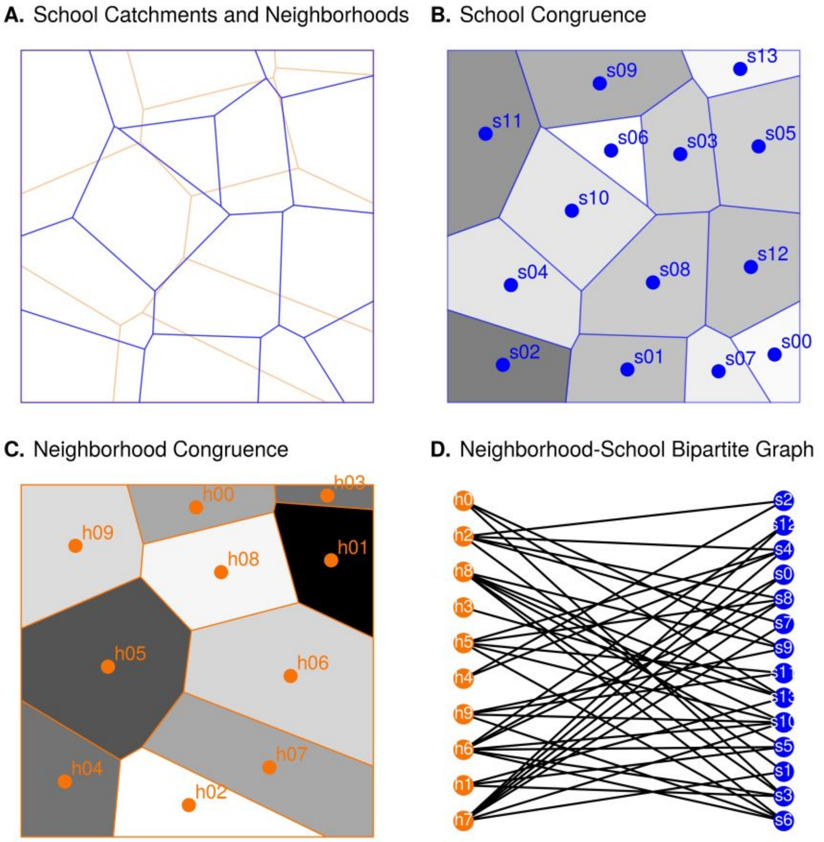
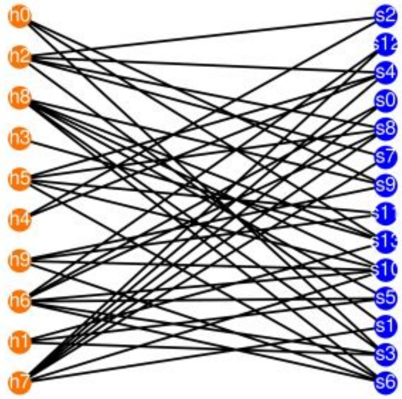


Figure 2: Incommensurate Spatial Partitions: [A] Catchments (blue), Neighborhoods (orange); [B] School congruence; [C] Neighborhood Congruence; [D] Bipartite Graph.

Bipartite Graph Analytics: Global

D. Neighborhood-School Bipartite Graph



$$\text{Density} = m/(n_S + n_H)$$

with m the number of edges in the graph, and n_S and n_H the number of school and neighborhood nodes, respectively. In the example from Figure 2 there are $|E| = 43$ edges, with 24 nodes (10 neighborhoods and 14 schools) resulting in a density of 0.56.

The mean degree of the nodes in the graph provides a measure of incomensurability between the catchments and the neighborhoods. Perfect comensurability between the two partitions would be reflected in a mean degree of 1. As incomensurability between the two partitions increases, so would the mean degree. The 24 nodes in the example graph have a mean degree of 3.58.

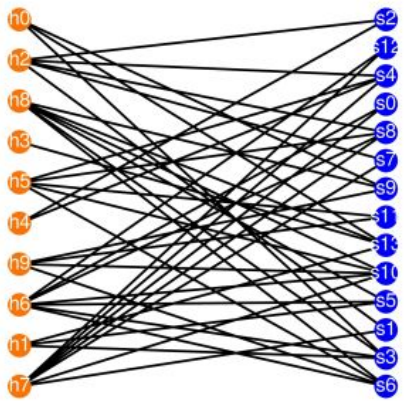
A global measure of clustering in a bipartite graph, due to Robins & Alexander (2004), is defined as four times the number of four cycles C_4 divided by the number of three paths L_3:

$$C_{RA} = \frac{4C_4}{L_3}$$

which gives a value of 0.39 for our example graph.

Bipartite Graph Analytics: Local

D. Neighborhood-School Bipartite Graph



Local measures of the bipartite graph can be used to identify key nodes in the spatial system. Table 1 reports three such measures, together with their ascending ranks, for our example. For node v in the two bipartite sets S (schools) with n_S nodes and H (neighborhoods) with n_H nodes, *betweenness centrality* is defined as the sum of the fraction of all-pairs shortest paths that pass through the node v

Closeness centrality for v is:

$$c_v = \frac{n_S + 2(n_H - 1)}{d}, \ \forall \ v \in H$$

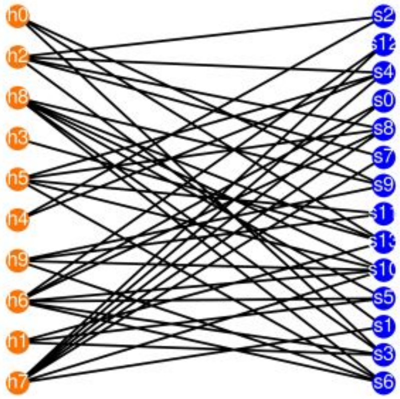
or

$$c_v = \frac{n_H + 2(n_S - 1)}{d}, \ \forall \ v \in S$$

where d is the sum of the distances from v to all other nodes.

Bipartite Graph Analytics: Local

D. Neighborhood-School Bipartite Graph



Degree centrality :

$$d_v = \frac{deg(v)}{m}, \forall\, v \in H$$

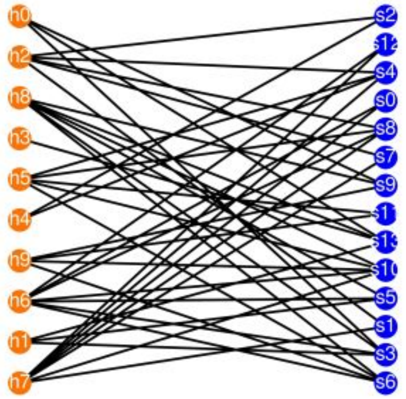
or

$$d_v = \frac{deg(v)}{n}, \forall\, v \in S$$

where $deg(v)$ is the degree of node v. Note that in the bipartitle case, the maximum possible degree of a node in a particular node set is the number of nodes in the opposite set. School 10 has degree centrality 0.5 as it has edges with 5 of the 10 neighborhoods, while neighborhood 2 has degree centrality of 0.357 as it has edges with 5 of the 14 schools.

Bipartite Graph Analytics: Local

D. Neighborhood-School Bipartite Graph



	betweenness	rank	closeness	rank	degree	rank
s0	0.02	7.50	0.58	15.50	0.20	5.50
s1	0.01	3.50	0.51	9.50	0.20	5.50
s2	0.01	2.00	0.42	3.00	0.20	5.50
s3	0.11	15.00	0.60	19.00	0.40	18.00
s4	0.13	18.00	0.58	15.50	0.40	18.00
s5	0.05	12.00	0.58	15.50	0.30	12.50
s6	0.09	14.00	0.67	22.50	0.40	18.00
s7	0.01	3.50	0.51	9.50	0.20	5.50
s8	0.14	19.00	0.67	22.50	0.40	18.00
s9	0.04	11.00	0.51	9.50	0.30	12.50
s10	0.24	22.00	0.75	24.00	0.50	23.00
s11	0.01	5.00	0.50	7.00	0.20	5.50
s12	0.02	7.50	0.58	15.50	0.20	5.50
s13	0.11	17.00	0.51	9.50	0.40	18.00
h0	0.02	10.00	0.44	5.00	0.21	9.50
h1	0.02	9.00	0.42	4.00	0.21	9.50
h2	0.11	16.00	0.48	6.00	0.36	14.50
h3	0.00	1.00	0.35	1.00	0.07	1.00
h4	0.01	6.00	0.39	2.00	0.14	2.00
h5	0.15	20.00	0.57	13.00	0.36	14.50
h6	0.27	24.00	0.64	21.00	0.50	23.00
h7	0.24	23.00	0.62	20.00	0.50	23.00
h8	0.23	21.00	0.59	18.00	0.43	21.00
h9	0.07	13.00	0.52	12.00	0.29	11.00

Table 1: Bipartite Local Centrality Measures

Bipartite Graph Based Spatial Weights

Rectangular

- Direct expression of B
- Two cases
  - W = B (hood-school)
  - W = B’ (school-hood)

Square

- Unipartite projection
- Two cases
  - W = BB’ (hood-hood)
  - W = B’B (school-school)

Bipartite Graph Based Spatial Weights

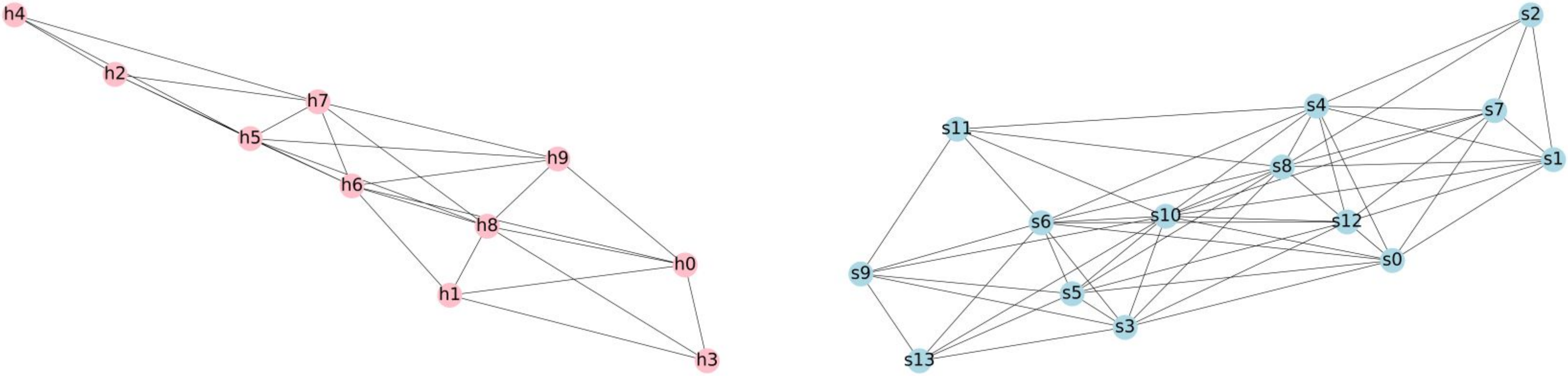
Rectangular

$$B' = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Square

$$BB' = \begin{bmatrix} 3 & 2 & 0 & 1 & 0 & 0 & 1 & 0 & 3 & 1 \\ 2 & 3 & 0 & 1 & 0 & 0 & 2 & 0 & 3 & 0 \\ 0 & 0 & 5 & 0 & 2 & 2 & 1 & 4 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 5 & 3 & 3 & 2 & 3 \\ 1 & 2 & 1 & 0 & 0 & 3 & 7 & 4 & 4 & 2 \\ 0 & 0 & 4 & 0 & 1 & 3 & 4 & 7 & 1 & 1 \\ 3 & 3 & 0 & 1 & 0 & 2 & 4 & 1 & 6 & 3 \\ 1 & 0 & 0 & 0 & 0 & 3 & 2 & 1 & 3 & 4 \end{bmatrix}$$

Bipartite Graph Based Spatial Weights



(a) Neighborhhoods (b) Schools

Figure 3: Unipartite graphs for neighborhoods and schools

Bipartite Graph Based Spatial Weights

Rectangular

- Direct expression of B
- Two cases
  - W = B (hood-school)
  - W = B’ (school-hood)

Square

- Unipartite projection
- Two cases
  - W = BB’ (hood-hood)
  - W = B’B (school-school)

Uses

- ESDA
- Confirmatory Modeling

Specifications for Incommensurate Spatial Partitions

From a confirmatory modeling perspective, the bipartite graph framing of this particular case of the incommensurate partitions problem offers several alternative specifications for the econometric modeling of spatial spillovers in this nexus. Consider a stylized educational outcomes model at the school level (Choi et al., 2005) that relates the performance of a school O_j to characteristics of the school S_j (e.g., student-teacher ratios, spending per student, teacher quality, etc.) as well as the neighborhood context (N_j) of the school:

$$O_j = \alpha + S_j \beta_S + N_j \beta_N + \epsilon_j$$

where α, β_S, and β_N are a scalar, and vectors of parameters, and ϵ_j is the disturbance term. We leverage the network-based representation of the school-neighborhood nexus, and specifically, the matrix representation of B and the related projections, to extend this model in several ways.

School-to-school spillovers

$$O_j = \alpha + S_j\beta_S + \rho W_{j,.}O + N_j\beta_N + \epsilon_j$$

where $W_{j,.}$ is row j of a spatial weights matrix (Anselin & Rey, 2014) and O is a vector of outcomes for schools in the community. We propose that the spatial weights W obtain from a projection of the bipartite graph, $W = B'B$. Such a projection represents the possible channels of diffusion of school-based exposure (e.g., curricula unique to a single school) from one school to another via unstructured neighborhood-based interactions.

Neighborhood-to-school spillovers

$$O_j = \alpha + S_j\beta_S + W_{j,.}N\beta_N + \epsilon_j$$

where $W_{j,.}$ is the jth row of a rectangular matrix capturing potential interactions between the catchment and the neighborhoods it draws from. A spatial lag of of these neighborhood influcences is WN where N is a vector (or matrix) of neighborhood characteristics. We propose that the spatial lag WN be formed by setting $W = B'$.

This gives rise to two forms of contextual mobility (Sharkey, 2013) in which the children of wealthy mover families experience an upward shift in their neighborhood attainment. In contrast, the children of the lower SES families, who remain, experience a relative drop in their neighborhood attainment due to the out-migration of wealthier peers. This suggests that school quality matters for neighborhood SES. A specification that considers this link might incorporate the school channel as:

$$P_i = \alpha + S_i\beta_S + N_i\beta_N + \epsilon_i$$

where P_i is the poverty rate in neighborhood i, S_i is a vector of school characteristics for students residing in i, and N_i is a set of neighborhood characteristics. We can extend this specification in a number of ways by drawing on the bipartite graph.

School-to-neighborhood spillovers When there is less than complete congruence between catchments and neighborhoods, children from the same neighborhood will be attending different schools. Indeed, a study of UK public schools found that the median neighborhood sent children to three different schools (Leckie, 2009). As such, the measure of the school context for the neighborhood can be expanded:

$$P_i = \alpha + W_{i,.} S\beta_S + N_i\beta_N + \epsilon_i$$

where $W_{i,.}$ is the ith row of a rectangular matrix specifying potential interactions between neighborhoods and the school catchments children attend, and S is a vector of school attributes. We propose that the spatial lag WS obtain from the product of the bipartite graph B and school characteristics S, such that $W = B$.

Neighborhood-to-neighborhood spillovers

$$P_i = \alpha + S_i\beta_S + \zeta WP + N_i\beta_N + \epsilon_i$$

where P is a vector of neighborhood poverty rates and W a spatial weights matrix. In this case, we propose that the spatial weights W obtain from the other projection of B, $W = BB'$ (Neal et al., 2021). This second projection captures the ties between students living in different neighborhoods that form through interactions at school (Burdick-Will, 2018). It therefore represents the possible channels of diffusion of neighborhood-based exposure (e.g., concentrated disadvantage) from one neighborhood to another via structured school-based interactions (Johnson, 2008, 2013).

Empirical Illustration: San Diego County

- 370 public elementary school catchments
- 198 neighborhoods (regionalization: spatially constrained k-means on 2058 blockgroups)
- Focus
  - Contiguity versus BPG-based spatial weights
  - ESDA
    - Global SA
    - Local SA
    - Global BV SA
    - Local BV SA
  - Spatial Segregation

Empirical Illustration: San Diego County

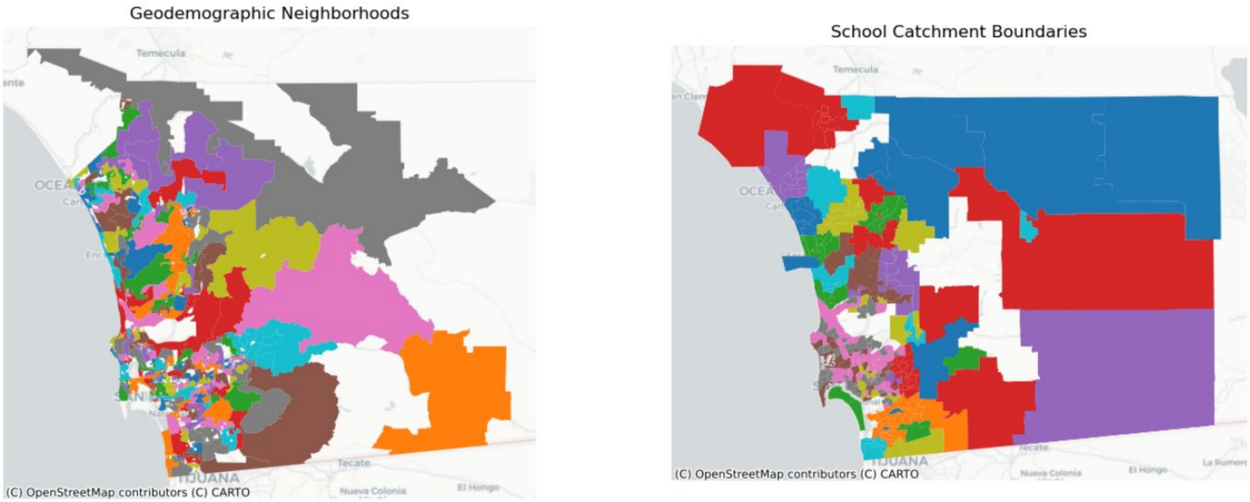
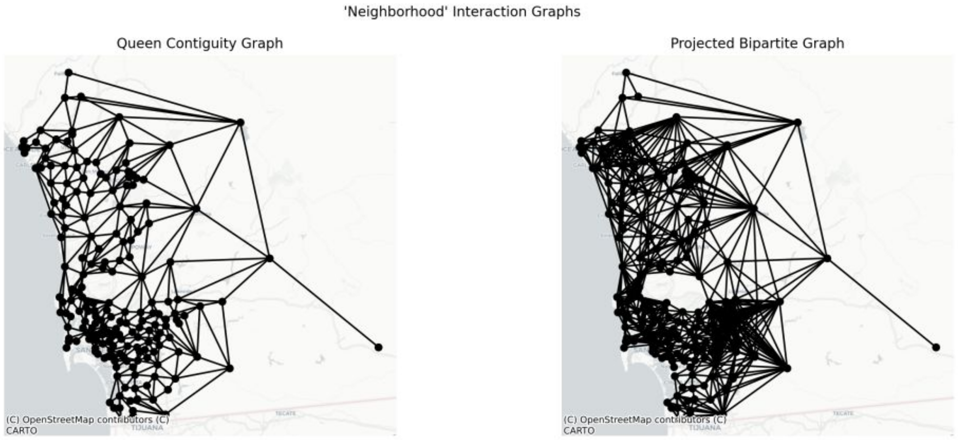
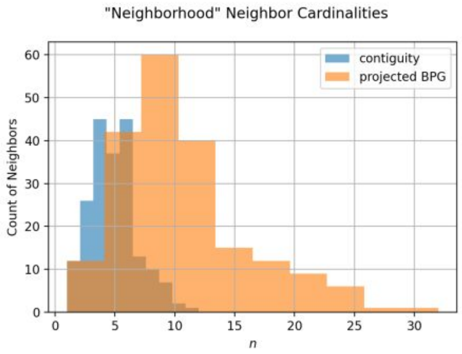


Figure 4: Geodemographic Neighborhoods & School Catchments

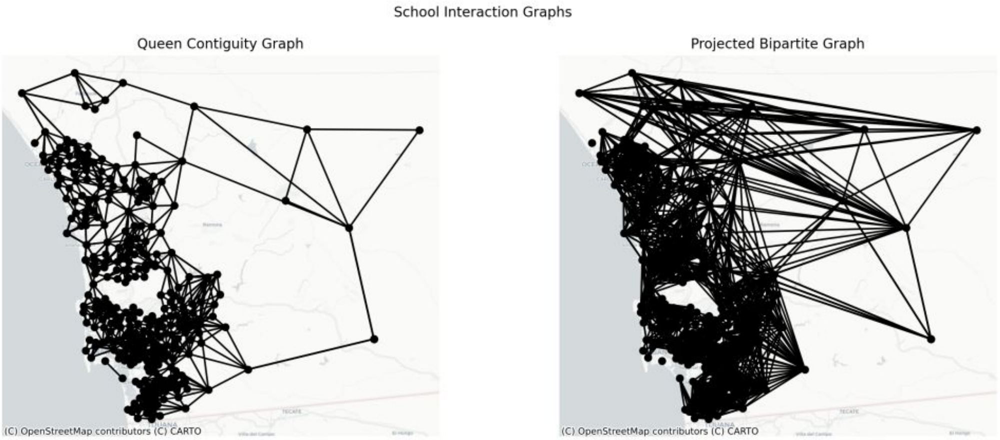


(a) Neighborhood Contiguity Graph vs. Projected BPG

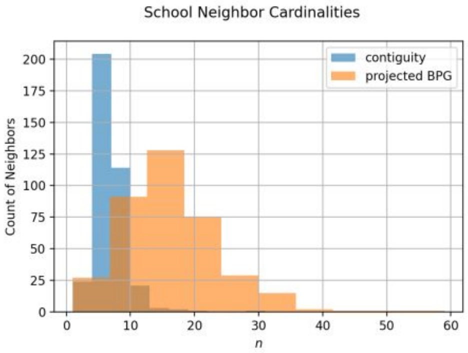


(b) Neighborhood Cardinality Histogram

Figure 5: Neighborhood Contiguity Graph vs. Projected BPG

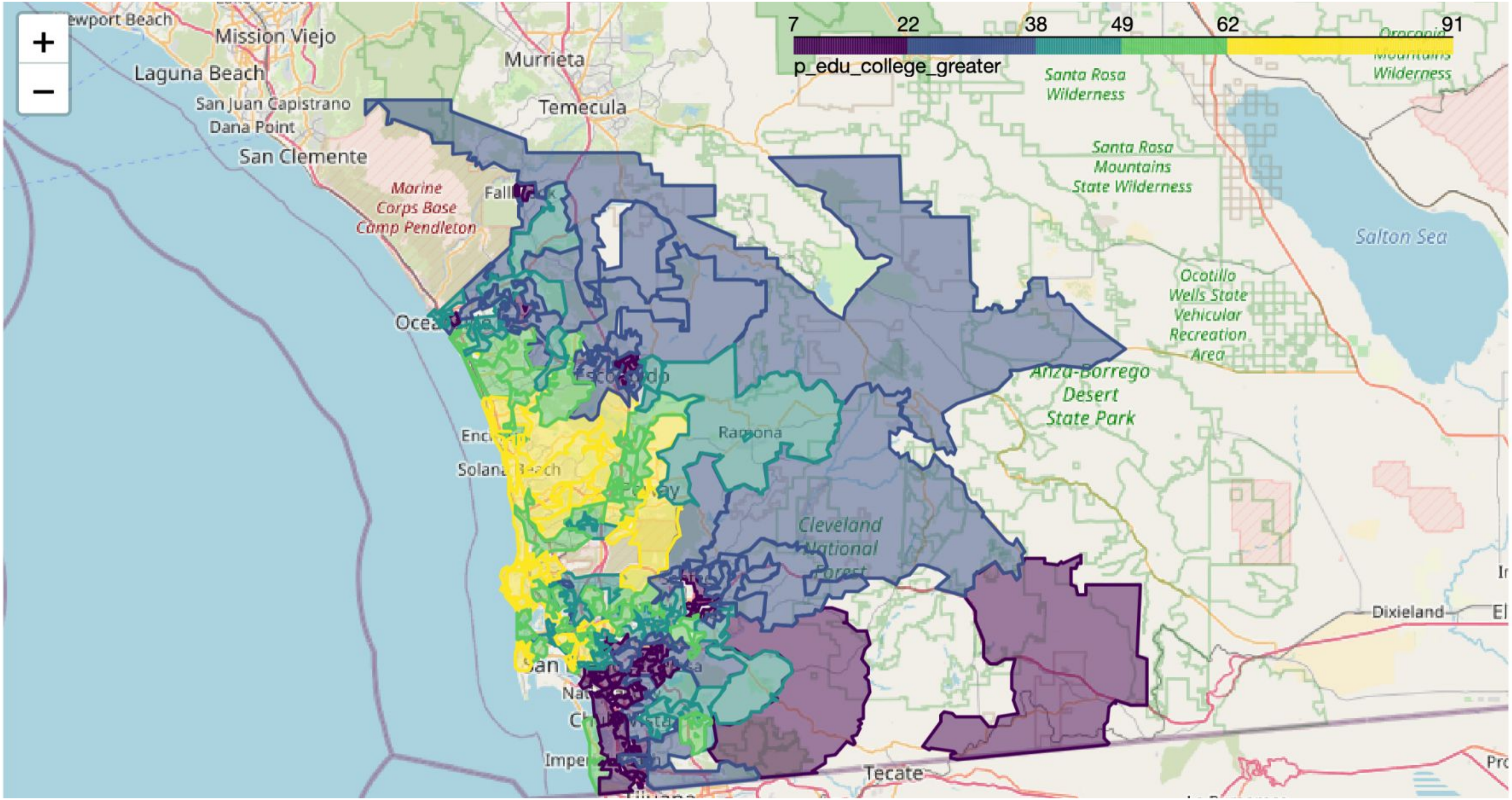


(a) School Contiguity Graph vs. Projected BPG

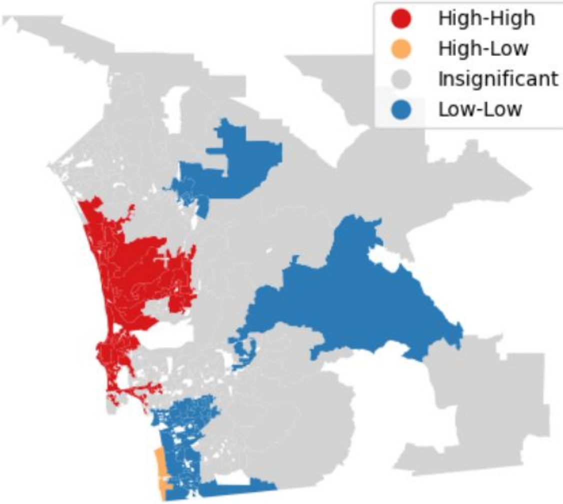


(b) School Cardinality Histogram

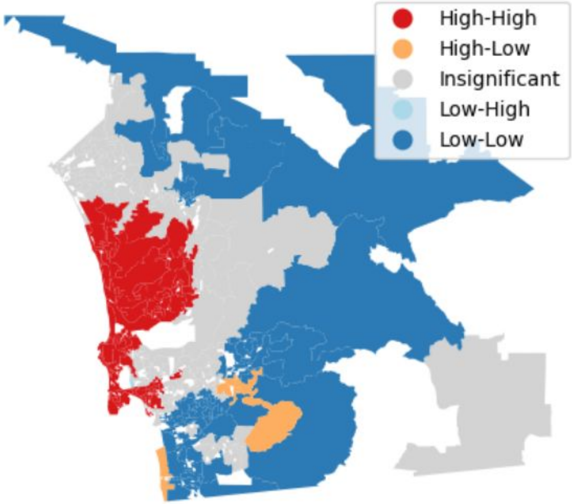
Figure 6: School Contiguity Graph vs. Projected BPG



7
22
38
49
62
91
p_edu_college_greater
Mission Viejo
Laguna Beach
San Juan Capistrano
Dana Point
San Clemente
Murrieta
Temecula
Marine Corps Base Camp Pendleton
Santa Rosa Wilderness
Santa Rosa Mountains State Wilderness
Salton Sea
Ocotillo Wells State Vehicular Recreation Area
Anza-Borrego Desert State Park
Ramona
Solana Beach
Cleveland National Forest
Dixieland
Tecate

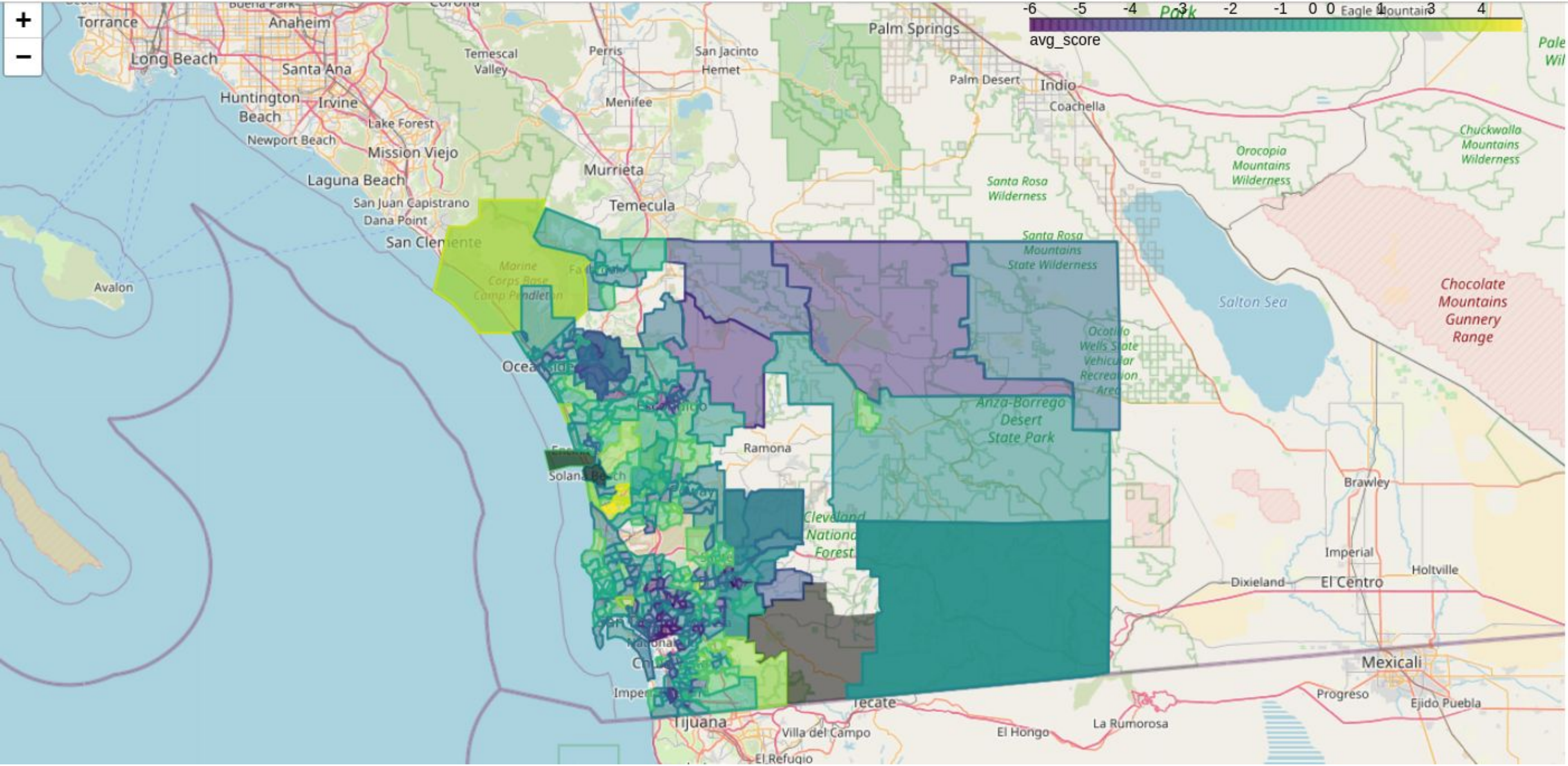


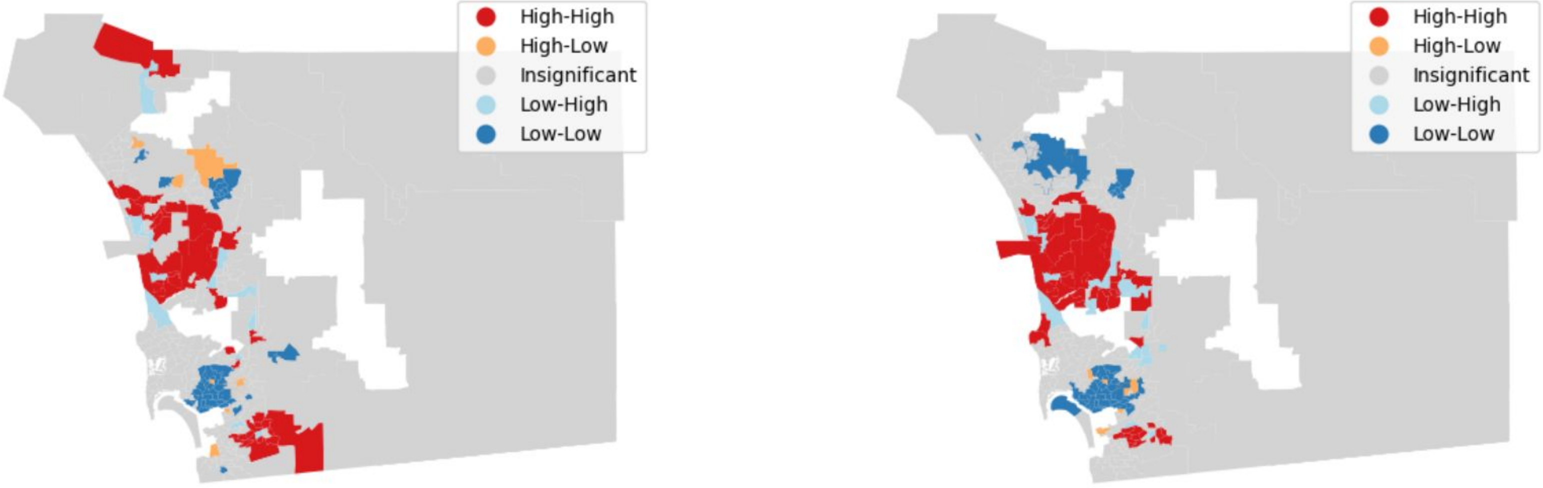
(a) Contiguity



(b) Projection

Figure 8: Cluster Maps for Neighborhood Adult Education Levels


-6 -5 -4 -3 -2 -1 0 0 1 3 4
avg_score



(a) Contiguity

(b) Projection

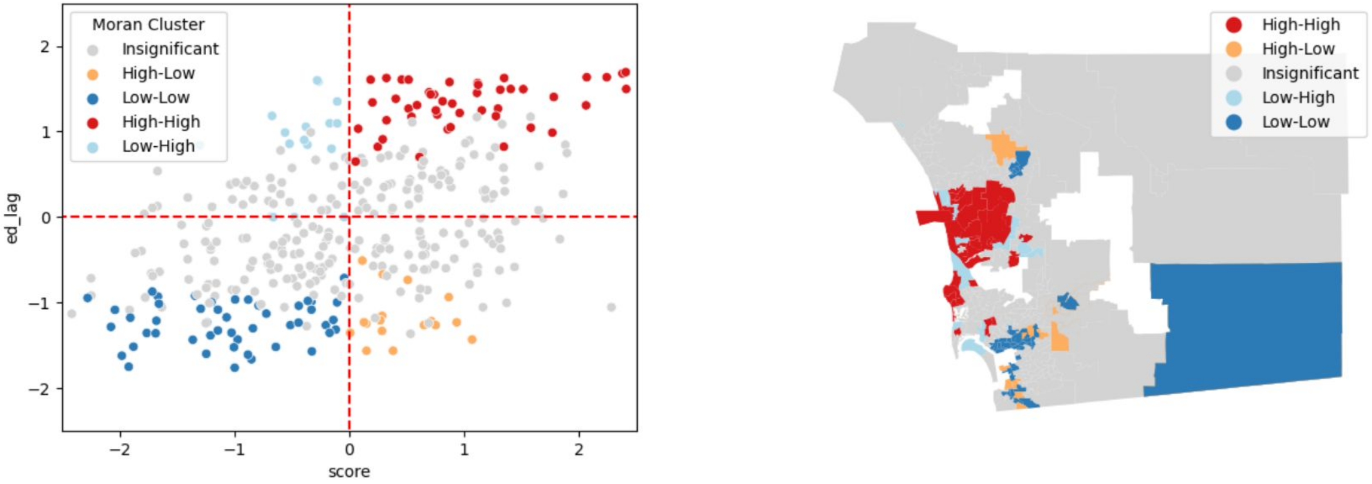
Figure 10: Cluster Maps for Average Test Scores by Catchment

Bivariate LISA

$$I_s = score_s B_{.,s}^{'} ed$$

where $score_s$ is the test score for school s, $B_{.,s}$ is column s of the bipartite adjacency matrix B and ed is a vector of neighborhood education levels. For inference on the bivariate LISA we rely on random permutations of the ed vector.

Bivariate LISA



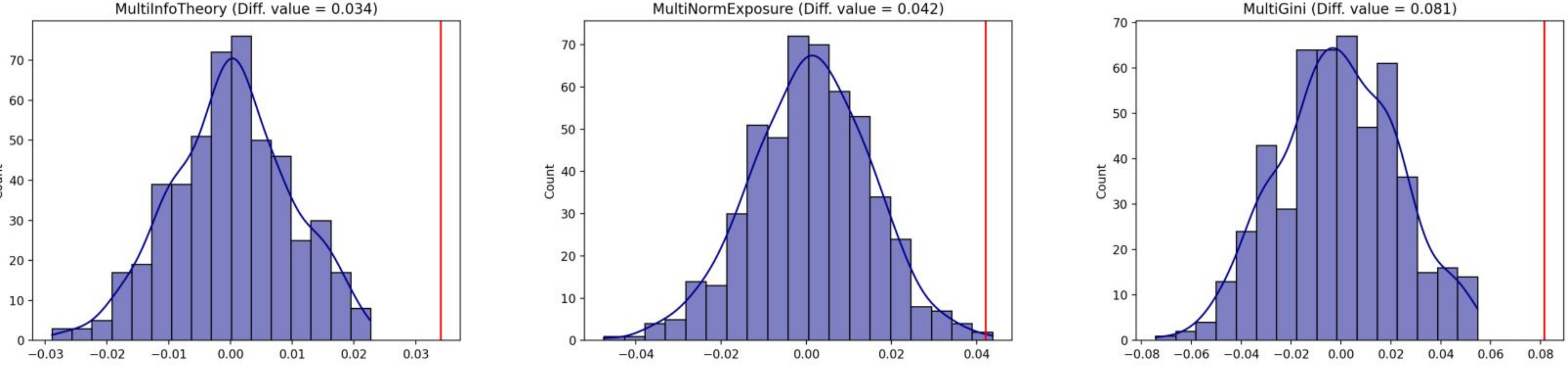
(a) Moran Scatterplot

(b) LISA Cluster Map

Figure 11: Bivariate Local Autocorrelation Average Test Score and Lag Parental Education (by school)

Spatial Segregation

To test the robustness of this difference, we apply the method described in Cortes, Rey, et al. (2019) and Rey et al. (2021) that provides for comparative inference between two segregation measures using the segregation module from the PySAL software ecosystem (Cortes, Knaap, et al., 2019; Rey et al., 2022). Specifically, the test uses a computational inference to develop a distribution of differences between segregation measures under the null hypothesis of no difference. The observed difference (in this case the difference between a measure computed with two different W specifications) is then tested against this reference distribution.



(a) Information Theory Index (b) Exposure Index (c) Gini Index

Figure 13: Comparative Segregation Analytics

Conclusion

- Bipartite graph framing of the ICP
  - New analytics for exploring the school-neighborhood spatial nexus
  - Comparative, global, local
- Bipartite based spatial weights
  - Rectangular or square
  - Multi-process or single process
- Confirmatory spillover specifications
- ESDA
  - Bivariate LISA Enhancements

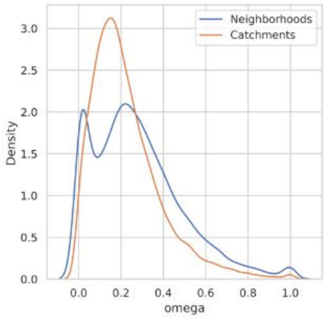
Next Steps

- Beyond San Diego
- Other cases of ISP problems
- Generalized weights for BPG
- Extraction of Backbone of graphs
- Extension from single to multi-equation systems

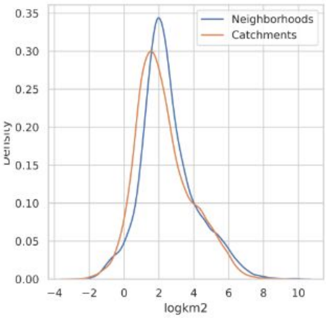
	Neighborhoods	Primary Catchments	Population (1,000)	Non-Hispanic White %
count	17608	24507	220763	
mean	160	223	20072	62.4
std	209	280	2620	17.4
min	34	45	501	6.3
25%	55	80	677	52.1
50%	77	118	981	65.9
75%	184	240	2179	76.0
max	1506	1867	19318	92.2

Table 1
Summary statistics for the 110 Core Based Statistical Areas.

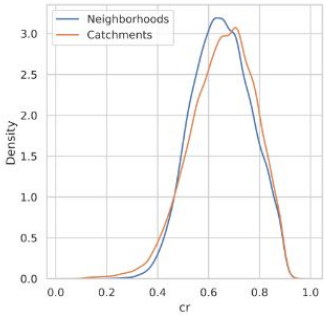
Study areas summary statistics



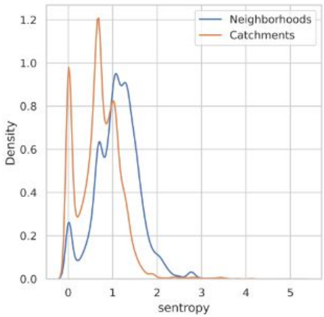
(a) Congruence



(b) Area (log(km2))

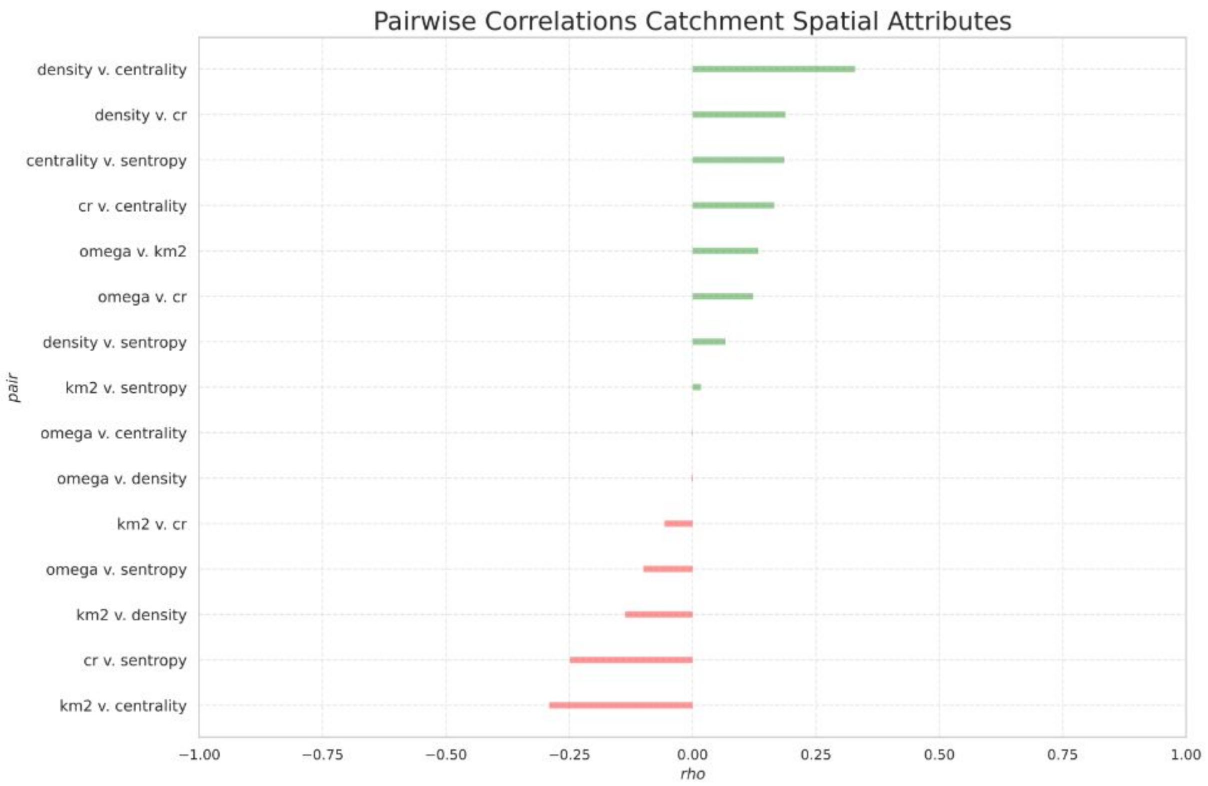


(c) Circularity Ratio

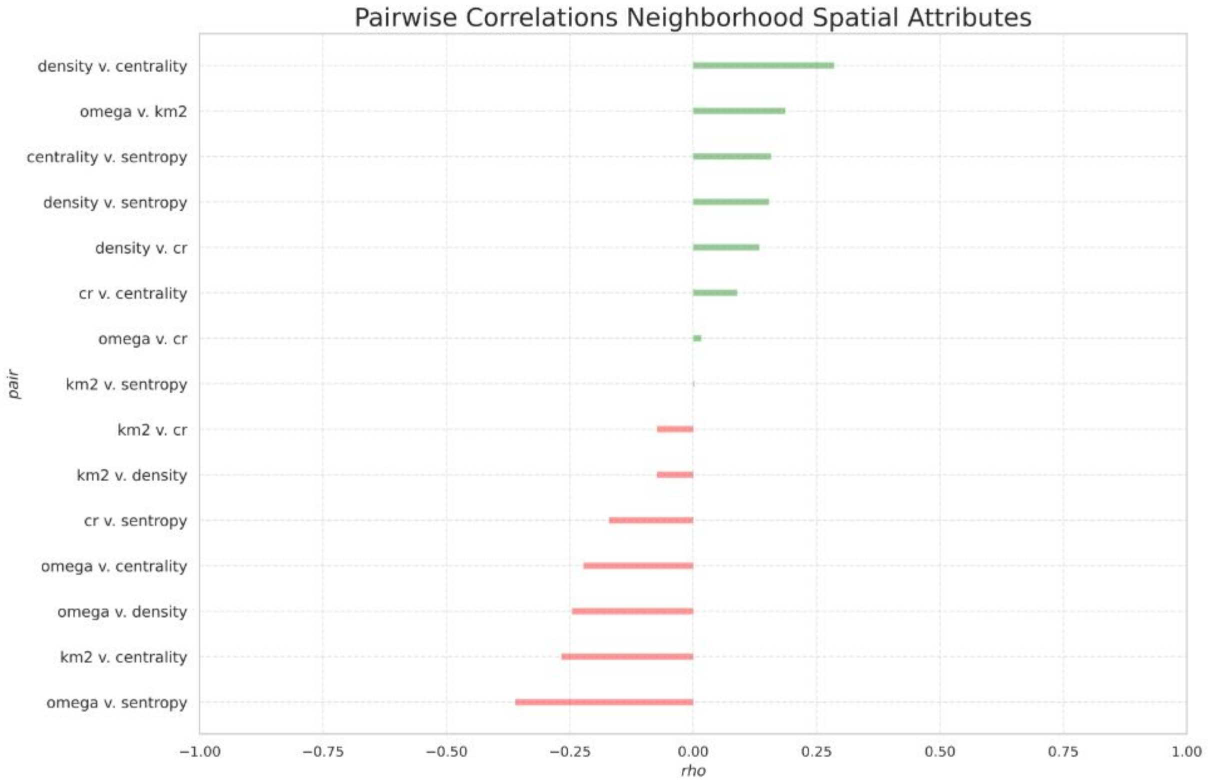


(d) Spatial Entropy

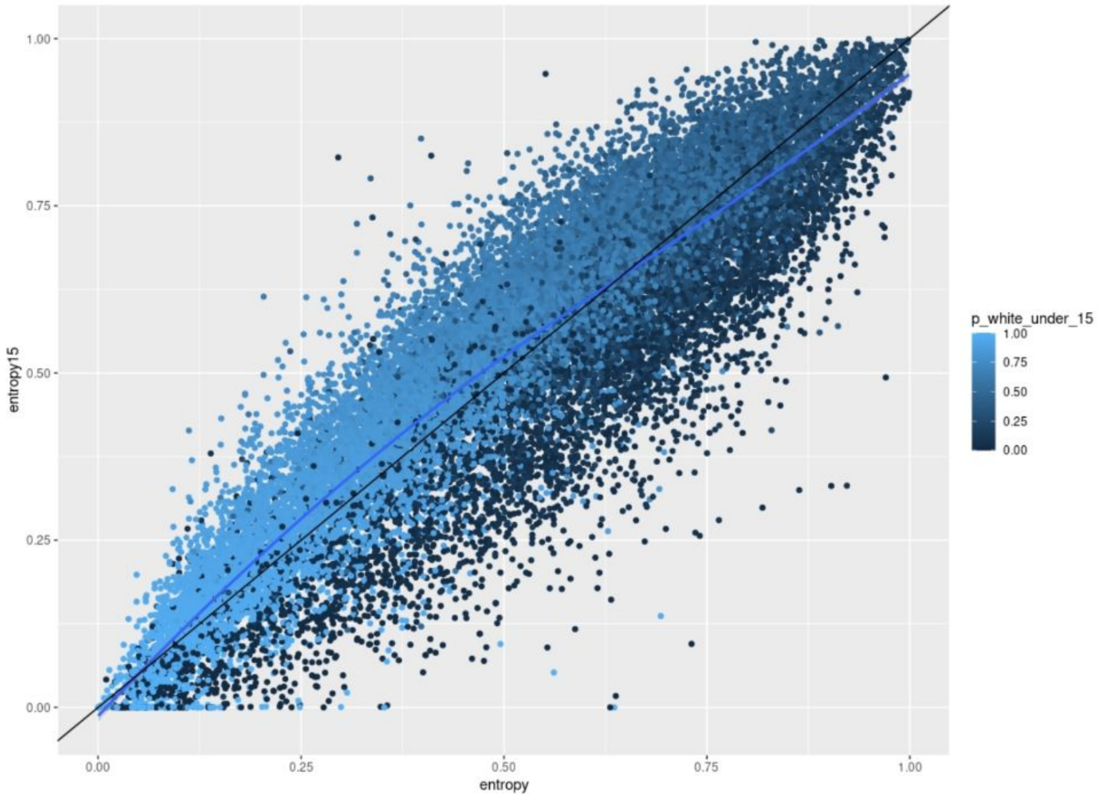
Summary measures: Catchments vs Neighborhoods



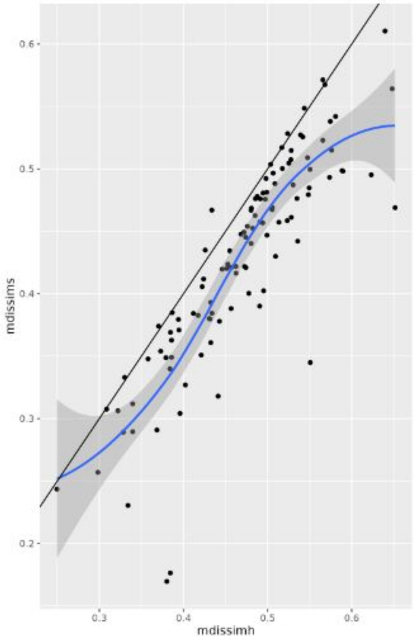
Spatial Attributes: Catchments



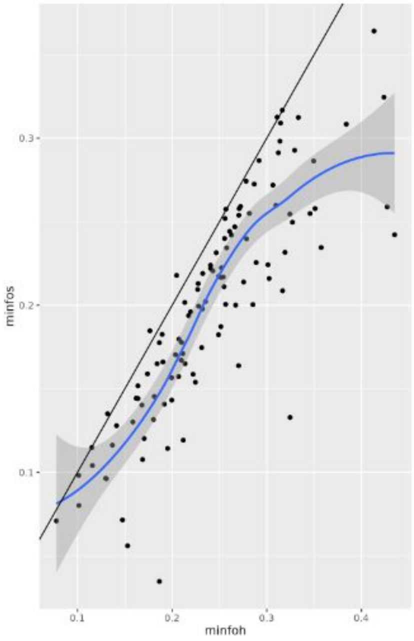
Spatial Attributes: Neighborhoods



Ethnoracial Entropy: Overall vs. School-Aged



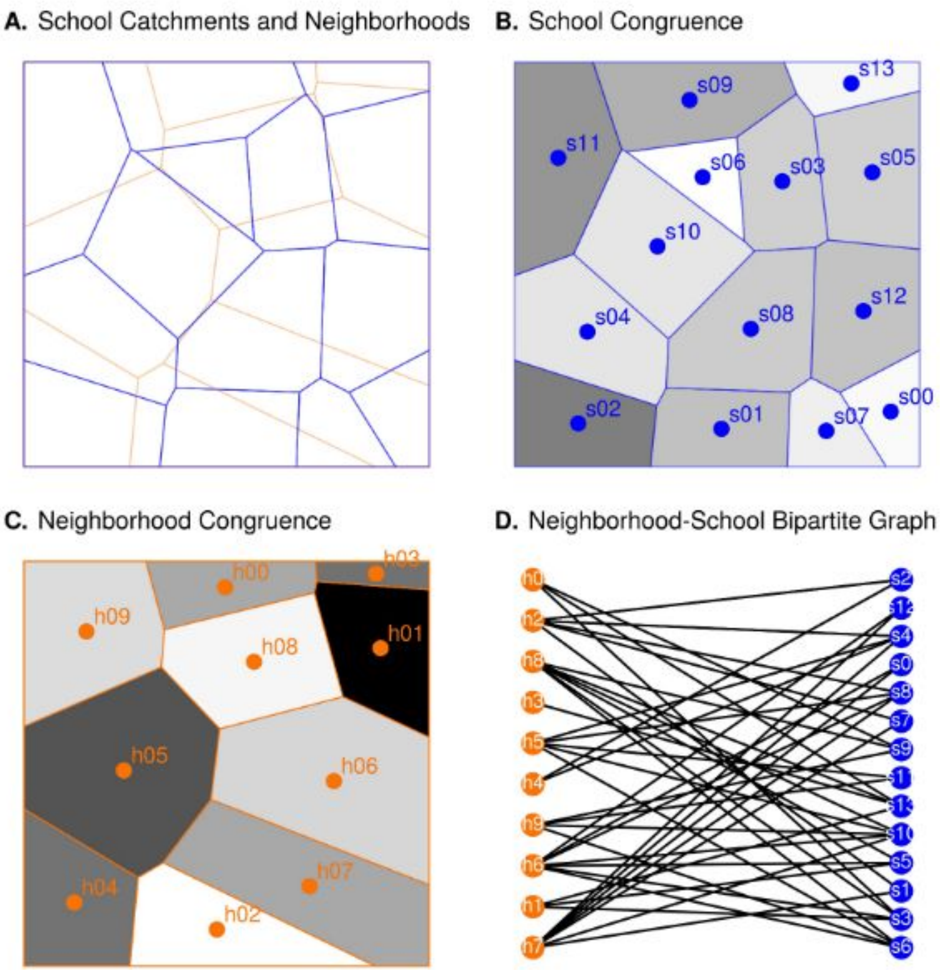
(a) Multigroup Dissimilarity Segregation Index



(b) Multigroup Information Segregation Index

Segregation: Schools vs. Neighborhoods

Extensions


A. School Catchments and Neighborhoods
B. School Congruence
s13
s09
s11
s06
s03
s05
s10
s12
s04
s08
s00
s02
s01
s07
C. Neighborhood Congruence
h03
h00
h09
h01
h08
h05
h06
h07
h04
h02
D. Neighborhood-School Bipartite Graph

let B be a bipartite graph where $B_{ik} = \frac{|i \cap k|}{|i \cup k|}$.

$$O_j = \alpha + S_j\beta_S + N_j\beta_N + \epsilon_j$$

Traditional school outcome specification

$$O_j = \alpha + S_j\beta_S + \rho W_{j,.}O + N_j\beta_N + \epsilon_j$$

School-to-school spillovers

$$O_j = \alpha + S_j\beta_S + \rho W_{j,.}O + N_j\beta_N + \epsilon_j$$

$$\rho W_{j,.}O$$

$$W = B'B$$

Diffusion of school-based exposure (curricula)
via
Unstructured neighborhood-based interactions

School-to-school spillovers

$$O_j = \alpha + S_j \beta_S + W_{j,.} \beta_N + \epsilon_j$$

Neighborhood-to-school spillovers

$$P_i = \alpha + S_i\beta_S + N_i\beta_N + \epsilon_i$$

Traditional neighborhood outcome specification

$$O_j = \alpha + S_j \beta_S + W_{j,.} \beta_N + \epsilon_j$$

Neighborhood-to-school spillovers

$$P_i = \alpha + W_{i,.}\beta_S + N_i\beta_N + \epsilon_i$$

$$W = BS$$

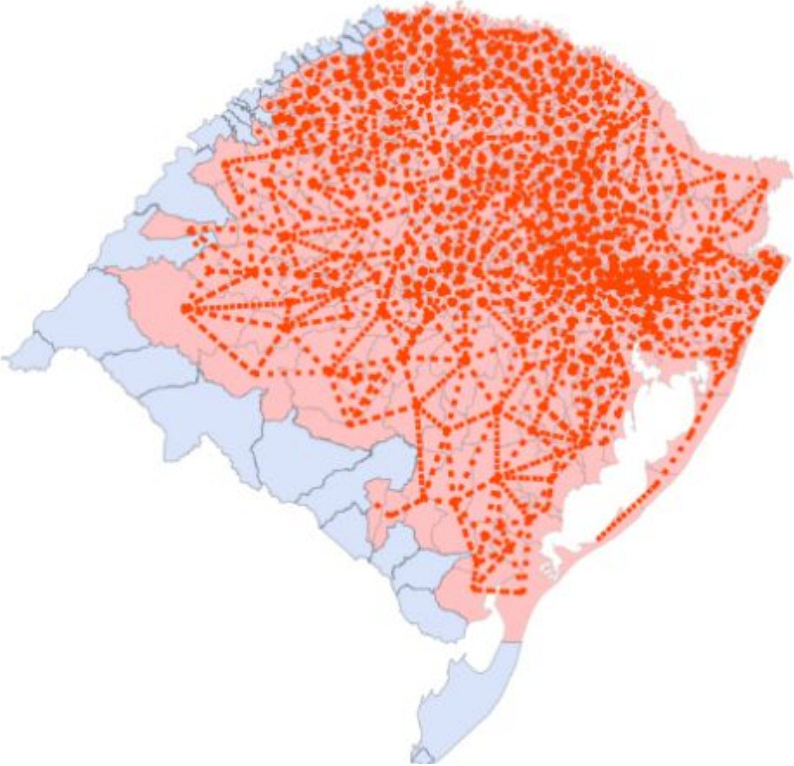
School-to-neighborhood spillovers

$$P_i = \alpha + S_i\beta_S + \zeta W P + N_i\beta_N + \epsilon_i$$

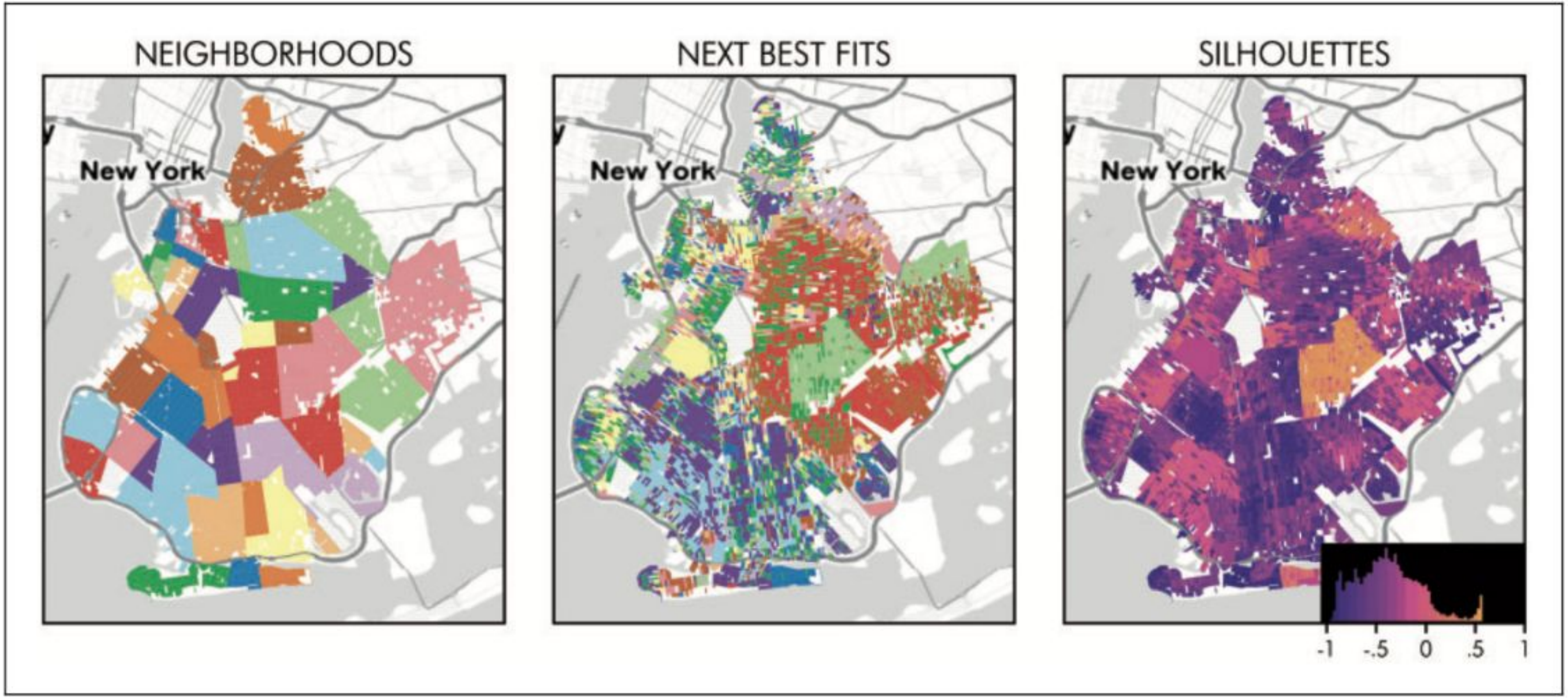
$$W = BB'$$

Ties between students living in different neighborhoods formed at school.

Neighborhood-to-neighborhood spillovers



Topological errors



Geosilhouettes for neighborhood delineation

Figure 3. School spatial congruence. School catchments are labeled with *sx* with black boundaries. Neighborhoods are labeled with *hx* with white boundaries.

Figure 5. Neighborhood (hx) spatial congruence boundaries (white), school (hx) boundaries (black).