Segregated by Design? The Effect of Street Network Topological Structure on the Measurement of Urban Segregation*

ELIJAH KNAAP, San Diego State University
SERGIO REY, San Diego State University

Racial residential segregation is a longstanding topic of focus across the disciplines of urban social science. Classically, segregation indices are calculated based on areal groupings (e.g. counties or census tracts), with more recent research exploring ways that spatial relationships can enter the equation. Spatial segregation measures embody the notion that proximity to one's neighbors is a better specification of residential segregation than simply who resides together inside the same arbitrarily-drawn polygon. Thus, they expand the notion of "who is nearby" to include those who are geographically close to each polygon rather than a binary inside/outside distinction. Yet spatial segregation measures often resort to crude measurements of proximity, such as the euclidean distance between observations, given the complexity and data requirements of calculating more theoretically-appropriate measures, such as distance along the pedestrian travel network. In this paper, we examine the ramifications of such decisions. For each metropolitan region in the U.S., we compute both Euclidean and network-based spatial segregation indices. We use a novel inferential framework to examine the statistical significance of the difference between the two measures and following, we use features of the network topology (e.g. connectivity, circuity, throughput) to explain this difference using a series of regression models. We show that there is often a large difference between segregation indices when measured by these two strategies (which is frequently significant). Further, we explain which topology measures reduce the observed gap and discuss implications for urban planning and design paradigms

Keywords: segregation, neighborhoods, spatial analysis, network analysis, spatial weights

INTRODUCTION

An exceedingly common abstraction in applied spatial analysis is the use of euclidean distance as a proxy measure for geographic proximity (which is, itself, often a proxy for the frequency of social interaction). It is the geographical science equivalent of physics' spherical cow,[1] or the economists' perfect market: a useful abstraction that helps partially explain a much more complex underlying process, however imperfectly. A major difference in spatial analysis, however, is that scientists from many disciplines often fail to realize how simplified the assumption of euclidean distance is when traversing the built or natural environment. While, in general, simple proximity is a reasonable heuristic for understanding Tobler's Law (Tobler, 1970), the behavioral realities of movement and social interaction in complex urban environments often require a more thoughtful model.

More directly, cities, regions, and neighborhoods are not featureless planes in which agents have perfect freedom of mobility. Rather, they are multifaceted environments populated by highways, canyons, rivers, mountains, railroad tracks, alleyways, and power plants. To facilitate movement in this environment, an interleaved transportation system provides passageways through discrete locations, and conditions how easy it is to move throughout the region and interact with individuals in other parts of the region. Although pure euclidean distance can proxy this system, the urban design

*This work is supported by NSF-SES Grant 1831615
[1] https://en.wikipedia.org/wiki/Spherical_cow

decisions that govern how and where networks are located, as well as the natural features like elevation or water features play an important albeit underexamined role in mediating social interactions.

One particular topic where a full understanding of space would provide significant benefits is segregation analysis. Racial residential segregation is a longstanding topic of focus across the disciplines of urban social science. Classically, segregation indices are calculated based on areal groupings (e.g. counties or census tracts), with more recent research exploring ways that spatial relationships can enter the equation. Spatial segregation measures embody the notion that proximity to one's neighbors is a better specification of residential segregation than simply who resides together inside the same arbitrarily-drawn polygon. Thus, they expand the notion of "who is nearby" to include those who are geographically close to each polygon rather than a binary inside/outside distinction. Yet spatial segregation measures often resort to crude measurements of proximity, such as the euclidean distance between observations, given the complexity and data requirements of calculating more theoretically-appropriate measures, such as distance along the pedestrian travel network.

In this project, we examine the relationship between pedestrian network characteristics and the measurement of metropolitan segregation. In doing so, the paper examines three research questions, in turn: First, how much does the operationalization of space matter for segregation measurement? More specifically, how large is the difference between euclidean-based and network-based measures of spatial segregation? Second, if reasonable differences exist between euclidean and network measures, are they large enough that they cannot be attributed to chance? Third, what characteristics of the pedestrian network explain the observed difference in measurement? If there is a large and/or systematic difference between traditional spatial measurements and those leveraging more realistic measurements of distance, then there may be much to learn about the contribution of network structure and design when seeking to reduce segregation in cities.

Since the inception of city planning, the relationship between social interactions and the built environment has been a topic of intense focus for both social scientists and urban designers (Talen, 2017). The normative concepts of urban utopias prescribed by architects like Ebeneezer Howard, Frank Lloyd Wright, and Le Corbusier included distinct visions for how densely populated and separated/integrated land uses could facilitate the ideal level of interaction between (a) a resident and her neighbors, and (b) her natural surroundings (Campbell & Fainstein, 1996; Corbusier, 1986; Howard & Osborn, 2001).

Olsen (2003);Wirth (1938)]

how do we represent space in social science research?

classics: - sociology uses groups. Neighborhoods or cities are discrete containers that condition social behaviors (Park, Burgess, McKenzie) - econ and regional science use distance from the city center - ultimately about transport of goods (von thunen was based on an *agricultural economy* and moving crops from the ag hinterlands into the marketplace where people actually lived). - Transport connectivity is implicit, but models are high-level in the 1950s, and the abstraction works conceptually. Neither the theory or computational power exist yet to examine the role of better measurements of W_{ij}

recents: - GIS, geography, and spatial econometrics concepts of spatial weights - multiscale and/or bespoke neighborhoods in geography and sociology (HIPP & BOESSEN (2013), van Ham,) - street networks in empirical work - grannis shows social interactions are more frequent inside T-communities defined by street networks - Roberto uses street networks to measure segregation in a small-scale case study.

Now we have both the tools and the logic to test these assumptions and understand the role of abstractions such as euclidean distance-based measures in our assessment of critical social processes such as residential segregation. Fast graph algorithms allow us to construct more realistic concepts of spatial weights matrices, and computational statistics allow us to construct and test realistic null hypotheses about the allocation of urban population groups. Here, we examine the role of street network topology in the appropriate measurement of urban segregation. Our goals are twofold. First, we aim to understand the implications of simple Euclidean distance- based abstractions when conducting formal spatial analyses; that is, do we find substantive differences in results when more realistic concepts of spatial relationships (e.g. network connectivity) are considered? Second, we aim to explore the elements of urban design (particularly the street network configuration) in widening the gap between analytical abstraction and empirical reality. More simply, we aim to understand whether certain elements of the street network are associated with a greater difference in measured segregation. With this knowledge, urban designers and planners can begin with more inclusive communities from the beginning.

URBAN INFRASTRUCTURE AND SOCIAL INTERACTIONS

Classically, space is treated as a discrete concept, by membership in a group (i.e. a school, classroom, neighborhood, or city), where any of these groupings is defined exogenously.

The Role of Space in Segregation Measurement

Making Space Explicit

Reardon & O'Sullivan (2004) Reardon et al. (2009) Wong (1997) Bailey (2012) Rey & Folch (2011) O'Sullivan & Wong (2007) Wong (2004) Dawkins (2004)

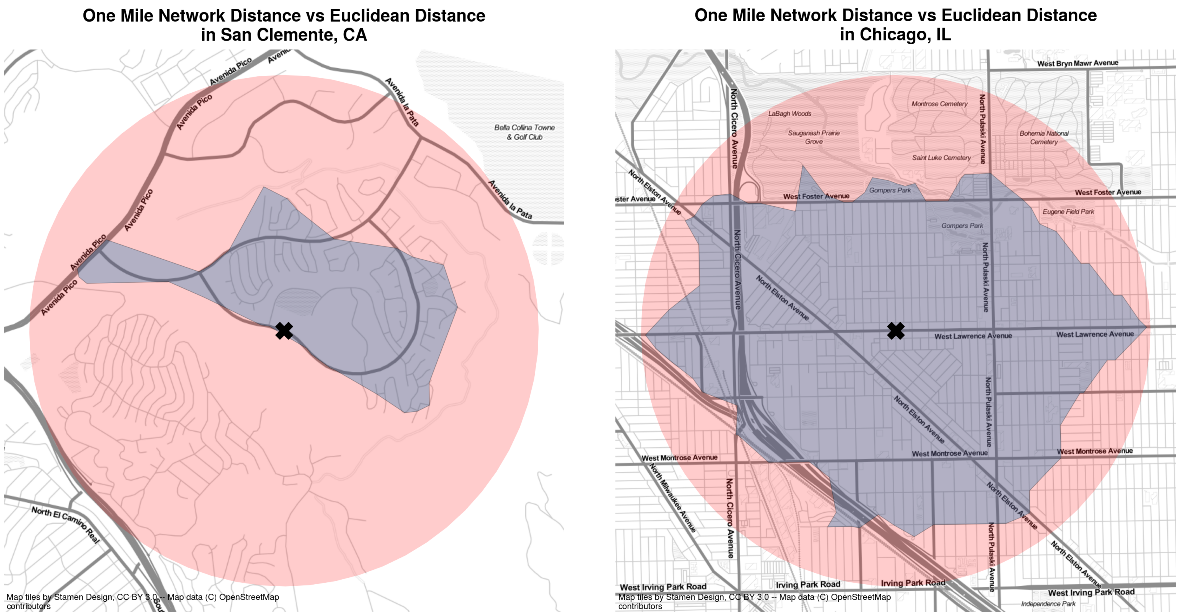
Interrogating Spatial Scale

Lee et al. (2008) Reardon et al. (2008) Bézenac et al. (2022) Olteanu et al. (2019) Östh et al. (2015) Clark et al. (2015)

Homogenous Communities and Urban Design

A long-recognized but understudied element of metropolitan segregation patterns is the role of transport networks, physical barriers, and other factors such as elevation or congestion that change travel behavior, and thus, the expected potential for social interaction in space. For example work in sociology has shown the importance of street network connectivity in fostering connected social networks inside small urban geographic zones (Grannis, 1998).

Grannis (2005)



(a) Network Distance vs Euclidean Distance

(b) Network Distance vs Euclidean Distance

Figure 1: Network Distance vs Euclidean Distance in Urban Environments

A depiction of the difference between network travel distance and “as the crow flies” distance is shown in Figure 1. The figure shows an origin marked with an X in the center, and two different polygons representing a one-mile travel distance using different methods. The small polygon depicts the total extent accessible from the origin point when traveling along the pedestrian network, whereas the larger polygon depicts the 1-mile buffer representing unconstrained travel. It is immediately apparent in the figure that network-constrained travel covers a much smaller footprint than euclidean distance in the depicted location. Furthermore, the pattern appears to be influenced strongly by the street network and urban design features that characterize the largely suburban region of the San Diego metro area. Instead of a regular grid that facilitates travel in all directions, the street network in Figure 1 includes several insular patterns, cul-de-sacs, and 3-way intersections that help channel traffic in certain directions rather than others. Furthermore, the fact that some subdivisions have only a single entrance makes clear how much further a person would need to travel to reach the homes in certain regions (versus how much easier they appear to be reached via the circular buffer).

Recent work by Roberto (2018) shows the importance of considering network distances when measuring segregation using both simulated data and an empirical example in Pittsburgh, PA. That study shows that segregation is consistently higher at all spatial scales when the measure accounts for local network connectivity. As Roberto (2018, p. 28) notes, “even small positive differences in the city-level results are meaningful and suggest that physical barriers facilitate greater separation between ethno-racial groups and higher levels of segregation." We agree with this assessment and in what follows, we examine the magnitude of differences between network and simple euclidean measures in detail for every metropolitan region in the United States. Specifically, we expand upon prior work in three different directions. First, we expand the geographic scope by considering every metropolitan region in the United States, rather than a case study of a single city. Second, we adopt a computational inference framework that allows us to assess whether the observed differences between the segregation measures are large enough that they could not happen by chance. Finally, we explore the relationship between differences in observed segregation and characteristics of the local travel network.

THE ROLE OF NETWORK DISTANCE IN SEGREGATION MEASUREMENT

We begin our analysis by computing two sets of segregation indices, adopting the spatial information theory statistic $\tilde{H}$ as our measure of segregation. As Reardon et al. (2008, p. 512) describe, "the index $\tilde{H}$ is a measure of how much less diverse individuals' local environments are, on average, than is the total population of region," and reaches its maximum of 1 only when "each individual's local environment is monoracial." Here, our goal is to test how sensitive the statistic is to different concepts of the "local environment," with one concept adopting the simplified assumption of euclidean-based distance measurements, and the other requiring that distance be measured along a pedestrian transport network.

Measuring Segregation in Space

Following Reardon & O'Sullivan (2004) we consider a spatial region R populated by M racial groups indexed by m, with τ and π as population density and proportion, respectively. Here we diverge from the classical notation in the segregation literature and instead adopt conventions more common in spatial econometrics and geographic analysis. Doing so allows us to strengthen the connection between similar concepts in different disciplines as well as gain finer control over the definition of spatial relationships. Since many spatial segregation measures are implemented in GIS and spatial analysis software designed by geographers, clarifying this connection can help ease interdisciplinary adoption and conversation around spatial segregation measures.

Thus, we index locations within R as i and j, and we operationalize the concept of spatial relationships using a spatial weights matrix W_{ij}. By focusing on W_{ij}, we are forced "to specify [our] underlying assumptions about socio-spatial proximity," following the call by Reardon & O'Sullivan (2004, p. 154) for analysis that "compares segregation levels based on different theoretical bases for defining spatial proximity." Conceptually, the spatial weights matrix W_{ij} is connectivity graph that defines the spatial relationship between nodes i and j, and the values w_{ij} encode the intensity of the edge $\bar{ij}$. Thus, the spatial weights matrix is a useful and flexible representation of the local neighborhood environment because it provides a generic data structure for encoding spatial relationships, where any link function (ϕ, following the notation of Reardon & O'Sullivan (2004)) can be used to specify the proximity between units. Formally,

$$W_{ij} = \phi(D_{ij}) \tag{1}$$

Where ϕ is a proximity weighting function and D is a matrix containing pairwise distances for i and j. Classically, W_{ij} is typically created via binary connectivity between adjacent units, but a wide variety of other continuous specifications are also used in practice (Getis, 2009; Halleck Vega & Elhorst, 2015; Rey & Anselin, 2010), such as the euclidean distance between observations, or various kernel or distance-decay functions. Critically, the distance-weighting function ϕ is distinct from the concept of *distance* (D), itself, which could be measured in Euclidean/geodesic distance, minutes of congested travel time, meters traveled along the sidewalk, or some generalized measure of utility. Separating these two concepts allows us to consider alternative distance metrics distinctly from alternative decay functions. The local environment for a given feature y at location i can then be measured by its *spatial lag*, SL, defined as

$$SL_i = \sum_j w_{ij} y_j \tag{2}$$

In the spatial econometrics literature, it is common to exclude the diagonal elements from W_{ij} to differentiate between focal effects and spatial spillovers in regression models, but when the diagonal is filled, then SL_i becomes a consummate measure of the local environment at location i.

To compute the spatial multigroup information theory index $\tilde{H}$, we first calculate local spatially-weighted population proportions as

$$\tilde{\pi}_{im} = \frac{SL_{im}}{\sum_{m=1}^{M} SL_{im}} \tag{3}$$

The density at location i is

$$\tilde{\tau}_i = \frac{\sum_{m=1}^{M} SL_{im}}{\sum_{m=1}^{M} \sum_{i=1}^{I} SL_{im}} \tag{4}$$

The entropy of the local environment at each location $\tilde{E}_i$ is

$$\tilde{E}_i = -\sum_{m=1}^{M} (\tilde{\pi}_{im}) \log_M (\tilde{\pi}_{im}) \tag{5}$$

where M indicates the number of groups in the population. Finally,

$$\tilde{H} = 1 - \frac{1}{TE} \sum^{I} \tilde{\tau}_i \tilde{E}_i \tag{6}$$

where $\tilde{H}$ is the spatial information theory index defined by Reardon & O'Sullivan (2004). We perform all calculations using the open-source Python package `segregation` (Cortes et al., 2020), distributed as part of the Python Spatial Analysis Library (PySAL) (Rey, Anselin, et al., 2021)

Assessing Difference Between Distance Metrics

To understand the implications of different parameterizations of space, we use data blockgroup-level from the US Census American Community Survey (ACS) 5-year sample (2013-2017) with four mutually-exclusive racial groups (non-Hispanic white, non-Hispanic Black, Hispanic, and Asian). Our sample contains data for 380 metropolitan Core Based Statistical Areas (CBSAs) in the United States. Blockgroups are the smallest geographic unit for which racial and ethnic data are available in the ACS. To compute euclidean-based spatial segregation measures, our distances are measured between blockgroup centroids; to compute network-based spatial segregation measures, we first attach the blockgroup centroids to the nearest intersection in the travel network, then compute the shortest network-based path between each pair of observations

Our data on street networks is collected from OpenStreetMap and the shortest network path is computed using the Python package `pandana` (Foti et al., 2012). To operate efficiently on metropolitan-scale street networks, the pandana package relies on a graph pre-processing technique known as contraction hierarchies that simplifies the computation by removing inconsequential nodes from consideration during the routing algorithm.

Constructing Comparable Indices

In each metropolitan region, we proceed by creating two different spatial weights matrices by varying the way distance is measured between observations. In both matrices, the proximity-weighting function ϕ is a simple linear decay (triangular kernel) encoding a spatial weight that decreases with distance up to a threshold of two kilometers, outside of which observations no longer have an effect, (that is, $r = 2000$):

$$\phi = \begin{cases} 1 - \left(\frac{d_{ij}}{r}\right), & \text{if } d_{ij} \leq r \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

Between the two W matrices, however, we vary the input distance matrix D, between two concepts, euclidean distance and network distance (where network distance is defined as the shortest path along the pedestrian transportation network), W_{net}, and W_{euc}. In both matrices the diagonal is set to one, indicating that there is no spatial discount for the value located at observation i. Using these weights matrices W_{net} and W_{euc} to build local environments for each metropolitan region in Equation 1 propagates the two constructs through Equations 2, 3, 4, 5, 6, yielding two segregation measures $\tilde{H}_{net}, \tilde{H}_{euc}$ and, implicitly, a difference between the two, $\Delta_{\tilde{H}} = \tilde{H}_{net} - \tilde{H}_{euc}$. The relative difference between segregation measures is the difference divided by the euclidean measure $\Delta_{pct} = \frac{\Delta_{\tilde{H}}}{\tilde{H}_{euc}}$

Inferential Framework

We assess the importance of considering network distance in segregation measurement by adopting the inferential framework outlined in Rey, Cortes, et al. (2021) and Cortes et al. (2020). The approach leverages a computational approach to statistical inference using random labelling to compare the observed difference between the two segregation measures (network versus euclidean) to a distribution of differences generated from the same data. More specifically, the measures $\tilde{H}_{net}$, $\tilde{H}_{euc}$ and $\Delta_{\tilde{H}}$ are computed and recorded for each metro region. As a result of this process, two "spatialized" versions of the metropolitan demographic composition are created, with one dataset representing euclidean distances and the other representing network-based distances.

We then create two synthetic datasets by pooling the input units from both original datasets and reassigning them at random. For each block-group, we randomly reassign the labels (net, euc) to the observed spatial lags from Equation 2. Once all units have been assigned to a group, the segregation measures are re-computed and their difference taken. This process is repeated 10,000 iterations. By comparing the observed difference between the two segregation measures against a distribution of differences generated via synthetic datasets, we can develop pseudo p-values based on a standard T-test. Our test in this case measures the empirical likelihood of obtaining the observed difference at random under the null hypothesis that the observed difference is within the standard range of differences[2]. The pseudo-p values represent probability of obtaining results in which the simulated difference was greater than the observed difference $\Delta_{\tilde{H}}$.

Network Distance is an Important Consideration

Although the correlation between planar and network based segregation measures is $\rho = 0.987$, our results provide clear evidence that the choice of appropriate distance metric plays an important role in the computation of a spatial segregation index. In all but four cases, we show that segregation is higher when measured according to network distance than by pure euclidean distance[3] (none of the four cases are significant different from a random pooling of the same data). Among the 380 CBAs in our dataset, 25.3% have a difference between euclidean and network-based segregation measures that is signficant at the $\alpha = 0.05$ level, and 14.2% of the CBSAs are significant at the $\alpha = 0.01$ level. Descriptive statistics of the differences between segregation measures in each metro are shown in

[2]Note this does not explicitly require the null $\Delta_{\tilde{H}} = 0$. Instead the "null value" is the mean of the simulated parameter distribution.

[3]For each CBSA in our sample, our euclidean distances are based on UTM coordinate systems, with each region's data projected into its appropriate UTM zone.

Table 1: Descriptive Statistics for Segregation Differences

	$\tilde{H}_{net}$	$\tilde{H}_{euc}$	$\Delta_{\tilde{H}}$	Δ_{pct}
count	380.000	380.000	380.000	380.000
mean	0.207	0.178	0.029	0.198
std	0.078	0.077	0.013	0.113
min	0.070	0.051	-0.053	-0.204
25%	0.141	0.114	0.023	0.118
50%	0.205	0.172	0.029	0.184
75%	0.254	0.224	0.036	0.260
max	0.489	0.454	0.077	0.694

Table **??**, and a list of the 54 CBSAs significant at the one percent level are listed in Table **??**. Among these 54 CBAS, eight metros are located in California–twice the number of the next-most prevalent state (Texas)

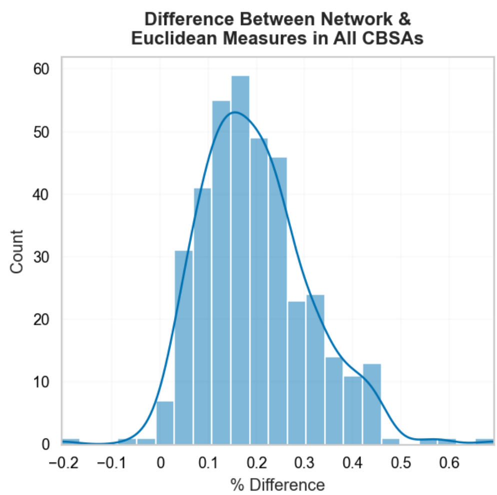


Figure 2: Histogram of % Differences in Segregation Measures

The shape of distribution of differences is approximately normal. While the absolute difference between the two segregation measures in each CBSA can appear small, the relative difference is often reasonably large, with the network-based segregation measure approximately 20% higher than the euclidean-based measure on average. The largest relative difference gets as high as 69% (Carson City, NV).

Table 2: CBSAs with Highly Significant $\Delta_{\tilde{H}}$

name	$\tilde{H}_{net}$	$\tilde{H}_{euc}$	$\Delta_{\tilde{H}}$	Δ_{pct}	pseudo-p
Anchorage, AK	0.135	0.092	0.043	0.470	0.002
Atlanta-Sandy Springs-Alpharetta, GA	0.321	0.293	0.028	0.095	0.001
Austin-Round Rock-Georgetown, TX	0.174	0.152	0.023	0.150	0.006
Baltimore-Columbia-Towson, MD	0.331	0.284	0.047	0.164	0.000
Boston-Cambridge-Newton, MA-NH	0.254	0.228	0.025	0.111	0.001
Bridgeport-Stamford-Norwalk, CT	0.223	0.181	0.042	0.235	0.001
Charlotte-Concord-Gastonia, NC-SC	0.264	0.233	0.032	0.136	0.000
Chicago-Naperville-Elgin, IL-IN-WI	0.386	0.365	0.021	0.057	0.000
Cincinnati, OH-KY-IN	0.315	0.262	0.054	0.205	0.000
Cleveland-Elyria, OH	0.412	0.380	0.033	0.086	0.004
Columbus, OH	0.273	0.234	0.039	0.168	0.001
Dallas-Fort Worth-Arlington, TX	0.255	0.218	0.037	0.171	0.000
Denver-Aurora-Lakewood, CO	0.197	0.176	0.021	0.120	0.002
Detroit-Warren-Dearborn, MI	0.489	0.450	0.038	0.085	0.000
Hartford-East Hartford-Middletown, CT	0.313	0.264	0.048	0.182	0.001
Houston-The Woodlands-Sugar Land, TX	0.270	0.243	0.028	0.114	0.000
Indianapolis-Carmel-Anderson, IN	0.314	0.279	0.034	0.123	0.008
Kansas City, MO-KS	0.294	0.266	0.028	0.104	0.007
Lansing-East Lansing, MI	0.210	0.166	0.044	0.268	0.003
Las Vegas-Henderson-Paradise, NV	0.139	0.115	0.023	0.202	0.000
Los Angeles-Long Beach-Anaheim, CA	0.284	0.264	0.019	0.072	0.000
Louisville/Jefferson County, KY-IN	0.328	0.269	0.060	0.222	0.000
Miami-Fort Lauderdale-Pompano Beach, FL	0.354	0.323	0.031	0.097	0.000
Milwaukee-Waukesha, WI	0.429	0.398	0.031	0.079	0.008
Minneapolis-St. Paul-Bloomington, MN-WI	0.209	0.177	0.032	0.182	0.000
New Haven-Milford, CT	0.225	0.183	0.042	0.230	0.001
New Orleans-Metairie, LA	0.260	0.227	0.032	0.142	0.009
New York-Newark-Jersey City, NY-NJ-PA	0.299	0.280	0.019	0.067	0.000
Oklahoma City, OK	0.253	0.219	0.034	0.155	0.006
Olympia-Lacey-Tumwater, WA	0.112	0.077	0.035	0.456	0.004
Peoria, IL	0.351	0.288	0.063	0.218	0.004
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.326	0.302	0.024	0.080	0.000
Phoenix-Mesa-Chandler, AZ	0.209	0.182	0.027	0.149	0.000
Pittsburgh, PA	0.280	0.225	0.055	0.245	0.000
Portland-Vancouver-Hillsboro, OR-WA	0.118	0.095	0.024	0.249	0.001
Redding, CA	0.120	0.084	0.036	0.431	0.004
Riverside-San Bernardino-Ontario, CA	0.169	0.144	0.025	0.171	0.000
Rochester, NY	0.316	0.274	0.042	0.152	0.003
Sacramento-Roseville-Folsom, CA	0.152	0.129	0.023	0.177	0.000
Salt Lake City, UT	0.146	0.121	0.026	0.212	0.007
San Antonio-New Braunfels, TX	0.231	0.200	0.031	0.154	0.000
San Diego-Chula Vista-Carlsbad, CA	0.204	0.184	0.019	0.105	0.009
San Francisco-Oakland-Berkeley, CA	0.167	0.154	0.014	0.089	0.008
Santa Rosa-Petaluma, CA	0.118	0.088	0.030	0.342	0.005
Seattle-Tacoma-Bellevue, WA	0.130	0.107	0.023	0.217	0.000
Springfield, MA	0.312	0.256	0.056	0.217	0.000
St. Louis, MO-IL	0.390	0.356	0.034	0.095	0.005
Stockton, CA	0.132	0.104	0.027	0.262	0.009
Tallahassee, FL	0.196	0.146	0.050	0.346	0.005
Tampa-St. Petersburg-Clearwater, FL	0.239	0.213	0.026	0.124	0.002
Toledo, OH	0.268	0.218	0.050	0.231	0.001
Virginia Beach-Norfolk-Newport News, VA-NC	0.206	0.161	0.045	0.277	0.000
Washington-Arlington-Alexandria, DC-VA-MD-WV	0.240	0.220	0.020	0.092	0.001
Worcester, MA-CT	0.200	0.156	0.044	0.281	0.001

NETWORK CHARACTERISTICS AND SEGREGATION DIFFERENCES

Metropolitan Travel Infrastructure as a Network Graph

The travel infrastructure in a metropolitan region serves as its skeleton for both urban development and social interactions. For decades, scholars have worked to quantify the aspects of urban form that help explain behaviors such as travel mode choice Ewing & Cervero (2010). A recent evolution of this work is the conception of a travel network as a formal graph structure (Boeing, 2018a, 2018b; Fleischmann, 2018; Fleischmann et al., 2021), and a set of software tools that facilitate its analysis (Boeing, 2016; Fleischmann, 2019).

Measuring Graph Structure

We use OSMNx and Momepy to create measures of the pedestrian travel network collected from OpenStreetMap.

Graph Topology and Segregation Differences

To understand how urban design decisions such as the topology of the travel network may impact the ability for residents to interact (as measured by the segregation index), we regress the difference in measured segregation on measures of the network graph structure.

Our two-value test is doing a good job (i.e., it is picking up a difference)

The ps_inter is an interaction term between the planar_measure and whether the two-value test was significant. This tells us that the slope is larger for those cities where the difference in the two-value test is significant.

The pct difference generally declines with the overall level of segregation and network size (as measured by street_length) althoug the latter association appears to be driven by the places with the significant two-value tests

DISCUSSION

There are two additional parameters worth exploring: the distance-decay function ϕ, and the radius that defines the extent of the local environment r.

CONCLUSION

In the segregation literature, the importance of *space* has long been recognized, but a full grasp of its implications still eludes researchers. In this paper, we show that when considering the role of transportation infrastructure in segregation measurement, we obtain substantially different results than classic spatial approaches that adopt euclidean measurements.

In future work, this research could be extended in several directions

One promising direction is the consideration of alternative impedance measures when calculating shortest-path distances along the travel network. In the present study, we assume a constant rate of travel consistent with the average walking pace, and that impedance is reflected by graph distance alone. Alternative constructs could include elevation along with distance to get a more complete measure of the effort required to traverse by foot or bicycle. Similarly, the travel network could also be extended to include public transportation or (potentially congested) automobile travel. These considerations would require extensive additional data, which may limit the capacity for cross-sectional comparisons, but would also provide insight into alternative concepts of space and distance.

Another important avenue for further work is the blending of multiple graphs for a more complete understanding of multi-contextual segregation. For example children who live in a given neighborhood are simultaneously embedded in local neighborhood contexts, school catchment boundaries, and other local institutions such as religious and community organizations. Each of these contexts have partially-overlapping, occasionally nested, and often imperfectly-defined geographic boundaries, a full synthesis of which requires the development of new methods that integrate across these contexts (Galster, 2001; Galster, 2019). As one example, Wolf (2021) provides a technique for blending multiple graphs together, one spatial and one aspatial, and similar methods could be possibly used to integrate multiple contexts. Work along these lines would also help address the call by Reardon & O'Sullivan (2004, p. 156) for metrics that help understand bridges across social networks

REFERENCES

Bailey, N. (2012). How Spatial Segregation Changes over Time: Sorting Out the Sorting Processes. *Environment and Planning A: Economy and Space*, *44*(3), 705–722. https://doi.org/10.1068/a44330

Bézenac, C., Clark, W. A. V., Olteanu, M., & Randon-Furling, J. (2022). Measuring and Visualizing Patterns of Ethnic Concentration: The Role of Distortion Coefficients. *Geographical Analysis*, *54*(1), 173–196. https://doi.org/10.1111/gean.12271

Boeing, G. (2016). OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks. *Computers, Environment and Urban Systems*, *65*, 126–139. https://doi.org/10.1016/j.compenvurbsys.2017.05.004

Boeing, G. (2018a). Planarity and street network representation in urban form analysis. *Environment and Planning B: Urban Analytics and City Science*, 239980831880294. https://doi.org/10.1177/2399808318802941

Boeing, G. (2018b). The Morphology and Circuity of Walkable and Drivable Street Networks. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3119939

Campbell, S., & Fainstein, S. S. (1996). Readings in Planning Theory. In *Building* (Vol. 2nd). https://doi.org/10.1080/14649357.2012.669994

Clark, W. A. V., Anderson, E., Östh, J., & Malmberg, B. (2015). A Multiscalar Analysis of Neighborhood Composition in Los Angeles, 2000–2010: A Location-Based Approach to Segregation and Diversity. *Annals of the Association of American Geographers*, *105*(6), 1260–1284. https://doi.org/10.1080/00045608.2015.1072790

Clifton, K., Ewing, R., Knaap, G.-J., & Song, Y. (2008). Quantitative analysis of urban form: A multidisciplinary review. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, *1*(1), 17–45. https://doi.org/10.1080/17549170801903496

Corbusier, L. (1986). *Towards a new architecture*. Dover Publications.

Cortes, R. X., Rey, S., Knaap, E., & Wolf, L. J. (2020). An open-source framework for non-spatial and spatial segregation measures: The PySAL segregation module. *Journal of Computational Social Science*, *3*(1), 135–166. https://doi.org/10.1007/s42001-019-00059-3

Crane, R. (2000). The Influence of Urban Form on Travel: An Interpretive Review. *Journal of Planning Literature*, *15*(1), 3–23. https://doi.org/10.1177/08854120022092890

Dawkins, C. J. (2004). Measuring the {Spatial} {Pattern} of {Residential} {Segregation}. *Urban Studies*, *41*(4), 833. https://doi.org/10.1080/0042098042000194133

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment. *Journal of the American Planning Association*, *76*(3), 265–294. https://doi.org/10.1080/01944361003766766

Ewing, R., & Handy, S. (2009). Measuring the Unmeasurable: Urban Design Qualities Related to Walkability. *Journal of Urban Design*, *14*(1), 65–84. https://doi.org/10.1080/13574800802451155

Fleischmann, M. (2018). Measuring Urban Form: Overcoming Terminological Inconsistencies for a Quantitative and Comprehensive Morphologic Analysis of Cities. *Environment and Planning B: Urban Analytics and City Science 1*, *0*(0), 1–19. https://doi.org/10.1177/2399808320910444

Fleischmann, M. (2019). Momepy: Urban Morphology Measuring Toolkit. *Journal of Open Source Software*, *4*(43), 1807. https://doi.org/10.21105/joss.01807

Fleischmann, M., Feliciotti, A., Romice, O., & Porta, S. (2021). Methodological foundation of a numerical taxonomy of urban form. *Environment and Planning B: Urban Analytics and City Science*, *0*(0), 239980832110598. https://doi.org/10.1177/23998083211059835

Foti, F., Waddell, P., & Luxen, D. (2012). A Generalized Computational Framework for Accessibility : From the Pedestrian to the Metropolitan Scale. *4th Transportation Research Board Conference on Innovations in Travel Modeling (ITM)*, 1–14.

Galster, G. (2001). On the nature of neighbourhood. *Urban Studies*, *38*(12), 2111. https://doi.org/10.1080/00420980120087072

Galster, G. C. (2019). *Making our neighborhoods, making our selves*. University of Chicago Press.

Getis, A. (2009). Spatial weights matrices. *Geographical Analysis*, *41*(4), 404–410. https://doi.org/10.1111/j.1538-4632.2009.00768.x

Grannis, R. (1998). The Importance of Trivial Streets: Residential Streets and Residential Segregation. *American Journal of Sociology*, *103*(6), 1530–1564. https://doi.org/10.1086/231400

Grannis, R. (2005). T-Communities: Pedestrian Street Networks and Residential Segregation in Chicago, Los Angeles, and New York. *City and Community*, *4*(3), 295–321. https://doi.org/10.1111/j.1540-6040.2005.00118.x

Halleck Vega, S., & Elhorst, J. P. (2015). THE SLX MODEL. *Journal of Regional Science*, *55*(3), 339–363. https://doi.org/10.1111/jors.12188

HIPP, J. R., & BOESSEN, A. (2013). EGOHOODS AS WAVES WASHING ACROSS THE CITY: A NEW MEASURE OF "NEIGHBORHOODS." *Criminology*, *51*(2), 287–327. https://doi.org/10.1111/1745-9125.12006

Howard, E., & Osborn, F. J. (2001). *Garden cities of to-morrow* (11. print). M.I.T. Pr.

Lee, B. A., Reardon, S. F., Firebaugh, G., Farrell, C. R., Matthews, S. A., & O'Sullivan, D. (2008). Beyond the Census Tract: Patterns and Determinants of Racial Segregation at Multiple Geographic Scales. *American Sociological Review*, *73*(5), 766–791. https://doi.org/10.1177/000312240807300504

O'Sullivan, D., & Wong, D. W. S. (2007). A Surface-Based Approach to Measuring Spatial Segregation. *Geographical Analysis*, *39*(2), 147–168. https://doi.org/10.1111/j.1538-4632.2007.00699.x

Olsen, J. (2003). *Better places, better lives: A biography of James Rouse*. ULI- the Urban Land Institute.

Olteanu, M., Randon-Furling, J., & Clark, W. A. V. (2019). Segregation through the multiscalar lens. *Proceedings of the National Academy of Sciences*, *116*(25), 12250–12254. https://doi.org/10.1073/pnas.1900192116

Östh, J., Clark, W. A. V., & Malmberg, B. (2015). Measuring the Scale of Segregation Using k -Nearest Neighbor Aggregates. *Geographical Analysis*, *47*(1), 34–49. https://doi.org/10.1111/gean.12053

Reardon, S. F., Farrell, C. R., Matthews, S. A., O'Sullivan, D., Bischoff, K., & Firebaugh, G. (2009). Race and space in the 1990s: Changes in the geographic scale of racial residential segregation, 1990–2000. *Social Science Research*, *38*(1), 55–70. https://doi.org/10.1016/j.ssresearch.2008.10.002

Reardon, S. F., Matthews, S. A., O'Sullivan, D., Lee, B. A., Firebaugh, G., Farrell, C. R., & Bischoff, K. (2008). The geographic scale of Metropolitan racial segregation. *Demography*, *45*(3), 489–514. http

s://doi.org/10.1353/dem.0.0019
Reardon, S. F., & O'Sullivan, D. (2004). Measures of Spatial Segregation. *Sociological Methodology*, *34*(1), 121–162. https://doi.org/10.1111/j.0081-1750.2004.00150.x
Rey, S. J., & Anselin, L. (2010). PySAL: A Python Library of Spatial Analytical Methods. In *Handbook of Applied Spatial Analysis* (Vol. 37, pp. 175–193). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-03647-7_11
Rey, S. J., Anselin, L., Amaral, P., Arribas-Bel, D., Cortes, R. X., Gaboardi, J. D., Kang, W., Knaap, E., Li, Z., Lumnitz, S., Oshan, T. M., Shao, H., & Wolf, L. J. (2021). The PySAL Ecosystem: Philosophy and Implementation. *Geographical Analysis*, gean.12276. https://doi.org/10.1111/gean.12276
Rey, S. J., Cortes, R., & Knaap, E. (2021). Comparative Spatial Segregation Analytics. *Spatial Demography*, *January*, 1–34. https://doi.org/10.1007/s40980-021-00075-w
Rey, S. J., & Folch, D. C. (2011). Impact of spatial effects on income segregation indices. *Computers, Environment and Urban Systems*, *35*(6), 431–441. https://doi.org/10.1016/j.compenvurbsys.2011.07.008
Roberto, E. (2018). The Spatial Proximity and Connectivity Method for Measuring and Analyzing Residential Segregation. *Sociological Methodology*, *48*(1), 008117501879687. https://doi.org/10.1177/0081175018796871
Talen, E. (2017). Social science and the planned neighbourhood. *Town Planning Review*, *88*(3), 349–372. https://doi.org/10.3828/tpr.2017.22
Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, *46*(May), 234. https://doi.org/10.2307/143141
Wirth, L. (1938). Urbanism as a Way of Life. *American Journal of Sociology*, *44*(1), 1–24. https://doi.org/10.1086/217913
Wolf, L. J. (2021). Spatially–encouraged spectral clustering: A technique for blending map typologies and regionalization. *International Journal of Geographical Information Science*, *35*(11), 2356–2373. https://doi.org/10.1080/13658816.2021.1934475
Wong, D. S. W. S. (1997). Spatial dependency of segregation indices. *Canadian Geographer*, *41*(2), 128–136. https://doi.org/10.1111/j.1541-0064.1997.tb01153.x
Wong, D. W. S. (2004). Comparing Traditional and Spatial Segregation Measures: A Spatial Scale Perspective. *Urban Geography*, *25*(1), 66–82. https://doi.org/10.2747/0272-3638.25.1.66