

Human interaction networks reveal that large cities facilitate segregation

Hamed Nilforoshan^{*,1}, Wenli Looi^{*,1}, Emma Pierson^{*,2}, Blanca Villanueva³, Nic Fishman¹, Yiling Chen¹, John Sholar¹, Beth Redbird^{4,5}, David Grusky⁶, Jure Leskovec^{†,1}

¹ Department of Computer Science, Stanford University, Stanford, CA 94305, USA

² Department of Computer Science, Cornell Tech, New York, NY 10044, USA

³ Department of Biomedical Data Science, Stanford University, Stanford, CA 94305, USA

⁴ Institute for Policy Research, Northwestern University, Evanston, IL 60208, USA

⁵ Department of Sociology, Northwestern University, Evanston, IL 60208, USA

⁶ Department of Sociology, Stanford University, Stanford, CA 94305, USA

* These authors contributed equally to this work

† Corresponding author. Email: jure@cs.stanford.edu

A long-standing expectation is that large, dense, and cosmopolitan areas will support diverse interactions and socioeconomic mixing^{1–6}. It has been difficult to assess this hypothesis because past approaches to measuring socioeconomic mixing have relied on static residential housing data rather than real-life interactions among people meeting at work, in places of leisure, and in home neighborhoods^{7,8}. Here we develop a new measure of interaction segregation (IS) that captures the economic diversity of the set of people that a given person meets in their everyday life. Leveraging cell phone mobility data to represent 1.6 billion interactions among 9.6 million people in the United States, we measure interaction segregation across 382 Metropolitan Statistical Areas (MSAs) and 2829 counties. When averaged across all MSAs, interaction segregation is 38% lower than a conventional static estimate, which means that people meet diverse others mostly when outside their home neighborhoods. But, we also find that interaction segregation is 67% higher in the 10 largest Metropolitan Statistical Areas (MSAs) than in small MSAs with fewer than 100,000 residents. We find evidence that because large cities can offer a greater choice of differentiated spaces targeted to specific socioeconomic groups, they end up promoting—rather than reducing—everyday economic segregation. We also discover that this segregation-increasing effect is countered when hubs of interaction (e.g. shopping malls) are positioned to bridge diverse neighborhoods and thus attract people of all socioeconomic statuses. Overall, our findings challenge a long-standing conjecture in human geography and urban design, and highlight how built environment can both prevent and facilitate diverse human interactions.

¹ Introduction

² In the U.S., economic segregation is very high, with income affecting where one lives⁹, who
³ one marries¹⁰, and who one meets and befriends¹¹. This extreme segregation is costly: it re-
⁴ duces economic mobility^{12–15}, fosters a wide range of health problems^{16,17}, and increases political
⁵ polarization¹⁸. Although there are all manner of reforms designed to reduce economic segregation
⁶ (e.g. subsidized housing), it has long been argued that one of the most powerful segregation-
⁷ reducing dynamics is rising urbanization¹⁹ and the happenstance mixing that it induces^{1–6}.

⁸ As plausible as this “cosmopolitan mixing hypothesis” might seem, big cities also provide
⁹ new opportunities for self-segregation, given that they’re large enough to allow people to seek out
¹⁰ and find others like themselves²⁰. These contrasting hypotheses about the effects of urbanization on
¹¹ interaction remain untested because it has been difficult to measure real-world interaction among
¹² individuals. In most cases, analysts have treated residential housing segregation as a proxy for
¹³ the diversity of interaction^{7,21–23}, an approach that rests on the implausible assumption that people
¹⁴ interact uniformly with those in their neighborhood and do not interact with anyone outside their
¹⁵ neighborhood.

¹⁶ To understand how urbanization affects segregation, we instead need a measure that captures
¹⁷ where people go, when they go there, and with whom they come into contact. We introduce a
¹⁸ measure of interaction segregation (IS) by leveraging anonymized cell phone geolocation data to
¹⁹ construct a fine-grained, dynamic network that captures 1.6 billion physical interactions (i.e. path-
²⁰ crossings) between 9.6 million people in the United States. These data are used to define a measure
²¹ of *interaction segregation* (IS) that extends a traditional static segregation measure by capturing
²² actual interactions in a given geographic area. In recent years, cell phone networks have been used
²³ for many research purposes^{24–32}, but a nationwide study of economic mobility and urbanization
²⁴ has not been undertaken because of difficulties in ascertaining individual-level economic status,
²⁵ determining when dyadic interactions occur, and amassing the data needed to compare across cities
²⁶ or counties^{24–26,28–32}. We undertake the first credible test of the “cosmopolitan mixing hypothesis”
²⁷ and the mechanisms underlying it by using cell phone data to link geolocated interactions with
²⁸ individual-level economic status in 382 MSAs and 2829 counties.

²⁹ **Results**

³⁰ We estimate *interaction segregation* (IS), defined as the extent of contact between individuals of
³¹ different economic statuses, for each geographic area in the U.S. (e.g. MSA, county). This en-
³² tails building a dynamic *interaction network* with 9,567,559 nodes (representing individuals) and
³³ 1,570,782,460 edges (representing interactions/path-crossings in the physical space).

³⁴ **Developing a more realistic measure of socioeconomic segregation.** To estimate each per-
³⁵ son's economic standing (ES), we first infer their night-time home location from cell phone mo-
³⁶ bility data, and we then recover the estimated monthly rent value of the home at this location
³⁷ (Methods M2). This method is more accurate than the convention of proxying individual ES with
³⁸ neighborhood-level Census averages^{26,27}. The economic segregation of each geographic region is
³⁹ measured by the correlation between a person's ES and the mean ES of everyone with whom they
⁴⁰ interact. This correlation is estimated by fitting a linear mixed effects model that eliminates atten-
⁴¹ uation bias (Methods M3). The resulting measure of interaction segregation (Figure 1a-c), which
⁴² ranges from 0 (perfect integration) to 1 (complete segregation), is a generalization of a widely
⁴³ used measure of socioeconomic segregation, the Neighborhood Sorting Index (NSI)⁷. The NSI is
⁴⁴ equivalent to the correlation between each person's ES and the mean ES of all people in their Cen-
⁴⁵ sus tract, whereas the IS is equivalent to the correlation between each person's ES and the mean
⁴⁶ ES of *all* people with whom they have interacted, either inside or outside their census tract. Like
⁴⁷ the NSI, the IS measures cross-class contact of any type, rather than ties that are persistent and
⁴⁸ possibly stronger (such as friendship ties)¹¹.

⁴⁹ **Interaction segregation is lower than previously estimated.** We find that the median inter-
⁵⁰ action segregation across all MSAs is 38% ($p < 10^{-4}$, 95% confidence interval 37%–41%) lower
⁵¹ than a conventional static estimate (NSI; Figure 1d top). To understand why, we examine how
⁵² segregation differs by interaction location (Figure 1d bottom). We find that interactions that oc-
⁵³ cur when both people are within their home Census tract are 41% ($p < 10^{-4}$, 95% CI 38%–44%)
⁵⁴ more segregated than under the hypothetical that residents interact uniformly with all people in the

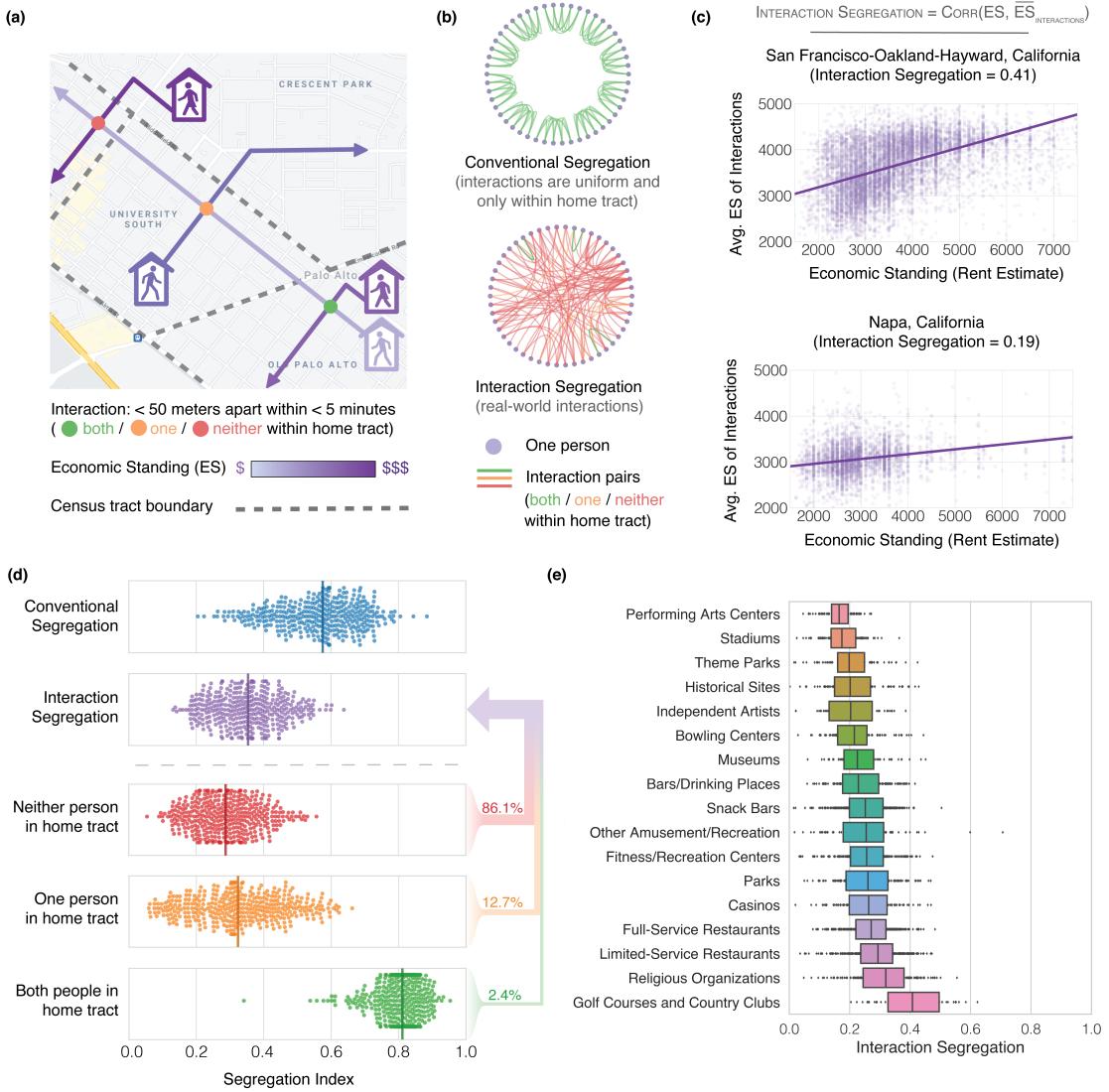


Figure 1: Interaction segregation (IS) captures the likelihood of contact between people of different economic backgrounds. (a) For 9.6 million individuals (i.e. cell phones), we infer economic standing (rent or rent equivalent) from home address, based on location at night. We then leverage anonymized cell phone mobility data to identify *interactions* between pairs of individuals (where two individuals are defined to interact if they were within D meters of each other within less than T minutes). $D = 50$ and $T = 5$ in our primary analysis; results are robust to the precise choice of D and T (Methods M1-M2, Supplementary Figure S5-S8). **(b)** The nationwide network of 1.6 billion interactions spans 2829 counties and 382 Metropolitan Statistical Areas (MSAs). Our interaction network contrasts with a conventional measure of economic segregation, the Neighborhood Sorting Index, which assumes that people interact uniformly and only with other residents of their home Census tract. Graphs based on a sample community of 50 individuals from San Francisco, CA residing in 10 different census tracts. Nodes are individuals; edges are interactions. As this sample illustrates, most interactions happen when both persons are away from their home tract. These cross-tract interactions are undetected by conventional segregation measures. **(c)** For each geographic region (e.g. MSA, county), we estimate interaction segregation, defined as the correlation between an individual's economic standing (ES) and the mean ES of those with whom they interact. 1 signifies perfect segregation; 0 signifies no segregation. This definition is equivalent to the conventional Neighborhood Sorting Index (NSI), but with the key difference that it leverages real-life interactions from mobility data instead of synthetic interactions from individuals grouped by Census tracts. For two MSAs, we show the raw data; each point represents one individual. San Francisco-Oakland-Hayward, CA is $2.2 \times$ ($p < 10^{-4}$, 95% CI 1.6–2.8×) more segregated than Napa, CA. **(d)** Top: Interaction segregation is 38% ($p < 10^{-4}$, 95% CI 37%–41%) lower than the conventional segregation measure NSI. Each point represents the interaction segregation estimate in one MSA; vertical colored lines represent median across MSAs. Bottom: breaking down interaction segregation into component parts. Interactions where both people are within their home Census tract (green) are most segregated, reflecting the *homophily effect* in which people preferentially interact with those of similar ES in their home tracts. Out-of-tract interactions (orange and red) are less segregated, reflecting the *visitor effect* in which visiting other tracts exposes individuals to economically diverse individuals. Because a small minority (2.4%, 95% CI 2.4%–2.4%) of interactions happen within home tract, the visitor effect dominates the homophily effect and thus interaction segregation is lower than conventional NSI. **(e)** Interaction segregation varies by location type. Each point represents segregation in one MSA using only interactions occurring in a given location type; boxes indicate the interquartile range across MSAs. Segregation is highest at golf courses and country clubs (median IS 0.42), and lowest at performing arts centers (median IS 0.16) and stadiums (median IS 0.17).

55 same tract (as the NSI assumes). This result illustrates the *homophily effect*: even within their own
56 neighborhood, people interact with neighbors who are socioeconomically most similar to them.
57 By contrast, interactions where one or both individuals are outside of their home tract are far less
58 segregated (i.e. 44% ($p < 10^{-4}$, 95% CI 41%–46%) less segregated when one person is outside
59 of home tract and 50% ($p < 10^{-4}$, 95% CI 48%–53%) less segregated when both people are out-
60 side home tract), meaning that people interact with more heterogeneous populations when they
61 visit non-home tracts for work, leisure, or other activities. We refer to this phenomenon as the
62 *visitor effect*. Because within-home-tract interactions constitute a small minority (2.4%, 95% CI
63 2.4%–2.4%) of interactions compared to out-of-tract interactions, the visitor effect dominates the
64 homophily effect and interaction segregation is lower than conventional static measures imply.

65 **Interaction segregation varies across leisure sites.** We validate and explore the IS metric by
66 measuring variability in interaction segregation across different points of interest (POIs). We do so
67 by filtering for interactions that occur within a single POI category and re-applying our same mixed
68 model (Methods M3; Figure 1e). The resulting estimates show, for instance, that golf courses and
69 country clubs have over $2.6 \times$ ($p < 10^{-4}$, 95% CI $2.2 \times$ – $2.9 \times$) higher interaction segregation than
70 performing arts centers (in the median MSA). We find that the degree to which POIs service small
71 and thereby socioeconomically homogeneous communities (measured by average travel distance
72 to nearest POI and # of POIs) explains much of this variability in POI-level segregation (Spearman
73 Corr. -0.75, $p < 0.001$ for travel distance, Spearman Corr. 0.69, $p < 0.01$ for # of POIs, Extended
74 Data Figure 2a,b). The POIs that are numerous, embedded locally within residential communities,
75 and thereby serve economically differentiated communities (e.g. religious organizations) tend to
76 be more segregated than larger POIs that are designed to serve the entire city (e.g. stadiums). For
77 instance, in the median MSA, religious organizations require 92% ($p < 10^{-4}$, 95% CI 92%–93%)
78 less travel distance and are $16 \times$ ($p < 10^{-4}$, 95% CI $8 \times$ – $18 \times$) more numerous than stadiums, and
79 are thus 75% ($p < 10^{-4}$, 95% CI 58%–87%) more segregated. In rare cases, a POI with few venues
80 may still be highly segregated (i.e. golf courses have significantly higher (all $p < 10^{-4}$) segregation
81 than all other POI categories in Figure 1e despite there being only 2 golf courses in the median

82 MSA) because cross-venue economic differentiation is generated through other mechanisms, such
83 as a public-private distinction (Extended Data Figure 2c).

84 **Large cities facilitate segregation.** We discover that interaction segregation is higher in large
85 MSAs (Figure 2), which directly undermines the “cosmopolitan mixing hypothesis”. The Spear-
86 man correlation between MSA population and MSA segregation is 0.62 ($p < 10^{-4}$), and the 10
87 largest MSAs by population size are 67% ($p < 10^{-4}$, 95% CI 49-87%) more segregated than small
88 MSAs with less than 100,000 residents. This result is robust: we validate it by recalculating
89 the correlation with a measure of density rather than population size (Spearman Correlation 0.45,
90 $p < 10^{-4}$, Supplementary Table S7), by controlling for relevant covariates (Extended Data Table
91 1 and Supplementary Table S7), by varying the granularity of the analysis (Figure 2b, Extended
92 Data Figure 3), and by testing a variety of specifications of interaction segregation (Supplementary
93 Table S6, Supplementary Figures S2-S8). The consistent result that larger, denser cities are more
94 segregated runs counter to the hypothesis that such cities facilitate diverse social interactions by
95 attracting liberal cosmopolitans and by constraining space in ways that oblige diverse individuals
96 to come into contact with each other¹⁻⁶. Our results support the opposite hypothesis: big cities
97 allow their inhabitants to seek out people who are more like themselves.

98 **Mechanisms producing higher interaction segregation in larger metropolitan areas.** To un-
99 derstand why large metropolitan areas support these homophilous tendencies, it is useful to ex-
100 plore interaction segregation within leisure POIs as a case study. Full-service restaurants provide
101 an illustrative example (Figure 2c,d,e) of a segregation-inducing dynamic that holds widely across
102 other leisure sites (Supplementary Figure S20) and scales (Extended Data Figure 4-5). We find
103 that larger MSAs offer their residents a greater number of leisure choices: the average resident of
104 one of the 10 largest MSAs has 22× ($p < 10^{-4}$, 95% CI 11-39×) more restaurants within 10 kilo-
105 meters of their home than an average resident of a small MSA (where a “small MSA” is defined as
106 one with less than 100,000 residents; Figure 2c). These choices are also more socioeconomically
107 differentiated. When a restaurant’s ES is defined as the median ES of all people who cross paths

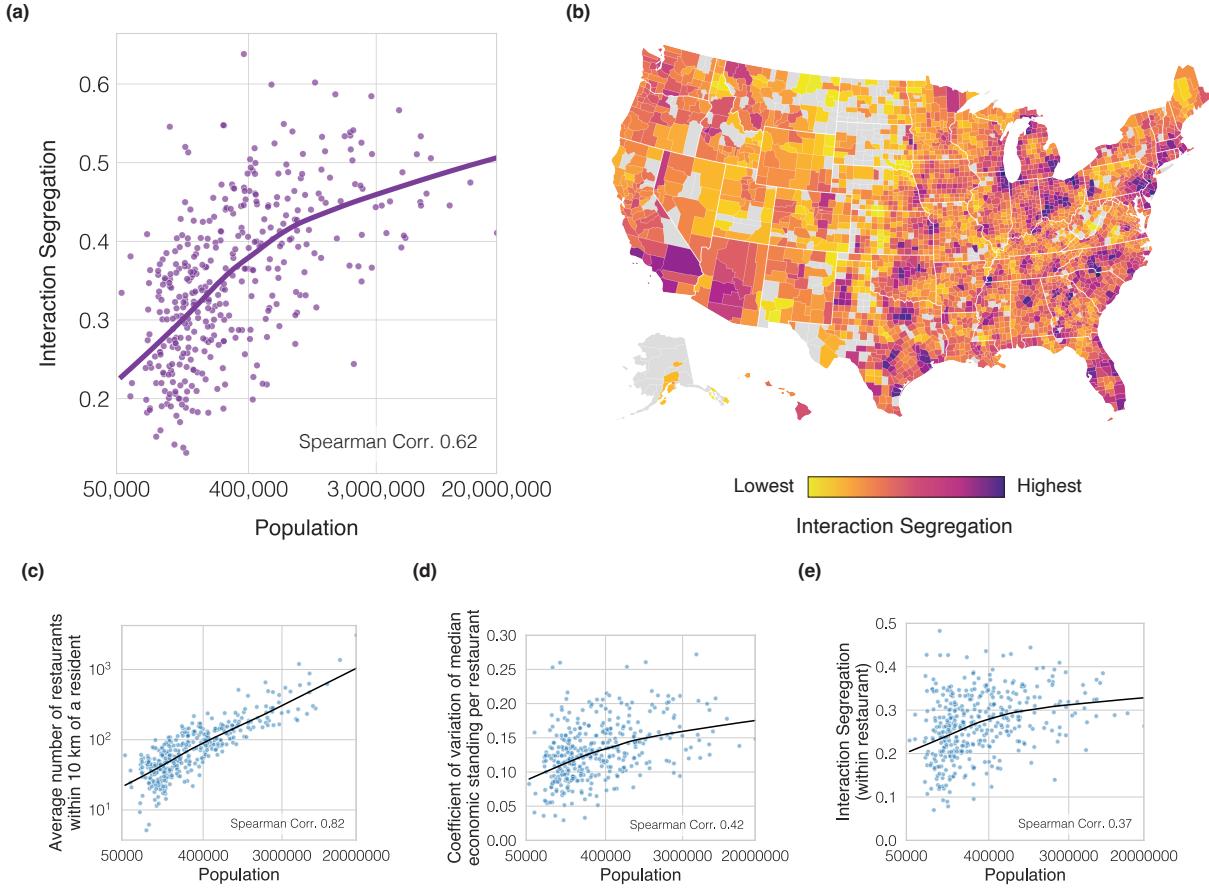


Figure 2: Highly-populated metropolitan areas are more segregated due to socioeconomic differentiation of spaces. Contrary to the hypothesis that highly-populated metropolitan areas support diverse interactions and socioeconomic mixing, we find that **(a)** Larger MSAs are more segregated. Interaction segregation presented as a function of population size; each dot represents one MSA; purple line indicates LOWESS fit. Upward trend reveals that urbanization is associated with higher interaction segregation (Spearman Correlation 0.62, N=382, $p < 10^{-4}$). The top 10 largest MSAs, by population size, are 67% more segregated than small MSAs with less than 100,000 residents. Associations are robust to controlling for potential confounders and are similar for population density and IS (Extended Data Table 1, Supplementary Table S7). **(b)** Interaction segregation across the 2829 US counties. Analysis limited to counties with at least 50 individuals. Interaction segregation varies significantly across counties in the United States. Moreover, as with MSA-level, county-level interaction segregation is also positively associated with both population size and population density (Extended Data Figure 3). **(c-e)** A case study of full-service restaurants illustrates the mechanism through which urbanization produces interaction segregation. Highly-populated metropolitan areas are more segregated not only because they offer a wider choice of venues but also because these venues are more socioeconomically differentiated (e.g. in New York City, one can spend \$10, \$100, or \$1000 on a meal, depending on the choice of restaurant^{33,34}). **(c)** Larger MSAs have more restaurants within 10 kilometers of the average resident, giving residents more options to self-segregate. **(d)** Moreover, restaurants in larger MSAs vary more in the median ES of their visitors, offering a greater choice of socioeconomically differentiated restaurants. The coefficient of variation across restaurant ES in 10 largest MSAs is 63% ($p < 10^{-4}$, 95% CI 37-100%) more than the coefficient of variation in small MSAs (with fewer than 100,000 residents). **(e)** Consequently, interaction segregation within restaurants is higher in larger MSAs. These relationships are also detectable at the scale of interaction hubs (i.e. higher-level clusters of POIs such as plazas and shopping malls) as well as at the neighborhood level (Extended Data Figures 4-5).

108 within it, the coefficient of variation of “restaurant ES” in the 10 largest MSAs is 63% ($p < 10^{-4}$,
109 95% CI 37-100%) larger than that in small ones (Figure 2d). Thus, large MSAs not only offer
110 their residents a larger choice of restaurants, but these restaurants are also more socioeconomically
111 differentiated. These processes combine to increase interaction segregation by 29% ($p < 10^{-3}$, 95%
112 CI 8-49%) at restaurants in the 10 largest MSAs relative to those in small MSAs (Figure 2e). We
113 also find analogous results at higher levels of scale: interaction hubs (e.g. plazas, malls, shopping
114 centers, boardwalks) (Extended Data Figure 4) as well as neighborhoods (Extended Data Figure
115 5) and across different many POI types (Supplementary Figure S20).

116 **Mitigating segregation via urban design.** Our results so far suggest that segregation could be
117 mitigated via urban design by placing POIs of high interaction to act as bridges between diverse
118 neighborhoods, which would allow residents of nearby high-ES and low-ES neighborhoods to eas-
119 ily visit and interact (Figure 3c)³⁵⁻³⁷. We develop the *Bridging Index* (BI; Methods M3) to assess
120 whether interaction hubs (i.e. highly-visited POIs) are located in such bridging positions. This in-
121 dex, which measures the economic diversity of the groups that would interact if everybody visited
122 only their nearest hub, is computed by clustering individuals by the nearest hub to their home and
123 then measuring the economic diversity within these clusters (Extended Data Figure 6). The result-
124 ing index ranges from 0 to 1, where 0 means that individuals near each hub have uniform ES, and
125 1 means that individuals near each hub are as diverse as the overall area (Extended Data Figure 7).
126 We compute BI for commercial centers (e.g. plazas, malls, shopping centers, boardwalks) because
127 we find that they are common hubs of interaction: the majority (56.9%, 95% CI 56.9%-56.9%) of
128 interactions across all 382 MSAs occur in close proximity (within 1km) of a commercial center,
129 even though only 2.5% of land area is within 1km of a commercial center. (see Figure 3c). The re-
130 sults show that BI is strongly associated with interaction segregation (Spearman Correlation -0.78,
131 $p < 10^{-4}$; Figure 3d). The top 10 MSAs with the highest BI are 53.1% ($p < 10^{-4}$, 95% CI 44-60%)
132 less segregated than the 10 MSAs with the lowest BI. This finding is again robust: the effect of BI
133 is strong and significant ($p < 10^{-4}$) even after including controls for race, population size, economic
134 inequality, and many other variables (Extended Data Tables 2 and 3; see also Supplementary Table

¹³⁵ **S6**; Supplementary Figures **S2** and **S8**; Supplementary Figure **S11**). It follows that policies that
¹³⁶ encourage developers to locate hubs such as commercial centers between diverse residential neigh-
¹³⁷ borhoods (e.g. zoning laws or subsidies) may reduce interaction segregation. We have identified
¹³⁸ several large cities that increase integration in this way (Supplementary Table S8) and present an
¹³⁹ illustrative example (Figure 3c-d) in which well-placed interaction hubs bridge diverse individuals
¹⁴⁰ in Fayetteville, North Carolina.

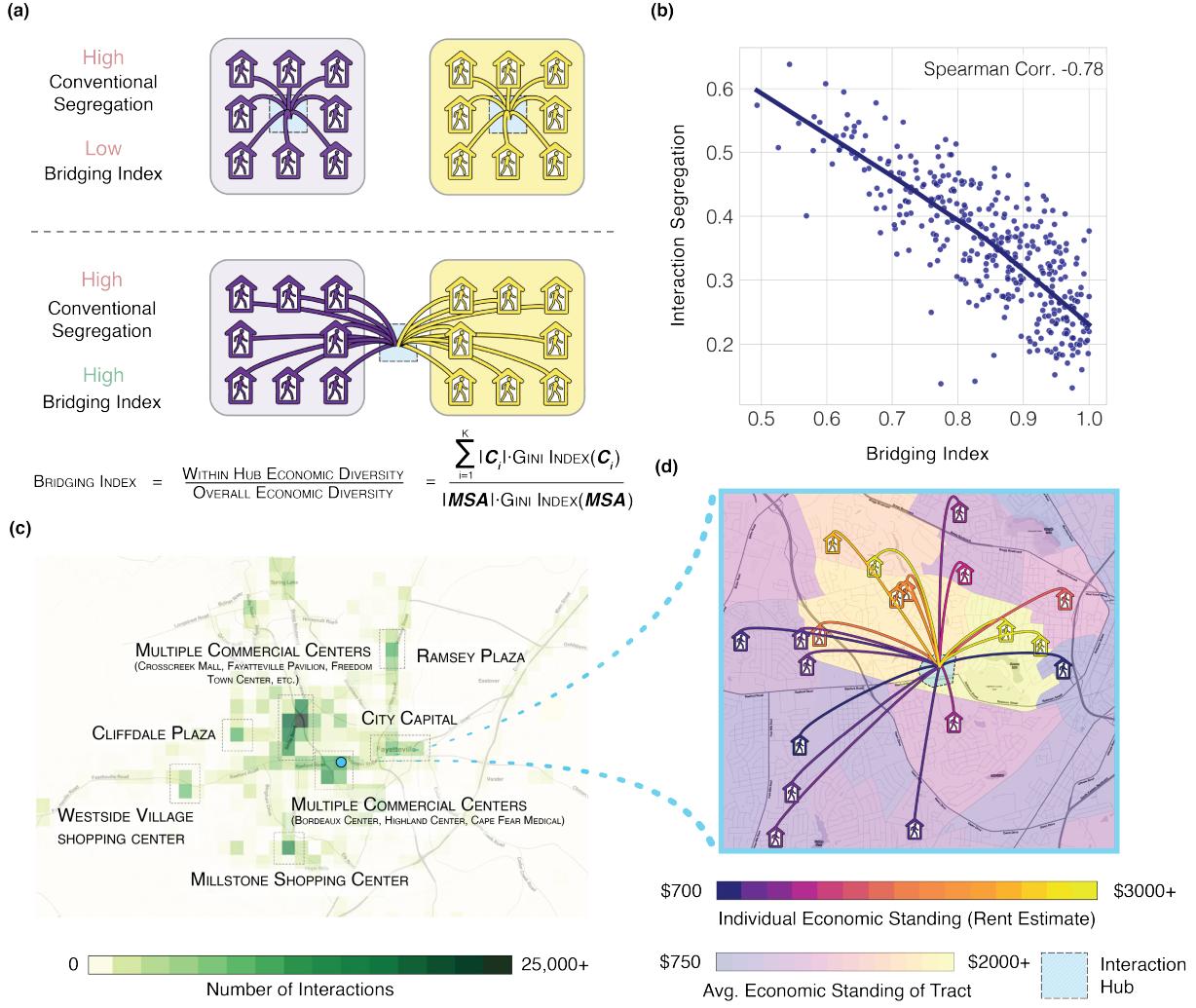


Figure 3: Interaction segregation is lower when interaction hubs bridge socioeconomically-diverse neighborhoods. (a) We develop a Bridging Index (BI) that quantifies the extent to which interaction hubs bridge socioeconomically diverse neighborhoods. The metric is constructed by clustering homes by nearest interaction hub, then measuring the within-cluster diversity of ES (Methods M3). Two plots illustrate that BI is distinct from a conventional residential measure of segregation (i.e. Neighborhood Sorting Index). BI ranges from 0.0 (no bridging; top) to 1.0 (perfect bridging; bottom) while residential segregation is constant (high and low-ES individuals are highly segregated by census tract, denoted by purple and yellow bounding boxes). We compute BI with hubs of interaction defined as commercial centers (e.g. shopping malls, plazas) because the majority (56.9%, 95% CI 56.9%-56.9%) of interactions across all 382 MSAs occur in close proximity (within 1km) of a commercial center, even though only 2.5% of land area is within 1km of a commercial center. (b) BI strongly predicts interaction segregation (Spearman Correlation -0.78 , N=382, $p < 10^{-4}$). The top 10 MSAs with the highest BI are 53.1% ($p < 10^{-4}$, 95% CI 44-60%) less segregated than the 10 MSAs with lowest BI. Bridging Index predicts segregation more accurately ($p < 10^{-4}$) than population size, ES inequality, NSI, and racial demographics, and is significantly ($p < 10^{-4}$) associated with interaction segregation after controlling for these variables and other potential confounders (Extended Data Tables 2-3). (c-d) A case study of Fayetteville, NC, an MSA with low interaction segregation (21st percentile) despite having above-median population size (64th percentile) and income inequality (60th percentile). (c) Interaction heat map of Fayetteville; all visually discernible hubs are associated with one or more commercial centers. (d) Interaction hubs are located in accessible proximity to both high and low ES census tracts (Bridging Index = 0.90, 62nd percentile), leading to diverse interactions. An illustrative example of one hub (Highland Center) in Fayetteville and a random sample of 10 interactions occurring inside of it. Home icons demarcate individual home location (up to 100m of random noise added for anonymity); colors denote individual and mean tract ES.

141 **Discussion**

142 As big cities continue to grow and spread, it is important to ask whether they are increasing socioe-
143 conomic mixing. Although it is often argued that big cities promote mixing by increasing density,
144 in fact we find that interaction diversity and city size are *negatively* related. We find that the key
145 mechanism here is scale. Because large cities can sustain venues that are targeted to thin socioe-
146 conomic slices of the population, they have become homophily-generating machines that are more
147 segregated than small cities. We also find that some cities are able to mitigate this segregative
148 effect because their interaction hubs are located in bridging zones that can draw in people from
149 diverse neighborhoods. We were able to detect these pockets of homophily (and the counteracting
150 effects of bridging hubs) because we have developed a dynamic measure of economic segregation
151 that captures everyday interactions at home, work, and leisure.

152 This new methodology for measuring interaction segregation, while an important improve-
153 ment over conventional static approaches, has limitations. We use close physical proximity as a
154 proxy for interaction. It is reassuring, however, that our core results persist under stricter time,
155 distance, and tie-strength thresholds (Supplementary Table S6, Supplementary Figures S5-S8).
156 It is likewise important to locate and analyze supplementary datasets that cover subpopulations
157 (e.g., homeless subpopulations) that aren't as well represented in our dataset³⁸. The available evi-
158 dence indicates that our sample is representative on many key racial, economic, and demographic
159 variables³⁹, but cellphone market penetration is still not complete. Lastly, our measure of eco-
160 nomic standing relies on housing consumption, an indicator that does not exhaust the concept
161 of economic status. It is again reassuring that our analytic approach, which improves on conven-
162 tional neighborhood-level imputations, is robust under a range of alternative measures of economic
163 standing (Supplementary Figure S3).

164 Our work advances the usual static approach to measuring economic segregation to using
165 large-scale mobility data and develops a dynamic notion of quantifying segregation and diversity
166 of mixing between people. The dynamic approach that we have taken here could further be ex-
167 tended to examine cross-population differences in the sources of segregation and to develop a more
168 complete toolkit of approaches to reducing segregation and improving urban design.

References

1. Jacobs, J. *The Death and Life of Great American Cities* (Random House, 1961).
2. Milgram, S. The experience of living in cities. *Science* **167**, 1461–1468 (1970).
3. Derex, M., Beugin, M.-P., Godelle, B. & Raymond, M. Experimental evidence for the influence of group size on cultural complexity. *Nature* **503**, 389–391 (2013).
4. Gomez-Lievano, A., Patterson-Lomba, O. & Hausmann, R. Explaining the prevalence, scaling and variance of urban phenomena. *Nature Human Behaviour* **1**, 1–6 (2016).
5. Wirth, L. Urbanism as a way of life. *American Journal of Sociology* **44**, 1–24 (1938).
6. Stier, A. J. *et al.* Evidence and theory for lower rates of depression in larger us urban areas. *Proceedings of the National Academy of Sciences* **118**, e2022472118 (2021).
7. Jargowsky, P. A. Take the money and run: Economic segregation in us metropolitan areas. *American Sociological Review* 984–998 (1996).
8. Massey, D. S. & Eggers, M. L. The spatial concentration of affluence and poverty during the 1970s. *Urban Affairs Quarterly* **29**, 299–315 (1993).
9. Reardon, S. F., Bischoff, K., Owens, A. & Townsend, J. B. Has income segregation really increased? bias and bias correction in sample-based segregation estimates. *Demography* **55**, 2129–2160 (2018).
10. Schwartz, C. R. Earnings inequality and the changing association between spouses' earnings. *American journal of sociology* **115**, 1524–1557 (2010).
11. Chetty, R. *et al.* Social capital: Determinants of economic connectedness. *Nature* **608**, 122–34 (2022).
12. Massey, D. & Denton, N. A. *American apartheid: Segregation and the making of the under-class* (Harvard university press, 1993).
13. Chetty, R., Hendren, N. & Katz, L. F. The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review* **106**, 855–902 (2016).
14. Chetty, R. & Hendren, N. The impacts of neighborhoods on intergenerational mobility ii: County-level estimates. *The Quarterly Journal of Economics* **133**, 1163–1228 (2018).
15. Chetty, R. *et al.* Social capital: Measurement and associations with economic mobility. *Nature* **608**, 108–21 (2022).
16. Bor, J., Cohen, G. H. & Galea, S. Population health in an era of rising income inequality: Usa, 1980–2015. *The Lancet* **389**, 1475–1490 (2017).

17. Do, D. P., Locklar, L. R. & Florsheim, P. Triple jeopardy: The joint impact of racial segregation and neighborhood poverty on the mental health of black americans. *Social psychiatry and psychiatric epidemiology* **54**, 533–541 (2019).
18. Brown, J. R., Enos, R. D., Feigenbaum, J. & Mazumder, S. Childhood cross-ethnic exposure predicts political behavior seven decades later: Evidence from linked administrative data. *Science Advances* **7**, eabe8432 (2021).
19. of Economic, U. N. D. *World urbanization prospects: The 2018 revision*, vol. 216 (United Nations Publications, 2018).
20. Fischer, C. S. *To dwell among friends: Personal networks in town and city* (University of chicago Press, 1982).
21. Jargowsky, P. A. & Kim, J. A measure of spatial segregation: The generalized neighborhood sorting index. *National Poverty Center Working Paper Series* (2005).
22. Massey, D. S. & Denton, N. A. The dimensions of residential segregation. *Social forces* **67**, 281–315 (1988).
23. Brown, J. R. & Enos, R. D. The measurement of partisan sorting for 180 million voters. *Nature Human Behaviour* 1–11 (2021).
24. Matthews, S. A. & Yang, T.-C. Spatial polygamy and contextual exposures (spaces) promoting activity space approaches in research on place and health. *American behavioral scientist* **57**, 1057–1081 (2013).
25. Zenk, S. N. *et al.* Activity space environment and dietary and physical activity behaviors: a pilot study. *Health & place* **17**, 1150–1161 (2011).
26. Moro, E., Calacci, D., Dong, X. & Pentland, A. Mobility patterns are associated with experienced income segregation in large us cities. *Nature communications* **12**, 1–10 (2021).
27. Athey, S., Ferguson, B. A., Gentzkow, M. & Schmidt, T. Experienced segregation. Tech. Rep., National Bureau of Economic Research (2020).
28. Candipan, J., Phillips, N. E., Sampson, R. J. & Small, M. From residence to movement: The nature of racial segregation in everyday urban mobility. *Urban Studies* **58**, 3095–3117 (2021).
29. Phillips, N. E., Levy, B. L., Sampson, R. J., Small, M. L. & Wang, R. Q. The social integration of american cities: Network measures of connectedness based on everyday mobility across neighborhoods. *Sociological Methods & Research* **50**, 1110–1149 (2021).
30. Wang, Q., Phillips, N. E., Small, M. L. & Sampson, R. J. Urban mobility and neighborhood isolation in america’s 50 largest cities. *Proceedings of the National Academy of Sciences* **115**, 7735–7740 (2018).

31. Levy, B. L., Phillips, N. E. & Sampson, R. J. Triple disadvantage: Neighborhood networks of everyday urban mobility and violence in us cities. *American Sociological Review* **85**, 925–956 (2020).
32. Levy, B. L., Vachuska, K., Subramanian, S. & Sampson, R. J. Neighborhood socioeconomic inequality based on everyday mobility predicts covid-19 infection in san francisco, seattle, and wisconsin. *Science advances* **8**, eabl3825 (2022).
33. Sutton, R. <https://ny.eater.com/2022/1/4/22867230/sushi-prices-omakase-supply-chain-luxury-nyc-restaurants> (2022). Accessed Jan 7th, 2022.
34. Sutton, R. <https://ny.eater.com/2022/3/10/22968257/masa-price-hike-most-expensive-restaurant-sushi-bar-nyc-restaurants> (2022). Accessed March 10th, 2022.
35. Zipf, G. K. The p 1 p 2/d hypothesis: on the intercity movement of persons. *American sociological review* **11**, 677–686 (1946).
36. Simini, F., González, M. C., Maritan, A. & Barabási, A.-L. A universal model for mobility and migration patterns. *Nature* **484**, 96–100 (2012).
37. Schläpfer, M. *et al.* The universal visitation law of human mobility. *Nature* **593**, 522–527 (2021).
38. Coston, A. *et al.* Leveraging administrative data for bias audits: Assessing disparate coverage with mobility data for covid-19 policy. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 173–184 (2021).
39. Squire, R. F. What about bias in the safegraph dataset? *Safegraph* (2019).

169 **Methods**

170 In Methods M1, we explain the datasets used in our analysis; in Methods M2, we explain data pro-
171 cessing procedures we leverage to infer economic standing and interactions; and in Methods M3,
172 we explain the analysis underlying our main results.

173 **M1 Datasets**

174 **SafeGraph**

175 Our primary mobility and location data comprise GPS locations from a sample of adult smartphone
176 users in the United States, provided by the company SafeGraph. The data are anonymous GPS
177 location pings from smartphone applications which are collected and transmitted to SafeGraph
178 by participating users⁴⁰. While the sample is not random sample, prior work has demonstrated
179 that SafeGraph data is geographically representative (e.g. an approximately unbiased sample of
180 different census tracts within each State), and well-balanced along the dimensions of race, in-
181 come, and education^{39,42}. Furthermore, SafeGraph data is a widely used standard in large-scale
182 studies of human mobility across many different areas including COVID-19 modeling⁴², polit-
183 ical polarization⁴³, and tracking consumer preferences⁴⁴. All data provided by SafeGraph was
184 anonymized, does not contain any identifying information, and was stored on a secure server be-
185 hind a firewall. Data handling and analysis was conducted in accordance with SafeGraph policies
186 and in accordance with the guidelines of the Stanford University Institutional Review Board.

187 The raw data consists of 91,755,502 users and 61,730,645,084 pings (one latitude and lon-
188 gitude for one user at one timestamp) from three evenly spaced months in 2017: March, July, and
189 November. The mean number of raw pings associated with a user is 667 and the median num-
190 ber of pings is 12. We apply several filters to improve the reliability of the SafeGraph data, and
191 subsequently link each user to an estimated rent (i.e. Zillow Zestimate) using their inferred home
192 location (i.e. CoreLogic address), as described in Methods M2.

193 We apply several filters to improve the reliability of the SafeGraph data. To ensure locations
194 are reliable, we remove pings whose location is estimated with accuracy worse than 100 meters as
195 recommended by SafeGraph⁴⁵. We filter out users with fewer than 500 pings, as these are largely

noise. Since we incorporate a user's home value and rent in measuring their economic standing, we filter out users for whom we are unable to infer a home. Finally, to avoid duplicate users, we remove users if more than 80% of their pings have identical latitudes, longitudes, and timestamps to those of another user; this could potentially occur if, for example, a single person in the real world carries multiple mobile devices. After the initial filters on ping counts and reliability, we are able to infer home locations for 12,805,490 users in the United States (50 states and Washington D.C.), leveraging the CoreLogic database. Of users for whom we can infer a home location, we are able to successfully link 9,576,650 to an estimated rent value via the Zillow API. Section Methods M2 provides full details on the use of CoreLogic database to infer home locations and the use of the Zillow API to link these home locations to estimated rent values. Finally, after removing users where > 80% of their pings are duplicates with another user, we reduce the number of users from 9,576,650 to 9,567,559 (i.e., we remove about 0.1% of users through de-duplication).

208 **CoreLogic**

209 We use the CoreLogic real estate database to link users to home locations⁴⁶. The database provides
210 information covering over 99% of US residential properties (145 million properties), over 99% of
211 commercial real estate properties (26 million properties), and 100% of US county, municipal, and
212 special tax districts (3141 counties). The CoreLogic real estate database includes the latitude and
213 longitude of each home, in addition to its full address: street name, number, county, state, and zip
214 code.

215 **Zillow**

216 We use the Zillow property database to query for rent estimates⁴⁷ (our primary measure of eco-
217 nomic standing). The Zillow database contains rent data ("rent Zestimate") for 119 million US
218 residential properties. We were able to determine a rent Zestimate, the primary measure of eco-
219 nomic standing (ES) used in our analysis, for 9,576,650 out of 12,183,523 inferred SafeGraph user
220 homes (a 79% hit rate).

221 **SafeGraph Places**

222 Our database of US business establishment boundaries and annotations comes from the SafeGraph
223 Places database⁴⁰, which indexes the names, addresses, categories, latitudes, longitudes, and geo-
224 graphical boundary polygons of 5.5 million US points of interest (POIs) in the United States. Safe-
225 Graph includes the NAICS (North American Industry Classification System) category of each POI,
226 which is standard taxonomy used by the Federal government to classify business establishments⁴⁸.
227 For instance, the NAICS code 722511 indicates full service restaurants. We identify relevant
228 leisure sites using the prefixes 7, which includes arts, entertainment, recreation, accommodation,
229 and food services, and supplement these POIs with the prefix 8131 to include religious organiza-
230 tions such as churches. We restrict our analysis of leisure sites to the top most frequently visited
231 POI categories within these NAICS code prefixes (Figure 1d): full service restaurants, snack bars,
232 limited-service restaurants, stadiums, etc. SafeGraph Places also includes higher-level “parent”
233 POI polygons which encapsulate smaller POIs. Specifically, we identified interaction hubs with
234 the NAICS code 531120 (lessors of non-residential real estate) which we find in practice corre-
235 sponds to commercial centers such as shopping malls, plazas, boardwalks, and other clusters of
236 businesses. We provide illustrative examples of such interaction hubs in Supplementary Figures
237 [S14-S16](#).

238 **US Census**

239 We extract demographic and geographic features from the 5-year 2013-2017 American Commu-
240 nity Survey (ACS)⁴⁹. This allows us, as described below, to link cell phone locations to geographic
241 areas including census block group, census tract, and Metropolitan Statistical Area (MSA), as
242 well as to infer demographic features corresponding to those demographic areas including median
243 household income.

244 A census block group (CBG) is a statistical division of a census tract. CBGs are generally
245 defined to contain between 600 and 3,000 people. A CBG can be identified on the national level
246 by the unique combination of state, county, tract, and block group codes.

247 A census tract is a statistical subdivision of a county containing an average of roughly 4,000

248 inhabitants. Census tracts range in population from 1,200 to 8,000 inhabitants. Each tract is
249 identified by a unique numeric code within a county. A tract can be identified on the national level
250 by the unique combination of state, county, and tract codes.

251 Census tracts and block groups typically cover a contiguous geographic area, though this is
252 not a constraint on the shape of the tract or block group. Census tract and block group boundaries
253 generally persist over time so that temporal and geographical analysis is possible across multiple
254 censuses.

255 Most census tracts and CBGs are delineated by inhabitants who participate in the Census
256 Bureau's Participant Statistical Areas Program. The Census Bureau determines the boundaries
257 of the remaining tracts and block groups when delineation by inhabitants, local governments, or
258 regional organizations is not possible ⁵⁰.

259 A Metropolitan Statistical Area (MSA) is a US geographic area defined by the Office of
260 Management and Budget (OMB) and is one of two types of Core Based Statistical Area (CBSA).
261 A CBSA comprises a county or counties associated with a core urbanized area with a population
262 of at least 10,000 inhabitants and adjacent counties with a high degree of social and economic
263 integration with the core area. Social and economic integration is measured through commuting
264 ties between the adjacent counties and the core. A Micropolitan Statistical Area is a CBSA whose
265 core has a population of between 10,000 and 50,000; a Metropolitan Statistical Area is a CBSA
266 whose core has a population of over 50,000. In our primary analysis, we follow Athey et al⁵¹
267 and focus on Metropolitan Statistical Areas, excluding Micropolitan Statistical Areas due to data
268 sparsity concerns.

269 **TIGER**

270 Road and transportation feature annotations come from the Census-curated Topologically Inte-
271 grated Geographic Encoding and Referencing system (TIGER) database⁵². The TIGER databases
272 are an extract of selected geographic and cartographic information from the U.S. Census Bureau's
273 Master Address File / Topologically Integrated Geographic Encoding and Referencing (MAF/TIGER)
274 Database (MTDB). We use the MAF/TIGER Feature Class Code (MTFCC) from the TIGER Roads

275 and TIGER Rails databases to identify road and railways. TIGER data is in the format of Shape-
276 files, which provide the exact boundaries of roads and railways as latitude/longitude coordinates.

277 M2 Data processing

278 For each individual, we first infer their home location and subsequently estimate economic standing
279 based on their home rent value (see *Inferring home location* and subsequently *Inferring economic*
280 *standing*). We then calculate all interactions between individuals (see *Constructing interaction*
281 *network*), which we then annotate based on the location, i.e. if the interaction occurred in both,
282 one, or neither individual’s home tract, and whether it occurred inside of a fine-grained POI such
283 as a specific restaurant or a “parent” POI such as an interaction hub (see *Annotating interactions*).
284 Details on all inferences and interaction calculations are provided below.

285 Inferring home location

286 We first infer a user’s home latitude and longitude using the latitude and longitude coordinates of
287 their pings during local nighttime hours, based on best practices established by SafeGraph⁵³. We
288 first remove users with fewer than 500 pings to ensure that we have enough data to reliably infer
289 home locations. We then interpolate each person’s location at each hour (eg, 6 PM, 7 PM, and 8
290 PM) using linear interpolation of latitudes and longitudes, to ensure we have timeseries at constant
291 time resolution. We filter for hours between 6 PM and 9 AM where the person moves less than 50
292 meters until the next hour; these stationary nighttime observations represent cases when the person
293 is more likely to be at home. We filter for users who have stationary nighttime observations on
294 at least 3 nights and with at least 60% of observations within a 50 meter radius. Finally, we infer
295 home latitude and longitude as the median latitude and longitude of these nighttime home locations
296 (after removing outliers outside the 50 meter radius). We choose the thresholds above because they
297 yield a good compromise between inferring the home location of most users and inferring home
298 locations with high confidence. Overall, we are able to infer home locations for 70% of users
299 with more than 500 pings, and these locations are inferred with high confidence; 89% of stationary
300 nighttime observations are within 50 meters of the inferred home latitude and longitude.

301 **Inferring economic standing from home latitude and longitude**

302 Having inferred home location from nighttime GPS pings, we link their latitude and longitude to
303 a large-scale housing database (Zillow) to infer the estimated rent of each user's home, which we
304 use as a measure of economic standing. We do this in two steps. First, we link the inferred user's
305 home latitude and longitude to the CoreLogic property database (Methods M1), a comprehensive
306 database of properties in the United States, by taking the closest CoreLogic residential property
307 (single family residence, condominium, duplex, or apartment) to the user's inferred home latitude
308 and longitude. Second, we use the CoreLogic address to query the Zillow database, which provides
309 estimated home rent and price for each user. (The Zillow database does not allow for queries using
310 raw latitude and longitude, which it is necessary to leverage to CoreLogic to obtain an address for
311 each user.) We use Zillow's estimated rent for the user's home as our main measure of economic
312 standing. We apply several quality control filters to ensure that the final set of users we use in
313 our main analyses have reliably inferred home locations and economic standings: 1) we remove
314 a small number of users whose inferred nighttime home latitude and longitude are identical to
315 another user's, since we empirically observe that these people have unusual ping patterns; 2) we
316 remove users for whom we are lacking an Zillow rent estimate, since this constitutes our primary
317 economic standing measure; 3) we winsorize Zillow rent estimates which are greater than \$20,000
318 to avoid spurious results from a small number of outliers; 4) we remove a small number of users
319 who are missing Census demographic information for their inferred home location; 5) we remove
320 users whose Zillow home location is further than 100 meters from their CoreLogic home location,
321 or whose CoreLogic home location is further than 100 meters from their nighttime latitude and
322 longitude; 6) we remove a small number of users in single family residences who are mapped to
323 the exact same single family residence as more than 10 other people, since this may indicate a data
324 error in the Zillow database.

325 The set of users who pass these filters constitute our final analysis set of 9,567,559 users. We
326 confirm that the Census demographic statistics of these users' inferred home locations are similar
327 to those of the US population in terms of income, age, sex, and race, suggesting that our inference
328 procedure yields a demographically representative sample.

329 Any individual quantitative measure provides only a partial picture of a person’s economic
330 standing. Recognizing this, we conduct robustness checks in which rather than using the Zillow
331 estimated rent of the user’s home as a proxy for economic standing, we use 1) the median Census
332 Block Group household income in that area; and 2) the percentile-scored rent of the home, to
333 account for long-tailed rent distributions. Our main results are robust to using these alternate
334 measures of economic standing (Supplementary Figure S3).

335 **Constructing interaction network**

336 We construct a fine-grained, dynamic interaction network \mathcal{G} between all 9,567,559 individuals
337 across 382 MSAs and 2829 counties, which is represented as an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$
338 with time-varying edges. Each node $v_i \in \mathcal{V}$ in the graph represents one of the $N = 9,567,559$
339 individuals in our study, such that the set of nodes is $V = \{v_1, v_2, \dots, v_N\}$. Each node v_i has a
340 single attribute x_i , representing the inferred economic standing (estimated rent) of the individual.

341 Individuals v_i and v_j are connected by one edge $e_{i,j,k} \in \mathcal{E}$ per interaction, with k indicating
342 the k th interaction between individuals v_i and v_j . Each edge $e_{i,j,k}$ has three attributes $t_{i,j,k}$, $lat_{i,j,k}$,
343 $lon_{i,j,k}$ indicating the timestamp, latitude, and longitude of the interaction respectively. We now
344 focus our discussion on explaining how each of the interactions edges of the network is calculated.

345 We define an *interaction* to occur when two users have GPS pings which are close (according
346 to a fixed threshold) in both physical proximity and time. Specifically if user v_i has a GPS ping
347 with t_i, lat_i, lon_i (indicating the timestamp, latitude, and longitude of the ping respectively), and
348 user v_j has a GPS ping with t_j, lat_j, lon_j , then we users are said to have shared an interaction if
349 $|t_i - t_j| < T$ and $distance((lat_i, lon_i), (lat_j, lon_j)) < D$, where T represents the time threshold
350 (i.e. maximum time distance the two pings can be apart to count as an interaction) and D represents
351 the distance threshold (i.e. maximum physical distance the two pings can be apart to count as an
352 interaction). We filter for both distance and time simultaneously to ensure that our interaction
353 network only includes pairs of users who are likely to have come into contact with each other. This
354 contrasts to other methods which consider all individuals to visit the same location, irrespective of
355 time²⁶, to have an equal likelihood of interaction, an assumption which may prove unrealistic in

many cities (e.g. demographics of individuals visiting public parks varies starkly by time of day⁵⁵). We use a threshold T of 5 minutes, which is a stringent threshold on time as the mean number of pings per person per hour during day time is approximately one ping; we use a distance threshold D of 50 meters, following prior work which shows that even exposure to individuals from afar is linked to long term outcomes¹⁸. Our network is validated by correlation to external, gold-standard datasets (Extended Data Figure 1). Furthermore, we show through a series of robustness checks that our key results in Figure 1, Figure 2, and Figure 3 are highly robust to varying thresholds (i.e. 1 minute or 2 minutes time threshold, as well as 10 meters or 25 meters distance threshold), as well as additional criteria to increase tie strength (i.e. requiring multiple consecutive interactions, or multiple interactions on unique days)—and under all observed circumstances the main findings remain consistent (see Supplementary Table S6, Supplementary Figures S2-S8).

To efficiently calculate the interactions between all users, we implement our interaction threshold as a k-d tree⁵⁷, a data structure which allows one to efficiently identify all pairs of points within a given distance of each other. In total, we identify 1,570,782,460 interactions. The timestamp $t_{i,j,k}$ of the interaction is the minimum ping timestamp in the pair of individuals' ping timestamps (t_i, t_j), and the location $lat_{i,j,k}, lon_{i,j,k}$ of the interaction is the average latitude and longitude of pair of pings belonging to the two individuals (lat_i, lat_j) and (lon_i, lon_j).

373 Annotating interactions

Interactions are annotated to indicate whether they occurred at or near features of interest: e.g., near a user's home. Annotations are not mutually exclusive in that an interaction may be simultaneously tagged as having occurred near multiple features from multiple data sources. We describe the specific annotations below.

We annotate a user's interaction as having occurred in their home if it occurs within 50 meters of the user's home location. An interaction is annotated with a TIGER road/railway if it occurs within 20 meters from that feature. An interaction is annotated as having occurred within a SafeGraph Places point-of-interest (POI) if the interaction occurs within the polygon defined for the POI. Polygons are provided by the SafeGraph Places database for both fine-grained POIs (e.g.

383 individual restaurants) as well as “parent” POIs (e.g. interaction hubs). We focus our analysis
384 of fine-grained POIs (Figure 1e, Extended Data Figure 2) on the most visited fine-grained POIs:
385 full-service restaurants, snack bars, limited-service restaurants (e.g. fast food), stadiums, etc (see
386 Figure 1e for full list). These categories roughly align with those used by prior work ⁵¹.

387 **M3 Analysis**

388 **Interaction Segregation**

389 We define the interaction segregation (IS) of a specified geographical area (i.e. Metropolitan Statis-
390 tical Area, County) as the Pearson correlation between the economic standing (ES) of individuals
391 residing in that geographical area, and the mean ES those that they come into contact with.

$$\text{Interaction Segregation} = \text{Corr}(ES, \overline{ES}_{\text{interactions}})$$

392 Our metric captures the extent to which an individual’s ES predicts the ES of their immediate
393 interaction network. Thus, in a perfectly integrated area in which individuals interact randomly
394 with others regardless of ES, interaction segregation would equal 0.0. In a perfectly segregated
395 area in which individuals interact with only those of the exact same ES, interaction segregation
396 would equal 1.0.

397 Interaction segregation nests a classic definition residential segregation, the Neighborhood
398 Sorting Index^{7,21} (NSI), which is equivalent to the Pearson correlation across between each per-
399 son’s ES and the mean ES in their Census tract. The NSI is widely used because it can be calculated
400 directly from Census data on the ES of people living in each tract. However, a fundamental limita-
401 tion of NSI as a measure segregation is that the Census tract in which people live is a weak proxy
402 for who they interact with. Census tracts are static and artificial boundaries which fail to capture
403 interactions as individuals move throughout the cityscape during work, leisure time, and schooling.

404 We design our interaction segregation (IS) metric such that it accommodates any interaction

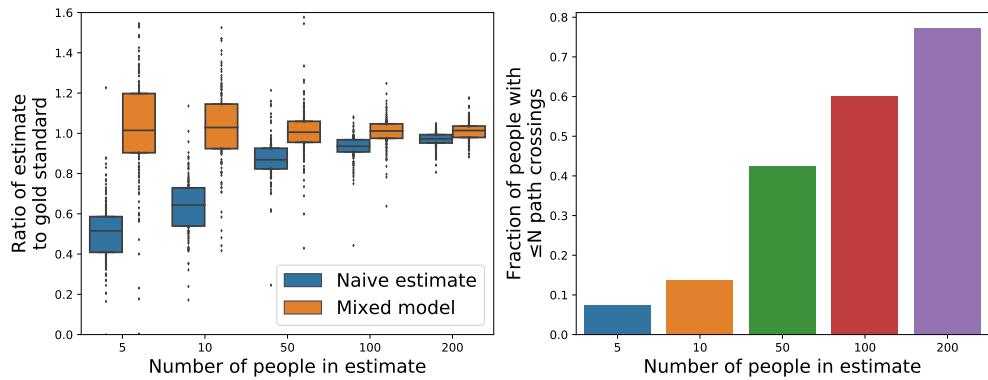
network, and thus NSI is a special case of our metric. Specifically, if interaction segregation is computed for a synthetic interaction network under the unrealistic assumptions that a) people only interact with those in their home Census tract and b) they do so uniformly at random—then it is equivalent to NSI (Supplementary Figure S17). However, constructing such a synthetic interaction network from Census tracts has limited applicability to measuring segregation in the real world, because people may also interact with more heterogeneous populations as they visit other Census tracts for work, leisure, or other activities, a phenomenon we refer to as the *visitor effect*. Furthermore, even within home tract, individuals may interact non-uniformly as they seek out people of similar economic standing; we refer to this as the *homophily effect*. Thus, we instead leverage dynamic mobility data from cell phones to capture the extent of contact between diverse individuals throughout the day, and apply our metric, interaction segregation (IS), to this real-world interaction network. An advantage of interaction segregation is that it allows for direct comparability to NSI, because both measures are of the same underlying statistical quantity, but differ in their definition of the interaction network. Our results indicate that this choice of interaction network matters; IS is a stronger predictor of upward economic mobility (Extended Data Figure 1) as the two metrics are shown to be distinct (Supplementary Figure S18).

To calculate the interaction segregation of a specified geographical area (i.e. Metropolitan Statistical Area, County), we first select the set of all individuals who reside in area: $\mathcal{V}_A \subset \mathcal{V}$. For instance, to calculate Interaction Segregation for Napa, California (Figure 1c Top), \mathcal{V}_A is the 3707 individuals with home locations inside the geographical boundary of the Napa, CA MSA. Subsequently, for each individual resident of the area $v_i \in \mathcal{V}_A$ we query the population interaction network ($\mathcal{G} = (\mathcal{V}, \mathcal{E})$) for the ES of the set of individuals they interact with, \mathcal{Y}_i : $\{x_j \in \mathcal{V} | e_{i,j,k} \in \mathcal{E}\}$. We then aim to estimate the Pearson correlation between the ES of each individual x_i and the mean ES of those they interact with, $y_i = \text{mean}(\mathcal{Y}_i)$.

429 Mixed model

We use a linear mixed effects model to accurately estimate interaction segregation: the Pearson correlation coefficient between a person's ES and the mean ES of the people they interact with. A

432 statistical model is required to estimate interaction segregation because naively computing the cor-
 433 relation based on limited data (in counties or MSAs with low population sizes) results in estimates
 434 that are downward biased.¹ By contrast, our linear mixed effects model is an unbiased estimator of
 435 the Pearson correlation. We compare the unbiased estimates from our linear mixed effects model
 436 to naive estimates of interaction segregation in Methods Figure 1. To illustrate why naive estimates
 437 of interaction segregation are downward biased, imagine that we compute the correlation between
 438 a person’s ES and the “true” mean ES of the people they interact with. Now, we add noise to the
 439 mean ES values, which represents the noisy mean estimates given limited data. As the noise is
 440 increased, the correlation is decreased. Thus, because estimates of each person’s mean ES will be
 441 more noisy in geographical areas with less data, there will be a downward bias to naive estimates
 442 of the Pearson correlation in these areas.



Methods Figure 1: Mixed model estimate compared with naive estimates of the Pearson correlation. We took people who interacted with at least 500 other people and computed the Pearson correlation coefficient (the “gold standard estimate”). Then, for each person we randomly sampled 5, 10, 50, 100, and 200 people from the 500+ people and computed segregation estimates based on the reduced sets of people. The left plot shows the ratio of the estimates to the gold standard, for each MSA. The right plot shows the overall number of people in the dataset with $\leq N$ interactions.

443 Our mixed model models the distribution of datapoints (x_i, y_{ij}) through the following equa-
 444 tion:

¹By “naive” estimation of the Pearson correlation, we intend to convey calculating the correlation using the sample:

$$\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}}$$

$$y_{ij} = ax_i + b + \epsilon_i^{(1)} + \epsilon_{ij}^{(2)}$$

where x_i = ES of person i

y_{ij} = ES of person j who has interacted with person i

a, b = model parameters

$\epsilon_i^{(1)}$ = person-specific noise term

$\epsilon_{ij}^{(2)}$ = noise for each data point

445 The Pearson correlation coefficient between person i 's ES and the mean ES of the people
 446 they interact with is then computed as follows. We assume that x_i has a variance of 1 through data
 447 preprocessing and that x_i is uncorrelated with $\epsilon_i^{(1)}$.

$$\begin{aligned} \text{corr}\left(x_i, ax_i + b + \epsilon_i^{(1)}\right) &= \text{corr}\left(x_i, ax_i + \epsilon_i^{(1)}\right) \\ &= \frac{\text{cov}\left(x_i, ax_i + \epsilon_i^{(1)}\right)}{\sqrt{\text{Var}(x_i) \text{Var}\left(ax_i + \epsilon_i^{(1)}\right)}} \\ &= \frac{\text{cov}\left(x_i, ax_i\right)}{\sqrt{\text{Var}\left(ax_i + \epsilon_i^{(1)}\right)}} \\ &= \frac{a}{\sqrt{a^2 + \text{Var}\left(\epsilon_i^{(1)}\right)}} \end{aligned}$$

448 We estimate a and $\text{Var}\left(\epsilon_i^{(1)}\right)$ by fitting the mixed model using the R lme4 package, optimiz-
 449 ing the restricted maximum likelihood (REML) objective.

450 Decomposing segregation by time

451 Each interaction edge $(e_{i,j,k})$ in our interaction network is timestamped with a time of interaction
 452 $t_{i,j,k}$. This allows us to decompose our overall interaction segregation into fine-grained estimates

453 of segregation during different hours of the day, by filtering for interactions that occurred within
454 a specific hour. In Supplementary Figures S19, we partition estimates of segregation by 3 hour
455 windows to illustrate how segregation varies throughout the day (see Supplementary Information).

456 **Decomposing segregation by activity**

457 Each interaction edge ($e_{i,j,k}$) in our interaction network occurs at a specific location $lat_{i,j,k}, lon_{i,j,k}$.
458 Thus, it is possible to annotate interactions by the fine-grained POI (e.g. specific restaurant) they
459 occurred in, as well as the by the higher-level “parent” POI (e.g. commercial center) in which the
460 POI was located (Methods M2). This allows us to decompose our overall interaction segregation
461 into fine-grained estimates of segregation by specific leisure activity. We do so by filtering the
462 network for all interactions that occurred in a specific POI category, and re-calculating interaction
463 segregation for the MSA or county, using only those interactions. In Figure 1e, we show the
464 variation in interaction segregation by leisure site, and further explain these variations in Extended
465 Data Figure 2.

466 **Bridging Index**

467 We seek to identify a modifiable, extrinsic aspect of a city’s built environment which may reduce
468 interaction segregation. One promising candidate is the location of a city’s hubs of interaction. We
469 define a new measure, the *Bridging Index* (BI), which measures the extent to which a particular set
470 of interaction hubs (i.e. high-interaction POIs, \mathcal{P}) facilitate the integration of individuals of diverse
471 economic standing within a geographic area (i.e. MSA or county). Specifically, BI measures the
472 economic diversity of the groups that would interact if everybody visited only their nearest hub
473 from \mathcal{P} —based on the observation that physical proximity significantly influences which hubs
474 individuals visit^{35–37}.

475 We compute the Bridging Index (BI) via two steps (Extended Data Figure 6).

476 1. Cluster all individuals who live in an area (i.e. MSA or county residents, \mathcal{V}_A) into K clusters
477 ($\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_K$) according to the interaction hub from \mathcal{P} closest to their home location. K is
478 the number of hubs in \mathcal{P} .

479 2. Bridging Index is computed as the weighted average of the economic diversity (i.e. Gini

480 Index) of these clusters of people, relative to the area's overall economic diversity.

$$\text{Bridging Index (BI)} = \frac{\text{Within Hub Economic Diversity}}{\text{Overall Economic Diversity}} = \frac{\sum_{i=1}^K |\mathcal{C}_i| \cdot \text{Gini Index}(\mathcal{C}_i)}{|\mathcal{V}_A| \cdot \text{Gini Index}(\mathcal{V}_A)}$$

481 We illustrate the intuition for BI and how it captures the relationship between home and hub
482 locations in Extended Data Figure 7. A BI of 1.0 indicates that if everybody visits their nearest
483 interaction hub, each person will be exposed to a set of people *as economically diverse as the*
484 *overall city they reside in*. Thus, a BI of 1.0 signifies perfect bridging, i.e. even if individuals live
485 in segregated neighborhoods, hubs are located such that individuals must leave their neighborhoods
486 and interact with diverse others. On the other hand, a BI of 0.0 signifies the opposite extreme; a
487 city with a BI of 0.0 is one in which, if everybody visits the nearest interaction hub, each person
488 will be exposed to only people of the exact same economic standing.

489 The economic diversity of each cluster \mathcal{C}_i is quantified using the Gini Index: $\text{Gini Index}(\mathcal{C}_i)$,
490 a well-established measure of economic statistical dispersion (Extended Data Figure 6c)⁶³, al-
491 though results are robust to choice of economic diversity measure such as using variance instead
492 of Gini Index (Supplementary Figure S12). The denominator of BI normalizes for the baseline
493 economic diversity observed in the city, allowing for direct comparisons between cities.

494 In our primary analysis, we identify hubs of interaction via commercial centers (e.g. shop-
495 ping malls, plazas, etc. which are higher-level clusters of individual POIs) because they are associ-
496 ated with a high density of interactions. Specifically, the majority (56.9%) of interactions happen
497 inside of or within 1km of a commercial center (e.g. shopping mall, plaza, etc.) even though only
498 2.5% of the land area of MSAs is within 1km of a commercial center. We thus compute BI using
499 the set \mathcal{P} of all commercial centers within each MSA. We discover that BI strongly predicts inter-
500 action segregation (Spearman Correlation -0.78 , Figure 3d). The top 10 MSAs with the highest
501 BI are 53.1% less segregated than the 10 MSAs with the lowest BI. BI predicts segregation more
502 accurately than population size, racial demographics ES inequality, NSI, and racial demographics,

503 and is significantly associated with segregation ($p < 10^{-8}$) after controlling for all aforementioned
504 variables (Extended Data Tables [2-3](#)).

505 **Hypothesis Testing and Confidence Intervals**

506 Unless otherwise noted, confidence intervals and hypothesis tests were conducted using a bootstrap
507 with 10,000 replications⁶⁴. Steiger's Z-test was used to compare different predictors of segregation
508 indices, and hypothesis tests for Spearman correlation coefficients were computed using two-sided
509 Student's t-tests⁶⁵⁻⁶⁷.

Methods References

40. Safegraph. Safegraph dataset (2017).
41. Squire, R. F. What about bias in the safegraph dataset? *Safegraph* (2019).
42. Chang, S. *et al.* Mobility network models of covid-19 explain inequities and inform reopening. *Nature* **589**, 82–87 (2021).
43. Chen, M. K. & Rohla, R. The effect of partisanship and political advertising on close family ties. *Science* **360**, 1020–1024 (2018).
44. Athey, S., Blei, D., Donnelly, R., Ruiz, F. & Schmidt, T. Estimating heterogeneous consumer preferences for restaurants and travel time using mobile location data. In *AEA Papers and Proceedings*, vol. 108, 64–67 (2018).
45. SafeGraph. Determining point-of-interest visits from location data: A technical guide to visit attribution. Tech. Rep. (2020).
46. CoreLogic, I. Corelogic real estate database (2017).
47. Zillow, I. Zillow rent zestimate (2017).
48. Kelton, C. M., Pasquale, M. K. & Rebelein, R. P. Using the north american industry classification system (naics) to identify national industry cluster templates for applied regional analysis. *Regional Studies* **42**, 305–321 (2008).
49. Bureau, U. C. American community survey 5-year estimates (2017). Accessed August 30, 2017.
50. US Census Bureau. Glossary.
51. Athey, S., Ferguson, B., Gentzkow, M. & Schmidt, T. Experienced segregation. *NBER* (2020).
52. Bureau, U. C. Tiger/line shapefiles (2017). Accessed August 30, 2017.
53. SafeGraph. Determining patterns. Tech. Rep. (2020). Available at: <https://docs.safegraph.com/docs/monthly-patterns>.
54. Moro, E., Calacci, D., Dong, X. & Pentland, A. Mobility patterns are associated with experienced income segregation in large us cities. *Nature communications* **12**, 1–10 (2021).
55. Madge, C. Public parks and the geography of fear. *Tijdschrift voor economische en sociale geografie* **88**, 237–250 (1997).
56. Brown, J. R., Enos, R. D., Feigenbaum, J. & Mazumder, S. Childhood cross-ethnic exposure predicts political behavior seven decades later: Evidence from linked administrative data. *Science Advances* **7**, eabe8432 (2021).
57. Bentley, J. L. Multidimensional binary search trees used for associative searching. *Communications of the ACM* **18**, 509–517 (1975).

58. Jargowsky, P. A. Take the money and run: Economic segregation in us metropolitan areas. *American Sociological Review* 984–998 (1996).
59. Jargowsky, P. A. & Kim, J. A measure of spatial segregation: The generalized neighborhood sorting index. *National Poverty Center Working Paper Series* (2005).
60. Zipf, G. K. The p 1 p 2/d hypothesis: on the intercity movement of persons. *American sociological review* **11**, 677–686 (1946).
61. Simini, F., González, M. C., Maritan, A. & Barabási, A.-L. A universal model for mobility and migration patterns. *Nature* **484**, 96–100 (2012).
62. Schläpfer, M. *et al.* The universal visitation law of human mobility. *Nature* **593**, 522–527 (2021).
63. Dorfman, R. A formula for the gini coefficient. *The review of economics and statistics* 146–149 (1979).
64. Efron, B. & Tibshirani, R. J. *An introduction to the bootstrap* (CRC press, 1994).
65. Myers, L. & Sirois, M. J. Spearman correlation coefficients, differences between. *Encyclopedia of statistical sciences* **12** (2004).
66. Steiger, J. H. Tests for comparing elements of a correlation matrix. *Psychological bulletin* **87**, 245 (1980).
67. Corder, G. W. & Foreman, D. I. *Nonparametric statistics: A step-by-step approach* (John Wiley & Sons, 2014).
68. Bailey, M., Cao, R., Kuchler, T., Stroebel, J. & Wong, A. Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives* **32**, 259–80 (2018).
69. Bailey, M., Farrell, P., Kuchler, T. & Stroebel, J. Social connectedness in urban areas. *Journal of Urban Economics* **118**, 103264 (2020).
70. Bailey, M. *et al.* International trade and social connectedness. *Journal of International Economics* **129**, 103418 (2021).
71. Kuchler, T., Li, Y., Peng, L., Stroebel, J. & Zhou, D. Social proximity to capital: Implications for investors and firms. Tech. Rep., National Bureau of Economic Research (2020).
72. Kuchler, T., Russel, D. & Stroebel, J. Jue insight: The geographic spread of covid-19 correlates with the structure of social networks as measured by facebook. *Journal of Urban Economics* 103314 (2021).
73. Chetty, R., Hendren, N., Kline, P. & Saez, E. Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics* **129**, 1553–1623 (2014).
74. Economist, T. The big mac index (2021). Accessed Sep 29th, 2021.

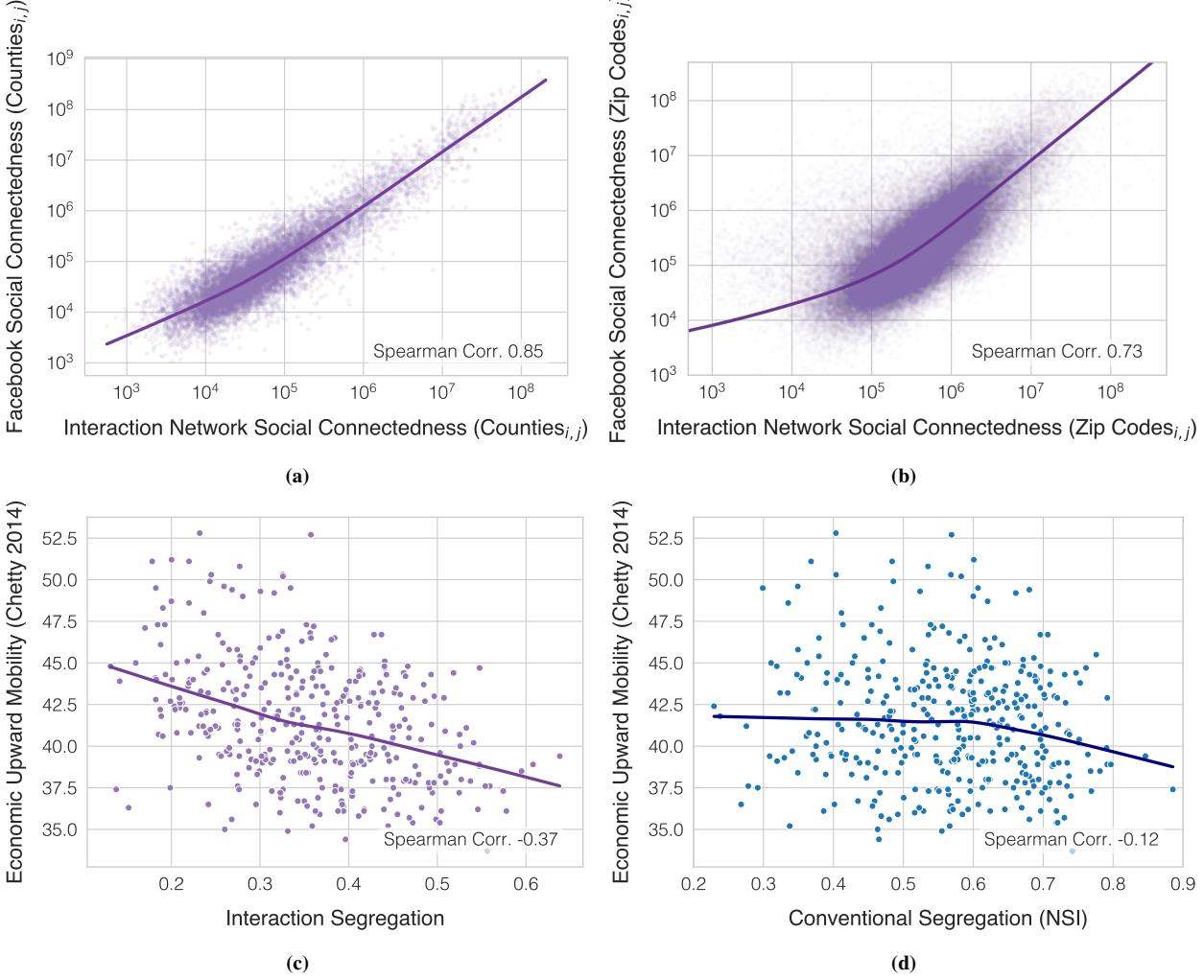
75. Pencil, C. How much to dine out at a top michelin-starred restaurant (2021). Accessed Sep 29th, 2021.
76. Walkscore. Walkscore walkability estimates (2020). Accessed August 30, 2020.
77. Merry, K. & Bettinger, P. Smartphone gps accuracy study in an urban environment. *PLoS one* **14**, e0219890 (2019).
78. Menard, T., Miller, J., Nowak, M. & Norris, D. Comparing the gps capabilities of the samsung galaxy s, motorola droid x, and the apple iphone for vehicle tracking using freesim_mobile. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 985–990 (IEEE, 2011).
79. Barabási, A.-L. Network science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **371**, 20120375 (2013).
80. Maslov, S. & Sneppen, K. Specificity and stability in topology of protein networks. *Science* **296**, 910–913 (2002).
81. Dunbar, R. *How many friends does one person need?: Dunbar's number and other evolutionary quirks* (Faber & Faber, 2010).

510 Extended Data

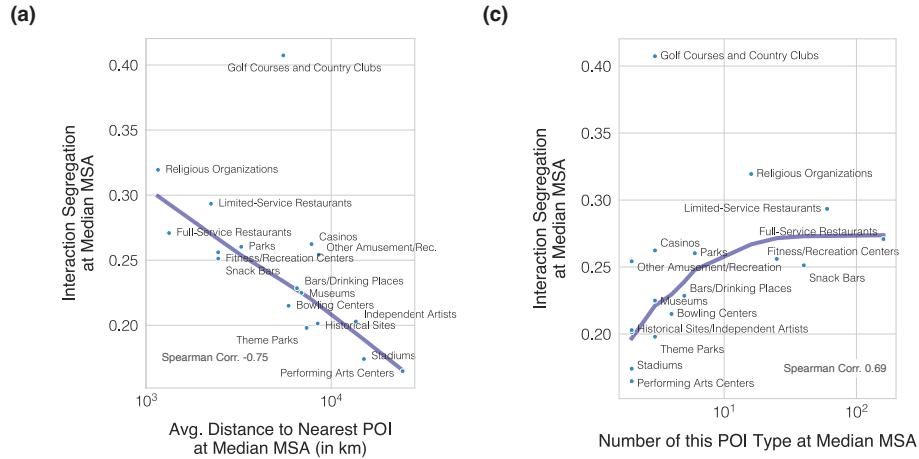
	<i>Dependent variable:</i>					Interaction Segregation
	(1)	(2)	(3)	(4)	(5)	
Intercept	0.355*** (0.004)	0.355*** (0.004)	0.355*** (0.003)	0.355*** (0.003)	0.355*** (0.003)	0.355*** (0.003)
Log(Population Size)	0.059*** (0.004)		0.041*** (0.004)	0.044*** (0.004)	0.026*** (0.004)	0.028*** (0.004)
Gini Index (Estimated Rent)		0.064*** (0.004)	0.050*** (0.004)	0.051*** (0.004)	0.045*** (0.003)	0.047*** (0.003)
Political Alignment (% Democrat in 2016 Election)				0.004 (0.004)		0.004 (0.004)
Racial Demographics (% non-Hispanic White)				0.001 (0.004)		0.006* (0.003)
Mean ES (Estimated Rent)				-0.012*** (0.004)		-0.005 (0.004)
Walkability (Walkscore)					0.002 (0.003)	0.001 (0.003)
Commutability (% Commute to Work)					-0.011*** (0.003)	-0.010*** (0.004)
Conventional Segregation (NSI)					0.042*** (0.003)	0.041*** (0.003)
Observations	382	382	382	376	382	376
<i>R</i> ²	0.350	0.419	0.567	0.578	0.704	0.705
Adjusted <i>R</i> ²	0.348	0.417	0.565	0.573	0.701	0.698

*p<0.1; **p<0.05; ***p<0.01

Extended Data Table 1: Population size is significantly associated with interaction segregation, after controlling for MSA income inequality (Gini Index), political alignment (% Democrat in 2016 election), racial demographics (% non-Hispanic White), mean ES, walkability (Walkscore⁷⁶), commutability (% of residents commuting to work), and residential segregation (NSI). Here we show the coefficients (after normalizing via z-scoring to have mean 0 and variance 1) from the primary specifications estimating the effect of population size on interaction segregation across all MSAs. Columns (1-5) are models specified with different subsets of covariates; Column 6 shows model specification with all covariates. Differences between sample size in models is due to missing data for several covariates in a small number of MSAs (Walkscores were not available for all MSAs). (*p < 0.1; **p < 0.05; *** p < 0.01).



Extended Data Figure 1: This studies' interaction network predicts population-scale friendship formation and upward economic mobility outcomes. We measure the external validity of our definition of interaction, by linking our interaction network to outcomes across two gold-standard, large-scale, datasets. We find at the zip code, county, and MSA-level, our interaction network mirrors population-scale outcomes resulting from dynamic human processes: **(a-b)** the Facebook Social Connectedness Index⁶⁸ measures the relative probability of a Facebook friendship link between a given Facebook user in location i and a given user in location j . FB Social Connectedness Index has been used social segregation⁶⁹, and has also been linked to economic^{70,71} and public health outcomes⁷². We reproduce the Social Connectedness Index using our interaction network ($\frac{\# \text{Interaction Pairs}_{i,j}}{\#\text{Individuals}_i \cdot \#\text{Individuals}_j}$) at the county **(a)** and zip code **(b)** level, and find strong correlations across county pairs (Spearman Correlation 0.85, $N = 121,595$, $p < 10^{-4}$) and zip code pairs (Spearman Correlation 0.73, $N = 1,053,539$, $p < 10^{-4}$). **(c-d)** The Chetty et al. Intergenerational Mobility dataset quantifies upward economic mobility from federal income tax records for each MSA as the mean income rank of children with parents in the bottom half of the income distribution⁷³. We find that interaction segregation at the MSA-level **(c)** correlates to (absolute) upward economic mobility (Spearman Correlation -0.37, $N = 379$, $p < 10^{-4}$), and does so significantly more strongly ($p < 10^{-4}$) than **(d)** the conventional segregation measure NSI (Spearman Correlation -0.12, $N = 379$, $p < 0.05$)

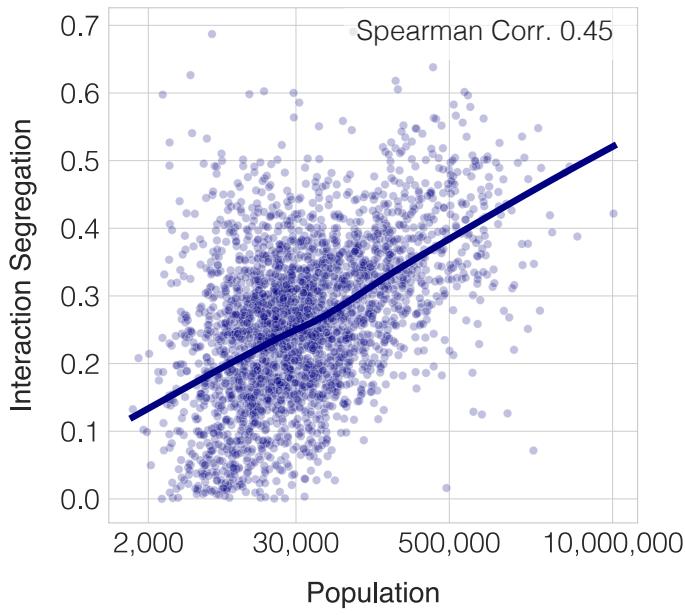


(c)

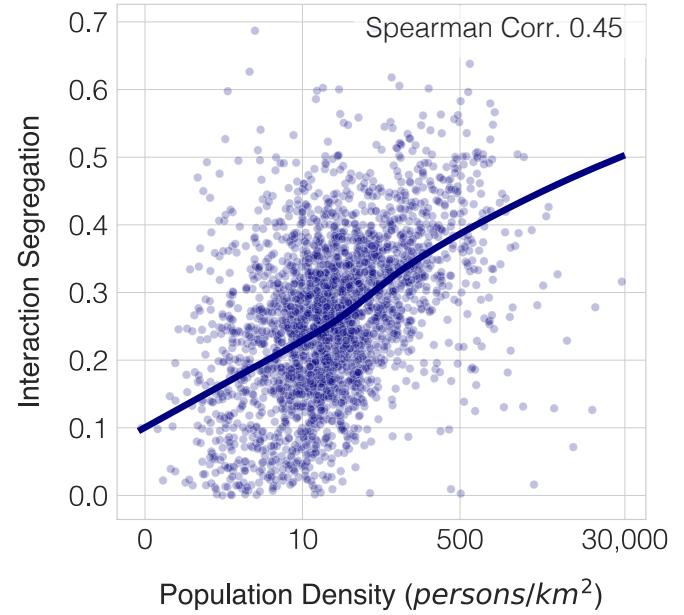
Metropolitan Statistical Area (MSA)	POI Name	ES	Minimum Cost of Entrance (\$)
Chicago-Naperville-Elgin, IL-IN-WI	Wynstone Golf Club	High	40,000
	Flagg Creek Golf Course	Low	19
Los Angeles-Long Beach-Anaheim, CA	Coto de Caza Golf & Racquet Club	High	45,000
	Rancho Vista Golf Club	Low	129
Miami-Fort Lauderdale-West Palm Beach, FL	Old Palm Golf Club	High	199,200
	Eco Golf Club	Low	17
New York-Newark-Jersey City, NY-NJ-PA	Scarsdale Golf Club	High	8,900
	Weequahic Park Golf Course	Low	35
Phoenix-Mesa-Scottsdale, AZ	The Estancia Club	High	150,000
	Peoria Pines	Low	28

Extended Data Figure 2: Understanding why interaction segregation varies significantly across leisure sites. We identify three primary facets of socioeconomic differentiation between POIs which explain the heterogeneous segregation levels of different leisure POIs (Figure 1e): **(a)** localization, **(b)** quantity, and **(c)** stratification. **(a)** Localization strongly predicts segregation across all POI categories (Spearman Correlation -0.75, N=17, p<0.001). POIs which are more locally embedded into neighborhoods (e.g. religious organizations) are more segregated than activities in which POIs serve multiple neighborhoods (e.g. stadiums). We operationalize localization as the average distance from each individual in the MSA to the nearest POI of that category. **(a)** The quantity of POIs also explains segregation (Spearman Correlation 0.69, N=17, p<0.01). Leisure activities with more options (e.g. restaurants) have differentiated venues catering to a specific economic standing (e.g. Michelin-star restaurants) compared to POIs which are small in number and cater to the overall city (e.g. stadiums) **(c)** Golf courses and country clubs (golf clubs) are an anomaly in that they have a small number of unlocalized POIs, but are highly segregated. We conduct a case study in which look at top and bottom golf clubs by mean visitor ES in 5 of the 10 largest MSAs. We find that the high segregation of golf clubs is due to extreme stratification between venues; for instance the minimum cost to play at the high-ES golf course in Miami, FL is 11717× higher than at the lowest-ES golf course. By contrast, the average cost of a MacDonalds Big Mac (\$5.65⁷⁴) is only 63× higher than the average cost of a Michelin 3-star restaurant (\$357⁷⁵). Finally, these findings foreshadow Bridging Index (BI), which captures POI localization, quantity, and stratification (Extended Data Figure 7).

(a)



(b)



Extended Data Figure 3: Large, dense counties are more segregated. We compute interaction segregation across 2829 USA counties (94% of the counties in the USA), excluding counties in which there are less than 50 individuals in our dataset. We find that at the county-level, interaction segregation is also positively correlated with population size (Spearman Correlation 0.45, N=2829, $p < 10^{-4}$) and population density (Spearman Correlation 0.45, N=2829, $p < 10^{-4}$). These correlations reveal that the association between large, dense cities and interaction segregation (Figure 2a) is not an artifact of city boundaries, and may in fact be an emergent property from dynamics of individuals residing highly populated, dense geographic areas, which persists across multiple scales of granularity.

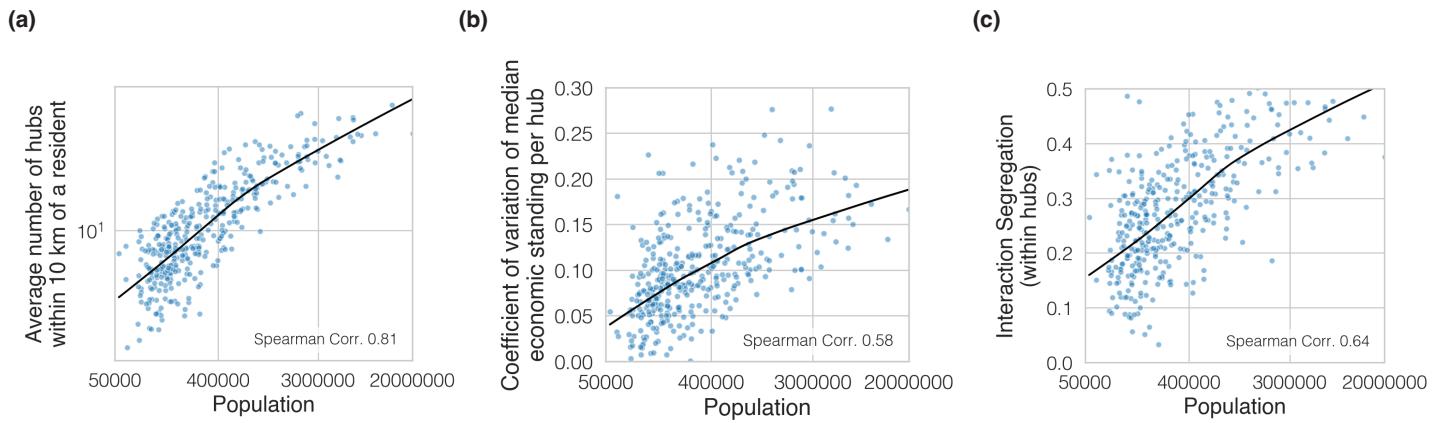
	<i>Dependent variable:</i>				
	(1)	(2)	(3)	(4)	(5)
Intercept	0.355*** (0.003)	0.355*** (0.003)	0.355*** (0.003)	0.355*** (0.003)	0.355*** (0.003)
Bridging Index	-0.078*** (0.003)	-0.059*** (0.005)	-0.058*** (0.005)	-0.035*** (0.006)	-0.036*** (0.006)
Log(Population Size)		0.003 (0.004)	0.008* (0.005)	0.010** (0.004)	0.017*** (0.006)
Gini Index (Estimated Rent)		0.031*** (0.003)	0.032*** (0.003)	0.035*** (0.003)	0.036*** (0.003)
Political Alignment (% Democrat in 2016 Election)			0.001 (0.004)		0.002 (0.004)
Racial Demographics (% non-Hispanic White)			0.003 (0.003)		0.005 (0.003)
Mean ES (Estimated Rent)			-0.009** (0.004)		-0.005 (0.003)
Walkability (Walkscore)				0.002 (0.003)	0.001 (0.003)
Commutability (% Commute to Work)				-0.011*** (0.003)	-0.009** (0.004)
Conventional Segregation (NSI)				0.028*** (0.004)	0.026*** (0.004)
# of Interaction Hubs					-0.006 (0.005)
Observations	382	382	376	382	376
R ²	0.620	0.686	0.693	0.733	0.736
Adjusted R ²	0.619	0.684	0.688	0.729	0.729

*p<0.1; **p<0.05; ***p<0.01

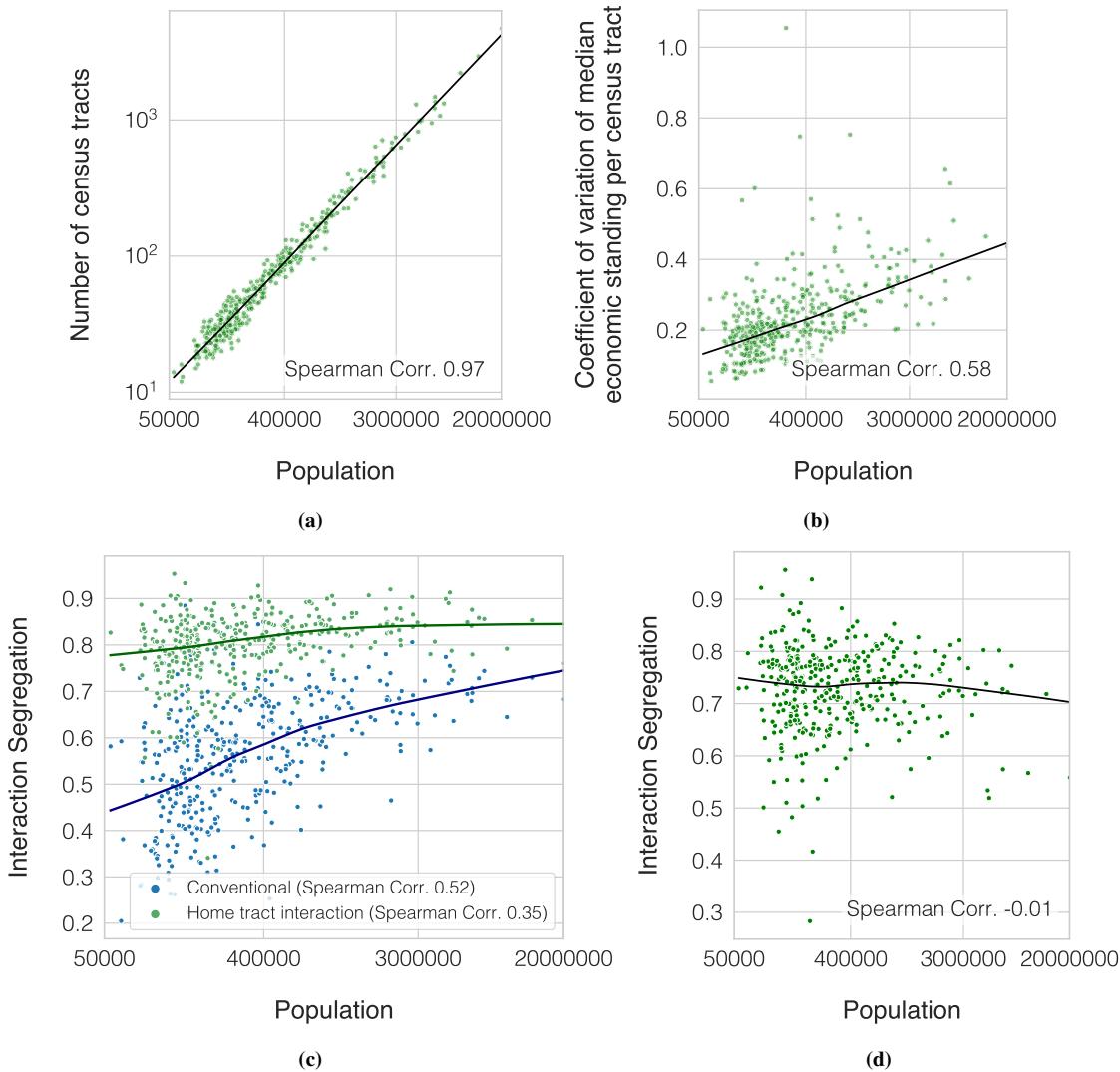
Extended Data Table 2: Bridging Index (BI) is significantly associated with interaction segregation, after controlling for population size, # of hubs, MSA income inequality (Gini Index), political alignment (% Democrat in 2016 election), racial demographics (% non-Hispanic White), mean ES, walkability (Walkscore⁷⁶), commutability (% of residents commuting to work), and residential segregation (NSI). Here we show the coefficients (after normalizing via z-scoring to have mean 0 and variance 1) from the primary specifications estimating the effect of population size on interaction segregation across all MSAs. Columns (1-4) are models specified with different subsets of covariates; Column 5 shows model specification with all covariates. Differences between sample size in models is due to missing data for several covariates in a small number of MSAs (Walkscores were not available for all MSAs). (*p < 0.1; **p < 0.05; *** p < 0.01).

Measure	Spearman ρ^2	Pearson R^2
Bridging Index	0.60	0.62
Log(Population Size)	0.39	0.35
Gini Index (Estimated Rent)	0.41	0.42
Political Alignment (% Democrat in 2016 Election)	0.06	0.05
Racial Demographics (% non-Hispanic White)	0.09	0.05
Mean ES (Estimated Rent)	0.09	0.05
Walkability (Walkscore)	0.01	0.02
Commutability (% Commute to Work)	0.04	0.03
Conventional Segregation (NSI)	0.44	0.42
# of Interaction Hubs	0.44	0.16

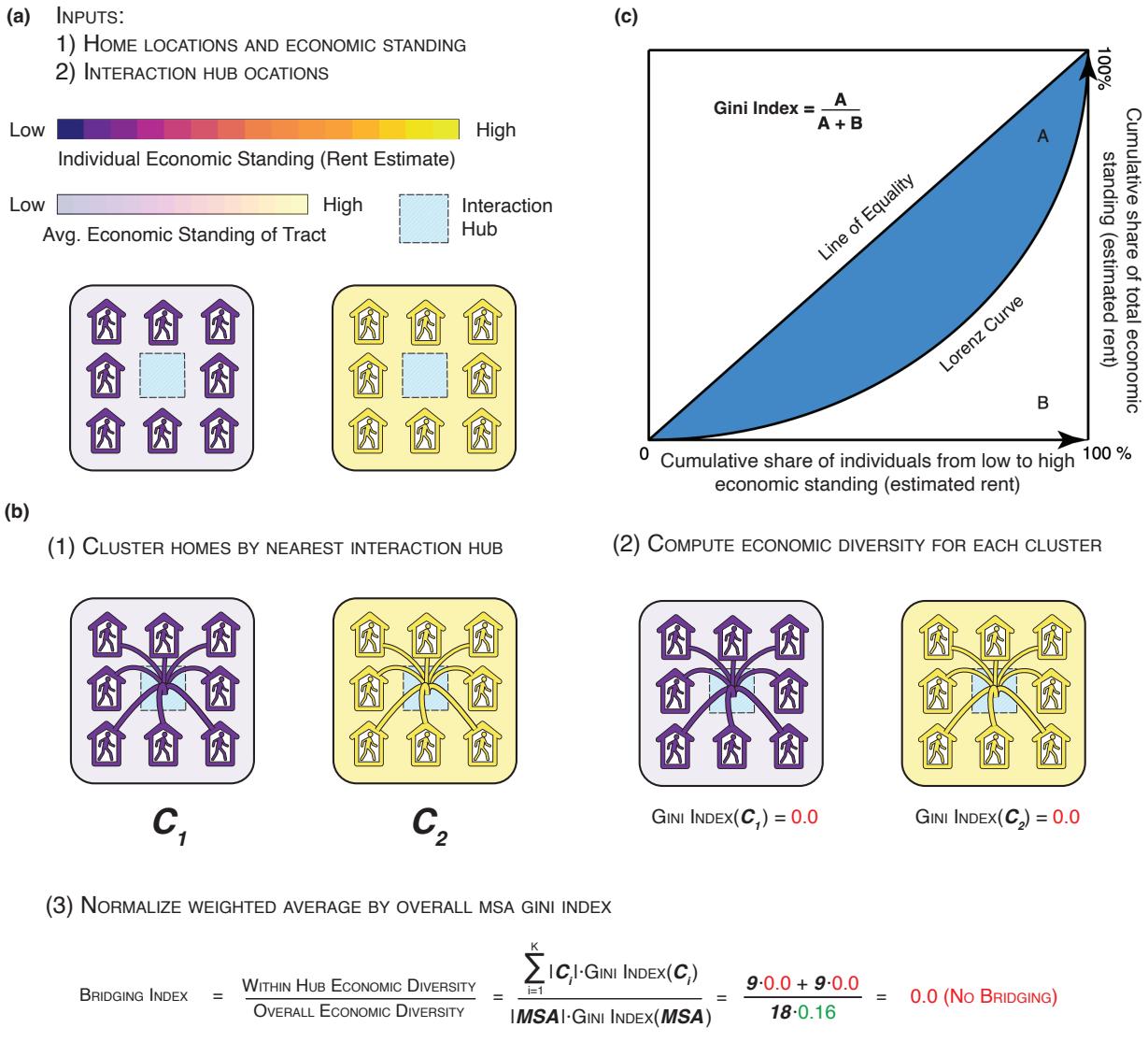
Extended Data Table 3: Bridging Index (BI) strongly predicts interaction segregation and does so more accurately ($p < 10^{-4}$, Steiger's Z-test) than population size, # of hubs, MSA income inequality (Gini Index), political alignment (% Democrat in 2016 election), racial demographics (% non-Hispanic White), mean ES, walkability (Walkscore⁷⁶), commutability (% of residents commuting to work), and residential segregation (NSI)



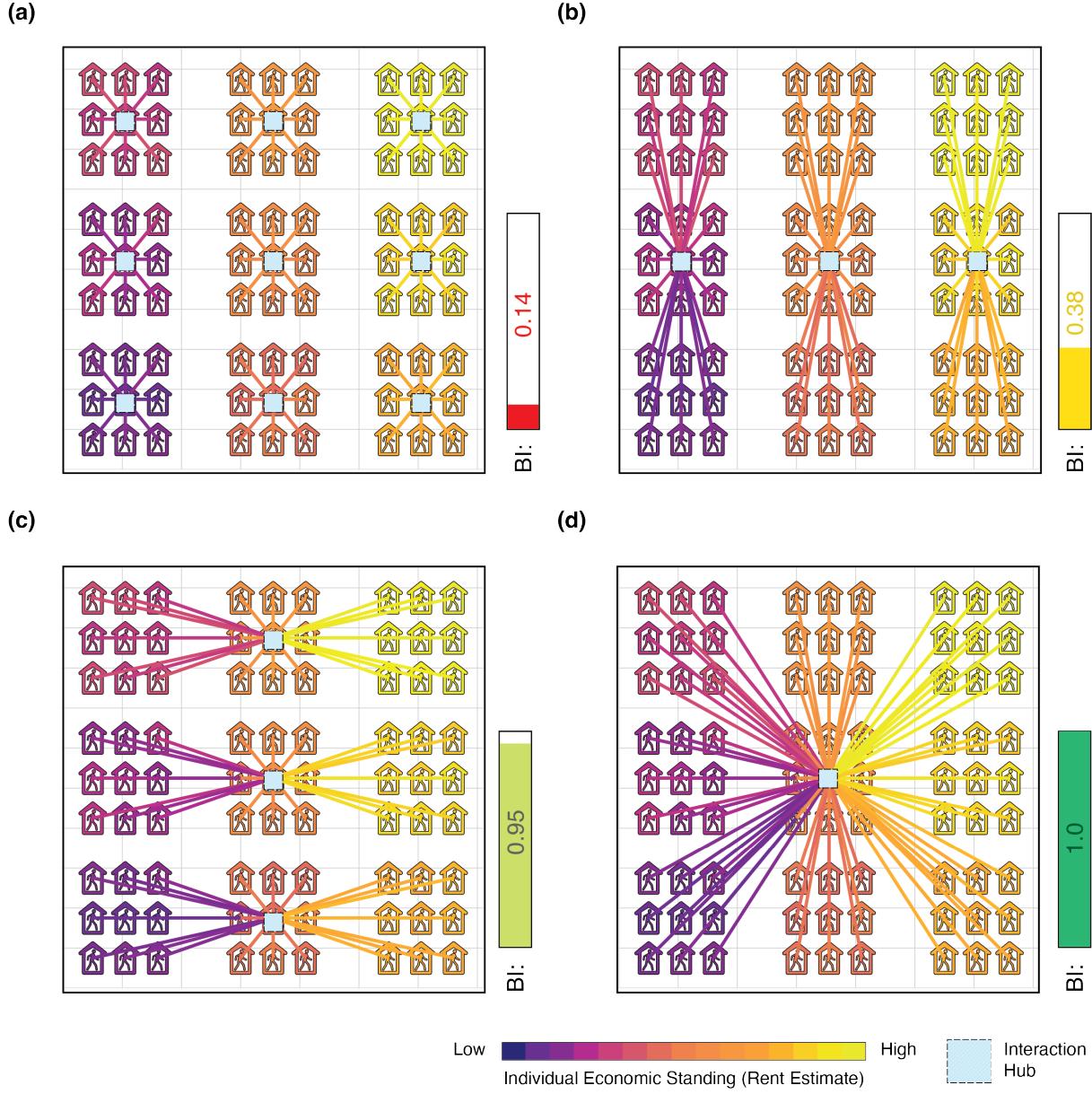
Extended Data Figure 4: At higher levels of scale, spaces in large cities are more differentiated and consequently segregated: interaction hubs. (a-c) Conducting an analogous analysis to that for restaurants in Figure 3c-e for interaction hubs (i.e. commercial centers which are higher-level clusters of restaurants, grocery stores, etc.). We find that higher segregation is driven by an increase in highly differentiated choice of interaction hubs in large cities: (a) Larger MSAs have more interaction hubs, giving residents more options to self-segregate (Spearman Correlation 0.81, N=382, $p < 10^{-4}$). (b) Consequently, hubs in larger MSAs vary more in terms of the mean ES of their visitors (Spearman Correlation 0.58, N=382, $p < 10^{-4}$) and as a result, (c) interaction segregation within hubs is higher in larger MSAs (Spearman Correlation 0.64, N=382, $p < 10^{-4}$). Overall, this analysis suggests that across multiple levels of scale, large cities offer a greater choice of differentiated spaces targeted to specific socioeconomic groups, promoting everyday segregation in interactions.



Extended Data Figure 5: At higher levels of scale, spaces in large cities are more differentiated and consequently segregated: home neighborhoods. (a-c) Conducting an analogous analysis to that for restaurants in Figure 3c-e, we find that higher segregation is driven by an increase in highly differentiated choice of neighborhoods in large cities: (a) Larger MSAs have more census tracts, giving residents more options to self-segregate (Spearman Correlation 0.97, N=382, $p < 10^{-4}$). (b) Consequently, census tracts in larger MSAs vary more in terms of the mean ES of their residents (Spearman Correlation 0.58, N=382, $p < 10^{-4}$) and as a result, (c) both conventional NSI and interaction segregation are higher (Spearman Correlations 0.52 and 0.35, N=382, $p < 10^{-4}$ and $p < 10^{-4}$). However, (c) also shows that interaction segregation (green series) rises more slowly with population than conventional segregation (blue series), suggesting that within-home-tract homophily, which increases interaction segregation but not conventional segregation, is *not* more pronounced in large MSAs. Substantiating this, (d) shows that when home tract interaction segregation is computed using an alternate ES measure so it captures only within-home-tract-homophily, it is no higher in large MSAs (Spearman Correlation -0.01, N=382, $p > 0.1$). (The alternative ES measure is computed by subtracting the mean ES in each Census tract; see Methods.) Overall, this analysis suggests that the higher home tract segregation in large MSAs is driven by people's greater choice of neighborhoods of varying ES in which to live, but not by a greater tendency to interact homophilously within their own neighborhood.



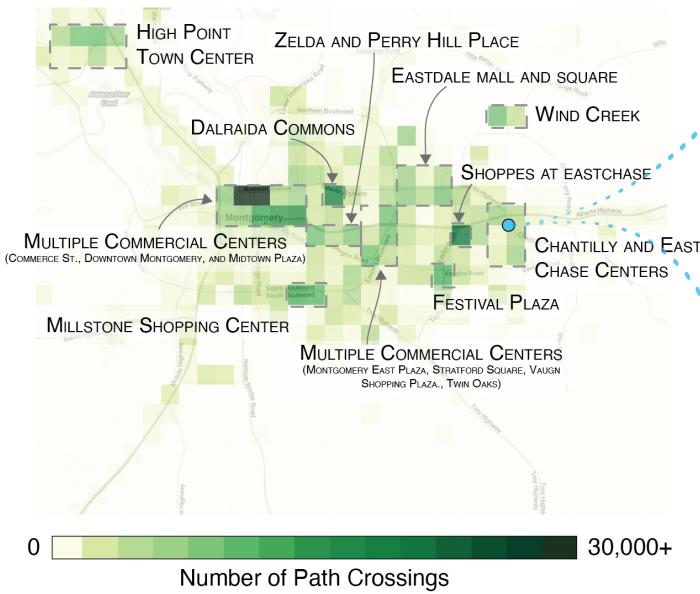
Extended Data Figure 6: Computing Bridging Index (BI). Illustration of our analytical pipeline for calculating BI. (a) BI is computed from the locations and number of POIs in the MSA which are expected to be hubs of interaction, as well the locations and economic standing values of all homes within MSA boundaries. We intentionally develop BI without using mobility data, with the intention of identifying a modifiable extrinsic aspect of an MSA that can be intervened on to *impact* mobility patterns and decrease interaction segregation (b) In order, we (1) cluster all homes by nearest interaction hub (using straight line distance from home to hub), partitioning all homes into K clusters, where K is the number of hubs in the MSA (2) compute the weighted average economic diversity (i.e. Gini Index) of the clusters, normalized by the overall economic diversity of the MSA to allow for comparisons between different MSAs of varying baseline levels of economic diversity (Extended Data Table 1) (c) The graphical definition of Gini Index is provided, which is a standard measure of economic dispersion⁶³. Results are robust to the definition of economic diversity, and holds true when using variance in ES instead of Gini Index (Supplementary Figure S12).



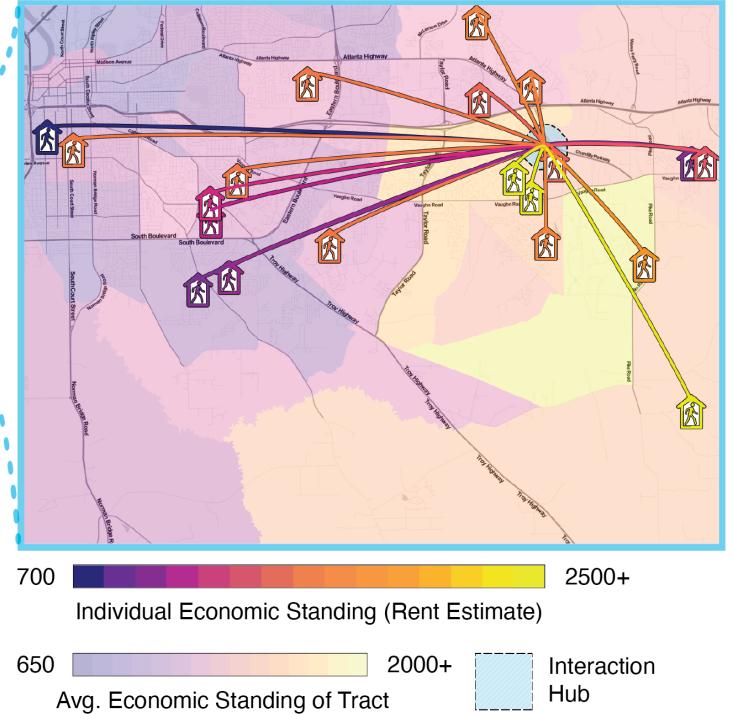
Extended Data Figure 7: Understanding the determinants of BI. The Bridging Index (BI) is a single metric which captures three important factors of built environment (see Supplementary Figure S11 for contributions of these factors to explaining interaction segregation):

- (1) The locations of interaction hubs — If hubs are located in between diverse neighborhoods, BI will be high as hubs will bridge together diverse individuals.
 - (2) The # of interaction hubs — as # of hubs decreases, BI increases (e.g if there is only 1 hub in a city, BI will be 1.0 as all individuals are unified by a single hub)
 - (3) Residential segregation, i.e. the locations of homes and their associated economic standing — as residential segregation decreases we can expect that individuals residing near each hub will be more diverse.
- This figure builds intuition for BI by showing how BI may vary for a single simulated city, consisting of highly segregated neighborhoods. We hold residential segregation (3) constant, and vary the location (1) and number (2) of interaction hubs across panels (a), (b), (c), (d), in order of increasing BI. Note that BI in (c) is substantially higher than BI in (b), because hubs in (c) are better positioned to bridge diverse neighborhoods—even though the number of hubs remains constant.

(a)



(b)

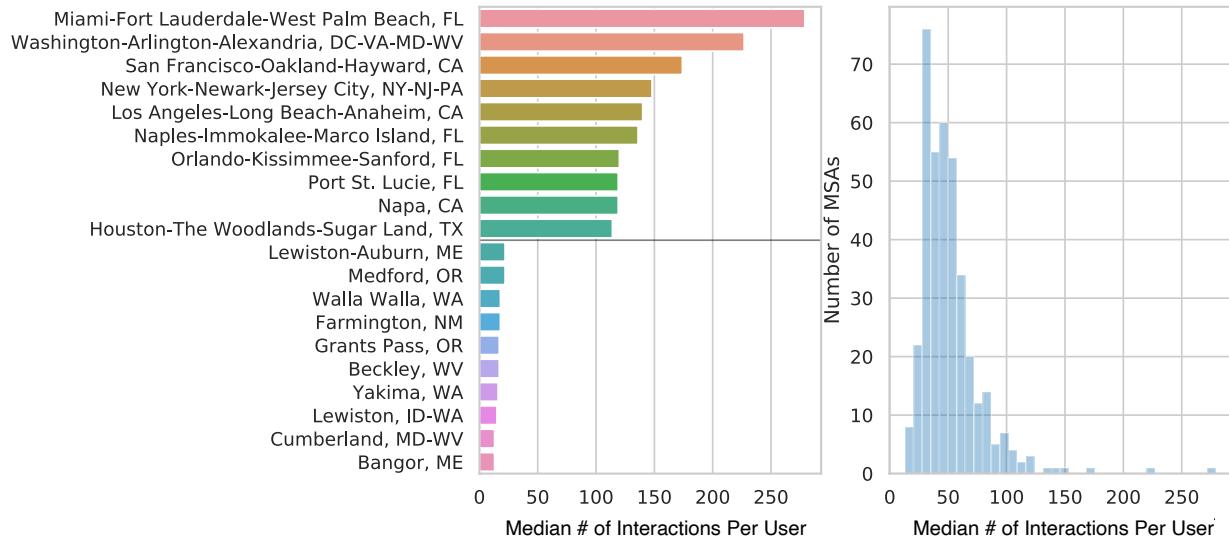


Extended Data Figure 8: Montgomery, AL. We conduct an analogous analysis to Figure 3a,b but for Montgomery, AL, which has nearly identical population (374K vs 385K residents) and income inequality (55th vs 60th percentile Gini Index) to Fayetteville, NC but is 74% more segregated (88th percentile vs. 21st percentile interaction segregation). We find that the difference in segregation is explained by Montgomery, AL having a significantly higher BI compared to Fayetteville, NC (65th vs. 13th percentile). In Montgomery, AL interaction hubs (i.e. commercial centers) are differentiated by ES which results in high-ES individuals and low-ES individuals visiting separate hubs and prevents them from engaging in cross-ES interactions. (a) shows that, as with all MSAs, commercial centers (e.g. shopping malls, plazas, etc.) are hubs of interaction. We illustrate that in Montgomery, AL all visually discernible hubs are associated with one or more commercial centers. (b) In Montgomery, AL, interaction hubs are located in different locations which cater separate to high and low ES residents, leading to segregated interactions. As an illustrative example, we show a zoomed-in map of one hub (Chantilly Center) in Montgomery, AL, and display a random sample of 10 interactions occurring inside of it. Chantilly Center in Montgomery, AL is located accessibly for high ES individuals but is far apart from low-ES tracts. As a result, the sample shows that the majority of interactions are middle-upper ES, and only a few low-ES individuals visit Chantilly Center and interact with these high-ES individuals. Home icons demarcate individual home location (up to 100 meters of random noise added to preserve anonymity); home colors denote individual ES; arcs indicate an interaction inside of the hub; background colors indicate mean census tract ES.

511 **Supplementary Information**

	Accurate pings	Unique days	Distinct user pairs interact	Interactions		Accurate pings	Distinct user pairs who interact	Interactions
count	8,609,406				382			
mean	3,273	35	184	363	73,757,695	2,577,322	4,845,144	
std	16,507	20	374	1,073	163,848,305	8,872,464	16,838,938	
min	11	2	1	1	2,196,084	27,326	53,350	
10%	570	13	8	17	8,398,875	140,251	313,803	
50%	1,471	30	76	141	22,054,930	504,525	1,031,691	
90%	5,857	63	436	785	175,295,175	4,573,152	8,954,800	
max	4,755,081	95	42,323	193,193	1,605,070,032	94,140,015	215,183,409	

Supplementary Table S1: Combined descriptive statistics for all individuals residing in 382 Metropolitan Statistical Areas (MSAs). 8,609,406 individuals reside in a Metropolitan Statistical Area (90% of the overall 9,567,559 individuals in our study). The remaining 958,153 users live outside of MSAs, influencing the interaction segregation of an MSA by coming into contact with MSA residents. Descriptive statistics are grouped by individual (left) and MSA (right). At least one of two users in each interaction pair must live in an MSA to be included in this table.



Supplementary Figure S1: Descriptive statistics of path crossings.

- (a) Ten Metropolitan Statistical Areas (MSAs) with the highest and lowest median users crossed per user.
- (b) Overall distribution of median users crossed per user over MSAs.

display_name	# POIs (25%)	# POIs (50%)	# POIs (75%)	# POIs (max)	# POIs (mean)	# POIs (min)	# POIs (std)
Full-Service Restaurants	75.5	160.0	424.0	24,689.0	609.8	12.0	1,820.05
Snack Bars	18.0	40.0	110.0	6,266.0	169.76	1.0	511.17
Limited-Service Restaurants	33.0	60.0	145.5	4,847.0	192.14	5.0	434.78
Stadiums	1.0	2.0	4.0	43.0	3.67	1.0	4.32
Performing Arts Centers	1.0	2.0	4.0	28.0	3.25	1.0	3.41
Fitness/Recreation Centers	10.0	25.0	72.0	4,877.0	126.6	1.0	414.26
Historical Sites	1.0	2.0	7.0	206.0	9.16	1.0	21.65
Theme Parks	1.0	3.0	6.0	158.0	7.64	1.0	16.77
Bars/Drinking Places	2.0	5.0	13.0	447.0	19.52	1.0	45.91
Parks	3.0	6.0	17.0	793.0	28.69	1.0	80.44
Religious Organizations	7.0	16.0	41.25	2,644.0	63.28	1.0	196.97
Bowling Centers	2.0	4.0	8.0	204.0	9.77	1.0	20.52
Museums	1.0	3.0	6.0	137.0	6.78	1.0	13.99
Casinos	1.0	3.0	7.0	188.0	8.05	1.0	17.22
Independent Artists	1.0	2.0	5.0	130.0	7.55	1.0	17.42
Other Amusement/Recreation	1.0	2.0	7.0	525.0	10.17	1.0	36.13
Golf Courses and Country Clubs	2.0	3.0	7.0	101.0	8.07	1.0	13.77

Supplementary Table S2: POI descriptive statistics (# of POIs in each MSA) for each of the fine-grained POI categories in Figure 1e.

display_name	POI ES (25%)	POI ES (50%)	POI ES (75%)	POI ES (max)	POI ES (mean)	POI ES (min)	POI ES (std)
Full-Service Restaurants	1,210.96	1,395.0	1,674.27	3,628.06	1,493.04	763.0	430.99
Snack Bars	1,229.73	1,412.35	1,684.69	3,621.34	1,513.61	788.12	433.86
Limited-Service Restaurants	1,174.84	1,351.64	1,587.4	3,501.19	1,440.15	771.34	410.95
Stadiums	1,310.0	1,500.0	1,775.0	3,585.25	1,593.21	795.0	424.77
Performing Arts Centers	1,395.0	1,583.1	1,832.4	3,632.78	1,659.56	875.0	431.06
Fitness/Recreation Centers	1,230.03	1,431.79	1,703.94	3,749.05	1,528.73	700.0	453.4
Historical Sites	1,325.0	1,527.94	1,793.75	3,618.58	1,627.62	757.5	452.96
Theme Parks	1,300.0	1,498.75	1,750.0	3,900.0	1,612.58	700.0	501.79
Bars/Drinking Places	1,220.02	1,420.25	1,676.4	3,656.17	1,505.86	750.0	440.08
Parks	1,279.82	1,470.15	1,748.12	3,748.11	1,562.62	725.0	454.13
Religious Organizations	1,269.27	1,459.86	1,677.08	3,670.38	1,529.02	754.0	428.42
Bowling Centers	1,180.08	1,368.75	1,621.15	3,504.36	1,457.96	725.0	434.56
Museums	1,275.0	1,490.83	1,775.36	3,606.66	1,585.92	800.0	474.37
Casinos	1,200.0	1,400.0	1,655.54	3,606.17	1,503.88	725.0	469.68
Independent Artists	1,374.38	1,611.5	1,904.6	3,691.68	1,725.42	850.0	528.33
Other Amusement/Recreation	1,266.0	1,450.0	1,700.74	4,053.39	1,549.13	758.0	462.03
Golf Courses and Country Clubs	1,399.06	1,648.4	1,964.19	4,248.5	1,765.92	900.0	542.13

Supplementary Table S3: POI descriptive statistics (average POI economic standing in an MSA) for each of the fine-grained POI categories in Figure 1e. POI economic standing is operationalized as the median visitor ES of the POI.

display_name	IS (25%)	IS (50%)	IS (75%)	IS (max)	IS (mean)	IS (min)	IS (std)
Full-Service Restaurants	0.22	0.27	0.32	0.48	0.27	0.08	0.07
Snack Bars	0.2	0.25	0.31	0.5	0.25	0.01	0.08
Limited-Service Restaurants	0.24	0.29	0.34	0.47	0.29	0.04	0.08
Stadiums	0.14	0.17	0.22	0.36	0.18	0.02	0.06
Performing Arts Centers	0.14	0.16	0.19	0.27	0.17	0.05	0.05
Fitness/Recreation Centers	0.2	0.26	0.31	0.47	0.25	0.03	0.08
Historical Sites	0.15	0.2	0.27	0.43	0.21	0.0	0.09
Theme Parks	0.16	0.2	0.25	0.42	0.2	0.02	0.08
Bars/Drinking Places	0.18	0.23	0.3	0.42	0.23	0.06	0.08
Parks	0.19	0.26	0.33	0.47	0.26	0.05	0.09
Religious Organizations	0.24	0.32	0.38	0.55	0.31	0.05	0.1
Bowling Centers	0.16	0.21	0.26	0.44	0.22	0.03	0.08
Museums	0.18	0.22	0.28	0.45	0.24	0.06	0.08
Casinos	0.2	0.26	0.32	0.47	0.26	0.02	0.09
Independent Artists	0.13	0.2	0.27	0.39	0.21	0.02	0.09
Other Amusement/Recreation	0.18	0.25	0.31	0.71	0.25	0.02	0.12
Golf Courses and Country Clubs	0.33	0.41	0.5	0.62	0.4	0.2	0.11

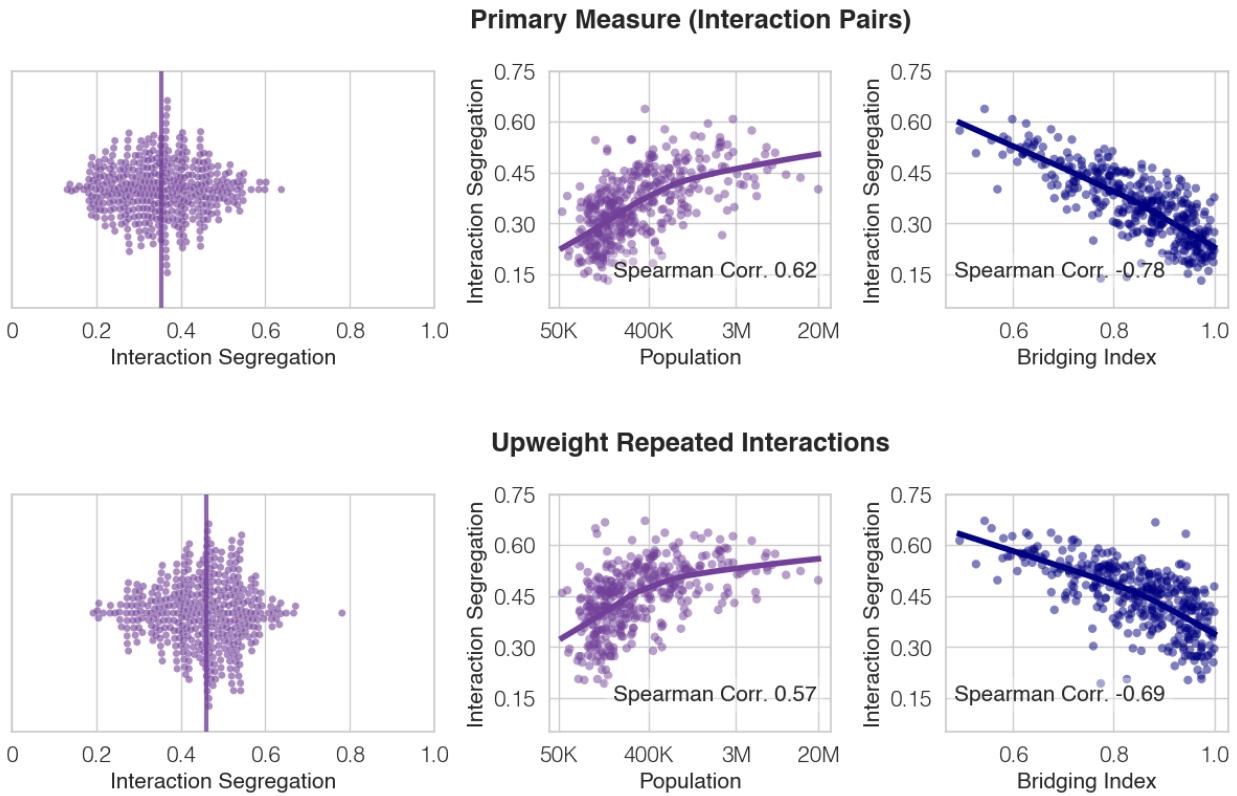
Supplementary Table S4: POI descriptive statistics (Interaction Segregation within-category) for each of the fine-grained POI categories in Figure 1e. Interaction Segregation is calculated for each POI category by filtering for only interactions which occurred inside of the POI category, before estimating interaction segregation (Methods).

display_name	# Interactions (25%)	# Interactions (50%)	# Interactions (75%)	# Interactions (max)	# Interactions (mean)	# Interactions (min)	# Interactions (std)
Full-Service Restaurants	23,060.75	54,219.5	156,645.0	19,540,673.0	398,304.04	6,112.0	1,634,147.68
Snack Bars	15,582.0	38,954.0	120,225.0	14,128,466.0	291,523.01	5,233.0	1,205,873.07
Limited-Service Restaurants	16,485.5	38,444.0	106,515.0	10,453,353.0	227,378.5	4,122.0	878,243.0
Stadiums	53,077.5	96,487.5	336,479.75	8,942,618.0	348,920.85	17,024.0	988,719.72
Performing Arts Centers	66,712.0	120,256.0	384,770.0	6,972,326.0	403,378.22	27,589.0	932,589.88
Fitness/Recreation Centers	11,165.25	21,541.5	62,380.5	5,630,299.0	158,025.53	3,740.0	573,788.93
Historical Sites	18,351.5	47,793.0	98,385.5	6,362,665.0	187,978.09	5,147.0	684,470.65
Theme Parks	34,989.0	61,553.0	135,744.0	1,883,136.0	157,622.73	14,460.0	290,329.93
Bars/Drinking Places	11,592.5	21,401.0	63,929.0	1,266,235.0	84,752.14	4,553.0	181,978.58
Parks	10,050.75	22,301.0	61,492.75	1,520,092.0	88,789.84	5,383.0	193,888.94
Religious Organizations	6,014.0	13,002.0	35,157.25	2,206,316.0	60,683.34	2,739.0	212,948.46
Bowling Centers	16,423.5	26,517.0	65,563.5	1,030,970.0	92,515.06	5,874.0	174,876.21
Museums	14,807.5	27,802.0	66,729.75	681,994.0	87,797.02	4,310.0	146,495.57
Casinos	13,844.0	23,109.0	60,222.0	826,676.0	68,012.52	7,474.0	124,063.14
Independent Artists	8,795.0	23,789.5	56,736.75	1,106,402.0	87,662.95	3,951.0	198,394.53
Other Amusement/Recreation	6,923.0	14,349.0	42,793.0	365,436.0	38,640.52	2,929.0	56,964.16
Golf Courses and Country Clubs	4,765.75	7,836.5	15,555.0	58,348.0	13,047.42	2,636.0	13,348.76

Supplementary Table S5: POI descriptive statistics (number of interactions occurring inside POI category) for each of the fine-grained POI categories in Figure 1e.

	Pearson Corr. w/ Primary	Spearman Corr. w/ Primary	Median	Mean
Interaction Segregation Measure				
Primary Measure	—	—	0.35	0.35
Primary Measure (+ Up-weight Multiple Interactions)	0.89	0.91	0.46	0.45
ES Definition: Rent Zestimate Percentile	0.78	0.80	0.46	0.45
ES Definition: Within-MSA Rent Zestimate Percentile	0.81	0.83	0.54	0.53
ES Definition: Census Median Household Income	0.75	0.77	0.47	0.46
Exclude Pri/Sec Roads	0.99	0.99	0.37	0.37
Exclude Roads	0.98	0.98	0.37	0.37
Exclude Same-home interactions	0.98	0.98	0.34	0.34
Work/Leisure (Neither in Home Tract)	0.93	0.93	0.31	0.31
Leisure (inside POI)	0.85	0.84	0.28	0.29
Minimum Distance Between Pings: < 25 meters	0.98	0.99	0.36	0.36
Minimum Distance Between Pings: < 10 meters	0.95	0.96	0.37	0.36
Minimum Time Between Pings: < 2 minutes	0.99	0.99	0.36	0.36
Minimum Time Between Pings: < 60 seconds	0.99	0.99	0.36	0.36
Minimum Tie Strength: 2 consecutive interactions	0.94	0.95	0.35	0.35
Minimum Tie Strength: 3 consecutive interactions	0.83	0.83	0.37	0.37
Minimum Tie Strength: 2 unique days of interaction	0.88	0.90	0.47	0.46
Minimum Tie Strength: 3 unique days of interaction	0.73	0.76	0.56	0.54
Dist. < 25 meters, Time < 2 min., >= 2 consec. interactions	0.93	0.94	0.35	0.35
Dist. < 25 meters, Time < 2 min., >= 2 unique days	0.88	0.89	0.46	0.45
Dist. < 10 meters, Time < 60 sec., >= 3 consec. interactions	0.80	0.80	0.38	0.37
Dist. < 10 meters, Time < 60 sec., >= 3 unique days	0.73	0.75	0.52	0.50

Supplementary Table S6: Robustness checks overview. We find that our definition of interaction segregation is robust to varying many parameters: weighting of repeated interactions between the same users, definition of economic standing, inclusion/exclusion of roads and same-home interactions, filtering location of interaction, minimum distance, minimum time, and minimum tie strength (as well as the intersection of distance, time, and tie strength). The above variants all are strongly correlated to our primary measure (all have Spearman Corr. ≥ 0.75). We also find that our primary findings that (1) large, dense cities facilitate segregation and (2) interaction hub locations accessible to diverse individuals may mitigate segregation are robust across all definitions of Interaction Segregation (Supplementary Figures S2-S7). Note that we exclude same-home interactions in robustness checks that vary minimum time, distance, or require repeated interactions, to ensure that results are not influenced by interactions with members of the same household (these interactions ordinarily have minimum influence on Interaction Segregation, as shown by the robustness check which excludes same-home interactions and results in virtually identical metric (Spearman Corr. 0.98); however, the influence of same-home interactions is higher after more conservative filters are applied to the definition of interactions, such as requiring a minimum tie strength of 3 consecutive interaction).

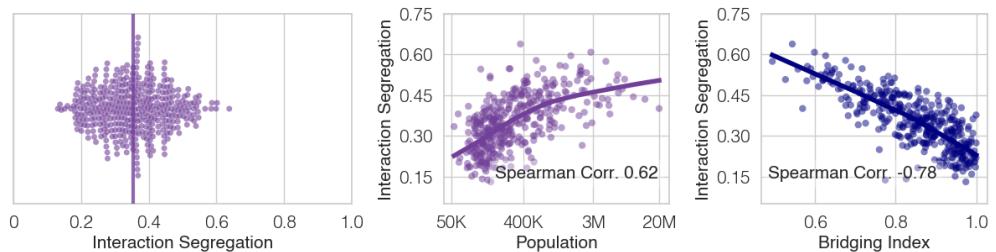


Supplementary Figure S2: Robustness of primary study findings to weighting of repeated interactions. We find that our primary study findings that, (1) large, dense cities facilitate segregation and (2) interaction hub locations accessible to diverse individuals may mitigate segregation, are robust to the choice of whether to upweight repeated interactions in our interaction network. We compare the results of:

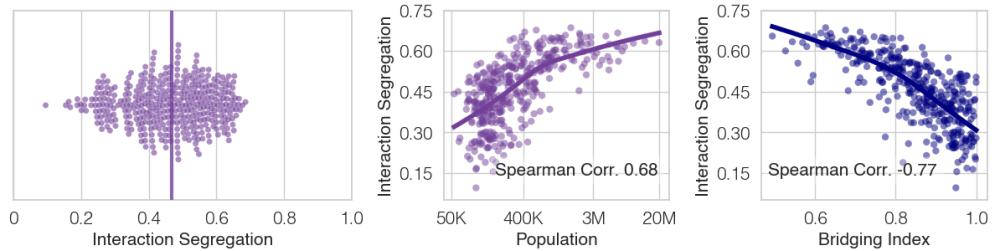
Primary Measure: Interactions are defined as pairs of users who have ever interacted within the study observation window (three months of 2017). We deduplicate repeated interactions, as frequency of pings varies across smartphone users, to reduce bias from users with a higher frequency of pings. For instance, if an individual A with an individual B (ES \$1000) two times and individual C (ES \$2000) one, we compute the mean ES of individual A's network as \$1500.

Upweight Repeated Interactions: Repeated interactions are unweighted when calculating the mean ES of an individual's interaction network. For instance, if an individual A with an individual B (ES \$1000) two times and individual C (ES \$2000) once, we compute the mean ES of individual A's network as \$1333.

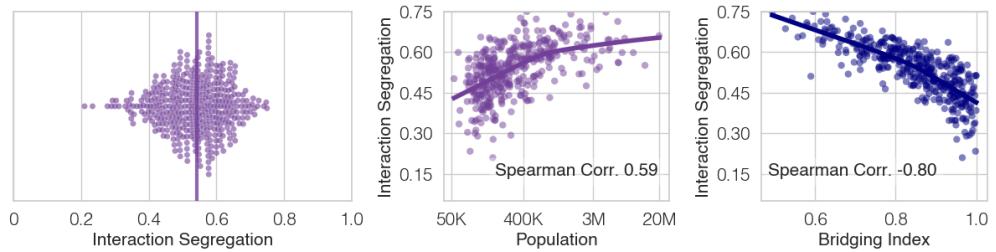
Primary Measure (Economic Standing: Rent Zestimate)



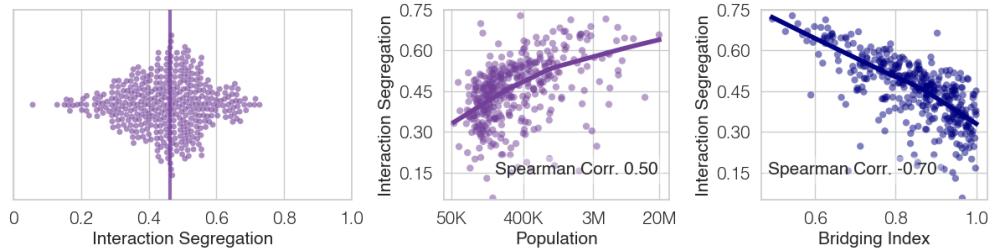
Economic Standing: Census Block Group Median Household Income



Economic Standing: Rent Zestimate Percentile



Economic Standing: Rent Zestimate Percentile Relative to MSA



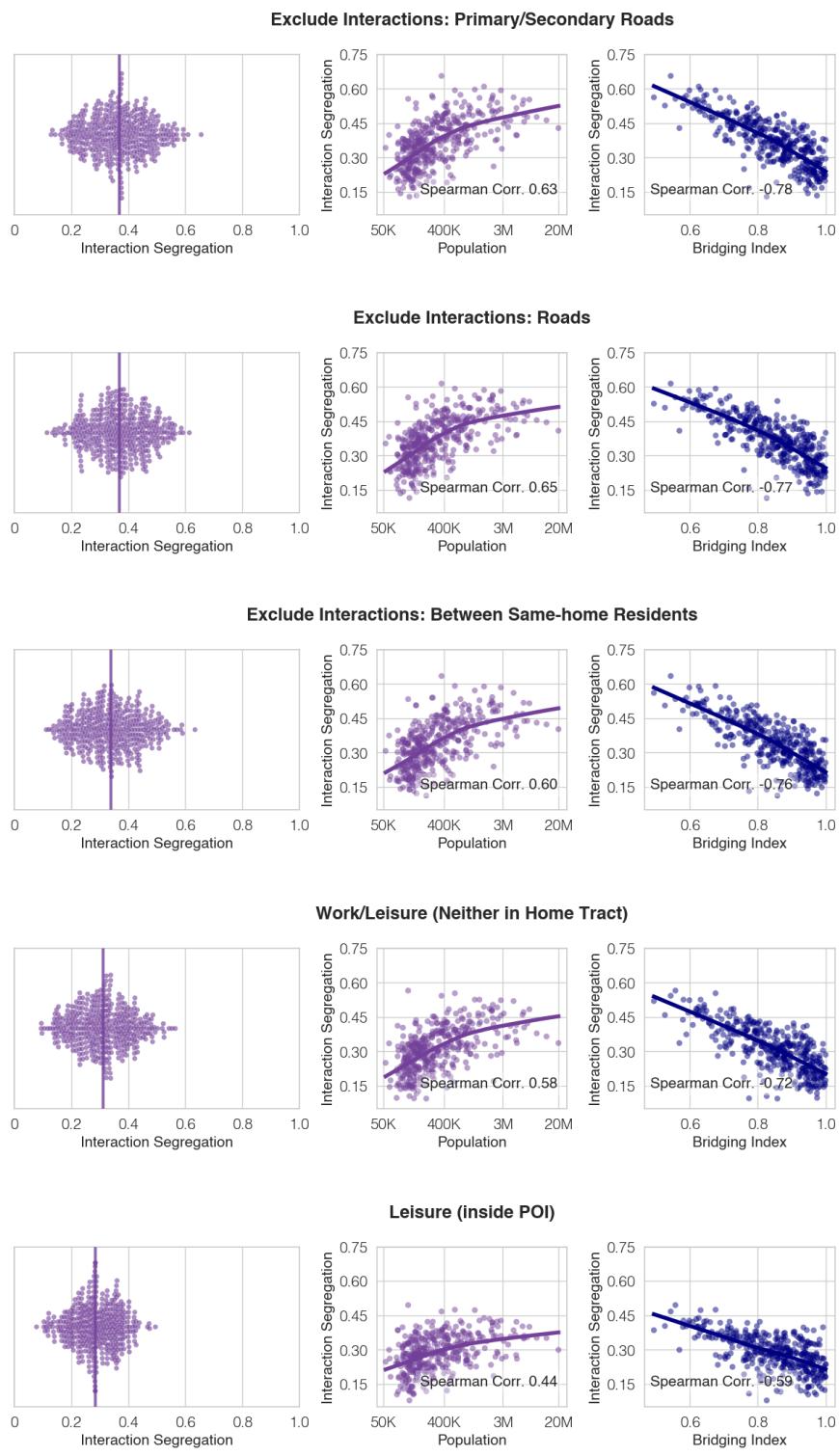
Supplementary Figure S3: Robustness of primary study findings to definition of economic standing. We find that our primary study findings that, (1) large, dense cities facilitate segregation and (2) interaction hub locations accessible to diverse individuals may mitigate segregation, are robust to the definition of economic standing. We compare the results of:

Primary Measure: Our primary measure leverages estimated monthly rent value (Zillow Rent Zestimate).

Census Block Group (CBG) Median Household Income: We define the ES of an individual is the median household income in the CBG in which they reside.

Rent Zestimate Percentile: We normalize Rent Zestimate values across all individuals.

Primary measure Relative to MSA: We normalize Rent Zestimate values across all individuals within an MSA, independent of other MSAs, to account for differences in cost of living across cities.



Supplementary Figure S4: Robustness of primary study findings to exclusion of interactions within roads, exclusion interactions with residents of the same home, and exclusion of non-work/leisure interactions. We find that our primary study findings that, (1) large, dense cities facilitate segregation and (2) interaction hub locations accessible to diverse individuals may mitigate segregation, are robust to filtering for a subset of interactions. We compare the results of:

Primary Measure: Our primary measure includes all interactions, aiming to give a complete account of an individual's interaction network including path crossings on roads as well as those they share a home with.

Excluding roads: We define the ES of an individual is the median household income in the CBG in which they reside.

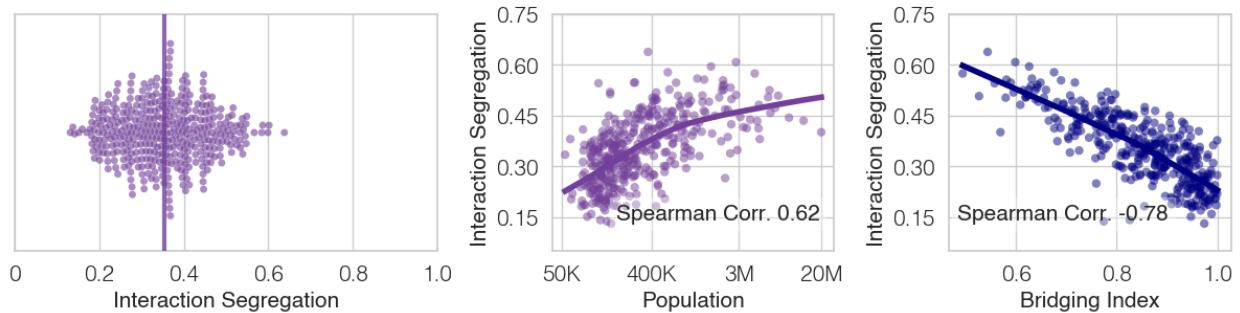
Rent Zestimate Percentile: We normalize Rent Zestimate values across all individuals.

Primary measure Relative to MSA: We normalize Rent Zestimate values across all individuals within an MSA, independent of other MSAs, to account for differences in cost of living across cities.

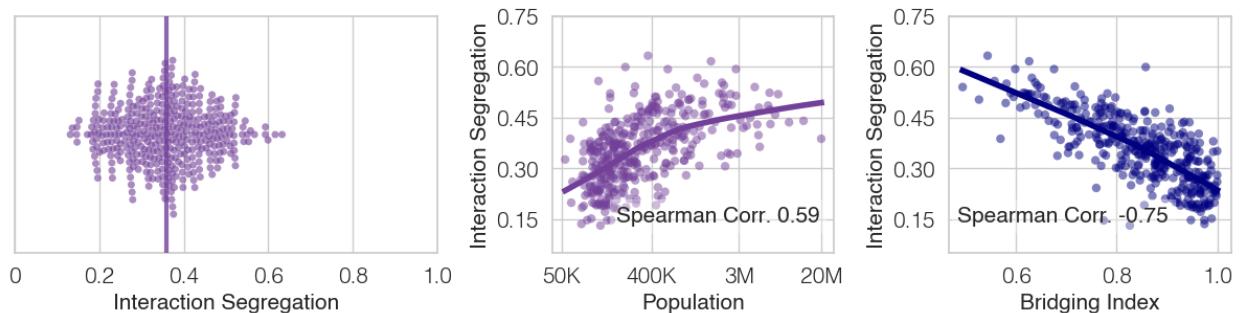
Work/Leisure: We filter to include only interactions likely⁴⁸ take place in the context of work or leisure, by excluding interactions which occurred when either individuals were located within their home tracts.

Leisure: We filter for leisure interactions by including only interactions occurring inside of the POIs categorized as related to leisure (Figure 1e).

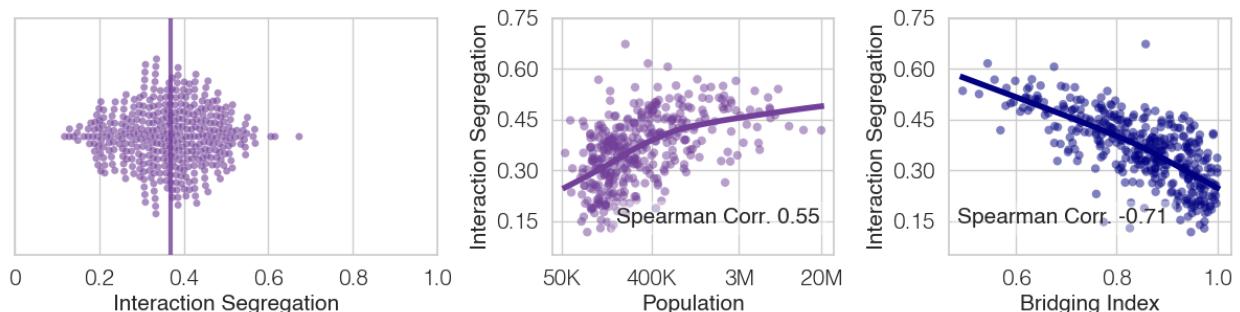
Primary Measure (Minimum Distance Between Pings: < 50 meters)



Minimum Distance Between Pings: < 25 meters



Minimum Distance Between Pings: < 10 meters

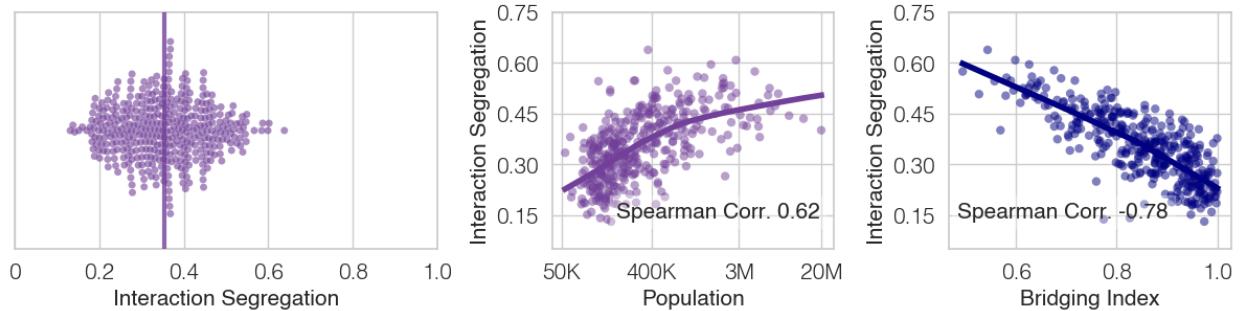


Supplementary Figure S5: Robustness of primary study findings to minimum distance required between two GPS pings for individuals to be considered interacting. We find that our primary study findings that, (1) large, dense cities facilitate segregation and (2) interaction hub locations accessible to diverse individuals may mitigate segregation, are robust to the time threshold used in our definition of interaction:

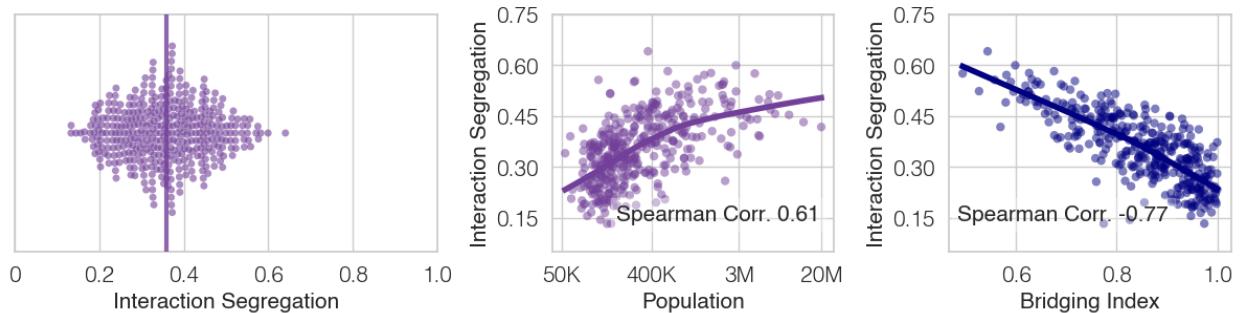
Primary Measure: Our primary measure uses a threshold of 50 meters, based on prior literature which shows that even distant exposure to diverse individuals is predictive of long-term behaviors¹⁸.

Alternative measures: We alternatively consider more conservative thresholds of 25 meters and 10 meters, with 10 meters being the lowest threshold due to limitations of GPS ping accuracy^{77, 78}.

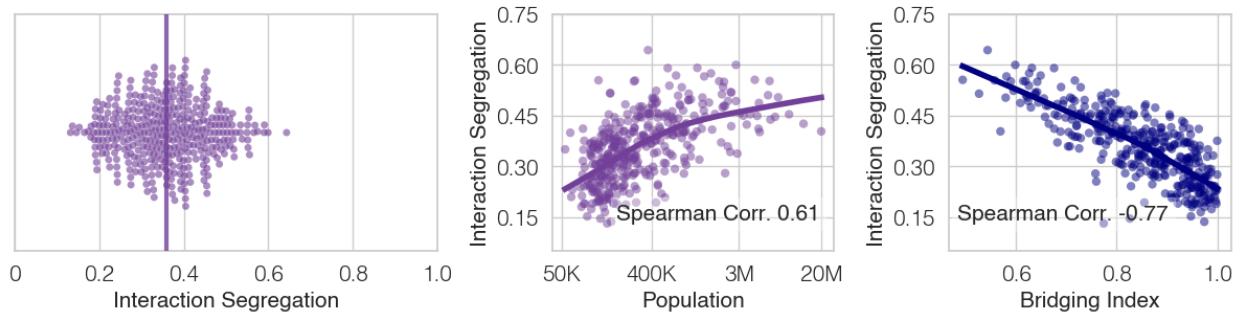
Primary Measure (Minimum Time Between Pings: < 5 minutes)



Minimum Time Between Pings: < 2 minutes



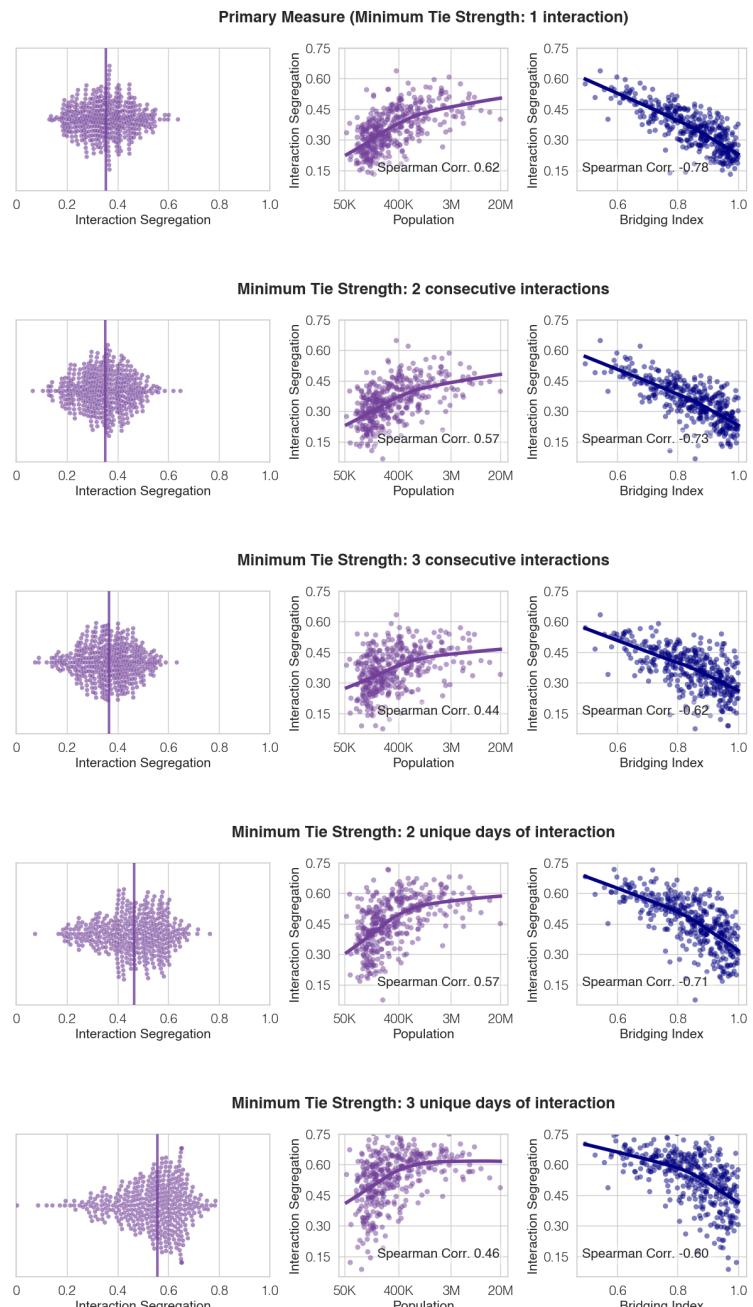
Minimum Time Between Pings: < 60 seconds



Supplementary Figure S6: Robustness of primary study findings to minimum time elapsed between two pings to constitute an interaction. We find that our primary study findings that, (1) large, dense cities facilitate segregation and (2) interaction hub locations accessible to diverse individuals may mitigate segregation, are robust to the time threshold used in our definition of interaction:

Primary Measure: Our primary measure uses a threshold of 5 minutes, to be inclusive of users with sparse pings (e.g., for a subset of users, we only have 1 ping per day, while for others we have 100+ pings per day) while maintaining a reasonable confidence that an interaction may have occurred.

Alternative measures: We alternatively consider more conservative thresholds of 2 minutes and 1 minute.

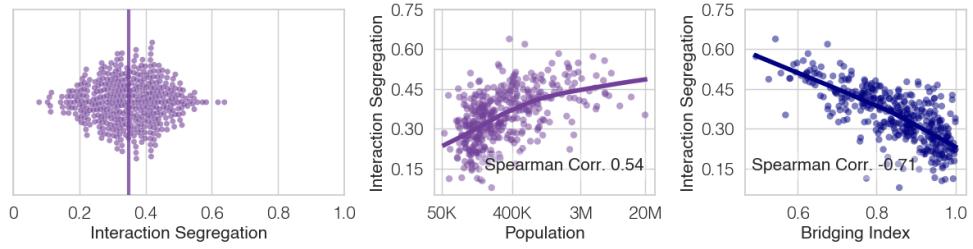


Supplementary Figure S7: Robustness of primary study findings to minimum tie strength required to constitute an interaction. We find that our primary study findings that, (1) large, dense cities facilitate segregation and (2) interaction hub locations accessible to diverse individuals may mitigate segregation, are robust regardless of the minimum tie strength threshold between two individuals to be constitute an interaction:

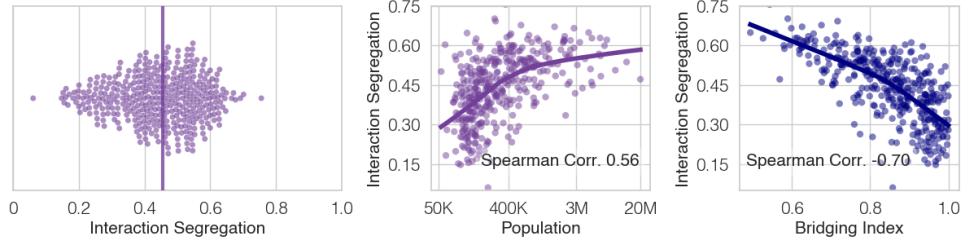
Primary Measure: Our primary measure only requires a single pair of pings between users to constitute an interaction, to be inclusive of users with sparse pings (e.g., for a subset of users, we only have 1 ping per day, while for others we have 100+ pings per day).

Alternative measures: We alternatively consider more conservative thresholds of 2 or 3 consecutive interactions, as well as 2 or 3 interactions across unique days. Requiring consecutive interactions increases the likelihood that individuals actually came into contact together; interactions across unique days increases the likelihood that interactions are not merely path crossings, but social interactions between individuals who are familiar with each other. **(continued on next page)**

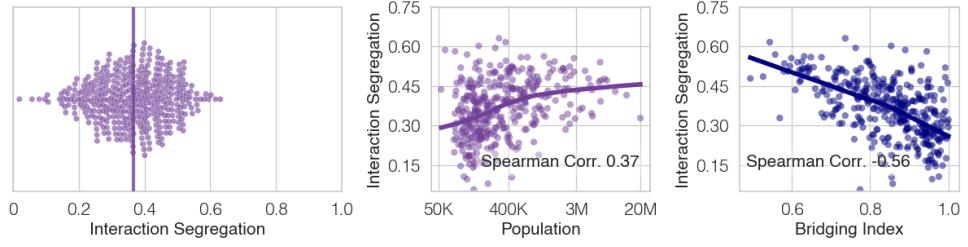
Distance: < 25 meters, Time: < 2 minutes, Minimum Tie Strength: 2 consecutive interactions



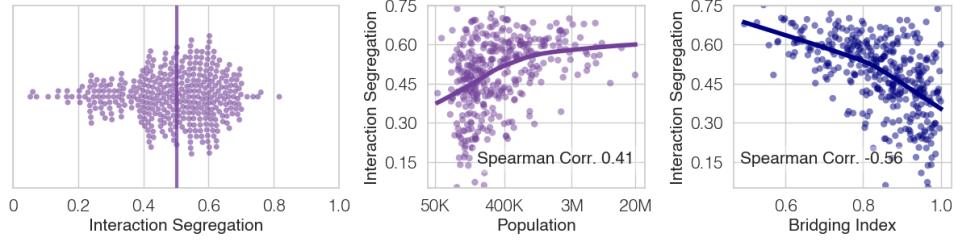
Distance: < 25 meters, Time: < 2 minutes, Minimum Tie Strength: 2 unique days of interaction



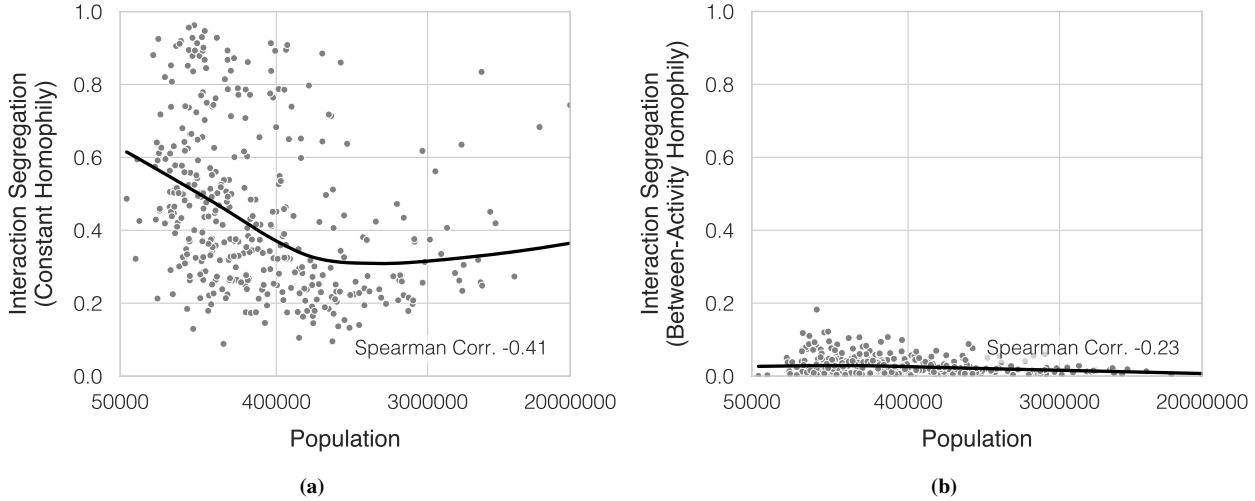
Distance: < 10 meters, Time: < 60 seconds, Minimum Tie Strength: 3 consecutive interactions



Distance: < 10 meters, Time: < 60 seconds, Minimum Tie Strength: 3 unique days of interaction



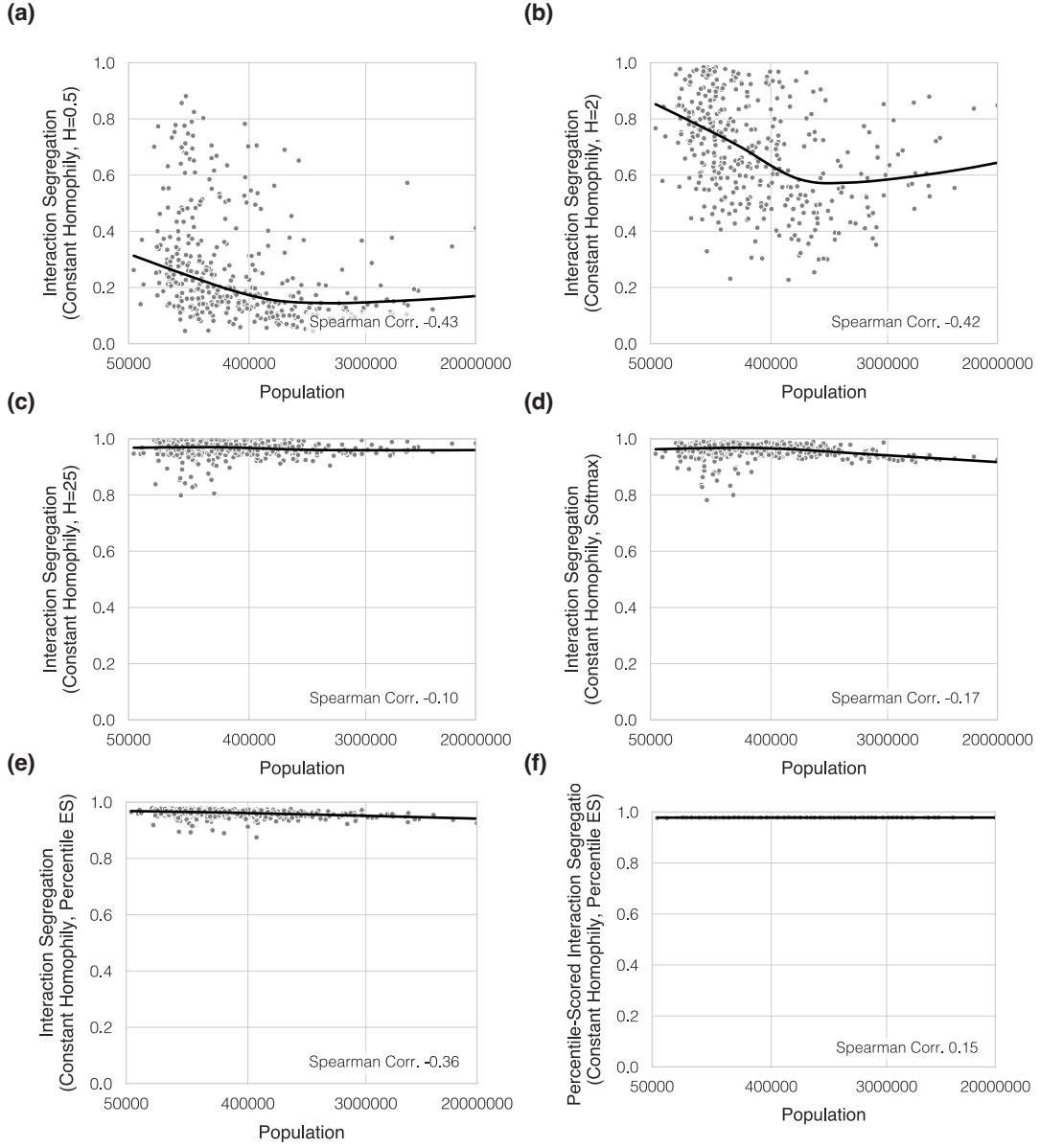
Supplementary Figure S8: (continued from previous page). We find that our primary study findings that, (1) large, dense cities facilitate segregation and (2) interaction hub locations accessible to diverse individuals may mitigate segregation, are *robust to the combination of the minimum time, minimum distance, and minimum tie strength threshold parameters*. To account for interactions between threshold parameters, we also consider combinations of parameter variants. For instance, the most conservative robustness check defines an interaction as two individuals being < 10 meters apart within a < 60 second window, and for this to have occurred either for either 3 consecutive minutes (second figure from the bottom) or across 3 unique days (bottom figure).



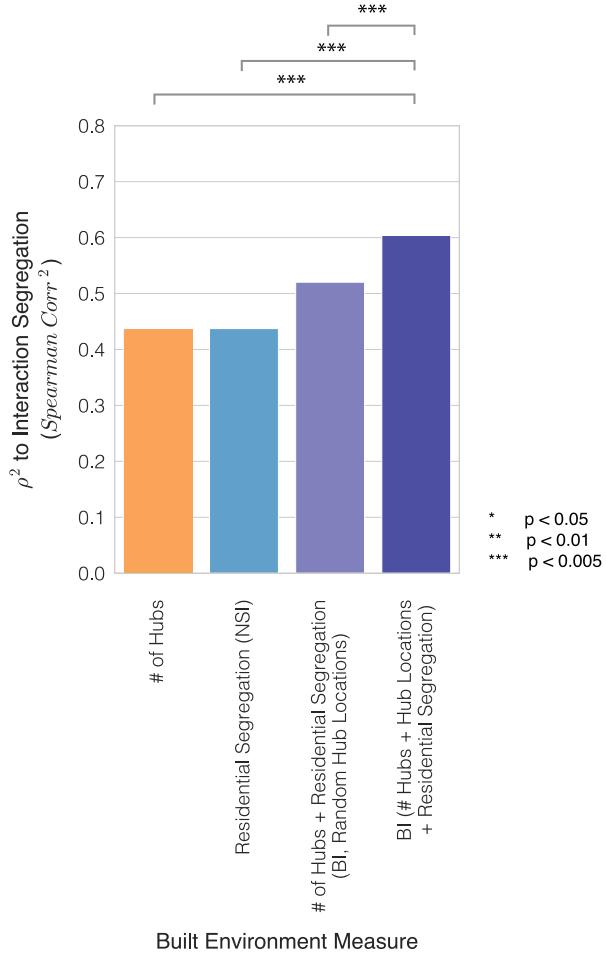
Supplementary Figure S9: Alternative homophily mechanisms do not explain segregation in large cities. We consider the possibility of two alternative hypotheses which may explain the trend towards high segregation in large cities

(a) Constant Homophily: i.e. individuals have the same proclivity for interacting with individuals of similar ES regardless of if they live in large or small cities, and it is instead change in distribution of economic standing that drives segregation in large cities (e.g. in large cities there may be a greater supply of people in the same economic class available to interact with). We test this hypothesis via a null network model, in which we preserve network nodes (individuals and their ES values) but randomize edges^{79,80}. We randomly assign interactions between pairs of people, weighting the likelihood of interaction between people of similar ES higher according to a constant homophily function. Specifically, the probability of interaction ($p_{i,j}$) between two individuals of ES_i and ES_j is weighted by their similarity in ES, defined as the complement of the normalized Euclidean distance in ES: $p_{i,j} \propto \text{Similarity}(ES_i, ES_j) = 1 - \frac{|ES_i - ES_j|}{\max(ES) - \min(ES)}$. We choose 75 interactions per person such that the mean number of interactions per person is 150, which corresponds to Dunbar's number⁸¹. We find that under this null model, there is no positive association between interaction segregation and population size; in fact, larger cities are less segregated on average, as there is an increase in supply of diverse individuals in economic standing in larger cities. These findings are also robust to a variety of null model specifications (Supplementary Figure S10).

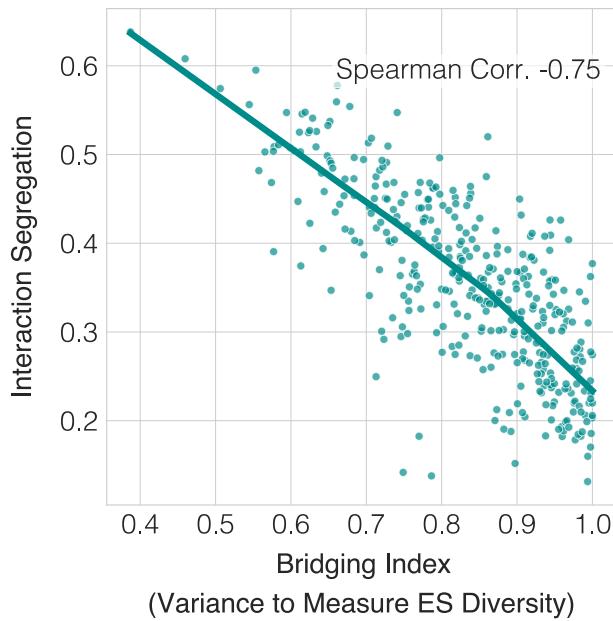
(b) Between Activity Homophily: i.e. it is not the differentiation of individual venues that drives segregation, but rather that in large cities individuals choose different categories of activities which results in segregation (e.g. in small cities, there are less country clubs so everybody visits restaurants to socialize, whereas in large cities high-ES individuals segregate by spending a higher proportion of time in exclusive venues such as country clubs). We test this hypothesis via a configuration model^{79,80}, a prominent null network model in which node degree is preserved. Specifically, by applying a configuration model to reconfigure network edges for each leisure category separately, we preserve network nodes (individuals and their ES values) as well as the number of interactions they had in each category of POI (node degree), but randomize the specific venue in which each interaction occurred. For instance, if an individual interacts with 5 people inside of restaurants and 100 people inside of a fitness center, they will be randomly assigned to interact with 5 people from all of those who visited restaurants, and 100 people from all of those who visited fitness centers. This null model preserves between-activity homophily which results from activity choices (e.g. whether to visit a country club or restaurant), but erases within-activity homophily (e.g. individuals who visit any restaurant are equally likely to interact). We find that under this null model, there is no positive association between interaction segregation and population size; in fact, there is minimal segregation across all cities as variation between activity categories is insufficient to retain segregation. This is further supported by Supplementary Table S3, which shows relatively small differences in ES between participants in different categories of leisure activity (e.g. the lowest ES activity, limited service restaurants has a median visitor ES of \$1,352, the highest ES activity, golf courses and country clubs has a median visitor ES of \$1,648.4).



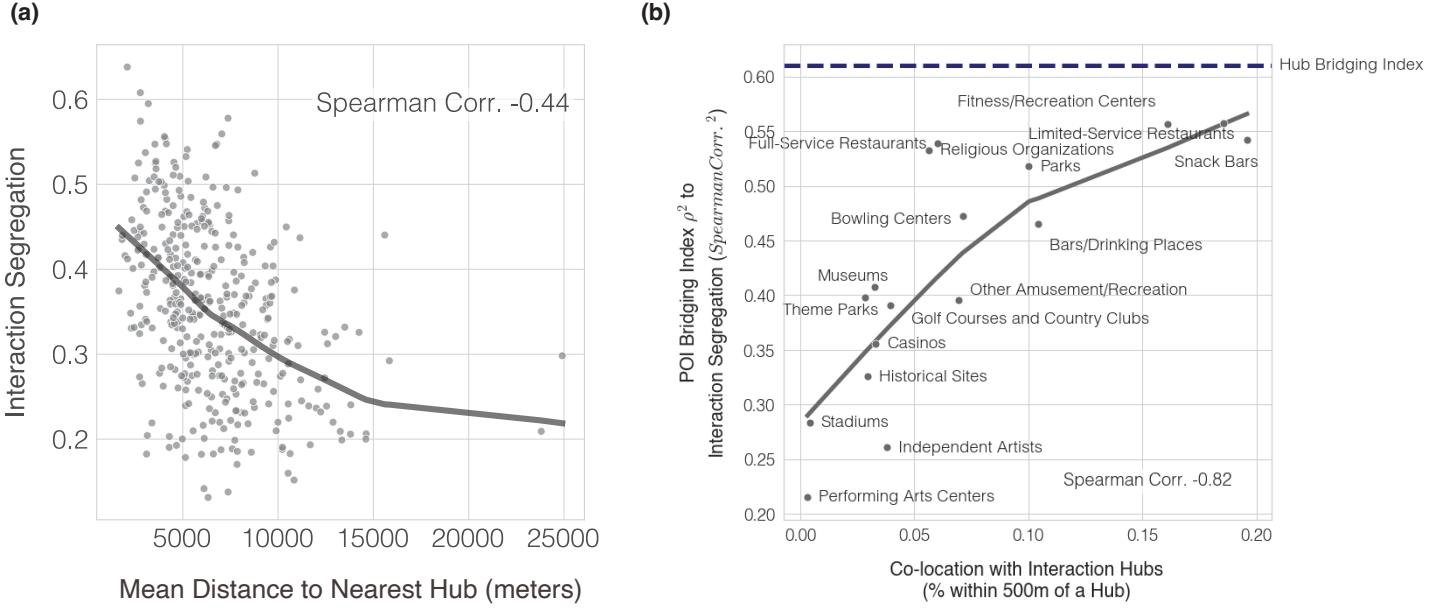
Supplementary Figure S10: Baseline Homophily null model results are robust to varying null model specifications. We re-run the analysis in Supplementary Figure S9a under a variety of null model specifications, and find that *in all cases there is no evidence to suggest that the Constant Homophily hypothesis explains the high segregation observed in large cities*. **(a-c)** We first consider varying the extent of homophily by adding a constant parameter H to the homophilous weighting of edges, to exponentially increase/decrease the extent of homophily in our null model for the probability of individuals i and j interacting: $p_{i,j} \propto \text{Similarity}(ES_i, ES_j)^H$. We find that regardless of if we **(a)** decrease homophily ($H=0.5$) or **(b-c)** increase homophily mildly ($H=2$) or strongly ($H=25$), there is no positive association between population size and segregation in our simulations. In fact, larger cities are less segregated on average, as there is an increase in supply of diverse individuals in economic standing in larger cities. We also consider alternative null model specifications such as **(d)** a softmax homophily function $p_{i,j} \propto \frac{e^{\text{Similarity}(ES_i, ES_j)}}{\sum_{k=1}^N e^{\text{Similarity}(ES_i, ES_k)}}$, **(e)** applying the original null model to percentile-scored values economic standing **(f)** applying the original null model to percentile-scored values economic standing, and calculating Interaction Segregation using percentile-scored values economic standing. This suggests that the high segregation in large cities is due to a change in resident behavior, facilitated by the built environment of large cities, and not an artifact the economic standing distribution in large cities.



Supplementary Figure S11: Understanding why BI explains Interaction Segregation. We show via an ablation study that interaction hub locations, in addition to number of hubs and residential segregation, contributes to the explanatory power of BI. As illustrated in Extended Data Figure 7, BI captures three factors of built environment: (1) locations of interaction hubs (2) number of interaction hubs and (3) residential segregation. In this analysis, we aim to disentangle how these three factors contribute to the ability of BI to explain interaction segregation (as measured by ρ^2 , the squared Spearman correlation with interaction segregation). We find that number of hubs (orange, $\rho^2 = 0.436$) and residential segregation (blue, $\rho^2 = 0.437$) are each correlated with interaction segregation. To measure the combined explanatory power of these two factors within BI, independent of interaction hub locations, we conduct an ablation study in which we calculate Bridging Index for each MSA, using the actual home location data and number of interaction hubs for each MSA, but randomize hub locations (light purple, $\rho^2 = 0.523$). For each MSA, we estimate this value over 1000 random trials. We find that calculating BI using randomized hub locations is a significantly weaker predictor ($p=0.0006 < 0.01$, Steiger's Z-test) compared to BI values computed using actual hub locations (dark purple, $\rho^2 = 0.604$). This demonstrates that hub locations contribute to the explanatory power of BI, i.e. BI explains interaction segregation because it captures the extent to which the locations of hubs in different cities facilitate the interaction of diverse individuals.

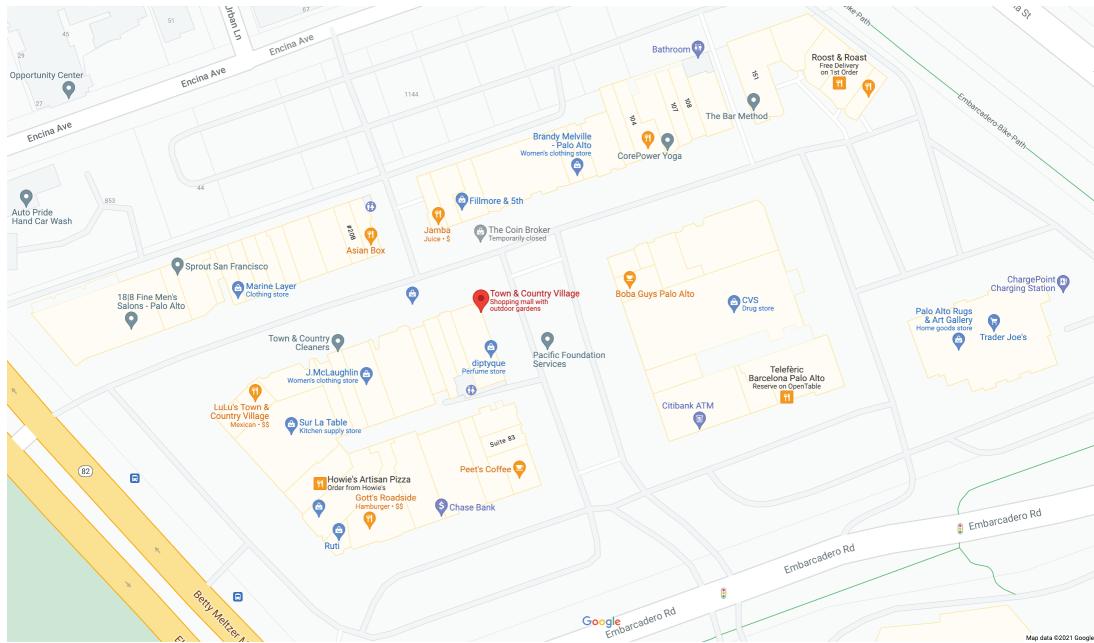


Supplementary Figure S12: Robustness of BI to definition of ES diversity. We calculate a version of Bridging Index which uses variance in to operationalize economic diversity: $Bridging Index (BI) = \frac{\sum_{i=1}^K |\mathcal{C}_i| \cdot Var(\mathcal{C}_i)}{|\mathcal{V}_{MSA}| \cdot Var(\mathcal{V}_{MSA})}$. This variant of BI explains interaction segregation comparably (Spearman Corr. -0.75 vs. -0.78, both N=382, both $p < 10^{-4}$) to our primary measure of BI which uses Gini Index to operationalize economic diversity. Thus, we find that the ability of Bridging Index to explain interaction segregation is robust to the definition of ES diversity.

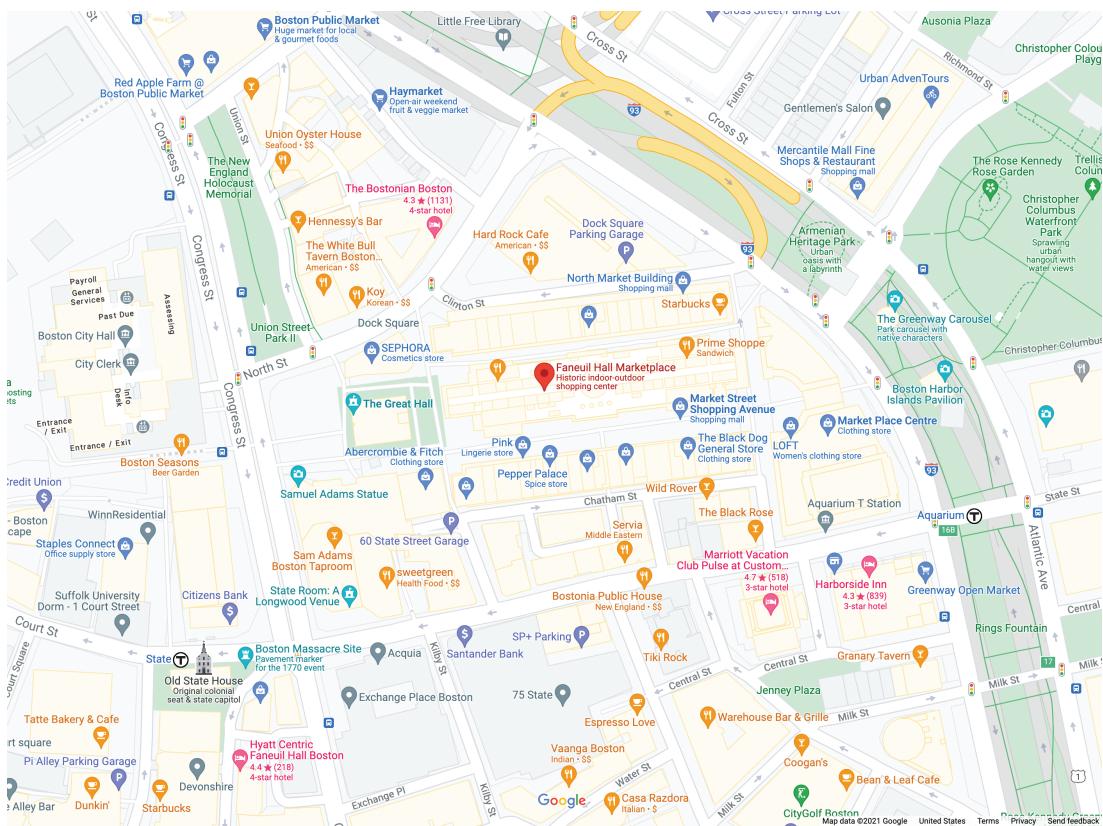


Supplementary Figure S13: We consider alternative mechanisms through which built environment may mitigate interaction segregation. (a) Following the inverse relationship between POI localization and segregation established in Extended Data Figure 2a, we consider whether the de-localization of interaction hubs alone can provide an alternative to BI. We compute mean distance to nearest interaction hub for each MSA, which is the same measure in Extended Data Figure 2a but calculated for interaction hubs. We find that while hub localization is inversely correlated with interaction segregation (Spearman Corr -0.44, N=382, $p < 10^{-4}$), this correlation is significantly less ($p < 10^{-4}$) than the correlation between BI and interaction segregation (Spearman Corr -0.78, N=382, $p < 10^{-4}$). This suggests that hub bridging, as quantified by BI may be a more promising direction to investigate as a potential mitigator of segregation. (b) We also consider whether fine-grained POIs may function as bridges between diverse individuals. For each of the fine-grained leisure POI categories in Figure 1e, we calculate a Bridging Index across all MSAs (using the same procedure to calculate BI as shown in Extended Data Figure S12, except using fine-grained POI locations instead of interaction hub locations). For instance, to calculate the Bridging Index for restaurants, we cluster all homes by the nearest restaurant location, and then calculate: $(Restaurant) \ Bridging \ Index = \frac{\sum_{i=1}^K |\mathcal{R}_i| \cdot Gini \ Index(\mathcal{R}_i)}{|\mathcal{V}_{MSA}| \cdot Gini \ Index(\mathcal{V}_{MSA})}$. After calculating the bridging index for all fine-grained POI categories and for each of the 382 MSAs, we then measure the correlation between each bridging index and interaction segregation across all MSAs (as measured by ρ^2 , the squared Spearman correlation). We find that BI for hubs provides a stronger correlation ($\rho^2 = 0.604$, horizontal line), than all other bridging indices which are plotted as points on the scatter-plot in (b). Further, we find that POI categories which are often located inside or near interaction hubs (co-location, X-axis) have bridging indices which are stronger predictors of interaction segregation (e.g. for fitness/recreation centers, snack bars etc.). The high correlation between (Spearman Correlation -0.82, N=17, $p < 0.001$) between co-location of POIs and bridging index predictive ability demonstrates asymptotic convergence between all other predictive bridging index metrics and our primary BI measure. This further suggests that bridging of interaction hubs should be the primary metric of interest for mitigators of segregation, because other bridging indexes computed for fine-grained POI locations are at best proxies for BI which leverages higher-level interaction hub locations. Supplementary Figures S14-S16 illustrate the frequent co-location between hubs and other fine-grained POIs.

a) Town & County Village, Palo Alto, California



b) Faneuil Hall Marketplace, Boston, MA



Supplementary Figure S14: Examples of interaction hubs in coastal cities (a) San Francisco Bay Area and (b) Boston, MA. Hubs frequently contain a diverse assortment of POIs including restaurants, fitness centers/gyms, grocery stores, etc. and are also frequently hubs around which other POIs are located nearby.

a) Cross Creek Mall and Surrounding Area, Fayetteville, NC



b) [Zoomed In] Cross Creek Mall, Fayetteville, NC

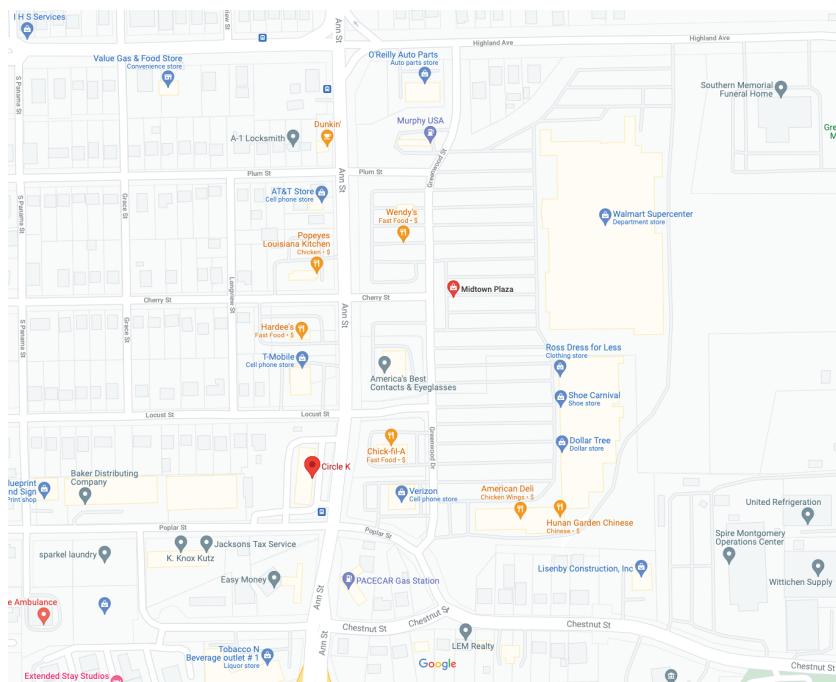


Supplementary Figure S15: Example of a major hubs in Fayetteville, NC (a) zoomed-out view of hub and surrounding co-located POIs (b) zoomed-in view of the hub core and businesses contained inside. We find that in Fayetteville, a city with a high Bridging Index, large hubs contain a variety of POIs which cater to diverse individuals of both high and low-ES.

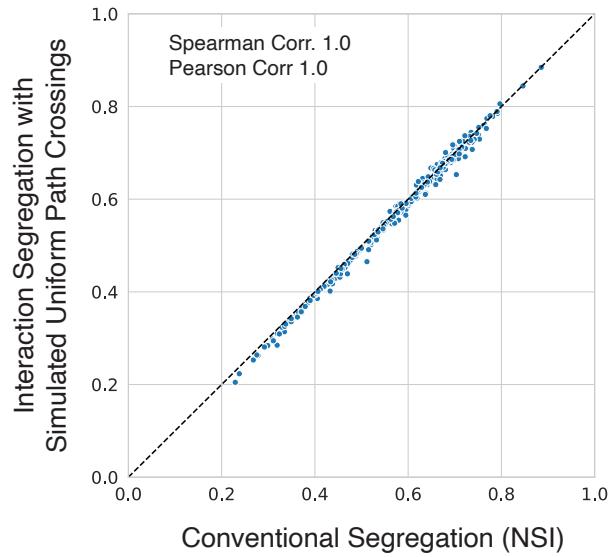
a) The Shoppes at Eastchase, Montgomery, AL



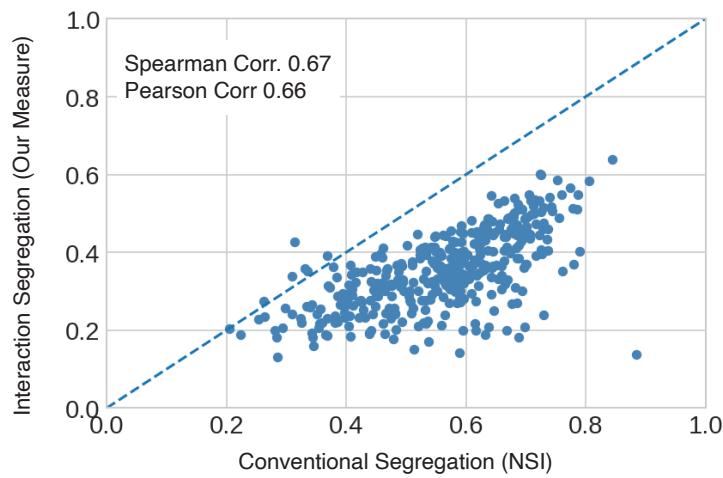
b) Midtown Plaza, Montgomery, AL



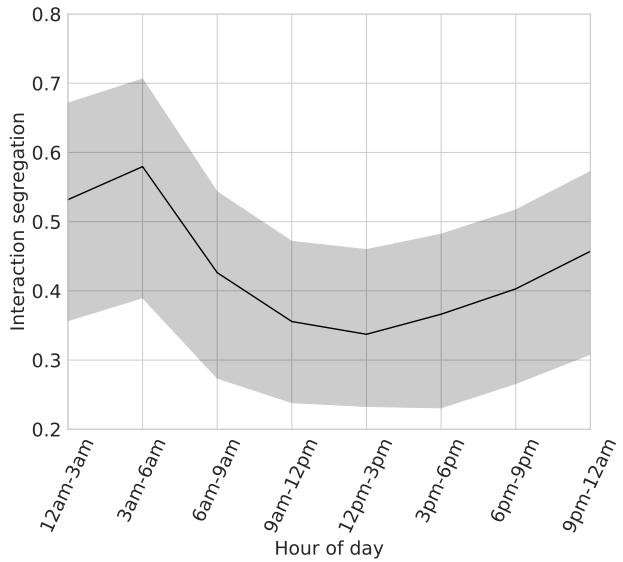
Supplementary Figure S16: Examples two hubs in Montgomery, AL which have visitors of predominantly (a) high economic standing (b) low economic standing. We find that in Montgomery, AL a city with a low Bridging Index, smaller hubs exist which contain POIs which cater to a narrow band of individuals in a specific economic stratum. For instance, we find that the nearby grocery store (a) is a Whole Foods Market in the high-ES hub, in contrast to the (b) Walmart Supercenter in the low-ES hub.



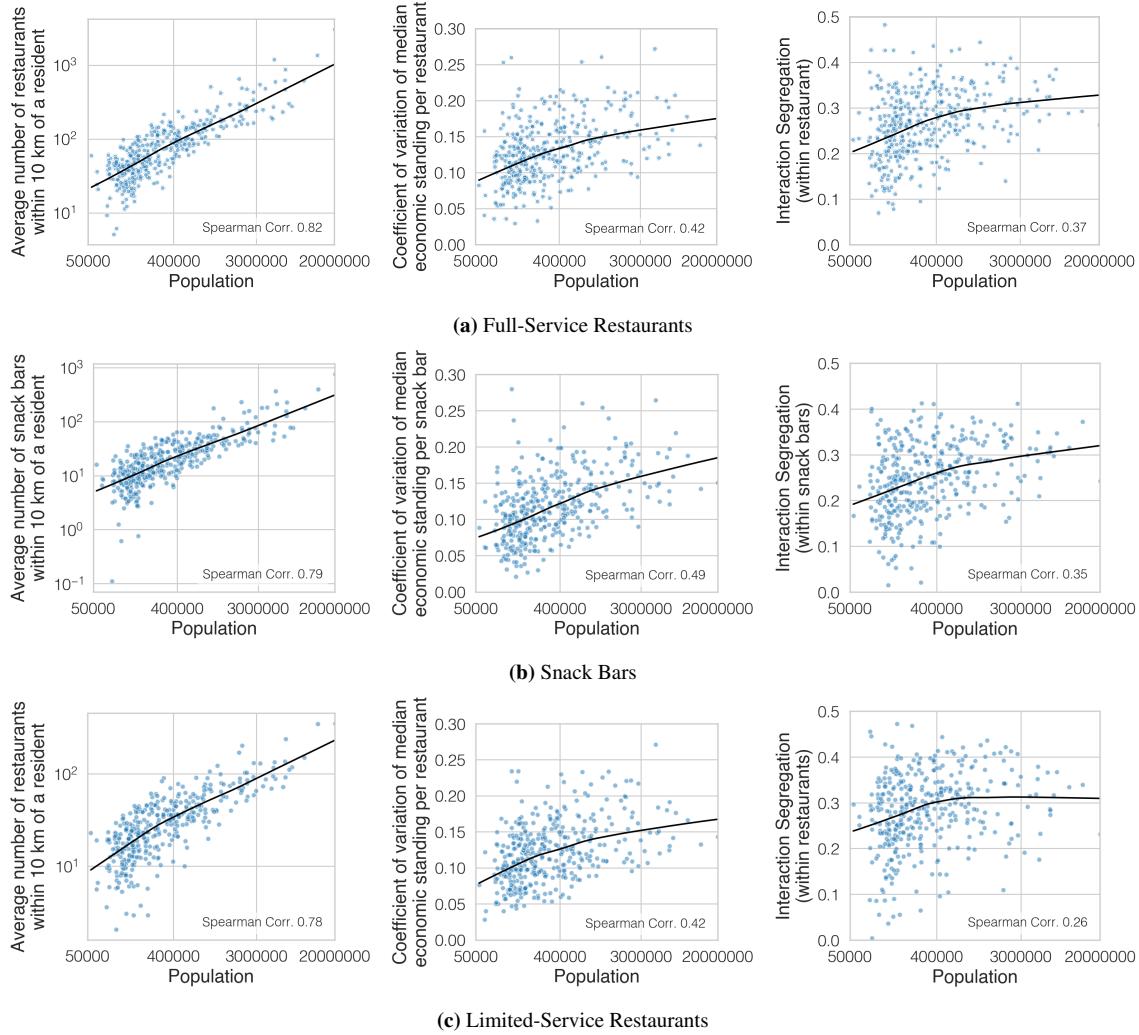
Supplementary Figure S17: Our model with simulated uniform within-tract crossings is equivalent to the conventional neighbourhood sorting index (NSI). Each point is a Metropolitan Statistical Area (MSA). The y-axis shows the interaction segregation estimate from the mixed model with a simulated path crossing between every person in a tract (in our dataset). The x-axis shows the correlation between a person's ES and the average ES of people in their tract, which is the neighbourhood sorting index (NSI). As these measures are equivalent, Spearman Corr = 1.0 and Pearson Corr. = 1.0.

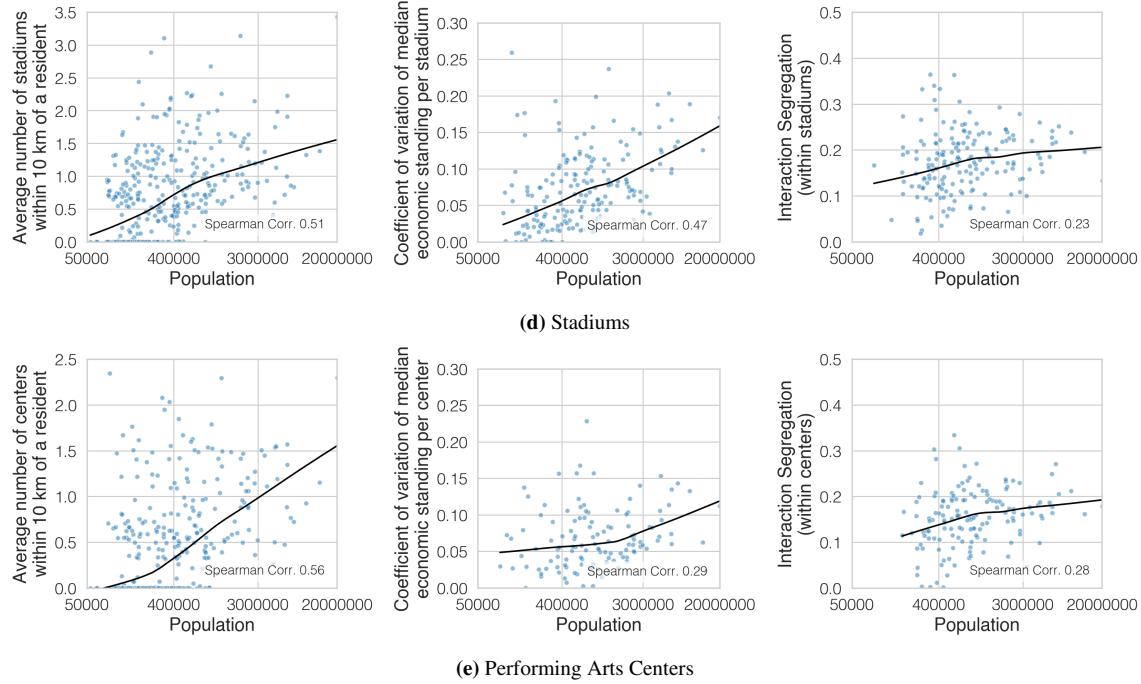


Supplementary Figure S18: Our segregation measure versus a conventional residential segregation, neighborhood sorting index (NSI). Each point is a Metropolitan Statistical Area (MSA). Regardless of whether we compare the numerical segregation values (Pearson Correlation 0.67) or the MSA ranking (Spearman Correlation 0.66), only moderate correlation indicates that our measure is different in kind from residential segregation as measured conventionally by NSI.



Supplementary Figure S19: Segregation decomposed by time. As described in Methods M3, our fine-grained interaction network allows us to decompose our overall interaction segregation into estimates of segregation during different hours of the day, by filtering for interactions that occurred within a specific hour. In Supplementary Figure S19, we partition estimates of segregation by 3 hour windows to illustrate how segregation varies throughout the day (see Supplementary Information). We observe that segregation increases by 61% between the afternoon and early morning hours. Segregation is lowest during commute and work hours, indicating higher levels of interaction with people of different SES while at work or otherwise away from home. Segregation is higher during nighttime hours. This is driven by individuals returning to their home neighborhoods, which are more homogeneous in economic standing, as well as mechanically a result of ES being defined by rent value, such that people who live in the same household will have the same ES (and thus will be highly segregated).





Supplementary Figure S20: Across many activities, POIs in large cities are more differentiated and consequently more segregated. This figure shows that the trend towards more options, increased differentiation, and consequently higher segregation is consistent across many prominent POI categories. Here we find similar results for the 5 most frequently visited fine-grained Safegraph place features. The analyses for full-service correspond to Figures 2c-e, and we additionally show the same trend for snack bars, limited-service restaurants, stadiums, and performing arts centers (ranked 2-5 after full-service restaurants in terms of most frequently visited POIs among Safegraph places). Across the board, large, densely populated metropolitan areas are associated with increased options and economic differentiation of POIs, which may facilitate higher self-segregation.

	<i>Dependent variable:</i>					Interaction Segregation
	(1)	(2)	(3)	(4)	(5)	
Intercept	0.355*** (0.005)	0.355*** (0.004)	0.355*** (0.004)	0.356*** (0.004)	0.355*** (0.003)	0.355*** (0.003)
Population Density	0.039*** (0.005)		0.024*** (0.004)	0.022*** (0.004)	0.017*** (0.004)	0.017*** (0.004)
Gini Index (Estimated Rent)		0.064*** (0.004)	0.058*** (0.004)	0.059*** (0.004)	0.049*** (0.003)	0.050*** (0.003)
Political Alignment (% Democrat in 2016 Election)				0.009* (0.005)		0.006 (0.004)
Racial Demographics (% non-Hispanic White)				-0.005 (0.004)		0.003 (0.003)
Mean ES (Estimated Rent)				-0.009* (0.005)		-0.003 (0.004)
Walkability (Walkscore)					0.002 (0.003)	0.001 (0.004)
Commutability (% Commute to Work)					-0.012*** (0.004)	-0.013*** (0.004)
Conventional Segregation (NSI)					0.048*** (0.003)	0.047*** (0.003)
Observations	382	382	382	376	382	376
R ²	0.151	0.419	0.475	0.490	0.682	0.680
Adjusted R ²	0.149	0.417	0.472	0.483	0.678	0.673

*p<0.1; **p<0.05; ***p<0.01

Supplementary Table S7: Population density is significantly associated with interaction segregation, after controlling for MSA income inequality (Gini Index), political alignment (% Democrat in 2016 election), racial demographics (% non-Hispanic White), mean ES, walkability (Walkscore⁷⁶), commutability (% of residents commuting to work), and residential segregation (NSI). This table is from an analogous regression to the regression shown in Extended Data Table 1, using population density instead of population size (we look at each separately due to co-linearity between population size and density). Here we show the coefficients (after normalizing via z-scoring to have mean 0 and variance 1) from the primary specifications estimating the effect of population density on interaction segregation across all MSAs. Columns (1-5) are models specified with different subsets of covariates; Column 6 shows model specification with all covariates. Differences between sample size in models is due to missing data for several covariates in a small number of MSAs (Walkscores were not available for all MSAs). (*p < 0.1; **p < 0.05; *** p < 0.01).

Supplementary Table S8: Interaction Segregation and related variables (i.e. # Interactions, Mean ES, NSI, Gini Index, Population Size, and Bridging Index (BI) by MSA

MSA	Interaction Segregation	# Interactions	Mean ES	NSI	Gini	Pop. Size	BI
Abilene, TX	0.44	561,896.00	1,245.47	0.63	0.21	170,516.00	0.79
Akron, OH	0.55	2,211,810.00	1,338.21	0.71	0.27	704,367.00	0.65
Albany, GA	0.40	355,999.00	1,077.89	0.47	0.26	151,293.00	0.79
Albany, OR	0.27	153,057.00	1,425.95	0.40	0.11	124,977.00	0.95
Albany-Schenectady-Troy, NY	0.40	2,058,079.00	1,651.16	0.62	0.18	882,130.00	0.81
Albuquerque, NM	0.35	1,979,325.00	1,316.79	0.59	0.19	912,897.00	0.78
Alexandria, LA	0.37	267,228.00	1,011.98	0.52	0.23	153,604.00	0.88
Allentown-Bethlehem-Easton, PA-NJ	0.44	2,629,311.00	1,647.92	0.62	0.19	838,081.00	0.79
Altoona, PA	0.14	199,365.00	799.18	0.59	0.08	123,175.00	0.83
Amarillo, TX	0.45	2,027,531.00	1,358.46	0.62	0.26	264,955.00	0.82
Ames, IA	0.22	178,625.00	1,255.36	0.37	0.19	97,260.00	0.97
Anchorage, AK	0.37	1,189,861.00	1,921.88	0.65	0.17	400,647.00	0.84
Ann Arbor, MI	0.42	984,188.00	2,087.55	0.69	0.19	369,208.00	0.76
Anniston-Oxford-Jacksonville, AL	0.27	498,621.00	949.79	0.43	0.19	114,664.00	0.96
Appleton, WI	0.28	798,727.00	1,201.77	0.60	0.10	236,058.00	0.94
Asheville, NC	0.36	1,449,906.00	1,634.66	0.44	0.18	455,255.00	0.88
Athens-Clarke County, GA	0.26	431,582.00	1,427.27	0.46	0.19	208,997.00	0.83
Atlanta-Sandy Springs-Roswell, GA	0.50	41,054,246.00	1,805.49	0.70	0.22	5,874,249.00	0.62
Atlantic City-Hammonton, NJ	0.55	788,439.00	1,993.59	0.79	0.27	266,328.00	0.59
Auburn-Opelika, AL	0.35	415,339.00	1,409.80	0.44	0.19	161,641.00	0.86
Augusta-Richmond County, GA-SC	0.39	1,698,887.00	1,257.25	0.57	0.21	600,006.00	0.79
Austin-Round Rock, TX	0.53	13,378,670.00	1,954.85	0.70	0.19	2,115,230.00	0.66
Bakersfield, CA	0.29	2,168,162.00	1,399.40	0.69	0.18	888,988.00	0.81
Baltimore-Columbia-Towson, MD	0.61	15,120,132.00	1,898.23	0.80	0.17	2,798,587.00	0.60
Bangor, ME	0.21	56,124.00	1,264.12	0.65	0.17	151,190.00	0.91
Barnstable Town, MA	0.32	556,457.00	2,483.81	0.43	0.18	213,482.00	0.90
Baton Rouge, LA	0.45	3,449,982.00	1,374.01	0.70	0.14	831,182.00	0.81
Battle Creek, MI	0.37	329,581.00	1,013.54	0.66	0.13	134,358.00	0.84
Bay City, MI	0.23	204,294.00	904.87	0.48	0.11	104,189.00	0.97
Beaumont-Port Arthur, TX	0.41	1,935,665.00	1,276.93	0.63	0.16	412,616.00	0.83
Beckley, WV	0.22	61,971.00	988.00	0.41	0.19	118,639.00	0.94
Bellingham, WA	0.18	300,129.00	1,946.31	0.35	0.16	221,650.00	0.95
Bend-Redmond, OR	0.38	257,079.00	1,983.21	0.53	0.17	186,807.00	0.92
Billings, MT	0.36	316,679.00	1,370.83	0.56	0.18	170,740.00	0.91
Binghamton, NY	0.21	325,247.00	1,125.73	0.56	0.16	241,609.00	0.90
Birmingham-Hoover, AL	0.56	7,522,699.00	1,392.29	0.73	0.26	1,149,685.00	0.63
Bismarck, ND	0.18	341,245.00	1,397.18	0.48	0.13	132,418.00	0.95
Blacksburg-Christiansburg-Radford, VA	0.33	260,509.00	1,340.74	0.57	0.17	182,692.00	0.82
Bloomington, IL	0.34	599,796.00	1,245.47	0.62	0.18	188,754.00	0.89
Bloomington, IN	0.35	405,186.00	1,517.20	0.47	0.22	167,513.00	0.91
Bloomsburg-Berwick, PA	0.19	80,784.00	994.00	0.64	0.14	83,924.00	0.94
Boise City, ID	0.41	1,299,521.00	1,584.28	0.60	0.16	710,080.00	0.81
Boston-Cambridge-Newton, MA-NH	0.38	22,163,903.00	2,655.52	0.63	0.16	4,844,597.00	0.70
Boulder, CO	0.36	744,121.00	2,570.91	0.57	0.21	324,073.00	0.84
Bowling Green, KY	0.43	541,630.00	1,227.13	0.59	0.22	174,962.00	0.93
Bremerton-Silverdale, WA	0.40	636,416.00	2,044.31	0.60	0.16	266,550.00	0.77
Bridgeport-Stamford-Norwalk, CT	0.51	3,009,778.00	2,840.03	0.79	0.32	943,457.00	0.53
Brownsville-Harlingen, TX	0.33	1,204,081.00	1,106.57	0.49	0.17	423,181.00	0.88
Brunswick, GA	0.55	345,233.00	1,939.59	0.65	0.33	117,728.00	0.56
Buffalo-Cheektowaga-Niagara Falls, NY	0.41	3,152,570.00	1,285.13	0.68	0.20	1,129,660.00	0.69
Burlington, NC	0.30	489,581.00	1,237.01	0.40	0.21	163,529.00	0.89
Burlington-South Burlington, VT	0.43	140,807.00	1,990.17	0.33	0.16	218,881.00	0.86
California-Lexington Park, MD	0.19	442,775.00	1,739.01	0.35	0.12	112,413.00	0.98
Canton-Massillon, OH	0.41	1,037,327.00	1,235.25	0.38	0.29	399,418.00	0.81
Cape Coral-Fort Myers, FL	0.39	6,067,007.00	1,990.09	0.57	0.27	739,506.00	0.74
Cape Girardeau, MO-IL	0.24	174,129.00	1,013.11	0.36	0.18	96,873.00	0.99
Carbondale-Marion, IL	0.21	182,895.00	855.65	0.37	0.13	125,065.00	0.99
Carson City, NV	0.33	124,126.00	1,700.73	0.59	0.18	54,608.00	0.98
Casper, WY	0.18	103,682.00	1,377.58	0.30	0.21	79,556.00	0.97
Cedar Rapids, IA	0.33	1,010,446.00	1,200.65	0.47	0.14	270,594.00	0.96
Chambersburg-Waynesboro, PA	0.26	158,476.00	1,080.40	0.48	0.10	154,487.00	0.96
Champaign-Urbana, IL	0.32	799,317.00	1,229.14	0.67	0.19	239,877.00	0.88
Charleston, WV	0.25	365,352.00	988.62	0.44	0.21	214,398.00	0.93
Charleston-North Charleston, SC	0.47	4,062,901.00	2,014.52	0.69	0.24	775,089.00	0.66
Charlotte-Concord-Gastonia, NC-SC	0.50	12,750,805.00	1,699.80	0.68	0.24	2,524,863.00	0.64
Charlottesville, VA	0.31	510,779.00	1,840.79	0.50	0.21	233,586.00	0.95
Chattanooga, TN-GA	0.46	2,432,138.00	1,376.30	0.62	0.18	556,081.00	0.86
Cheyenne, WY	0.33	234,335.00	1,450.51	0.67	0.15	98,460.00	0.96
Chicago-Naperville-Elgin, IL-IN-WI	0.44	61,552,971.00	1,943.77	0.66	0.21	9,520,784.00	0.68
Chico, CA	0.29	324,613.00	1,772.67	0.55	0.17	229,207.00	0.92
Cincinnati, OH-KY-IN	0.47	10,110,144.00	1,533.13	0.66	0.22	2,179,858.00	0.76
Clarksville, TN-KY	0.30	989,270.00	1,100.48	0.56	0.16	285,691.00	0.83
Cleveland, TN	0.23	421,419.00	1,072.17	0.33	0.15	122,082.00	0.92
Cleveland-Elyria, OH	0.54	6,830,481.00	1,385.15	0.68	0.25	2,058,549.00	0.63
Coeur d'Alene, ID	0.13	243,473.00	1,680.34	0.32	0.15	157,485.00	0.97
College Station-Bryan, TX	0.40	1,243,139.00	1,430.02	0.59	0.18	258,825.00	0.87
Colorado Springs, CO	0.42	2,666,493.00	1,758.07	0.64	0.16	725,438.00	0.72
Columbia, MO	0.36	425,486.00	1,194.35	0.50	0.20	178,523.00	0.86
Columbia, SC	0.42	3,047,549.00	1,390.00	0.59	0.21	825,110.00	0.81
Columbus, GA-AL	0.47	724,780.00	1,143.38	0.72	0.24	303,436.00	0.74
Columbus, IN	0.41	250,666.00	1,411.07	0.55	0.22	82,429.00	0.94
Columbus, OH	0.55	9,849,191.00	1,623.39	0.71	0.23	2,082,475.00	0.69

Continued on next page

Supplementary Table S8: (cont'd) Interaction Segregation and related variables (i.e. # Interactions, Mean ES, NSI, Gini Index, Population Size, and Bridging Index (BI) by MSA

MSA	Interaction Segregation	# Interactions	Mean ES	NSI	Gini	Pop. Size	BI
Corpus Christi, TX	0.50	2,288,424.00	1,487.67	0.72	0.16	453,684.00	0.79
Corvallis, OR	0.24	104,179.00	1,936.53	0.48	0.15	91,567.00	0.97
Crestview-Fort Walton Beach-Destin, FL	0.43	1,868,711.00	1,852.07	0.53	0.24	271,959.00	0.76
Cumberland, MD-WV	0.26	30,347.00	936.53	0.51	0.14	98,566.00	0.98
Dallas-Fort Worth-Arlington, TX	0.51	48,228,424.00	1,996.27	0.73	0.22	7,407,944.00	0.62
Dalton, GA	0.14	235,701.00	876.66	0.89	0.14	143,872.00	0.77
Danville, IL	0.24	70,716.00	693.62	0.73	0.05	77,776.00	0.96
Daphne-Fairhope-Foley, AL	0.34	1,066,095.00	1,602.47	0.41	0.16	212,619.00	0.96
Davenport-Moline-Rock Island, IA-IL	0.45	945,921.00	1,287.37	0.70	0.22	381,854.00	0.76
Dayton, OH	0.49	2,617,342.00	1,278.37	0.70	0.25	803,713.00	0.73
Decatur, AL	0.29	403,755.00	981.61	0.42	0.12	151,888.00	0.94
Decatur, IL	0.35	233,031.00	919.63	0.76	0.15	105,533.00	0.86
Deltona-Daytona Beach-Ormond Beach, FL	0.33	5,092,950.00	1,686.87	0.46	0.18	648,117.00	0.82
Denver-Aurora-Lakewood, CO	0.35	11,589,449.00	2,312.40	0.67	0.17	2,892,979.00	0.77
Des Moines-West Des Moines, IA	0.46	1,966,610.00	1,577.32	0.53	0.22	645,100.00	0.76
Detroit-Warren-Dearborn, MI	0.57	15,495,989.00	1,554.43	0.77	0.26	4,321,704.00	0.49
Dothan, AL	0.33	466,788.00	1,142.03	0.45	0.23	147,923.00	0.89
Dover, DE	0.26	462,113.00	1,471.35	0.34	0.13	176,445.00	0.93
Dubuque, IA	0.36	201,954.00	1,313.10	0.57	0.18	97,009.00	0.86
Duluth, MN-WI	0.44	458,229.00	1,373.15	0.67	0.21	278,659.00	0.74
Durham-Chapel Hill, NC	0.44	2,019,082.00	1,713.50	0.57	0.19	566,491.00	0.83
East Stroudsburg, PA	0.24	452,426.00	1,526.37	0.39	0.12	168,089.00	0.95
Eau Claire, WI	0.18	377,072.00	1,114.77	0.41	0.11	167,436.00	0.95
El Centro, CA	0.18	271,797.00	1,340.14	0.69	0.12	181,574.00	0.96
El Paso, TX	0.32	1,667,796.00	1,157.84	0.60	0.17	845,145.00	0.76
Elizabethtown-Fort Knox, KY	0.19	245,608.00	1,241.61	0.24	0.22	150,531.00	0.91
Elkhart-Goshen, IN	0.37	567,181.00	1,186.01	0.41	0.17	204,310.00	0.89
Elmira, NY	0.36	121,786.00	1,173.73	0.58	0.20	84,874.00	0.78
Enid, OK	0.38	320,003.00	1,102.25	0.59	0.24	61,492.00	0.89
Erie, PA	0.34	571,583.00	1,129.58	0.61	0.21	273,892.00	0.85
Eugene, OR	0.29	514,167.00	1,600.88	0.49	0.15	375,617.00	0.93
Evansville, IN-KY	0.47	751,625.00	1,240.07	0.60	0.26	314,960.00	0.80
Fairbanks, AK	0.16	60,754.00	1,619.53	0.36	0.12	99,725.00	0.99
Fargo, ND-MN	0.30	816,028.00	1,333.79	0.52	0.12	241,619.00	0.92
Farmington, NM	0.28	67,876.00	1,265.57	0.52	0.17	126,902.00	0.98
Fayetteville, NC	0.27	1,253,869.00	1,082.27	0.48	0.19	385,380.00	0.90
Fayetteville-Springdale-Rogers, AR-MO	0.43	1,854,285.00	1,371.30	0.60	0.18	538,412.00	0.91
Flagstaff, AZ	0.31	184,120.00	2,012.92	0.42	0.15	141,107.00	0.96
Flint, MI	0.50	981,799.00	1,063.23	0.72	0.22	407,673.00	0.58
Florence, SC	0.37	312,804.00	1,335.06	0.49	0.21	205,546.00	0.93
Florence-Muscle Shoals, AL	0.22	331,068.00	912.14	0.50	0.14	147,100.00	0.97
Fond du Lac, WI	0.19	220,305.00	873.95	0.47	0.09	102,371.00	0.92
Fort Collins, CO	0.28	926,669.00	1,916.57	0.43	0.13	343,993.00	0.91
Fort Smith, AR-OK	0.32	638,594.00	887.85	0.60	0.16	281,990.00	0.93
Fort Wayne, IN	0.50	1,475,260.00	1,342.98	0.69	0.25	434,001.00	0.66
Fresno, CA	0.35	2,137,796.00	1,471.73	0.69	0.18	986,542.00	0.73
Gadsden, AL	0.30	465,122.00	894.00	0.59	0.19	102,937.00	0.89
Gainesville, FL	0.35	1,165,180.00	1,593.03	0.56	0.22	284,685.00	0.82
Gainesville, GA	0.34	758,361.00	1,756.64	0.32	0.22	199,439.00	0.85
Gettysburg, PA	0.22	205,160.00	1,410.01	0.34	0.10	102,367.00	0.99
Glens Falls, NY	0.27	131,519.00	1,486.70	0.59	0.21	125,917.00	0.84
Goldsboro, NC	0.27	275,443.00	1,088.76	0.28	0.18	123,257.00	0.87
Grand Forks, ND-MN	0.33	184,078.00	1,228.24	0.57	0.21	102,277.00	0.98
Grand Island, NE	0.25	236,404.00	1,138.90	0.40	0.14	84,862.00	1.00
Grand Junction, CO	0.32	248,844.00	1,410.36	0.59	0.16	151,406.00	0.88
Grand Rapids-Wyoming, MI	0.40	2,808,054.00	1,540.98	0.65	0.16	1,060,326.00	0.71
Grants Pass, OR	0.19	84,482.00	1,601.76	0.38	0.14	86,653.00	1.00
Great Falls, MT	0.27	156,642.00	1,210.84	0.51	0.13	81,604.00	1.00
Greeley, CO	0.44	749,425.00	1,912.90	0.57	0.14	305,274.00	0.79
Green Bay, WI	0.44	1,141,954.00	1,416.76	0.71	0.21	319,786.00	0.86
Greensboro-High Point, NC	0.48	2,269,305.00	1,226.69	0.67	0.26	763,486.00	0.69
Greenville, NC	0.36	702,454.00	1,259.95	0.34	0.22	178,617.00	0.94
Greenville-Anderson-Mauldin, SC	0.46	2,888,574.00	1,383.12	0.56	0.21	895,422.00	0.82
Gulfport-Biloxi-Pascagoula, MS	0.35	1,349,098.00	1,223.72	0.49	0.18	394,322.00	0.92
Hagerstown-Martinsburg, MD-WV	0.32	599,583.00	1,349.39	0.54	0.16	265,295.00	0.96
Hammond, LA	0.31	354,092.00	1,182.56	0.38	0.14	132,322.00	0.94
Hanford-Corcoran, CA	0.22	253,730.00	1,343.94	0.62	0.15	149,696.00	0.96
Harrisburg-Carlisle, PA	0.39	1,555,132.00	1,467.19	0.54	0.19	571,101.00	0.78
Harrisonburg, VA	0.26	290,643.00	1,278.25	0.42	0.16	134,220.00	0.95
Hartford-West Hartford-East Hartford, CT	0.43	2,241,050.00	1,710.35	0.66	0.18	1,206,719.00	0.77
Hattiesburg, MS	0.37	311,274.00	1,156.01	0.70	0.16	148,719.00	0.83
Hickory-Lenoir-Morganton, NC	0.41	640,647.00	1,244.75	0.47	0.18	367,004.00	0.91
Hilton Head Island-Bluffton-Beaufort, SC	0.39	732,993.00	2,115.09	0.53	0.21	214,890.00	0.82
Hinesville, GA	0.20	138,926.00	1,134.33	0.29	0.13	80,518.00	0.97
Homosassa Springs, FL	0.30	528,334.00	1,417.05	0.46	0.21	145,512.00	0.94
Hot Springs, AR	0.34	274,972.00	1,162.56	0.39	0.23	98,444.00	0.89
Houma-Thibodaux, LA	0.29	511,209.00	1,159.51	0.65	0.12	209,893.00	0.83
Houston-The Woodlands-Sugar Land, TX	0.47	63,151,024.00	1,866.72	0.72	0.22	6,905,695.00	0.66
Huntington-Ashland, WV-KY-OH	0.34	860,612.00	1,069.95	0.49	0.18	355,582.00	0.87
Huntsville, AL	0.45	1,623,341.00	1,313.85	0.66	0.20	455,741.00	0.81
Idaho Falls, ID	0.25	212,821.00	1,219.23	0.54	0.14	145,792.00	0.95
Indianapolis-Carmel-Anderson, IN	0.52	10,182,520.00	1,466.08	0.68	0.24	2,026,723.00	0.64

Continued on next page

Supplementary Table S8: (cont'd) Interaction Segregation and related variables (i.e. # Interactions, Mean ES, NSI, Gini Index, Population Size, and Bridging Index (BI) by MSA

MSA	Interaction Segregation	# Interactions	Mean ES	NSI	Gini	Pop. Size	BI
Iowa City, IA	0.36	473,387.00	1,499.89	0.44	0.20	171,470.00	0.97
Ithaca, NY	0.23	130,902.00	1,607.18	0.45	0.14	102,678.00	0.95
Jackson, MI	0.40	294,167.00	1,049.19	0.67	0.15	158,690.00	0.96
Jackson, MS	0.58	1,725,331.00	1,409.20	0.72	0.24	581,552.00	0.68
Jackson, TN	0.40	324,665.00	1,075.77	0.74	0.16	129,186.00	0.82
Jacksonville, FL	0.49	10,861,594.00	1,742.93	0.61	0.24	1,504,841.00	0.64
Jacksonville, NC	0.30	544,185.00	1,162.30	0.45	0.17	194,838.00	0.90
Janesville-Beloit, WI	0.25	328,187.00	1,002.54	0.65	0.10	162,320.00	0.76
Jefferson City, MO	0.24	349,139.00	1,010.56	0.45	0.16	151,298.00	0.96
Johnson City, TN	0.33	357,946.00	1,075.54	0.55	0.17	201,844.00	0.84
Johnstown, PA	0.20	232,779.00	741.42	0.62	0.11	133,054.00	0.98
Jonesboro, AR	0.36	308,755.00	1,106.28	0.64	0.17	131,158.00	0.94
Joplin, MO	0.21	356,202.00	864.53	0.41	0.12	178,330.00	0.97
Kahului-Wailuku-Lahaina, HI	0.22	246,346.00	3,043.60	0.37	0.16	166,491.00	0.97
Kalamazoo-Portage, MI	0.37	910,106.00	1,412.34	0.58	0.14	338,347.00	0.88
Kankakee, IL	0.30	374,176.00	1,249.33	0.70	0.10	110,544.00	0.81
Kansas City, MO-KS	0.54	8,835,941.00	1,541.78	0.75	0.25	2,127,259.00	0.64
Kennewick-Richland, WA	0.37	373,182.00	1,575.53	0.61	0.15	290,570.00	0.88
Killeen-Temple, TX	0.37	1,587,760.00	1,116.71	0.58	0.16	443,653.00	0.87
Kingsport-Bristol-Bristol, TN-VA	0.29	373,431.00	1,097.78	0.46	0.17	306,253.00	0.93
Kingston, NY	0.31	290,633.00	1,750.29	0.47	0.13	178,723.00	0.94
Knoxville, TN	0.43	2,704,521.00	1,467.11	0.60	0.23	875,797.00	0.76
Kokomo, IN	0.31	288,583.00	944.72	0.55	0.16	82,311.00	0.88
La Crosse-Onalaska, WI-MN	0.17	200,045.00	1,144.24	0.54	0.09	136,778.00	0.97
Lafayette, LA	0.39	1,650,845.00	1,126.34	0.67	0.16	490,107.00	0.76
Lafayette-West Lafayette, IN	0.33	768,405.00	1,226.26	0.57	0.16	220,337.00	0.88
Lake Charles, LA	0.27	207,541.00	1,286.99	0.62	0.13	209,256.00	0.88
Lake Havasu City-Kingman, AZ	0.30	321,719.00	1,255.77	0.59	0.18	207,114.00	0.83
Lakeland-Winter Haven, FL	0.31	4,246,971.00	1,435.22	0.43	0.18	685,830.00	0.90
Lancaster, PA	0.28	1,111,168.00	1,304.06	0.54	0.11	541,054.00	0.86
Lansing-East Lansing, MI	0.42	1,146,004.00	1,249.29	0.65	0.19	480,353.00	0.80
Laredo, TX	0.55	682,291.00	1,308.82	0.76	0.18	273,982.00	0.72
Las Cruces, NM	0.27	340,688.00	1,194.16	0.44	0.18	216,186.00	0.97
Las Vegas-Henderson-Paradise, NV	0.27	10,258,483.00	1,703.98	0.51	0.20	2,183,310.00	0.82
Lawrence, KS	0.19	423,826.00	1,417.50	0.47	0.20	120,629.00	0.94
Lawton, OK	0.30	337,308.00	933.32	0.58	0.21	127,589.00	0.88
Lebanon, PA	0.43	276,347.00	1,155.40	0.70	0.13	139,566.00	0.88
Lewiston, ID-WA	0.20	43,087.00	1,309.23	0.23	0.13	62,881.00	0.99
Lewiston-Auburn, ME	0.21	110,078.00	1,111.21	0.53	0.07	107,569.00	0.98
Lexington-Fayette, KY	0.42	2,391,314.00	1,321.56	0.62	0.21	512,732.00	0.85
Lima, OH	0.35	228,949.00	983.20	0.60	0.16	103,069.00	0.95
Lincoln, NE	0.35	1,808,658.00	1,395.51	0.57	0.16	331,179.00	0.87
Little Rock-North Little Rock-Conway, AR	0.51	3,070,717.00	1,192.95	0.72	0.20	737,991.00	0.77
Logan, UT-ID	0.20	154,902.00	1,210.21	0.60	0.10	138,052.00	0.92
Longview, TX	0.35	262,120.00	1,243.45	0.61	0.18	218,594.00	0.86
Longview, WA	0.38	173,605.00	1,459.50	0.52	0.18	106,900.00	1.00
Los Angeles-Long Beach-Anaheim, CA	0.44	110,526,499.00	2,970.24	0.75	0.20	13,298,709.00	0.66
Louisville/Jefferson County, KY-IN	0.51	4,567,106.00	1,436.51	0.67	0.23	1,292,809.00	0.71
Lubbock, TX	0.45	1,549,243.00	1,381.12	0.55	0.22	316,588.00	0.83
Lynchburg, VA	0.31	496,432.00	1,201.66	0.58	0.18	261,954.00	0.84
Macon-Bibb County, GA	0.46	616,989.00	1,232.02	0.58	0.26	229,081.00	0.81
Madera, CA	0.30	249,720.00	1,396.67	0.47	0.14	155,904.00	0.96
Madison, WI	0.37	1,737,217.00	1,628.86	0.60	0.17	654,577.00	0.89
Manchester-Nashua, NH	0.46	929,901.00	2,027.43	0.69	0.18	413,157.00	0.88
Manhattan, KS	0.31	268,899.00	1,285.52	0.38	0.19	97,954.00	0.99
Mankato-North Mankato, MN	0.26	274,393.00	1,407.12	0.35	0.13	100,945.00	0.97
Mansfield, OH	0.30	203,318.00	872.54	0.50	0.12	120,543.00	0.88
McAllen-Edinburg-Mission, TX	0.36	2,672,266.00	1,165.08	0.45	0.22	858,323.00	0.87
Medford, OR	0.32	250,898.00	1,584.27	0.48	0.14	216,761.00	0.93
Memphis, TN-MS-AR	0.56	5,217,305.00	1,409.82	0.74	0.27	1,347,576.00	0.58
Merced, CA	0.24	471,172.00	1,489.37	0.57	0.13	271,340.00	0.92
Miami-Fort Lauderdale-West Palm Beach, FL	0.44	147,998,127.00	2,642.33	0.67	0.28	6,149,687.00	0.70
Michigan City-La Porte, IN	0.31	262,677.00	1,150.07	0.58	0.16	109,911.00	0.93
Midland, MI	0.35	156,234.00	1,161.45	0.63	0.18	83,245.00	0.96
Midland, TX	0.33	821,156.00	2,759.87	0.61	0.19	170,948.00	0.91
Milwaukee-Waukesha-West Allis, WI	0.60	5,452,737.00	1,428.80	0.77	0.24	1,575,151.00	0.63
Minneapolis-St. Paul-Bloomington, MN-WI	0.41	17,181,042.00	1,970.38	0.56	0.18	3,592,669.00	0.78
Missoula, MT	0.20	144,440.00	1,486.18	0.49	0.17	117,863.00	0.96
Mobile, AL	0.28	1,700,477.00	1,102.96	0.52	0.16	414,515.00	0.85
Modesto, CA	0.20	1,396,841.00	1,673.04	0.53	0.12	545,267.00	0.91
Monroe, LA	0.39	420,225.00	1,057.93	0.56	0.25	178,211.00	0.81
Monroe, MI	0.20	298,001.00	1,123.01	0.45	0.08	149,592.00	0.99
Montgomery, AL	0.47	933,055.00	1,116.61	0.73	0.18	374,042.00	0.71
Morgantown, WV	0.28	144,020.00	1,375.31	0.53	0.22	139,739.00	0.99
Morristown, TN	0.35	127,639.00	1,123.38	0.35	0.16	117,843.00	0.95
Mount Vernon-Anacortes, WA	0.29	164,737.00	1,848.65	0.46	0.13	126,026.00	0.99
Muncie, IN	0.34	325,604.00	951.96	0.58	0.19	115,389.00	0.91
Muskegon, MI	0.32	338,931.00	1,059.59	0.50	0.15	173,656.00	0.86
Myrtle Beach-Conway-North Myrtle Beach, SC-NC	0.24	1,937,451.00	1,614.84	0.27	0.22	463,386.00	0.87
Napa, CA	0.19	813,681.00	3,152.52	0.45	0.18	140,386.00	0.95
Naples-Immokalee-Marco Island, FL	0.45	3,239,165.00	3,865.02	0.70	0.36	372,345.00	0.77
Nashville-Davidson-Murfreesboro-Franklin, TN	0.50	10,766,763.00	1,845.97	0.74	0.22	1,900,584.00	0.62

Continued on next page

Supplementary Table S8: (cont'd) Interaction Segregation and related variables (i.e. # Interactions, Mean ES, NSI, Gini Index, Population Size, and Bridging Index (BI) by MSA

MSA	Interaction Segregation	# Interactions	Mean ES	NSI	Gini	Pop. Size	BI
New Bern, NC	0.32	309,000.00	1,270.00	0.47	0.19	125,010.00	0.89
New Haven-Milford, CT	0.37	2,382,587.00	1,669.24	0.53	0.19	857,794.00	0.76
New Orleans-Metairie, LA	0.40	6,489,654.00	1,556.45	0.68	0.16	1,270,465.00	0.81
New York-Newark-Jersey City, NY-NJ-PA	0.40	168,755,438.00	2,597.55	0.69	0.21	19,998,951.00	0.57
Niles-Benton Harbor, MI	0.51	301,883.00	1,248.76	0.78	0.19	154,362.00	0.80
North Port-Sarasota-Bradenton, FL	0.40	6,601,216.00	2,284.53	0.61	0.27	805,139.00	0.73
Norwich-New London, CT	0.27	286,633.00	1,478.91	0.60	0.15	267,826.00	0.87
Ocala, FL	0.30	1,669,292.00	1,370.81	0.39	0.21	353,717.00	0.90
Odessa, TX	0.28	687,081.00	2,187.43	0.54	0.13	157,173.00	0.88
Ogden-Clearfield, UT	0.38	1,513,318.00	1,609.07	0.63	0.15	664,589.00	0.84
Oklahoma City, OK	0.46	11,029,080.00	1,417.61	0.63	0.27	1,383,249.00	0.74
Olympia-Tumwater, WA	0.19	644,378.00	1,873.77	0.37	0.14	280,289.00	0.98
Omaha-Council Bluffs, NE-IA	0.49	3,944,289.00	1,583.10	0.64	0.22	932,217.00	0.69
Orlando-Kissimmee-Sanford, FL	0.42	25,094,242.00	1,870.76	0.57	0.21	2,512,917.00	0.72
Oshkosh-Neenah, WI	0.36	624,252.00	1,196.65	0.58	0.20	170,375.00	0.90
Owensboro, KY	0.36	203,069.00	1,064.78	0.48	0.21	118,543.00	0.95
Oxnard-Thousand Oaks-Ventura, CA	0.34	3,449,664.00	3,029.11	0.67	0.17	850,802.00	0.81
Palm Bay-Melbourne-Titusville, FL	0.43	4,331,998.00	1,826.84	0.52	0.20	588,265.00	0.77
Panama City, FL	0.28	1,089,950.00	2,071.06	0.40	0.20	200,168.00	0.93
Parkersburg-Vienna, WV	0.25	140,256.00	1,096.22	0.37	0.18	90,873.00	0.95
Pensacola-Ferry Pass-Brent, FL	0.42	3,672,000.00	1,405.47	0.55	0.21	487,527.00	0.78
Peoria, IL	0.46	914,882.00	1,178.67	0.63	0.22	371,810.00	0.80
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.53	24,822,104.00	1,802.57	0.74	0.20	6,078,451.00	0.61
Phoenix-Mesa-Scottsdale, AZ	0.48	17,152,709.00	1,746.59	0.75	0.19	4,761,694.00	0.61
Pine Bluff, AR	0.15	171,555.00	759.22	0.52	0.10	90,923.00	0.95
Pittsburgh, PA	0.47	7,756,479.00	1,348.23	0.70	0.24	2,330,283.00	0.72
Pittsfield, MA	0.37	128,317.00	1,477.61	0.52	0.15	126,485.00	0.86
Pocatello, ID	0.28	116,547.00	1,158.04	0.60	0.18	85,641.00	0.97
Port St. Lucie, FL	0.44	4,896,670.00	2,082.29	0.62	0.23	473,192.00	0.72
Portland-South Portland, ME	0.30	493,445.00	1,896.31	0.52	0.16	532,280.00	0.80
Portland-Vancouver-Hillsboro, OR-WA	0.34	6,614,427.00	2,016.66	0.58	0.16	2,456,462.00	0.86
Prescott, AZ	0.38	384,045.00	1,627.89	0.60	0.17	228,055.00	0.88
Providence-Warwick, RI-MA	0.40	5,763,343.00	1,744.07	0.63	0.16	1,617,057.00	0.79
Provo-Orem, UT	0.32	1,219,235.00	1,546.11	0.66	0.13	617,751.00	0.84
Pueblo, CO	0.36	443,458.00	1,270.67	0.65	0.18	166,426.00	0.85
Punta Gorda, FL	0.34	1,184,701.00	1,829.41	0.60	0.21	181,537.00	0.84
Racine, WI	0.42	548,195.00	1,383.27	0.63	0.16	195,949.00	0.79
Raleigh, NC	0.46	8,986,021.00	1,697.93	0.62	0.16	1,334,342.00	0.79
Rapid City, SD	0.32	250,301.00	1,441.12	0.41	0.20	146,869.00	0.88
Reading, PA	0.49	1,036,022.00	1,344.18	0.68	0.18	417,524.00	0.71
Redding, CA	0.23	258,289.00	1,573.52	0.43	0.19	179,539.00	0.99
Reno, NV	0.47	1,550,552.00	1,953.99	0.63	0.18	461,336.00	0.70
Richmond, VA	0.49	4,524,531.00	1,626.59	0.71	0.20	1,292,911.00	0.68
Riverside-San Bernardino-Ontario, CA	0.43	19,908,134.00	2,103.34	0.70	0.17	4,570,427.00	0.68
Roanoke, VA	0.38	811,899.00	1,291.22	0.58	0.20	313,488.00	0.80
Rochester, MN	0.33	648,812.00	1,540.01	0.58	0.16	217,828.00	0.94
Rochester, NY	0.45	2,870,914.00	1,536.09	0.66	0.19	1,071,589.00	0.77
Rockford, IL	0.40	946,786.00	1,225.98	0.67	0.18	338,252.00	0.72
Rocky Mount, NC	0.33	366,999.00	977.64	0.55	0.16	146,769.00	0.81
Rome, GA	0.36	264,276.00	1,078.32	0.39	0.16	97,427.00	0.99
Sacramento-Roseville-Arden-Arcade, CA	0.39	7,101,248.00	2,048.69	0.66	0.15	2,320,381.00	0.70
Saginaw, MI	0.37	355,021.00	927.43	0.70	0.18	191,996.00	0.77
Salem, OR	0.27	670,775.00	1,603.26	0.45	0.12	424,968.00	0.94
Salinas, CA	0.36	952,169.00	2,642.69	0.65	0.17	435,477.00	0.77
Salisbury, MD-DE	0.48	875,922.00	1,462.24	0.68	0.14	404,067.00	0.78
Salt Lake City, UT	0.33	3,468,862.00	1,763.34	0.60	0.15	1,205,238.00	0.77
San Angelo, TX	0.34	380,590.00	1,321.06	0.66	0.14	119,200.00	0.83
San Antonio-New Braunfels, TX	0.53	14,354,046.00	1,596.19	0.70	0.20	2,474,274.00	0.64
San Diego-Carlsbad, CA	0.42	13,807,983.00	2,854.26	0.73	0.20	3,325,468.00	0.71
San Francisco-Oakland-Hayward, CA	0.41	37,492,367.00	3,925.90	0.71	0.21	4,710,693.00	0.68
San Jose-Sunnyvale-Santa Clara, CA	0.37	8,012,471.00	3,766.32	0.78	0.16	1,993,582.00	0.68
San Luis Obispo-Paso Robles-Arroyo Grande, CA	0.22	656,686.00	2,601.39	0.47	0.13	282,838.00	0.94
Santa Cruz-Watsonville, CA	0.27	611,337.00	3,306.01	0.52	0.13	275,105.00	0.85
Santa Fe, NM	0.45	182,572.00	2,075.86	0.61	0.23	149,617.00	0.91
Santa Maria-Santa Barbara, CA	0.52	1,124,975.00	3,039.10	0.75	0.28	445,606.00	0.59
Santa Rosa, CA	0.22	1,144,462.00	2,793.94	0.54	0.11	503,246.00	0.92
Savannah, GA	0.39	1,601,410.00	1,599.94	0.57	0.19	386,337.00	0.83
Scranton-Wilkes-Barre-Hazleton, PA	0.30	1,064,769.00	1,077.93	0.62	0.16	555,645.00	0.90
Seattle-Tacoma-Bellevue, WA	0.44	18,136,495.00	2,474.85	0.64	0.20	3,884,469.00	0.73
Sebastian-Vero Beach, FL	0.52	1,148,601.00	2,259.56	0.71	0.31	154,314.00	0.78
Sebring, FL	0.26	352,266.00	1,291.88	0.31	0.21	104,060.00	0.98
Sheboygan, WI	0.35	220,669.00	1,204.87	0.58	0.10	115,235.00	0.86
Sherman-Denison, TX	0.39	648,078.00	1,321.20	0.46	0.14	131,214.00	0.95
Shreveport-Bossier City, LA	0.50	1,110,198.00	1,247.36	0.65	0.26	439,631.00	0.79
Sierra Vista-Douglas, AZ	0.21	153,651.00	992.99	0.70	0.14	124,990.00	0.93
Sioux City, IA-NE-SD	0.24	383,415.00	1,137.14	0.41	0.14	168,218.00	0.97
Sioux Falls, SD	0.27	733,324.00	1,239.36	0.68	0.13	260,521.00	0.98
South Bend-Mishawaka, IN-MI	0.48	1,021,877.00	1,289.97	0.67	0.26	321,447.00	0.86
Spartanburg, SC	0.39	913,944.00	1,295.96	0.46	0.23	334,130.00	0.93
Spokane-Spokane Valley, WA	0.34	944,039.00	1,472.98	0.58	0.16	563,958.00	0.88
Springfield, IL	0.40	797,160.00	1,152.17	0.79	0.17	209,175.00	0.87
Springfield, MA	0.45	1,422,143.00	1,632.08	0.69	0.16	629,506.00	0.82

Continued on next page

Supplementary Table S8: (cont'd) Interaction Segregation and related variables (i.e. # Interactions, Mean ES, NSI, Gini Index, Population Size, and Bridging Index (BI) by MSA

MSA	Interaction Segregation	# Interactions	Mean ES	NSI	Gini	Pop. Size	BI
Springfield, MO	0.39	1,300,346.00	1,100.41	0.59	0.23	462,300.00	0.85
Springfield, OH	0.39	361,791.00	895.90	0.61	0.18	134,649.00	0.89
St. Cloud, MN	0.23	518,337.00	1,273.69	0.40	0.14	198,106.00	0.95
St. George, UT	0.27	243,300.00	1,715.91	0.35	0.17	165,859.00	0.94
St. Joseph, MO-KS	0.27	258,535.00	916.56	0.69	0.14	126,598.00	0.88
St. Louis, MO-IL	0.51	11,016,511.00	1,413.67	0.72	0.25	2,805,850.00	0.62
State College, PA	0.35	215,274.00	1,600.27	0.54	0.18	162,250.00	0.89
Staunton-Waynesboro, VA	0.24	282,626.00	1,324.76	0.28	0.17	121,984.00	0.94
Stockton-Lodi, CA	0.42	2,116,634.00	1,865.80	0.73	0.15	742,516.00	0.72
Sumter, SC	0.35	237,013.00	1,050.69	0.46	0.21	106,514.00	0.92
Syracuse, NY	0.37	1,932,194.00	1,483.28	0.63	0.19	651,048.00	0.82
Tallahassee, FL	0.45	1,973,991.00	1,428.85	0.71	0.22	383,467.00	0.75
Tampa-St. Petersburg-Clearwater, FL	0.45	24,564,393.00	1,805.89	0.62	0.22	3,091,225.00	0.73
Terre Haute, IN	0.32	369,258.00	930.48	0.48	0.18	170,022.00	0.94
Texarkana, TX-AR	0.30	240,478.00	956.50	0.61	0.12	150,254.00	0.89
The Villages, FL	0.39	786,309.00	1,711.00	0.69	0.12	124,933.00	0.90
Toledo, OH	0.53	1,775,752.00	1,217.56	0.69	0.25	603,830.00	0.63
Topeka, KS	0.43	1,205,672.00	1,084.47	0.64	0.20	233,153.00	0.93
Trenton, NJ	0.64	1,240,113.00	2,005.70	0.85	0.21	368,602.00	0.54
Tucson, AZ	0.39	2,564,383.00	1,362.93	0.73	0.17	1,027,502.00	0.72
Tulsa, OK	0.49	5,223,272.00	1,242.38	0.72	0.20	991,610.00	0.77
Tuscaloosa, AL	0.38	1,468,839.00	1,332.43	0.60	0.18	242,700.00	0.91
Twin Falls, ID	0.27	182,971.00	1,187.85	0.41	0.13	109,037.00	0.98
Tyler, TX	0.27	593,804.00	1,432.30	0.41	0.19	227,460.00	0.87
Urban Honolulu, HI	0.33	1,368,021.00	2,616.52	0.62	0.19	986,429.00	0.85
Utica-Rome, NY	0.31	374,846.00	1,123.98	0.66	0.17	292,336.00	0.80
Valdosta, GA	0.32	380,832.00	1,136.00	0.52	0.23	145,403.00	0.94
Vallejo-Fairfield, CA	0.18	1,878,258.00	2,372.07	0.67	0.11	443,877.00	0.85
Victoria, TX	0.37	374,597.00	1,529.97	0.55	0.20	99,651.00	0.94
Vineland-Bridgeton, NJ	0.37	354,594.00	1,371.71	0.63	0.09	151,748.00	0.90
Virginia Beach-Norfolk-Newport News, VA-NC	0.45	6,944,774.00	1,666.43	0.62	0.20	1,724,876.00	0.75
Visalia-Porterville, CA	0.25	854,866.00	1,309.53	0.48	0.16	463,097.00	0.95
Waco, TX	0.41	1,245,450.00	1,334.52	0.55	0.20	268,550.00	0.89
Walla Walla, WA	0.22	53,561.00	1,448.66	0.39	0.13	64,675.00	1.00
Warner Robins, GA	0.40	600,913.00	1,271.38	0.54	0.20	191,227.00	0.86
Washington-Arlington-Alexandria, DC-VA-MD-WV	0.46	127,482,444.00	2,461.96	0.70	0.18	6,200,001.00	0.71
Waterloo-Cedar Falls, IA	0.38	326,152.00	1,111.17	0.62	0.18	169,553.00	0.84
Watertown-Fort Drum, NY	0.24	113,127.00	1,273.60	0.32	0.15	113,063.00	0.97
Wausau, WI	0.24	299,543.00	1,069.93	0.49	0.12	135,415.00	0.93
Weirton-Steubenville, WV-OH	0.24	235,799.00	823.09	0.52	0.12	118,181.00	0.94
Wenatchee, WA	0.21	126,600.00	1,748.88	0.31	0.14	118,646.00	1.00
Wheeling, WV-OH	0.26	102,043.00	1,100.65	0.53	0.15	141,228.00	0.87
Wichita Falls, TX	0.41	729,999.00	1,253.35	0.53	0.26	151,180.00	0.93
Wichita, KS	0.46	3,213,430.00	1,206.35	0.64	0.23	644,949.00	0.78
Williamsport, PA	0.23	148,520.00	1,025.19	0.38	0.14	113,930.00	0.99
Wilmington, NC	0.46	1,684,931.00	1,859.12	0.58	0.26	289,425.00	0.77
Winchester, VA-WV	0.33	296,124.00	1,560.81	0.50	0.16	138,107.00	0.95
Winston-Salem, NC	0.44	1,872,668.00	1,256.11	0.57	0.22	666,746.00	0.79
Worcester, MA-CT	0.48	2,650,313.00	1,777.27	0.71	0.17	942,303.00	0.74
Yakima, WA	0.33	222,165.00	1,236.55	0.54	0.15	250,377.00	0.88
York-Hanover, PA	0.41	1,303,401.00	1,430.52	0.56	0.18	445,722.00	0.85
Youngstown-Warren-Boardman, OH-PA	0.34	1,336,389.00	951.78	0.64	0.19	541,875.00	0.71
Yuba City, CA	0.29	309,905.00	1,574.77	0.57	0.13	173,213.00	0.89
Yuma, AZ	0.22	271,294.00	1,079.08	0.55	0.17	209,756.00	1.00