The Pandemic and Changing Patterns of Segregation in America

Andrew Renninger[1]

[1]University of Pennsylvania, Wharton School of Business, Weitzman School of Design

February 15, 2021

Summary

The following presents patterns of "experienced segregation"—interaction between groups in daily life—across time and space based on mobile phone data. Constructing a spatial interaction network for the Bay Area and New York City, we identify changing structure—marked by falling degree centrality and network density—and rising segregation. Expanding to the 100 largest metropolitan areas in America, we find mixed results, but a subtle relationship between significant changes to the network and segregation.

KEYWORDS: segregation, mobility, GPS

1. Introduction

The pandemic altered mobility patterns across the world, replacing previously dense networks of interactions between commercial and residential areas with comparably sparse ones. It has created a laboratory for understanding spatial interaction networks in cities as mobility data show changing travel patterns, which influence how communities interact. What impact has the pandemic had on integration in cities, and is there evidence of lasting changes to network structure?

In this project, we use phone tracking data containing origin-destination flows between home neighborhoods and points of interest to identify properties and variations of urban socio-spatial networks, using structural measurements to compare cities across time and space. For each metropolitan area in the United States, we compute a series of network statistics. Performing community detection, we explore the consequences of these changes: community size along with racial profile contribute to the relative sorting of the city as a whole, which allows us to compare the integration of cities across time. Across our sample, network density fell in April and has not yet returned to January or February levels, as did degree centrality; while segregation—measured as the homogeneity of detected communities—rose in New York City and the Bay area, we find little evidence that it rose systematically across the sample.

In measuring demographic composition, we build on work by economists exploring the concept of experienced segregation: the notion that residences—how we typically capture segregation—are only one aspect of daily life, and others may warrant consideration. Davis et al. (2019) use restaurant reviews to determine the degree to which different communities mix, estimating that individuals are less likely to visit restaurants in demographically dissimilar neighborhoods. Athey et al. (2020) use mobile phone data to monitor interactions between individuals, finding that residential segregation is greater than an experiential measure of it in all American cities studied, and that cities with high residential segregation have high segregation of activities.

Numerous studies apply theories of experienced segregation to the pandemic (see: Bassolas, Sousa, Nicosia 2021 and Chang et al. 2020), attempting to understand the relationship between experienced segregation and risk. Bonaccorsi et al. (2020) document differences in mobility reduction during Italy's first wave, finding that lower and income and higher inequality in an area corresponded to

[1] arenn@design.upenn.edu

stronger contractions.

2. Data and Methods

The following work develops rigorous approaches to understanding urban network structure by constructing social graphs with the growing body of mobility data. We use data from SafeGraph, which provides locations for 5 million points of interest in the United States along with a representative sample of visits made to them by American devices and with data on origin neighborhoods.[2] With these records, we construct urban networks for the 100 largest metropolitan areas by population.

Following Batty (2013), the basic construction of each city's network (*G*) is an origin Census block group (*V1*), a destination Census block group (*V2*), reducing each point of interest to its constituent neighborhood so our graph has a single model, and a weight (*w*) that represents the number of visitors moving between those vertices, or the weight of the edge. We then present metrics for assessing changes across and between networks, including structural correlation (Butts and Carley 2001), network density and aggregate measures—median, variance—of centrality to compare mobility networks.[3] Below we show this representation by month for New York City and the Bay Area.

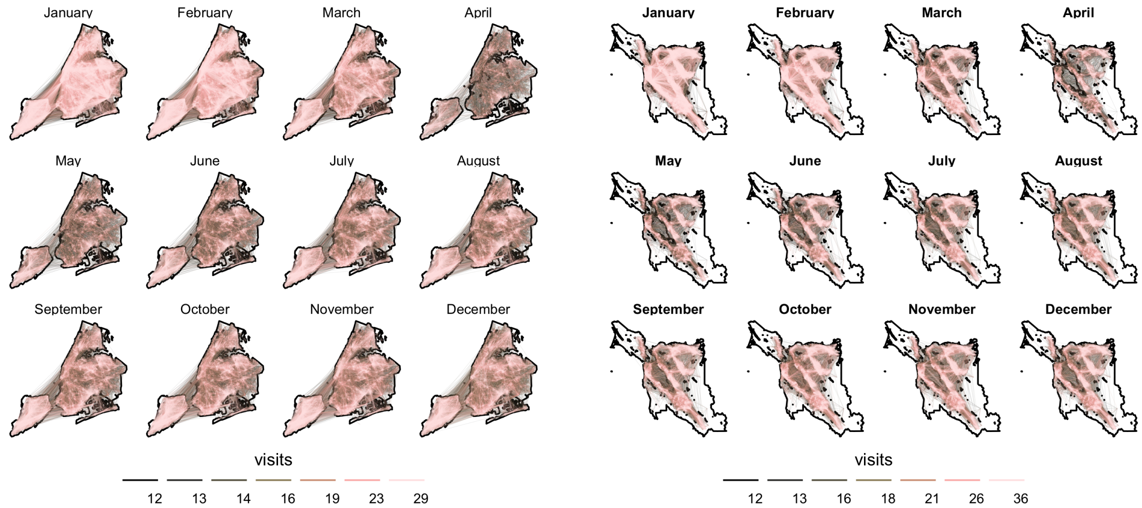

Figure 1 Mobility flows for New York City (left) and the Bay Area (right). Visits represent the number of people flowing along that particular desire line.

We also use community detection to inform our understanding of the network. Several community detection approaches inform our understanding of these dense networks. Prestby et al. (2019) use the Louvain method to demarcate communities from mobility data; a recent study of mobility patterns during the pandemic by Gibbs et al. (2020) leverages the InfoMap approach—which partitions the movement of a random walker (Bohlin et a. 2014)—to identify changing communities in the United Kingdom. Because mobility data is dense and large, the InfoMap method is suitable, according to Yang, Algesheimer, and Tessone (2016). We employ it below, running it multiple times and assigning neighborhoods based on modal label, following Gao et al. (2018). With these communities we can then track changes to community size—mean, standard deviation—along with demographic

[2] Concerns about the sample are valid: while Safegraph attempts to match the demographic composition of the users it tracks to estimates from the Census, the sample is representative nationally (Squire 2019), but may indeed suffer from biases at fine spatial resolutions such as we employ.
[3] For an explanation of centrality and density, see Jackson (2010).

composition.

To understand the composition of these communities we generate statistics for each community and summary statistics for each city, extending Prestby et al. (2019). Harris and Owen (2018) develop a dissimilarity index, which says the proportion of the population that would need to move to create an even demographic composition across areal units, with which we quantify the impact of changes to community structure.

3. Results

3.1. The Five Boroughs in Focus

We begin by exploring changes to network structure in New York City's five counties. As Figure 1 shows, the volume of traffic dropped in April and is still down compared to January and February, yet we see in Figure 2 a lower correlation between November and January than between April and January. No month is as similar to January as February, implying that the pandemic is still influencing mobility, but that its influence is changing as people adapt—and perhaps as travel over the holidays in November and December provide an additional shock.

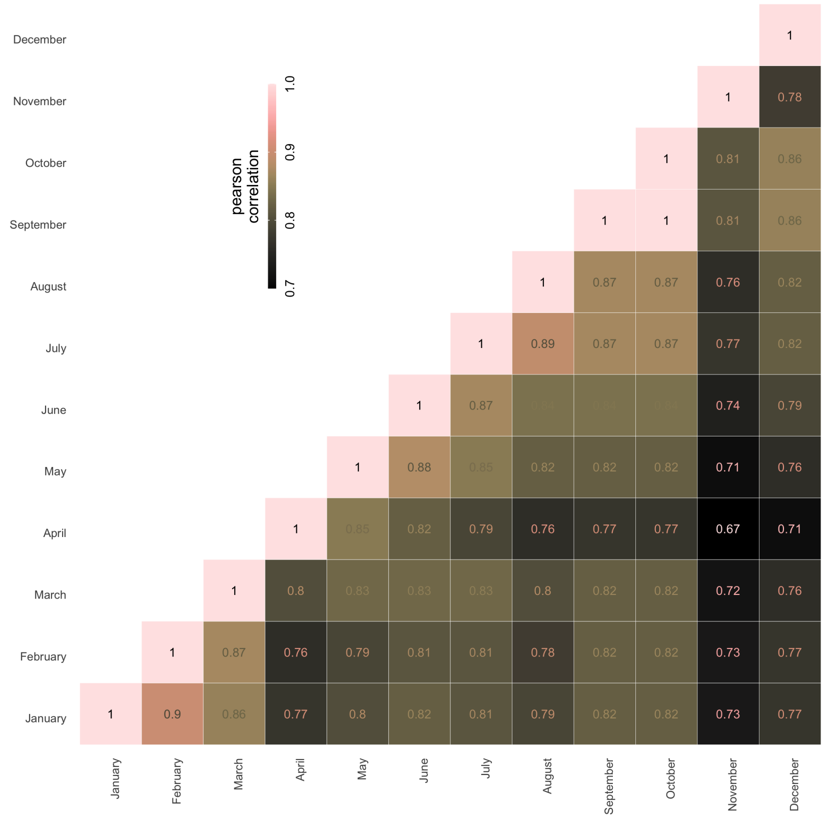

Figure 2 Graph correlations between months in New York City.

We now look at the consequences of these changes in Figure 3. Community size and network density fell in April; these communities were more dissimilar, indicating that neighborhoods interacted with similar neighborhoods—demographically—and did not interact with dissimilar ones. This suggests that experienced segregation may have increased.[4]

[4] Although, because we evaluate neighborhoods rather than individuals, we cannot say so definitively.

month	mean community size	network density	dissimilarity index
January	21.288525	0.004519251	0.528
February	22.389655	0.004149286	0.537
March	18.290141	0.002826970	0.546
April	8.378065	0.001079175	0.553
May	10.679276	0.001444752	0.545
June	11.391228	0.001762764	0.526
July	12.046382	0.001960142	0.542
August	12.068773	0.001957865	0.532
September	15.570743	0.002068700	0.532
October	15.914216	0.002083670	0.529
November	13.443064	0.001836994	0.528
December	12.908549	0.001779489	0.518

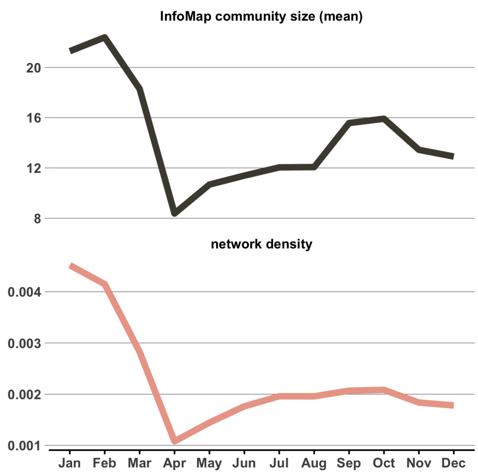

Figure 3 Dissimilarity index again changes to community structure.

3.2. The Bay Area in Focus

Next we document these changes within the Bay Area, which includes six counties adjacent to the San Francisco Bay. In Figure 4, We find stronger correlations from one month to the next, even during April and May, compared to New York City. Yet performing the same community detection and segregation assessment in Figure 5 as above, we find cleaner results in California than in New York: network density decreases to its lowest levels in April, where dissimilarity increases to its highest, and subsequently increases—though not to its highest levels—in the following months, as dissimilarity decreases.

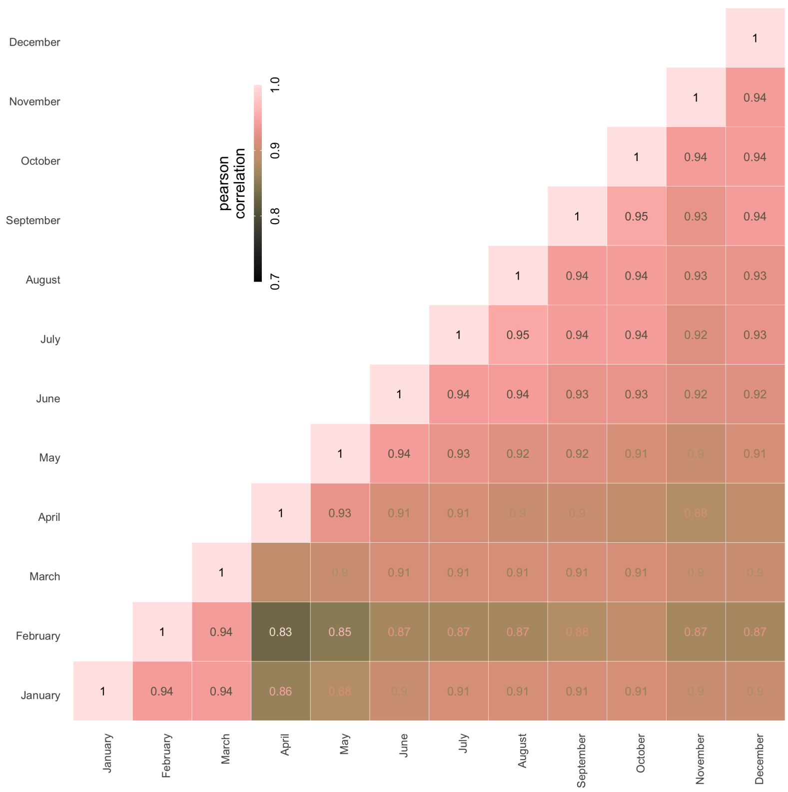

Figure 4 Graph correlations between months in the Bay Area.

month	mean community size	network density	dissimilarity index
January	117.00000	0.009403745	0.305
February	90.40909	0.008473496	0.312
March	58.50000	0.005690020	0.326
April	22.34831	0.002799278	0.335
May	24.40491	0.003430042	0.323
June	28.01408	0.004097846	0.326
July	25.50000	0.004192281	0.324
August	33.15000	0.004260673	0.325
September	33.71186	0.004208209	0.322
October	36.49541	0.004619574	0.319
November	29.90977	0.004195125	0.319
December	30.83721	0.004125658	0.321

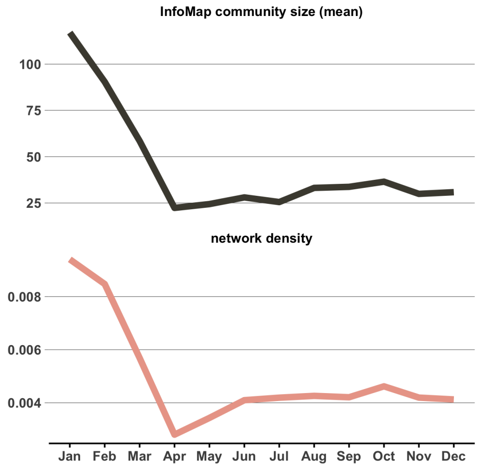

Figure 5 Dissimilarity index against changes to community structure.

3.3. Expanded Results

In this section we calculate the same statistics for the 100 largest metropolitan areas in America. The maps in Figure 6 show a clear spatial relationship, with the Southeast showing comparably little change while the Northeast and West show strong differences between January and April. There are interacting factors: Miami, as a coastal metropolis, is in this former group but behaves more like the latter. When we arrange correlations on a North-South axis in Figure 9, we see that Miami is anomalous: the lowest latitudes have the highest January-April correlation while the opposite holds true for the highest latitudes. Many of these disrupted cities are in the Acela corridor, stretching from Washington to Boston; the only exceptions are rural cities in Iowa and Idaho.

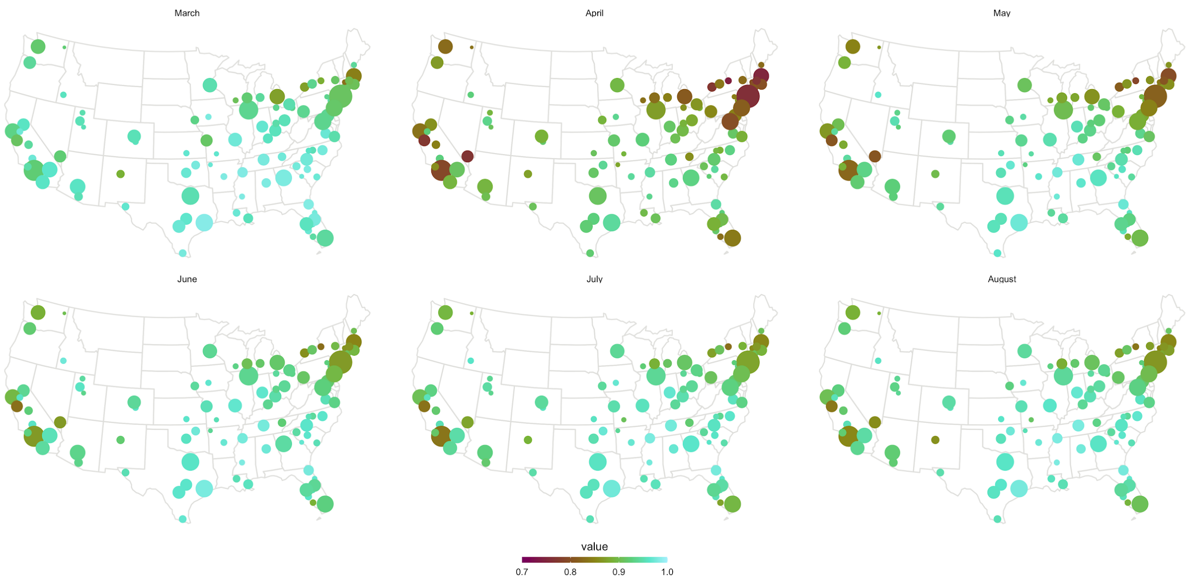

Figure 6 Correlation to January of month named.

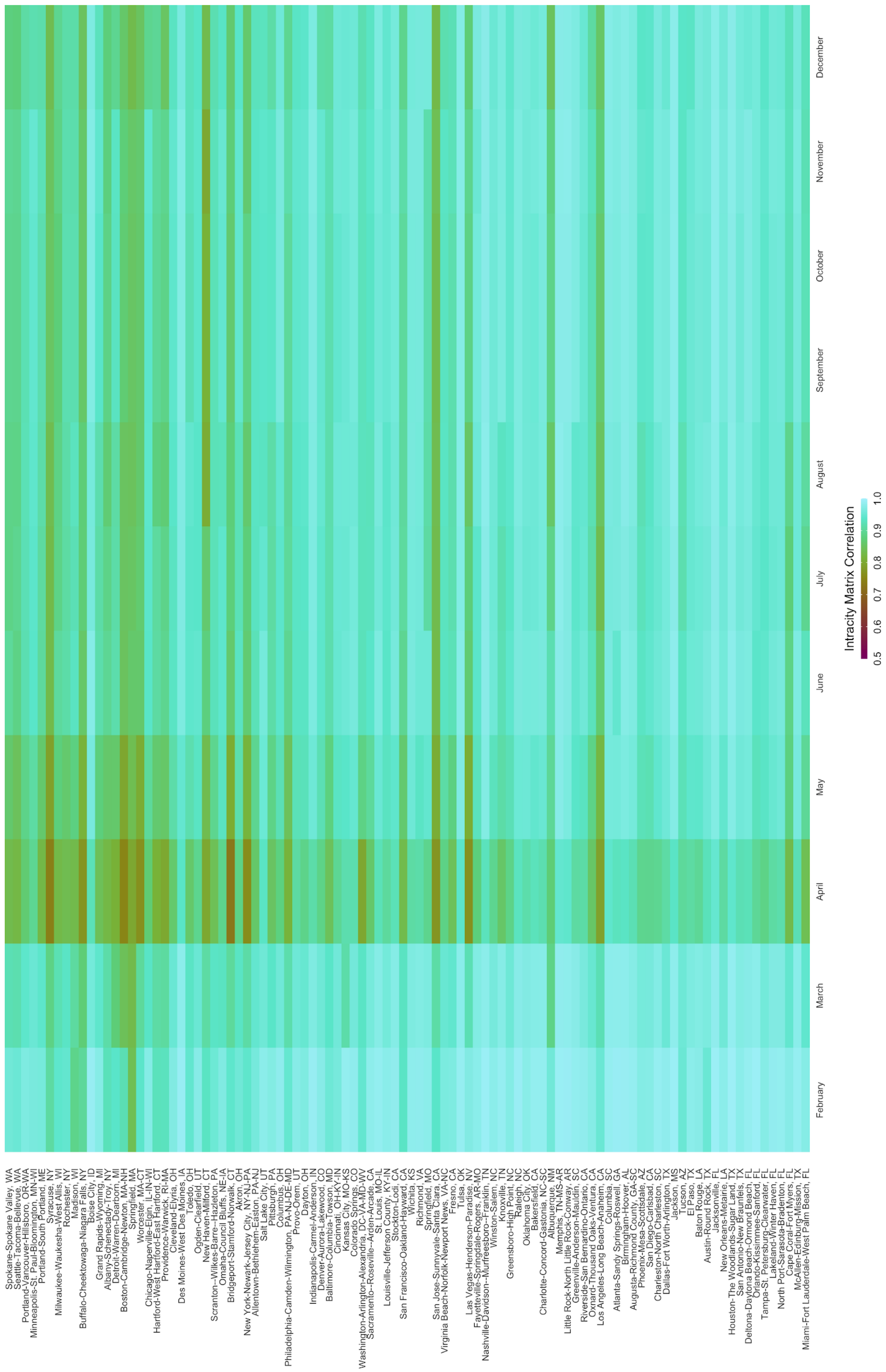

Figure 7 Correlation to January of month named, ordered from North to South.

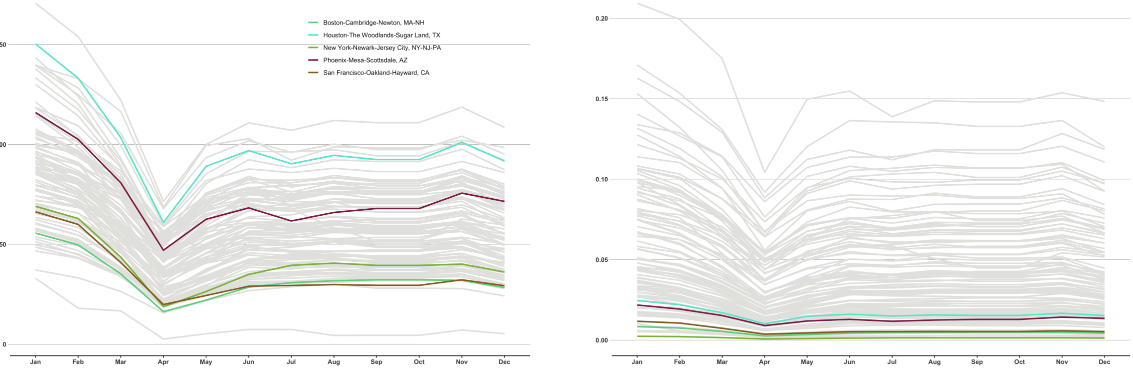

Figure 8 Mean degree centrality (left) and network density (right), with select cities highlighted.

Degree centrality—the number of other neighborhoods residents visit—and network density changed substantially during the pandemic, according to Figure 8. With this collection of metropolitan areas, we see a subtle relationship between changing networks and changing dissimilarity in Figure 9, which spikes in 50 out of 100 cities in April—and another 27 in May, June or July. In just 8 cities are the prepandemic months of February and April the most segregated (see Figure 10). Note that Boston and San Francisco, which saw the virus arrive and spread early, have spikes in March.

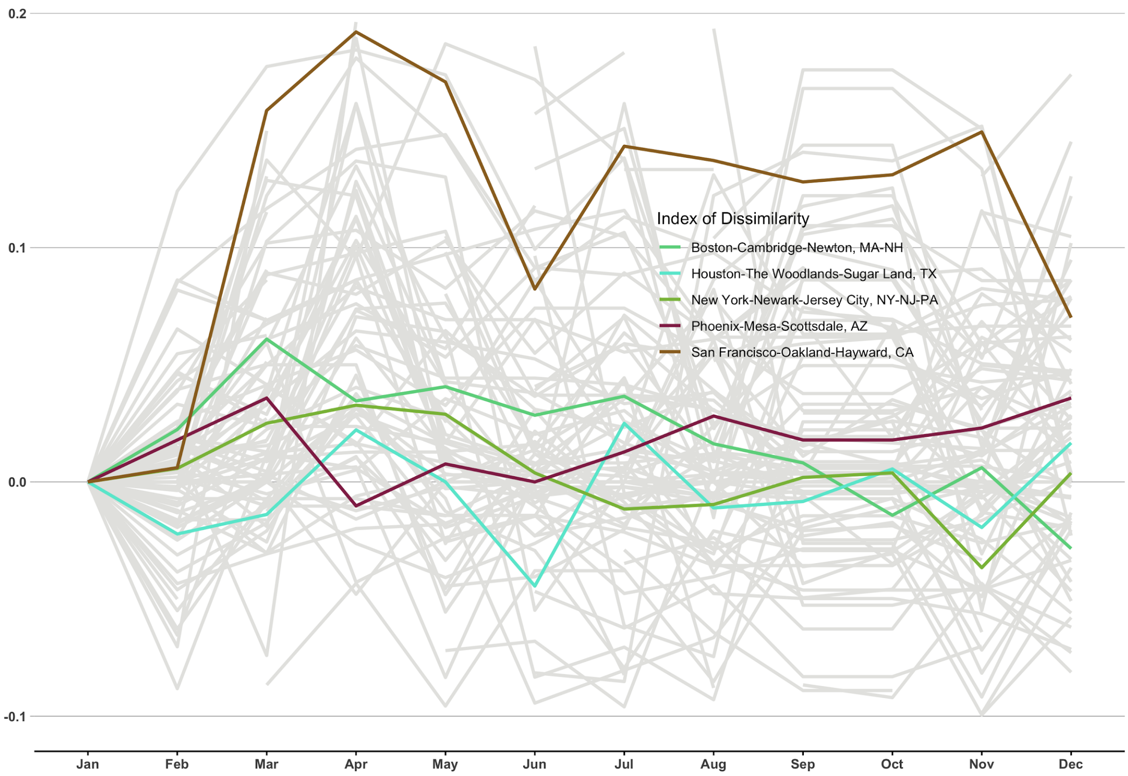

Figure 9 Index of dissimilarity, chained to January, over time with same highlighted cities.

4. Discussion

This process has introduced a method for monitoring changes to spatial interaction networks and shows much change during the early months of viral spread—with little recovery since. Advanced study is needed to understand what factors predict heightened experienced segregation during the pandemic. Geography alone is not sufficient to explain the fact that some cities saw segregation while others saw it fall or remain flat. Further research could also shed light on the differences between urban and suburban populations. The cases of New York and the Bay Area showed rising segregation while excluding suburban and exurban counties; when we included the entire metropolitan region, results for both conurbations along with many others showed a diverse array of network structures.

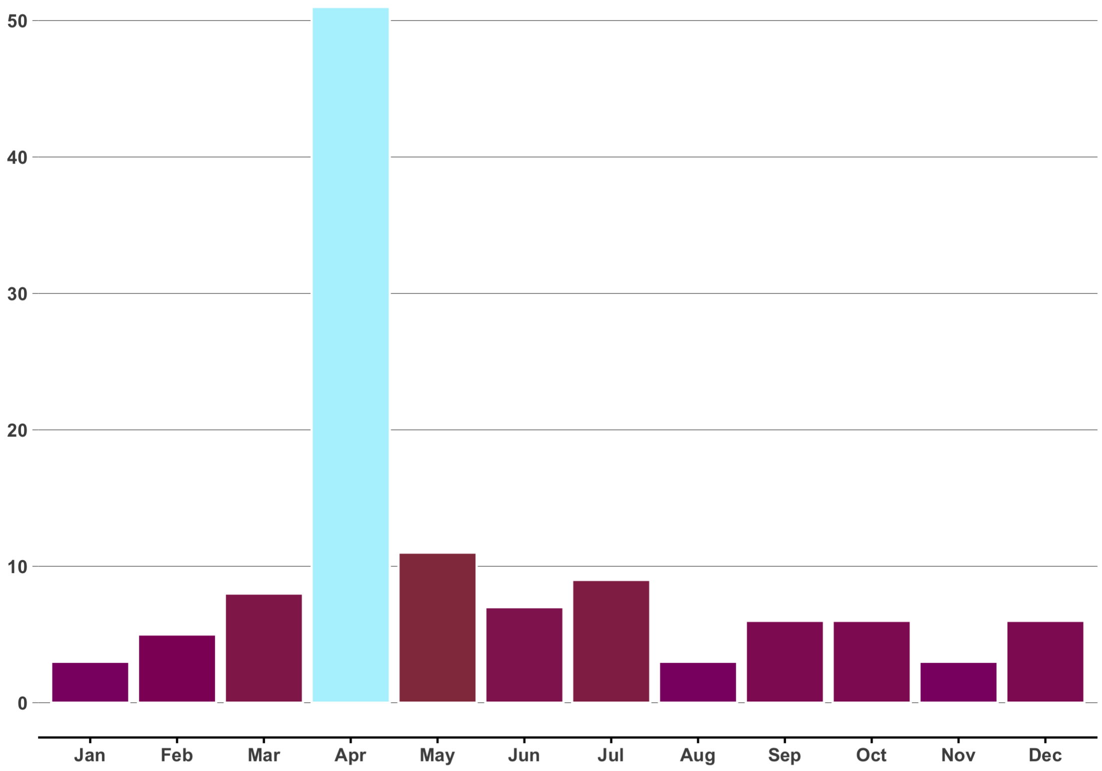

Figure 10 Most segregated month counts.

5. Acknowledgements

We thank SafeGraph and the SafeGraph Data Consortium for providing the data used in this analysis.

References

Athey, S., Ferguson, B.A., Gentzkow, M. and Schmidt, T., 2020. Experienced segregation (No. w27572). National Bureau of Economic Research.

Bonaccorsi, G., Pierri, F., Cinelli, M., Flori, A., Galeazzi, A., Porcelli, F., Schmidt, A.L., Valensise, C.M., Scala, A., Quattrociocchi, W. and Pammolli, F., 2020. Economic and social consequences of human mobility restrictions under COVID-19. Proceedings of the National Academy of Sciences, 117(27), pp.15530-15535.

Bassolas, A., Sousa, S. and Nicosia, V., 2021. Diffusion segregation and the disproportionate incidence of COVID-19 in African American communities. Journal of the Royal Society Interface, 18(174), p.20200961.

Batty, M., 2013. The new science of cities. MIT press.

Bohlin, L., Edler, D., Lancichinetti, A. and Rosvall, M., 2014. Community detection and visualization of networks with the map equation framework. In Measuring scholarly impact (pp. 3-34). Springer, Cham.

Butts, C.T. and Carley, K.M., 2001. Multivariate methods for interstructural analysis. Unpublished paper.

Chang, S., Pierson, E., Koh, P.W., Gerardin, J., Redbird, B., Grusky, D. and Leskovec, J., 2021. Mobility network models of COVID-19 explain inequities and inform reopening. Nature, 589(7840), pp.82-87.

Davis, D.R., Dingel, J.I., Monras, J. and Morales, E., 2019. How segregated is urban consumption?. Journal of Political Economy, 127(4), pp.1684-1738.

Gao, Y., Zhu, Z., Kali, R. and Riccaboni, M., 2018. Community evolution in patent networks: technological change and network dynamics. Applied network science, 3(1), pp.1-23.

Gibbs, H., Nightingale, E., Liu, Y., Cheshire, J., Danon, L., Smeeth, L., Pearson, C.A., Grundy, C., Kucharski, A.J., Eggo, R.M. and LSHTM CMMID COVID-19 Working Group, 2020. Human movement can inform the spatial scale of interventions against COVID-19 transmission. medRxiv.

Harris, R. and Owen, D., 2018. Implementing a multilevel index of dissimilarity in R with a case study of the changing scales of residential ethnic segregation in England and Wales. Environment and Planning B: Urban Analytics and City Science, 45(6), pp.1003-1021.

Jackson, M.O., 2010. Social and economic networks. Princeton university press.

Prestby, T., App, J., Kang, Y. and Gao, S., 2020. Understanding neighborhood isolation through spatial interaction network analysis using location big data. Environment and Planning A: Economy and Space, 52(6), pp.1027-1031.

Yang, Z., Algesheimer, R. and Tessone, C.J., 2016. A comparative analysis of community detection algorithms on artificial networks. Scientific reports, 6(1), pp.1-18.

Biographies

Andrew Renninger a researcher at the University of Pennsylvania, focusing on the geography of elections and urban economics; he holds degrees in urban planning and spatial data science.