Intracounty modeling of COVID-19 infection with human mobility: Assessing spatial heterogeneity with business traffic, age, and race

Xiao Hou[a], Song Gao[b,1], Qin Li[a,1], Yuhao Kang[b], Nan Chen[a], Kaiping Chen[c], Jinmeng Rao[b], Jordan S. Ellenberg[a], and Jonathan A. Patz[d]

[a]Department of Mathematics, University of Wisconsin–Madison, Madison, WI 53706; [b]Geospatial Data Science Lab, Department of Geography, University of Wisconsin–Madison, Madison, WI 53706; [c]Department of Life Sciences Communication, University of Wisconsin–Madison, Madison, WI 53706; and [d]School of Medicine and Public Health, University of Wisconsin–Madison, Madison, WI 53706

Edited by Lynn Goldman, George Washington University, Washington, DC, and accepted by Editorial Board Member Susan Hanson April 26, 2021 (received for review September 30, 2020)

The COVID-19 pandemic is a global threat presenting health, economic, and social challenges that continue to escalate. Metapopulation epidemic modeling studies in the susceptible–exposed–infectious–removed (SEIR) style have played important roles in informing public health policy making to mitigate the spread of COVID-19. These models typically rely on a key assumption on the homogeneity of the population. This assumption certainly cannot be expected to hold true in real situations; various geographic, socioeconomic, and cultural environments affect the behaviors that drive the spread of COVID-19 in different communities. What's more, variation of intracounty environments creates spatial heterogeneity of transmission in different regions. To address this issue, we develop a human mobility flow-augmented stochastic SEIR-style epidemic modeling framework with the ability to distinguish different regions and their corresponding behaviors. This modeling framework is then combined with data assimilation and machine learning techniques to reconstruct the historical growth trajectories of COVID-19 confirmed cases in two counties in Wisconsin. The associations between the spread of COVID-19 and business foot traffic, race and ethnicity, and age structure are then investigated. The results reveal that, in a college town (Dane County), the most important heterogeneity is age structure, while, in a large city area (Milwaukee County), racial and ethnic heterogeneity becomes more apparent. Scenario studies further indicate a strong response of the spread rate to various reopening policies, which suggests that policy makers may need to take these heterogeneities into account very carefully when designing policies for mitigating the ongoing spread of COVID-19 and reopening.

stochastic COVID-19 spread modeling | spatial epidemiology | neighborhood disparities | human mobility | data assimilation

The COVID-19 pandemic is a global threat presenting health, economic, and social challenges that continue to escalate. As of September 30, 2020, the Centers for Disease Control and Prevention had reported 7,168,077 total confirmed cases and 205,372 total deaths in the United States (1), and community transmission of COVID-19 remains widespread in many parts of the world. COVID-19 transmits mainly through close contact with infected patients (2). Thus, human mobility flows play a crucial role in the spatial spread of the virus; the heterogeneity of mobility patterns and social distancing behavior can largely explain the geographic heterogeneity of transmission (3–11). In the absence of vaccines or pharmaceutical agents to reduce the transmission of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that causes COVID-19, it is essential to understand the effects of nonpharmaceutical epidemic control and intervention measures. These include, but are not limited to, social (physical) distancing, travel restrictions, closures of schools and nonessential business services, limiting gathering size, mandated face coverings, testing, isolation, contact tracing, and timely quarantine to delay COVID-19 spread, all of which have been intensely investigated (12–18). The difference in the effectiveness of these interventions in mitigating COVID-19 across geography is substantial (19, 20). In the end, the same combination of interventions may have different effects on the overall progress of the epidemic in different contexts, and the dependence of outcomes on local conditions is typically complex. It is thus essential to develop accurate models that can incorporate spatial heterogeneity to make principled predictions about the effect of nonpharmaceutical interventions (NPIs) on the spread of COVID-19 under various combinations of local circumstances. For example, during the early outbreak in China, scholars developed a model using observed data about human mobility and COVID-19 infection and found that early detection and isolation of confirmed cases could prevent more infections than travel restrictions, but that combined NPIs achieved the strongest and most rapid effect (14). In Italy, scholars reconstructed the COVID-19 spatial spread dynamics and investigated the effects of population-wide movements and interventions

Significance

A human mobility flow-augmented stochastic susceptible–exposed–infectious–removed–style epidemic modeling framework is developed, which combines with data assimilation and machine learning to reconstruct the historical growth trajectories of COVID-19 infection in the two largest counties in Wisconsin. The associations between the spread of COVID-19 and human mobility/demographics are investigated. The results reveal different types of spatial heterogeneities (e.g., varying peak infection timing in different subregions) even within a county. Our scenario studies suggest that policy makers may need to take these heterogeneities into account very carefully when designing regionalization-based policies (e.g., testing and vaccination resource allocation) for mitigating the ongoing spread of COVID-19 and better preparing for future epidemics.

Author contributions: X.H., S.G., Q.L., N.C., J.S.E., and J.A.P. designed research; X.H., S.G., Q.L., Y.K., N.C., K.C., and J.R. performed research; S.G. and Y.K. contributed new reagents/analytic tools; X.H., S.G., Q.L., Y.K., N.C., K.C., and J.R. analyzed data; and X.H., S.G., Q.L., Y.K., N.C., K.C., J.R., J.S.E., and J.A.P. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. L.G. is a guest editor invited by the Editorial Board.

[1]To whom correspondence may be addressed. Email: song.gao@wisc.edu or qinli@math.wisc.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2020524118/-/DCSupplemental.

Published May 28, 2021.

(21, 22). In addition, researchers analyzed mobility flow data from smartphones across Europe before and after the implementation of NPIs and assessed the impact of coordination efforts of European countries (7).

The large body of modeling work on COVID-19 can be grouped into four broad categories: mechanistic models (12, 14, 16, 18, 21–25), agent-based models (26, 27), generalized linear regression models (4, 28), and machine learning models (29, 30). Both mechanistic and agent-based models typically rely on susceptible–infectious–removed (SIR) or susceptible–exposed–infectious–removed (SEIR) compartmental-style models in epidemiology, which separate the population into four compartments (susceptible, exposed, infectious, recovered or removed) and model the dynamics of transitions between the compartments. In the United States, the Institute for Health Metrics and Evaluation at the University of Washington employed a SEIR-style framework to project state-level hospital bed days, intensive care unit days, ventilator days, and deaths, and to model possible trajectories of SARS-CoV-2 infections (23, 24). Considering the inflows and outflows of interstate travelers, a travel network-based SEIR model system was developed to project state-specific COVID-19 infection information and to assess the impact of NPIs at the state level in the United States (31). Another study integrated fine-resolution mobility networks with the SEIR model to reconstruct the growth trajectory of COVID-19 infection in 10 large US metropolitan areas (8). Regarding the time-varying state-specific control measures, a temporal extended susceptible–antibody–infectious–removed model was developed for the projection of county-level COVID-19 prevalence (32). Under the agent-based modeling framework, several studies modeled the impact of testing, mobility, timing of social distancing and adherence level, contact tracing, and household quarantine on the multiple waves of COVID-19 infection in different counties and metropolitan areas (26, 27).

However, a straightforward SEIR metapopulation analysis, useful as it may be in revealing and informing public health and shaping policy making, relies on a key assumption of homogeneity of the population across the entire region. Each compartment is treated as an aggregate of indistinguishable individuals from different neighborhoods. This is not realistic. Recent studies have shown evidence of spatial heterogeneity and disparities in COVID-19 transmissions within major US cities (8, 33, 34). Socioeconomic and cultural environments, which differ within small geographic communities, affect the behaviors that drive the spread of COVID-19. What's more, geographic and transportation factors create spatial heterogeneity of infection spread in different subregions. Accurate and fine-grained modeling requires a more refined approach to address the heterogeneity issue.

Examining the spatial heterogeneity of infection spread is particularly important for modeling COVID-19, due to its complex dependence on social factors. COVID-19 presents a highly difficult policy issue because people need to weigh various trade-offs when they make health behavior decisions, trade-offs which may involve political ideologies, socioeconomic status, and lifestyles (35). Extensive research from social sciences has demonstrated that people's assessments of these trade-offs and their resulting health and risk behavior are affected by a variety of factors, such as age, gender, race, political ideology, economic status, culture, and religion (36–40). Despite a scientific consensus on the effectiveness of social distancing, precautionary actions like facial masking and social distancing are not universally adhered to (41), and adherence to these measures is by no means uniform across populations and geography. Thus, modeling the spread of COVID-19 requires scholars to attend to the heterogeneity in people from different backgrounds and in different locales. The place (i.e., the type of neighborhoods and communities) people live in is a reflection of race and ethnicity, socioeconomic status, and age structure (42–44). Segregation along these lines is not just an aggregate of individual choices but a consequence of public polices (45). Emerging evidence has shown disparities in COVID-19 outcomes between social groups. For instance, scholars found that income, along with race and other socioeconomic factors, has become a major predictor of COVID-19 infections and deaths (46, 47). A team of international scholars has addressed the importance of using social and behavioral sciences to support COVID-19 pandemic responses (48), such as the study of social inequalities and political polarization in response to social distancing and stay-at-home mandates (49–51).

In this paper, we aim to capture the effect of social and geographic heterogeneity within a population, and to reveal multiple aspects of the influences from different heterogeneities, by building models with finer-resolution mobility and epidemic information within a small geographic region (within a single county). To this end, we have developed a human mobility flow-augmented stochastic SEIR-style epidemic modeling framework that is able to distinguish different subregions and their corresponding epidemic growth characteristics. This modeling framework is combined with data assimilation and machine learning techniques to reconstruct the historical growth trajectories of COVID-19 confirmed cases in each subregion. We then investigate the association between the spread of COVID-19 and multiple factors that we believe are relevant and important, such as business foot traffic, race and ethnicity, and age information, based on existing studies (8, 44, 52). Finally, we perform scenario studies in order to understand the social and policy implications as well as to facilitate the development of guidelines to inform future policy making in public health crisis. The two geographic regions we consider are Dane County, WI, and Milwaukee County, WI, which contain the two largest cities in Wisconsin—respectively, Madison (a college town and home of the state government) and Milwaukee [a large city subject to severe racial and ethnic segregation (53, 54)].

Materials and Methods

Data. The official daily updated testing results of COVID-19 confirmed cases between March 11, 2020 and August 14, 2020 (as shown in Fig. 1) were obtained from the Public Health Offices of City of Madison & Dane County and Milwaukee County (55, 56) and aggregated to the census tract level using the case's residence information by the Wisconsin Department of Health Services (WDHS) (57). As shown in *SI Appendix*, Fig. S2, the COVID-19 testing capacity in Wisconsin was sufficient and kept up with evolving demands over time. The census tract geographic boundaries with demographics and socioeconomic attributes were obtained from the US Census Bureau (58, 59). We collected over 3.6 million points of interest (POIs) with aggregated and deidentified human travel patterns in the United States from SafeGraph, and the data were further spatially filtered by our study area. SafeGraph's data are demographically distributed in a way that matches the overall demographics in the United States Census populations (60). These mobile location data consist of "pings" identifying the coordinates of a smartphone over time. To enhance privacy, SafeGraph excludes census block group (CBG) information if fewer than two devices visited a place in a week from a given CBG. For each POI, the records of aggregated visitor patterns record the number of unique visitors and the number of total visits to each venue during a specified time window (i.e., hourly, daily, weekly, and monthly); this allows us to estimate the foot traffic of each venue and the origin-to-destination (O-D) spatial interaction flow patterns during the study period (Fig. 1) (61). We then aggregate the O-D flow matrices to the census tract level to match the COVID-19 testing result data. After the spatial clustering of census tracts based on their spatial connectivity (O-D flows) in each county, we further aggregate all of the above-mentioned data to cluster (region) granularity for modeling.

A Human Mobility Flow-Augmented Stochastic SEIR Model. The feature that differentiates our approach from the classical SEIR model is our incorporation of spatial interactions between different regions and population

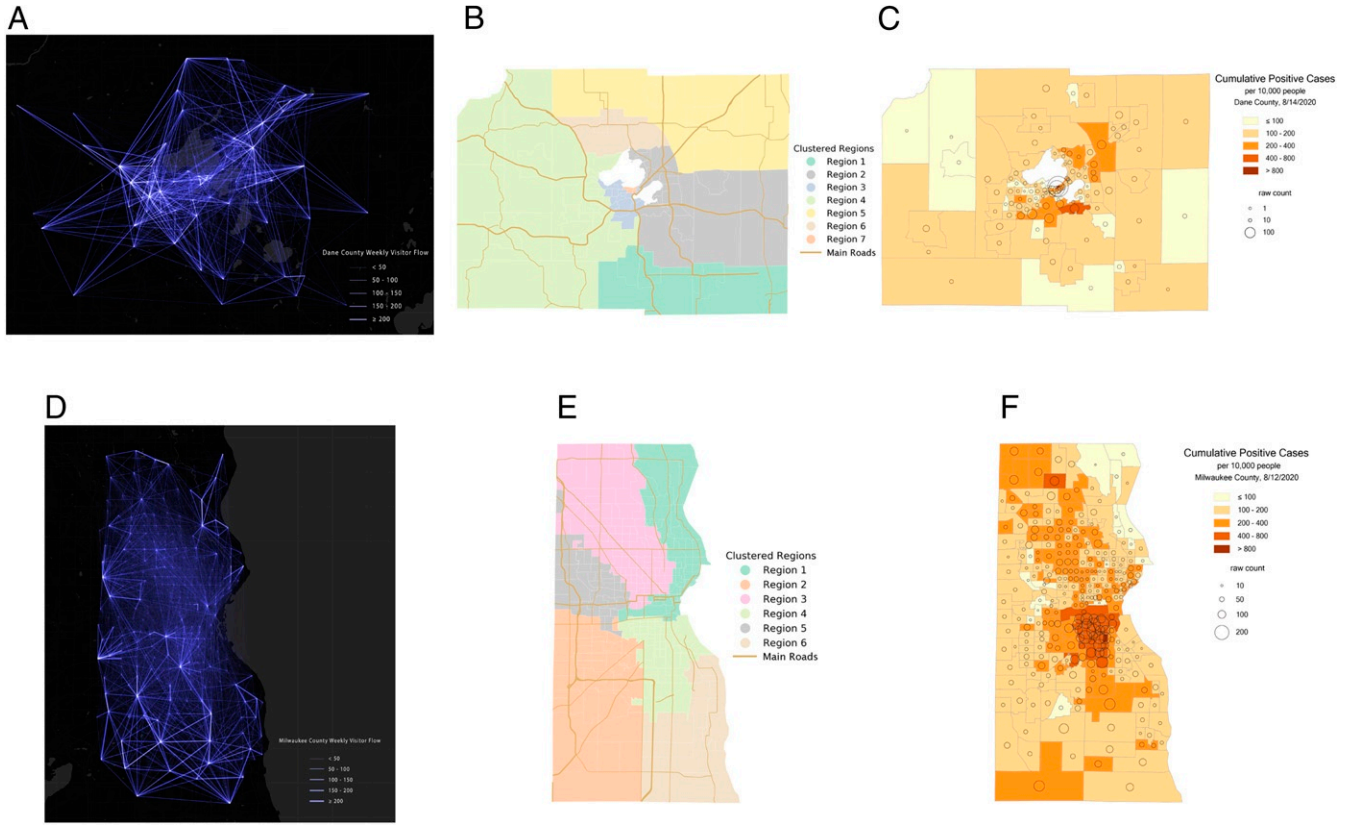

Fig. 1. The spatial distributions of (*A*) human movement O-D flows between census tracts in Dane County, (*B*) Dane County spatial clustering results using the Walktrap network community detection method, (*C*) the raw cumulative confirmed COVID-19 cases and ratio of per 10,000 people at the census tract level in Dane County by August 14, 2020, (*D*) human movement O-D flows between census tracts in Milwaukee County, (*E*) Milwaukee County spatial clustering results using the Walktrap network community detection method, and (*F*) the raw cumulative confirmed COVID-19 cases and ratio of per 10,000 people at the census tract level in Milwaukee County by August 12, 2020. (COVID-19 confirmed cases data were retrieved from the Wisconsin Department of Health Services.)

heterogeneities (see model details in *SI Appendix*). This feature enables our human mobility flow-augmented stochastic SEIR model to simulate and analyze the important differentiation between multiple communities (subregions) within one county, and the connection between the spread of COVID-19 and local mobility and demographics, and social activities within these communities. To be more specific, a traditional SEIR model divides the population into four compartments and uses an ordinary differential equation (ODE) system to describe the dynamics and flows between the compartments. This treats the individuals within each compartment as indistinguishable, and thus misses important distinctions which are relevant to epidemic dynamics. By contrast, our model begins by using an unsupervised machine learning algorithm to partition a county into local clusters (subregions), using the observed human mobility flow data. This allows us to divide the county into multiple distinct communities (based on spatial connectivity), each having its unique geographic and demographics features. Once these communities have been generated, we build a local SEIR model for each region, which takes into account region-specific factors. All these local SEIR models are then coupled together using the intracounty mobility flow data, which allows us to quantify how much daily transmission occurs within the regions as opposed to between the regions (an example is plotted in *SI Appendix*, Fig. S26, which shows that 51.4% of infection took place within the same regions in Dane County). The division of the county into communities (subregions) allows us to identify the time-varying characteristics in the spreading of COVID-19 in different subregions (e.g., peak infection timing and the leading factors), and enhances the forecasting capability of the model, which further facilitates the design of region-specific policies.

Another special feature of our model is that it allows a dynamical evolution of the key parameters (such as the transmission rate). In this model, the transmission rate is described by a simple but effective stochastic (Ornstein–Uhlenbeck) process (62), with the parameters to be tuned by the daily reported COVID-19 infection data fetched from WDHS. This modeling strategy facilitates an efficient and effective online learning procedure for the dynamical evolution of the ODE system that is capable of capturing the temporal variability of COVID-19 transmission. We stress that such time-dependent behavior of the transmission rate plays a vital role in describing the nonstationary characteristics of the overall SEIR model dynamics, resonating the uncertainties of the effective reproduction number (i.e., the actual average number of secondary cases per primary case at a specific time) described in ref. 63. Since the method is relatively sensitive to the input data (daily infection data), it is able to quickly detect the changes in the spread of the disease, and reveal the possible association with business traffic or notable social events (e.g., superspreading events) (44, 64).

An Efficient Data Assimilation Strategy for Reconstructing Infection Trajectory. The application of the data assimilation method for reconstruction of the remaining parameters in the model is another desirable feature of our approach. In each iteration, the data assimilation method combines the observed data with prior model information to generate a posterior distribution of the system state variables and parameters. In particular, we apply the ensemble Kalman filter method (65) to estimate the hidden parameters in the model and fit the observed confirmed case data to recover the historic trajectory of the spread of the virus. Such a model calibration procedure is completely data driven, with no ad hoc parameter adjustment. Since no assumption is made on the transmission process, the model is independent from unnecessary a priori presumptions, and the reconstructed parameters can be viewed completely objectively. These objective quantities are then used to draw association with possible influencing factors such as age, race and ethnicity, and business reopening policies. Before using this calibration process to perform different scenario studies and predictions, we first tested its validity. As shown in *SI Appendix*, Figs. S14 and S15, our calibration

Table 1. Population density and ratio of different age groups by region in Dane County

	Percent								
Region	0 y to 17 y	18 y to 24 y	25 y to 29 y	30 y to 39 y	40 y to 49 y	50 y to 59 y	60 y to 69 y	70 y and above	Density (people per km^2)
1	23.95	7.10	5.16	12.34	15.32	15.28	11.71	9.13	87
2	20.64	8.49	8.38	16.52	13.04	14.40	10.89	7.64	233
3	16.02	22.87	12.04	15.64	9.59	9.81	8.08	5.95	1,536
4	23.54	7.25	8.08	14.78	13.13	13.45	11.22	8.56	122
5	26.71	6.65	6.15	15.83	14.02	13.31	9.33	8.01	99
6	25.56	6.24	3.31	11.02	14.66	15.51	11.04	12.66	146
7	0.63	92.20	3.30	1.04	0.80	1.25	0.72	0.06	9,306

process can almost perfectly recover the observed COVID-19 confirmed cases time series data with very small relative errors (about 4% for Dane and 2% for Milwaukee) and RMS errors (RMSE) in *SI Appendix*, Table S3.

Results

Origin-to-Destination Flow-Based Spatial Clusters. The traditional SEIR model assumes homogeneity over its population. This assumption, however, is not valid in either of the two counties studied here. To have a finer-scale study, we divide both counties into several smaller intracounty regions using the Walktrap network-based community detection method (66). The travel O-D flow data of the first week of March (just before the initial widespread outbreak of COVID-19 in the United States) is used to generate the graph weights for clustering. The spatial clustering results for the two counties under investigation are shown in Fig. 1. Dane County is partitioned into seven regions (the two white areas present Lake Mendota and Lake Monona). Each region has its unique demographic features (e.g., race and ethnicity composition, age structure). The spatial distribution of the demographic information in both counties is presented in *SI Appendix*, Fig. S3. The maps show that there is high age group heterogeneity in Dane County but high racial and ethnic heterogeneity in Milwaukee County. In this respect, it is worth specifically singling out region 7, the area containing downtown Madison and the University of Wisconsin–Madison campus. As one might expect, region 7 has significantly higher population density (especially younger people) than the other regions. Region 3 is a residential region adjacent to region 7, and it also has relatively high population density. The population density and age group summary by region is shown in Table 1. Milwaukee County is partitioned into six regions. While the population density in each region is relatively uniform, the region is heavily segregated by race and ethnicity. The population density and the distribution of different race and ethnic backgrounds by region in Milwaukee are listed in Table 2. Region 3 has a predominantly Black population, while region 4 has a high concentration of Hispanic residents.

Intracounty Modeling Results

Dane County. The population within each clustered region is regarded as a homogeneous group. We then build a separate SEIR model for each region, and couple these models together using the interregional mobility flow traffic information. Each region then has its own effective reproduction number R_e, a function of time representing the expected number of new secondary cases directly infected by each existing case. The effective reproduction number is a critical indicator that quantifies the spread of disease, and its dependence on time clearly is related to the county's own "stay-at-home" order and subsequent reopening procedure. In Dane County, phase 1 reopening was enacted on May 26. During this period, all business were allowed to open with 25% limit capacity. Phase 2 reopening was enacted on June 15, and the capacity was increased to 50%. On July 2, the county, after seeing an unusual increase of infection, rolled back the phase 2 reopening, and the capacity limit was reduced to 25%. On July 13, the county imposed an indoor face mask requirement.

In Fig. 2, we plot the estimated effective reproduction number R_e and daily confirmed cases (with a 7-d average for smoothing out the possible day-of-week reporting/testing bias) in each region and its corresponding interquartile range (IQR, the range that shows 25th to 75th percentiles). The dates of the reopening and rollback are also indicated in the plots. Region 7 (the downtown Madison and university campus area) has the highest R_e, spiking as high as over five at the end of June, during the brief period of relaxed capacity limits for indoor businesses. By contrast, region 3 (the adjoining residential area), despite also possessing a high population density, has a substantially lower R_e, without a major late-June spike. In Fig. 3*A*, we plot the flow traffic normalized effective reproduction number in Dane County. This measure controls for the fact that some regions have higher intraregional spatial interactions (mobility flows) than other regions and are thus expected to have a higher R_e. However, region 7, even upon this normalization, still has a significantly higher R_e than all other regions, suggesting that there are social characteristics beyond population mobility contributing to rapid spread there. In particular, the normalized R_e has a peak in late June, approximately 7 d after the county started phase 2 reopening on June 15.

To further study the reason for region 7 having unproportionally high R_e, we also pull its correlation with different kinds of business foot traffic by investigating the SafeGraph mobile

Table 2. Population density and ratio of different race and ethnicity by region in Milwaukee County

	Percent							
Region	White	Black	Native	Asian	Hawaii	Other	Hispanic	Density (people per km^2)
1	75.51	14.59	0.46	3.94	0.03	2.33	6.82	1,623
2	87.52	3.94	0.60	3.45	0.01	1.74	11.59	1,030
3	17.56	71.88	0.28	5.27	0.01	1.72	4.32	2,033
4	66.49	6.77	0.95	2.58	0.04	19.04	48.22	2,474
5	81.26	9.70	0.73	3.29	0.01	1.63	6.92	1,720
6	89.09	2.48	0.92	3.78	0.04	1.06	9.59	739

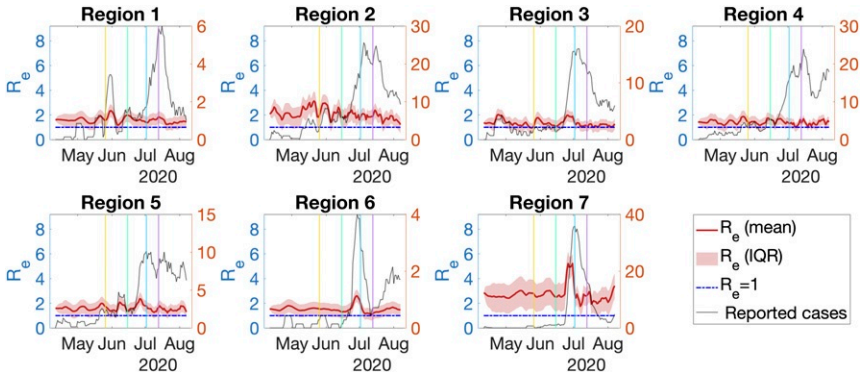

Fig. 2. Effective reproduction number (left *y* axis) and daily confirmed cases (right *y* axis) with 7-d average values for different regions in Dane County. The vertical lines indicate the dates of phase 1 reopening, phase 2 reopening, rollback reopening, and face covering order.

device visit tracking data. In Table 3, we document the Pearson's correlation of R_e in each region with different categories of business service visits. Clearly, the increase of R_e in region 7 has a strong cross-correlation with the visits to "Drinking Places (Alcoholic Beverages)" with a time lag of 5 d. Such a data-driven analysis may complement the local official report showing that more than 21% of the newly COVID-19–infected cases reported recent trips to bars and taverns (67). From June 13 through June 26, 49% of reported positive cases out of 614 people who tested positive were between the ages of 18 and 25 y, and 28% of cases were associated with infection clusters, including 132 cases from bars in Dane County (68).

We also perform scenario studies (by modifying the region-specific effective reproduction number) to evaluate the possible consequences under different policies to be put in place. In particular, we study what would have happened in the following scenarios:

1) In scenario 1, there is no phase 2 reopening enacted on June 15, and the effective reproductive number, R_e, stays unchanged afterward. Under this assumption, we run the model forward from June 15 for 2 wk, and compare it with the actual infection data.
2) In scenario 2, the county does not roll back the phase 2 reopening on July 2. We once again run the model forward for another 2 wk with the high value of R_e obtained during the reopening time (before July 2).
3) In scenario 3, the county further reopens on August 4 at different transmission levels. We run the model forward for 10 d from August 4 with three different configurations: using the R_e value obtained on August 4 and assuming it stays the same in the future, R_e doubles its value, and R_e triples its value.

In Fig. 4*A*, we plot the predicted infection of selected regions supposing the June phase 2 reopening had not taken place. Our model predicts clearly that the number of cases would be drastically reduced in regions 3, 6, and 7 if the county did not enter phase 2 reopening; in other words, those are the regions most severely affected by that decision. In particular, in region 7, had there been no reopening, the projected number of cases should have been only 3.4 per thousand people by June 30, as compared to the actual count of 11.6, more than 3 times as high. Regions 1, 2, 4, and 5 see small impacts from the reopening (see the results plotted in *SI Appendix*, Figs. S22 and S23). In Fig. 4*B*, we plot the predicted infection assuming there was no rollback from phase 2 reopening in the same three regions. It is clear that the rollback of the reopening reduced the number of infection cases significantly. Without the rollback, the predicted total number of infection in Dane County during July 2–17 would have been 7.3 per thousand people. This is almost 3 times 2.5, the actual number of confirmed cases per thousand people in the same time interval. In Fig. 4*C*, we plot the predicted number of infected people from August 4 into the future 10 d. If the effective reproduction number gets tripled, the predicted number of infected people in Dane County would increase from 0.9 per thousand to 2.8 within 10 d. We also plot the predicted cumulative and daily infection cases for each region in *SI Appendix*, Figs. S22–S25.

To sum up, a strong cross-correlation is observed between the number of visits to "Drinking places (Alcoholic Beverages)" and increased infections in region 7. At the same time, the scenario study suggests that, without phase 2 reopening, the infection in Dane County would have been significantly lower (reduce about two-thirds), and, without the subsequent rollback from phase 2 reopening, the infection would have continued to increase dramatically. All this together provides strong evidence that the phase 2 reopening allowing 50% capacity in drinking places played a major role in the drastic increase of the number of cases observed in Dane County during the summer of 2020.

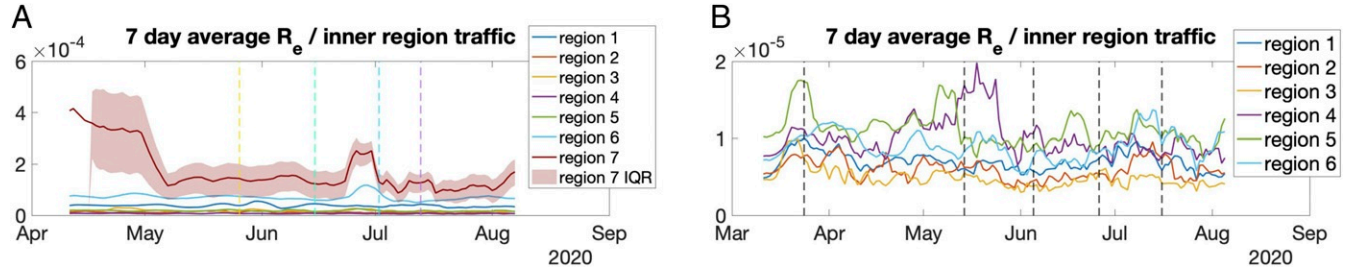

Fig. 3. The time-varying effective reproduction number normalized by inner region traffic frequency. (*A*) The 7-d average R_e normalized in Dane County. The yellow, green, blue, and purple vertical lines indicate the date of reopening phase 1, reopening phase 2, rollback of phase 2, and face covering order, respectively. (*B*) The 7-d average R_e normalized in Milwaukee County. The vertical lines indicate the starting date of state of Wisconsin Safer at Home Order, phase 2 reopening, phase 3 reopening, phase 4 reopening, and mask ordinance (from left to right).

Table 3. The cross-correlation of effective reproduction number (3-d average) of different regions and the number of service visits (3-d average) within each region

Categories	Region 1	Region 2	Region 3	Region 4	Region 5	Region 6	Region 7
Snack and nonalcoholic beverage bars	−0.21 (0.095)	−0.18 (0.147)	−0.22 (0.081)	−0.17 (0.187)	−0.21 (0.095)	−0.07 (0.557)	−0.02 (0.850)
Drinking places (alcoholic beverages)	−0.35 (0.005)	−0.09 (0.457)	0.16 (0.219)	−0.13 (0.293)	0.04 (0.751)	–	0.41 (0.001)
Supermarkets and Grocery stores	−0.03 (0.801)	0.15 (0.238)	0.43 (0.000)	0.17 (0.185)	0.10 (0.421)	−0.11 (0.407)	0.17 (0.189)
Nature parks and other similar institutions	−0.12 (0.339)	0.26 (0.039)	−0.21 (0.095)	0.21 (0.090)	0.11 (0.374)	0.32 (0.011)	0.01 (0.929)
Fitness and recreational sports centers	−0.19 (0.124)	−0.36 (0.004)	0.13 (0.309)	0.03 (0.801)	0.12 (0.326)	0.06 (0.625)	0.16 (0.198)
Full-service restaurants	−0.27 (0.033)	−0.20 (0.119)	0.24 (0.061)	−0.06 (0.662)	0.10 (0.446)	0.09 (0.488)	0.36 (0.003)
Limited-service restaurants	−0.01 (0.936)	−0.16 (0.207)	0.12 (0.354)	−0.14 (0.287)	−0.14 (0.256)	−0.10 (0.438)	0.28 (0.023)

The cross-correlations are computed using 5 d of time lag, and the corresponding *P* values are listed in parentheses.

Milwaukee County. We perform similar studies in Milwaukee County. Milwaukee started its phase 2 reopening on May 14. During this period, restaurants and bars were allowed to have take-out or delivery service only. On June 5, phase 3 reopening started, and restaurants and bars opened with 25% limit capacity. The limit capacity was lifted to 50% on June 26, when phase 4 began. On July 16, Milwaukee instituted a mask ordinance requiring face covering in public spaces, both indoors and outdoors.

In Fig. 5, we plot the effective reproduction number in different regions, and its IQR. Among the six regions, regions 3 and 4 see the highest R_e, averaging 1.8 and 2, respectively. In particular, in region 4, in mid-May, R_e rose as high as 3.76. As mentioned above, regions 3 and 4 are the two regions most subject to racial and ethnic segregation. Such segregated neighborhood patterns have also been identified previously through human mobility-based spatial interaction network analysis using location big data (69). The COVID-19 spread and deaths in Milwaukee shows the racial disparities in health. According to the official summary statistics in Milwaukee County (56), about 31% of the total confirmed COVID-19 cases are Hispanic and 28% are Black in our study period.

As we did in the Dane County study, we also consider the effective reproduction number normalized to control for

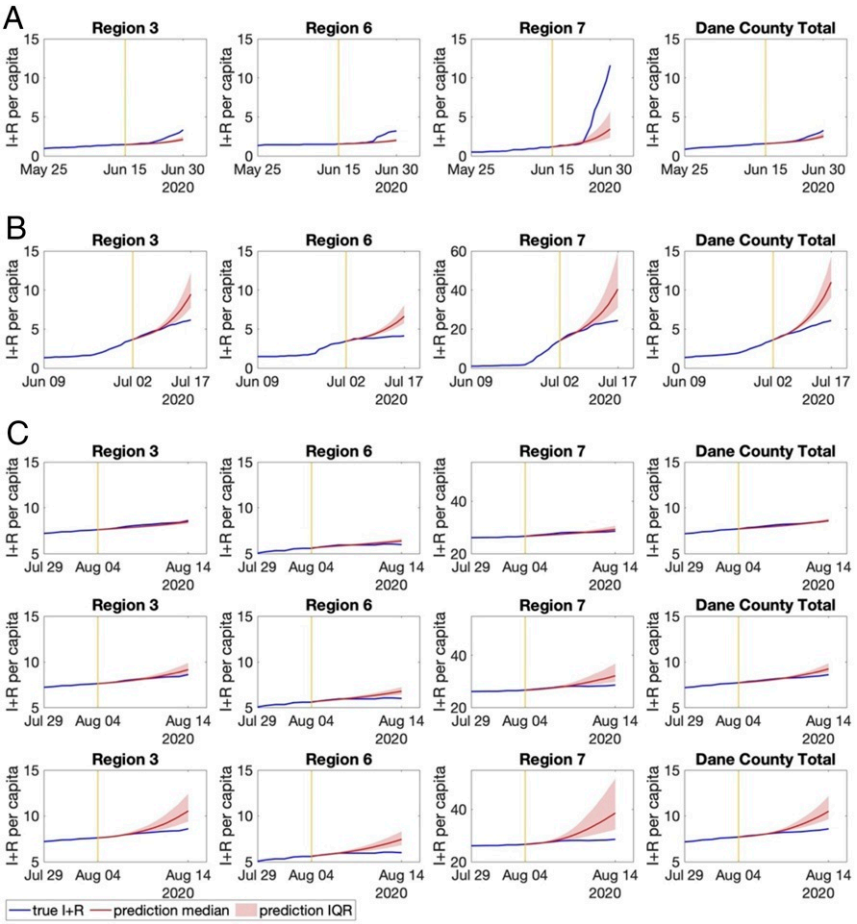

Fig. 4. Prediction of cumulative infections (I+R) per 1,000 people in selected regions if Dane County, (*A*) scenario 1, did not have phase 2 reopening; (*B*) scenario 2, did not have rollback from phase 2 reopening on July 2; or (*C*) scenario 3, further reopens on August 4. In *C*, shown are (*Top*) reopening with the current effective reproduction number, (*Middle*) reopening with a doubled effective reproduction number, and (*Bottom*) reopening with a tripled effective reproduction number.

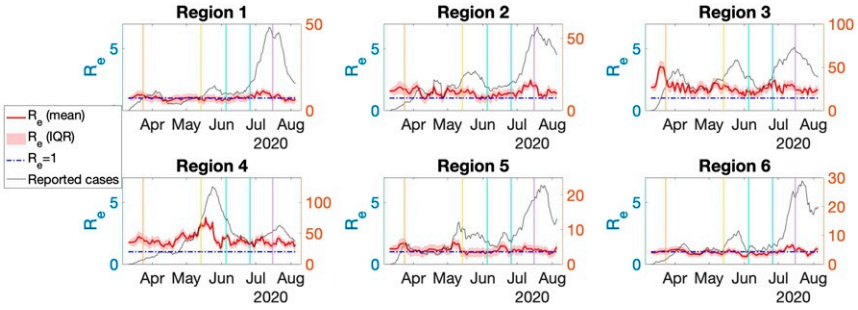

Fig. 5. Effective reproductive number (7-d average, left *y* axis) and reported cases (7-d average, right *y* axis) of region 1 to region 6 in Milwaukee County (left to right, top to bottom). The vertical lines indicate the starting date of state of Wisconsin Safer at Home Order, phase 2 reopening, phase 3 reopening, phase 4 reopening, and mask ordinance (from left to right).

intraregional flow traffic. As seen in Fig. 3*B*, even after this normalization, region 4 still demonstrates the highest transmission. That is, even with high mobility frequency within this region being discounted, the local infection rate is still very high, with the mid-May peak persisting. This suggests that policy makers need to explicitly consider local transmission variation contexts even after restricting intraregional flow traffic, to prevent more health disparities in ongoing and future pandemics.

Assessing Spatial Heterogeneity with Age, Race, and People's Self-Reported Social Distancing. The disparities in COVID-19 spread rate between different regions of the two counties studied reflect important interaction between demographic divisions and residents' reaction to COVID-19. In Dane County, we observed an appreciable difference in the outcomes between people who lived near the campus area and those farther away, that is, spatial heterogeneity (caused by age structure difference; *SI Appendix*, Fig. S3*A*). In Milwaukee County, we observed stark differences in the burden of COVID-19 between different race and ethnicity groups, especially between predominantly White regions and regions with a higher Black and Hispanic population (*SI Appendix*, Fig. S3*D*). One author in this research team was also on the data analysis team of the "COVID-19 and Social Distancing" survey conducted between March 19 and April 1, 2020 by a group of interdisciplinary scholars and nongovernmental organizations in Wisconsin (70). The survey received responses from over 30,000 Wisconsin residents. We conducted data analysis on a behavioral question asking people to what extent they practice social distancing, and compared different age groups' answers and different race and ethnic groups' answers. We found that, among respondents aged 20 y to 29 y, a substantially higher proportion of participants (4.25%) said they did not perform social distancing at all or performed little social distancing, compared to the other two age groups (30 y to 59 y: 1.26%; above 60 y: 1.03%) (Fig. 6, *Left*). For Hispanic and Black groups, there are still a certain number of respondents who reported that they did not perform social distancing at all (the top red bar in Fig. 6, *Right*). These self-reported behavioral differences in precautionary behavior might partially explain the heterogeneous spread patterns we observed in the two counties. More importantly, these disparities among age groups and race and ethnicity groups in the COVID-19 outcome and in the self-reported social distancing behavior flags an urgent policy need. Policy makers and health communicators must investigate the underlying reasons for demographically specific barriers to adoption of social distancing and other mitigation measures (71).

Discussion

Sensitivity Test of Clustering and Modeling Results. The results of clustering are subjected to the weekly travel O-D flow data and to the network-based community detection algorithms. In general, since each region in Dane County and Milwaukee County has its own unique employment and business spatial distribution and socioeconomic features, the clustering results in Milwaukee County and Dane County are rather stable with respect to the change of clustering methods and data from different weeks. In the simulation, we utilize the Walktrap method on the travel O-D flow data of the week of March 2–8. Here we discuss different clustering methods and using different periods of data and the impacts on modeling results as a test of the sensitivity.

The Walktrap algorithm uses short random walks to detect communities in a large graph. In comparison, we also apply the Louvain method on the same data. The Louvain method is a popular community detection algorithm that aims to maximize the modularity of a graph partition (72). *SI Appendix*, Figs. S4*B* and S5*B* show the clustering results of Dane County and Milwaukee County using the Louvain method with the same travel O-D flow data. For Milwaukee County, the two clustering results are very similar apart from a small section in the lower left corner of the map, and both methods result in six clustering groups. The modularities of clustering using Walktrap and Louvain are 0.365 and 0.376, respectively, indicating stable performance. For Dane County, the two methods provide slightly different results. Walktrap divides the county into seven areas, while the Louvain gives five. Comparing the panels in *SI Appendix*, Fig. S4, the Louvain method clusters the northeastern and central eastern areas as a single region, and the Walktrap method combines regions 7 and 3. Since region 7 is the downtown area where the campus of University of Wisconsin–Madison is located, it plays an important role in our analysis of human mobility flows, demographics, and spread of COVID-19. Therefore, it is meaningful to identify it as a single region, and we chose to use the Walktrap method. The modularities of clustering using Walktrap and Louvain algorithms are very close: 0.318 and 0.334, respectively. In addition, we quantitatively study the structure similarity between the two clustering results, and list the adjusted Rand index (ARI) (73) in *SI Appendix*, Table S1 and S2. Both Dane and Milwaukee counties have a relatively high structure similarity when two clustering approaches are used, with ARI being 0.68 and 0.81, respectively. For Dane County, only 3 out of 105 census tracts change group assignment. In Milwaukee County, 13 out of 296 census tracts change group assignment when the Louvain method is used, and 17 change group assignment when the Walktrap method is used.

We also study the sensitivity of the clustering results using O-D flow data from different time intervals. In *SI Appendix*, Figs. S4 *C* and *D* and S5 *C* and *D*, we show the results for the clustering of two counties using the O-D flow data from the weeks of March 9–15 and April 6–12, and compared them with that computed from the week of March 2–8. The corresponding modularity and the ARI values for the two counties for the first week of each

SOCIAL SCIENCES

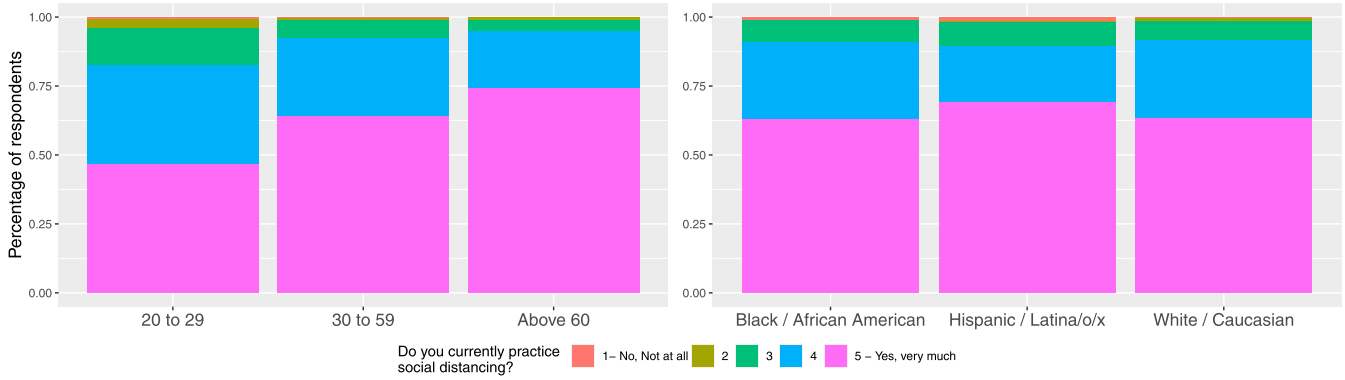

Fig. 6. (*Left*) Comparison of self-reported social distancing behavior among different age groups. (*Right*) Comparison of self-reported social distancing behavior among different race and ethnicity groups.

month from March to August 2020 are provided in *SI Appendix*, Tables S1 and S2, respectively. We obtain a good consistency of clustering results throughout different time periods.

Furthermore, we then test the model's sensitivity regarding the change in clustering. To do so, we run our model twice, using the clustering results based on the data of March 9–15 and April 6–12, respectively. The two new trajectories of effective reproductive numbers are plotted in *SI Appendix*, Figs. S6–S9. Comparing to the effective reproductive number presented in Figs. 2 and 5 using the data from March 2–8, the difference is very limited. In other words, even with small changes in the clustering step, our epidemic model results remain with regard to the spatial heterogeneity and the time-varying R_e curve in each region.

Policy Implications. Our modeling analysis results suggest several policy implications that can be used in practice. First, the findings lead to a better understanding of substantial local variations in COVID-19 peak timing (reflected by the effective reproduction number R_e) in different subregions and can inform further investigation of possible superspreading events (44) (captured by the very high R_e values on certain days). Second, the findings of scenario studies can help in assessing different business reopening policies. Consider Dane County, which contains a large "college town" environment with a unique age structure (as shown in *SI Appendix*, Fig. S3); the area experienced spikes in the effective reproduction number when the county entered the phase 2 reopening in mid-June (as shown in Fig. 3*A*). Several restriction policies need to be in place, such as reducing business seating capacity and limiting indoor gathering size. We also observed that the reproduction number decreased significantly after the county rolled back its reopening policy in early July. The pattern of reproduction rate following the policies is especially substantial for the campus region. These results suggest that policy makers need to design reopening in a more nuanced manner. Instead of implementing a one-size-fits-all reopening policy for all of the regions within a county, reopening policies that attend to the diverse nature of regions within a county should be considered. For a geographically heterogeneous county, a regionalization-based policy is needed. For instance, we demonstrated that certain types of businesses and certain subregions have a much higher infection rate. In Dane County, people's visits to drinking places around the downtown area have the strongest association with infection rates, and the campus region (with more young people) has a higher infection rate than other parts of the county do. In Milwaukee County, the racial and ethnic heterogeneity associated with the infection is more apparent. These empirical evidences flag the crucial importance of designing reopening policies which take these heterogeneities into account when choosing what types of businesses and what regions to reopen first. Public policies also need to adapt to different regions that consist of very different demographic populations, in order to control the reproduction rate of COVID-19. The COVID-19 is not only a public health problem but also a problem that reflects existing disparities among various socially distinct population groups and neighborhoods. Last but not least, our research can guide regionalization-based COVID-19 response resources allocation (e.g., testing and vaccination) and public health communication strategies tailored to specific demographics and time-varying infection characteristics in different neighborhoods. For example, the prioritizing strategy of resource supply and distribution (e.g., rapid COVID-19 testing and vaccination) is required when a county needs to adapt its community-specific containment strategy as COVID-19 evolves.

Data Availability. All the data that support the findings of this study are publicly available on the GitHub repository under the MIT license: https://github.com/GeoDS/IntraCounty-Mobility-SEIR.

ACKNOWLEDGMENTS. We thank the WDHS for providing daily census tract-level COVID-19 confirmed case data and thank SafeGraph Inc. for providing the anonymous and aggregated place visits and human mobility flow data. S.G., Q.L., K.C., and J.A.P. acknowledge funding support provided by the NSF Grant (BCS-2027375). J.S.E. acknowledges funding support provided by the NSF Grant (DMS-1700884). S.G. acknowledges support from the National Geospatial Fellowship for Advancing COVID-19 Research & Education provided by the NSF Geospatial Software Institute Conceptualization Project Grant (OAC-1743184). Q.L., N.C., and X.H. are also supported by Data Science Initiative, provided by the University of Wisconsin–Madison Office of the Chancellor and the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF and other funders.

1. Centers for Disease Control and Prevention, COVID-19 cases in U.S. Centers for disease control and prevention. https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html. Accessed 24 August 2020.
2. M. Klompas, M. A. Baker, C. Rhee, Airborne transmission of SARS-CoV-2: Theoretical considerations and available evidence. *J. Am. Med. Assoc.* **324**, 441–442 (2020).
3. D. Balcan *et al.*, Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 21484–21489 (2009).
4. M. U. Kraemer *et al.*, The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497 (2020).
5. K. H. Grantz *et al.*, The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. *Nat. Commun.* **11**, 1–8 (2020).
6. S. Gao *et al.*, Association of mobile phone location data indications of travel and stay-at-home mandates with COVID-19 infection rates in the US. *JAMA Netw. Open* **3**, e2020485 (2020).

7. N. W. Ruktanonchai *et al.*, Assessing the impact of coordinated COVID-19 exit strategies across Europe. *Science* **369**, 1465–1470 (2020).
8. S. Chang *et al.*, Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* **589**, 82–87 (2020).
9. C. Xiong, S. Hu, M. Yang, W. Luo, L. Zhang, Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 27087–27089 (2020).
10. F. Schlosser *et al.*, COVID-19 lockdown induces disease-mitigating structural changes in mobility networks. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 32883–32890 (2020).
11. C. Zachreson *et al.*, Risk mapping for COVID-19 outbreaks in Australia using mobility data. *J. R. Soc. Interface* **18**, 20200657 (2021).
12. A. Pan *et al.*, Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. *JAMA* **323**, 1915–1923 (2020).
13. D. M. Hartley, E. N. Perencevich, Public health interventions for COVID-19: Emerging evidence and implications for an evolving public health crisis. *JAMA* **323**, 1908–1909 (2020).
14. S. Lai *et al.*, Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature* **585**, 410–413 (2020).
15. M. Chinazzi *et al.*, The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395–400 (2020).
16. H. Tian *et al.*, An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* **368**, 638–642 (2020).
17. R. Zhang, Y. Li, A. L. Zhang, Y. Wang, M. J. Molina, Identifying airborne transmission as the dominant route for the spread of COVID-19. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 14857–14863 (2020).
18. D. Duque *et al.*, Timing social distancing to avert unmanageable COVID-19 hospital surges. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 19873–19878 (2020).
19. J. M. Brauner *et al.*, Inferring the effectiveness of government interventions against COVID-19. *Science* **371**, eabd9338 (2020).
20. N. Haug *et al.*, Ranking the effectiveness of worldwide COVID-19 government interventions. *Nat. Hum. Behav.* **4**, 1303–1312 (2020).
21. G. Giordano *et al.*, Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nat. Med.* **26**, 855–860 (2020).
22. M. Gatto *et al.*, Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 10484–10491 (2020).
23. IHME COVID-19 Forecasting Team, Modeling COVID-19 scenarios for the United States. *Nat. Med.* **27**, 94–105 (2020).
24. IHME COVID-19 Forecasting Team, Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months. MedRxiv [Preprint] (2020). https://doi.org/10.1101/2020.03.27.20043752. Accessed 30 August 2020.
25. Y. Zhou *et al.*, Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: A modelling study using mobile phone data. *Lancet Digit. Health* **2**, e417–e424 (2020).
26. A. Aleta *et al.*, Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nat. Hum. Behav.* **4**, 964–971 (2020).
27. O. Alagoz, A. K. Sethi, B. W. Patterson, M. Churpek, N. Safdar, Effect of timing of and adherence to social distancing measures on COVID-19 burden in the United States: A simulation modeling approach. *Ann. Intern. Med.* **174**, 50–57 (2021).
28. H. Fang, L. Wang, Y. Yang, Human mobility restrictions and the spread of the novel coronavirus (2019-nCoV) in China (Technical report 26906, National Bureau of Economic Research, 2020).
29. L. Yan *et al.*, An interpretable mortality prediction model for COVID-19 patients. *Nat. Med. Intel.* **2**, 283–288 (2020).
30. P. Wang, X. Zheng, J. Li, B. Zhu, Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics. *Chaos Solit. Fractals* **139**, 110058 (2020).
31. S. Chen, Q. Li, S. Gao, Y. Kang, X. Shi, State-specific projection of COVID-19 infection in the United States and evaluation of three major control measures. *Sci. Rep.* **33**, 22429 (2020).
32. Y. Zhou *et al.*, A spatiotemporal epidemiological prediction model to inform county-level COVID-19 risk in the United States. *Harvard Data Sci. Rev.*, 10.1162/99608f92.79e1f45e (2020).
33. L. J. Thomas *et al.*, Spatial heterogeneity can lead to substantial local variations in COVID-19 timing and severity. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 24180–24187 (2020).
34. S. Adhikari *et al.*, Assessment of community-level disparities in coronavirus disease 2019 (COVID-19) infections and deaths in large us metropolitan areas. *JAMA Netw. Open* **3**, e2016938 (2020).
35. D. A. Scheufele, N. M. Krause, I. Freiling, D. Brossard, How not to lose the COVID-19 communication war. *Issues in Science and Technology* (2020). https://issue.org/covid-19-communication-war/. Accessed 20 August 2020.
36. J. Flynn, P. Slovic, C. K. Mertz, Gender, race, and perception of environmental health risks. *Risk Anal.* **14**, 1101–1108 (1994).
37. D. Berrigan, K. Dodd, R. P. Troiano, S. M. Krebs-Smith, R. B. Barbash, Patterns of health behavior in us adults. *Prev. Med.* **36**, 615–623 (2003).
38. D. A. Scheufele, Science communication as political communication. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 13585–13592 (2014).
39. N. M. Baum, P. D. Jacobson, N. D. Goold, "Listen to the people": Public deliberation about social distancing measures in a pandemic. *Am. J. Bioeth.* **9**, 4–14 (2009).
40. M. Andersen, Early evidence on social distancing in response to COVID-19 in the United States. *SSRN* [Preprint] (2020). https://ssrn.com/abstract+3569368. Accessed 10 July 2020.
41. S. Farber, J. Johnson, New data shows young people need to take social distancing seriously. ABC News (2020). https://abcnews.go.com/Health/data-shows-young-people-social-distancing/story?id=712833844. Accessed 30 July 2020.
42. W. A. Clark, Residential segregation in American cities: A review and interpretation. *Popul. Res. Pol. Rev.* **5**, 95–127 (1986).
43. A. Chevan, Age, housing choice, and neighborhood age structure. *Am. J. Sociol.* **87**, 1133–1149 (1982).
44. M. S. Lau *et al.*, Characterizing superspreading events and age-specific infectiousness of SARS-CoV-2 transmission in Georgia, USA. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 22430–22435 (2020).
45. R. Rothstein, *The Color of Law: A Forgotten History of How Our Government Segregated America* (Liveright, 2017).
46. A. Goldstein, Income emerges as a major predictor of coronavirus infections, along with race. *The Washington Post*, 22 June 2020, https://www.washingtonpost.com/health/income-emerges-as-a-major-predictor-of-coronavirus-infections-along-with-race/2020/06/22/9276f31e-b4a3-11ea-a510-55bf26485c93_story.html. Accessed 10 August 2020.
47. D. Rubin *et al.*, Association of social distancing, population density, and temperature with the instantaneous reproduction number of SARS-CoV-2 in counties across the United States. *JAMA Netw. Open* **3**, e2016099 (2020).
48. J. J. Van Bavel *et al.*, Using social and behavioural science to support COVID-19 pandemic response. *Nat. Hum. Behav.* **4**, 460–471 (2020).
49. J. A. Weill, M. Stigler, O. Deschenes, M. R. Springborn, Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 19658–19660 (2020).
50. J. M. Barrios, Y. V. Hochberg, Risk perception through the lens of politics in the time of the COVID-19 pandemic (Technical report 27008, University of Chicago, 2020).
51. H. Allcott *et al.*, Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *J. Public Econ.* **191**, 104254 (2020).
52. P. Karaca-Mandic, A. Georgiou, S. Sen, Assessment of COVID-19 hospitalizations by race and ethnicity in 12 states. *JAMA Internal Medicine* **181**, 131–134 (2020).
53. The Brookings Institution, Black-white segregation in U.S. metro areas. https://www.brookings.edu/blog/the-avenue/2018/12/17/black-white-segregation-edges-downward-since-2000-census-shows/. Accessed 27 September 2020.
54. Robert Wood Johnson Foundation and University of Wisconsin Population Health Institute, The 2020 county health Rankings. https://www.countyhealthrankings.org/explore-health-rankings. Accessed September 27, 2020.
55. Public Health Office of Dane County, Dane County of Wisconsin COVID-19 dashboard. https://publichealthmdc.com/coronavirus. Accessed 15 August 2020.
56. Public Health Office of Milwaukee County, Milwaukee County of Wisconsin COVID-19 dashboard. https://county.milwaukee.gov/EN/COVID-19. Accessed 15 August 2020.
57. Wisconsin Department of Health Services, Wisconsin COVID-19 confirmed cases data by census tract boundary. https://publichealthmdc.com/coronavirus. Accessed 15 August 2020.
58. US Census Bureau, Cartographic boundary files - Shapefile. https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html. Accessed 12 August 2020.
59. US Census Bureau American Community Survey, Cartographic boundary files - Shapefile. https://www.census.gov/programs-surveys/acs. Accessed 12 August 2020.
60. What about bias in the SafeGraph dataset? https://www.safegraph.com/blog/what-about-bias-in-the-safegraph-dataset. Accessed 12 April 2020.
61. Y. Kang *et al.*, Multiscale dynamic human mobility flow dataset in the US during the COVID-19 epidemic. *Sci. Data* **7**, 390 (2020).
62. C. Gardiner, *Stochastic Methods* (Springer, 2009), vol. 4.
63. S. W. Park *et al.*, Reconciling early-outbreak estimates of the basic reproductive number and its uncertainty: Framework and applications to the novel coronavirus (SARS-CoV-2) outbreak. *J. R. Soc. Interface* **17**, 20200144 (2020).
64. J. Dehning *et al.*, Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* **15**, eabb9789 (2020).
65. A Stuart, K. Zygalakis, *Data Assimilation: A Mathematical Introduction* (Oak Ridge National Laboratory, Oak Ridge, TN, 2015).
66. P. Pons, M. Latapy, "Computing communities in large networks using random walks" in *International Symposium on Computer and Information Sciences*, P. Yolum, T. Güngör, F. Gürgen, C. Özturan, Eds. (Springer, 2005), pp. 284–293.
67. C. Rickert, D. Wahlberg, Amid rise in COVID-19 cases, Dane County tightens restrictions on bars, restaurants, indoor gatherings. *Wisconsin State Journal*, 1 July 2020. https://madison.com/wsj/news/local/govt-and-politics/amid-rise-in-covid-19-cases-dane-county-tightens-restrictions-on-bars-restaurants-indoor-gatherings/article_dec5a031-8a71-5746-9e1a-5f0d14996165.html. Accessed 20 August 2020.
68. City of Madison, Dane County issues new order limiting gatherings, bar and restaurant activity. https://cityofmadison.com/news/dane-county-issues-new-order-limiting-gatherings-bar-and-restaurant-activity. Accessed 15 August 2020.
69. T. Prestby, J. App, Y. Kang, S. Gao, Understanding neighborhood isolation through spatial interaction network analysis using location big data. *Environ. Plann. A Econ. Space* **52**, 1027–1031 (2020).
70. C. Wirz, J. Schwakopf, D. Brossard, L. DiPrete Brown, M. Brauer, Self-reported compliance and attitudes about social distancing during the COVID-19 outbreak. OSF Preprints [Preprint] (2020). 10.31219/osf.io/bv28d. Accessed 30 July 2020.
71. K. Chen *et al.*, How public perceptions of social distancing evolved over a critical time period: Communication lessons learnt from the American state of Wisconsin. *J. Sci. Commun.* **19**, A11 (2020).
72. V. D. Blondel, J. L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. *J. Stat. Mech. Theor. Exp.* **2008**, P10008 (2008).
73. L. Hubert, P. Arabie, Comparing partitions. *J. Classif.* **2**, 193–218 (1985).