Journal Pre-proof

Multiscale Dimensions of Spatial Process: COVID-19 Fully Vaccinated Rates in U.S. Counties

Tse-Chuan Yang PhD , Stephen A. Matthews PhD ,
Feinuo Sun PhD

PII: S0749-3797(22)00331-2
DOI: https://doi.org/10.1016/j.amepre.2022.06.006
Reference: AMEPRE 6900

To appear in: *American Journal of Preventive Medicine*

Please cite this article as: Tse-Chuan Yang PhD , Stephen A. Matthews PhD , Feinuo Sun PhD , Multiscale Dimensions of Spatial Process: COVID-19 Fully Vaccinated Rates in U.S. Counties, *American Journal of Preventive Medicine* (2022), doi: https://doi.org/10.1016/j.amepre.2022.06.006

Multiscale Dimensions of Spatial Process: COVID-19 Fully Vaccinated Rates in U.S. Counties

Tse-Chuan Yang, PhD,[1] Stephen A. Matthews, PhD,[2,3] Feinuo Sun, PhD[4]

From the [1]Department of Preventive Medicine and Population Health, University of Texas Medical Branch, Galveston, Texas; [2]Department of Sociology and Criminology, Pennsylvania State University, University Park, Pennsylvania; [3]Department of Anthropology, Pennsylvania State University, University Park, Pennsylvania; and [4]Global Aging and Community Initiative, Mount Saint Vincent University, Halifax, Nova Scotia, Canada

Address correspondence to: Tse-Chuan Yang, PhD, Department of Preventive Medicine and Population Health, University of Texas Medical Branch, 301 University Boulevard, Maurice Ewing Hall, Suite 1.128D, Galveston TX 77555. E-mail: tsyang@utmb.edu.

Introduction: To examine the heterogeneity of the associations between social determinants and COVID-19 fully vaccinated rate.

Methods: This study proposes three multiscale dimensions of spatial process, including “level of influence” (the percentage of population affected by a certain determinant across the entire area), “scalability” (the spatial process of a determinant into global, regional, and local process), and “specificity” (the determinant that has the strongest association with the fully vaccinated rate). The multiscale geographically weighted regression was applied to the COVID-19 fully vaccinated rates in U.S. counties (N=3,106) as of October 26, 2021 and the analyses were conducted in May 2022.

Results: The results suggest that: (1) percent of Republican votes in 2020 presidential election is a primary influencer as 84% of the U.S. population lived in counties where this determinant is found the most dominant; (2) demographic compositions (e.g., percentages of racial/ethnic minorities) play a larger role than socioeconomic conditions (e.g., unemployment) in shaping fully vaccinated rates; (3) the spatial process underlying fully vaccinated rates is largely local.

Conclusions: The findings challenge the one-size-fits-all approach to designing interventions promoting COVID-19 vaccination and highlight the importance of a place-based perspective in ecological health research.

INTRODUCTION

Ecological approaches to health research have been found to help produce well-specified individual models and improve population health[1] and they have been facilitated by the rapid development in ecological and spatial analysis methods.[2] A spatial perspective has been used to understand the geographical patterning of the ongoing novel coronavirus disease 2019 (COVID-19) pandemic.[3] The commonly used methods include, but are not limited to, data visualization,[4] spatial econometrics,[5] and geographically weighted regression (GWR).[6]

Though these methods have generated nuanced insight into the geography of COVID-19 in the U.S., the literature is limited in 1 major way. Explicitly, little attention has been paid to spatial heterogeneity or non-stationarity, which refers to the phenomenon when the direction and/or magnitude of the relationship between an independent and dependent variable varies by location. That is, the ecological approaches used in the literature mostly assume that changing the value of an independent variable will invoke the same change or response in the dependent variable regardless of location. This assumption is unrealistic for many reasons, such as differential responses to precaution measures (e.g., voluntary social distancing)[7] and different racial/ethnic composition.[8] Though some scholars have used GWR to address this issue,[9,10] they have not examined whether the spatial process operates at the same spatial scale across multiple social determinants.

Going beyond the extant literature, this study focuses on the COVID-19 fully vaccinated rates and argue that it is critical to investigate 3 dimensions of spatial process of a social determinant, namely level of influence, scalability, and specificity. These dimensions (see methods section for details) have not been proposed or examined in previous research. To examine these dimensions, this study first assembles a county-level dataset where the fully vaccinated rate (as of October 26, 2021) serves as the dependent variable and various political, demographic, and socioeconomic conditions are treated as the independent variables. The multiscale geographically weighted regression (MGWR) is used to identify the 3 dimensions of spatial process of each independent variable. The findings suggest that the 3 dimensions of spatial process vary across the independent variables. Moreover, these dimensions allow researchers to identify the important associations with fully vaccinated rates in U.S. counties and to facilitate discussions around place-based interventions that aim to increase vaccination rates.

METHODS

Study Sample

The analytical dataset is derived from multiple national sources and includes data on the U.S. counties in the lower 48 states (N=3,106). The dependent variable is the percentage of population aged 18 years and over who are fully vaccinated (i.e., having received 2-dose COVID-19 vaccine series or 1 dose of the single-shot vaccine) in a county as of October 26, 2021. These estimates are drawn from the overall U.S. COVID-19 vaccine administration and vaccine equity data, maintained by the Centers for Disease Control and Prevention.[11]

Measures

The independent variables include the percentage of votes for the Republican Party in the 2020 presidential election, and demographic composition and socioeconomic conditions. The percentage of Republican votes (i.e., total Republican votes divided by the total votes) in the 2020 presidential election is drawn from public data.[12] With respect to demographic and socioeconomic characteristics, the 2015–2019 American Community Survey 5-year estimates[13] were used to calculate the following variables (including people in both housing units and group quarters): percentage of older adults (aged ≥65 years), percentage of males, percentage of non-Hispanic Blacks, percentage of Hispanics, percentage of population aged ≥15 years who are married, and percentage of population aged ≥25 years who hold at least a bachelor's or professional degree. For socioeconomic conditions of a county, the following variables are considered: poverty rate (i.e., the percentage of households whose income in the past 12 months falls below the poverty level), unemployment rate (i.e., the percentage of people aged ≥16 years who are in the labor force but unemployed), public assistance reliance (i.e., the proportion of total income in the past 12 months for households with public assistance), and median household income in the past 12 months (a continuous variable measured in dollars).

Statistical Analysis

The multiscale geographically weighted regression (MGWR)[14] serves as the main analytic technique. The MGWR is an extension of GWR,[15] and both are discussed below. A GWR can be formulated as[15]:

$$y_i = \sum_{j=1}^{k} \beta_{ij} x_{ij} + \varepsilon_i \qquad (1)$$

where y_i is the response variable for location $i \in \{1, 2, \ldots, n\}$, x_{ij} refers to the *j*th independent variable ($j \in \{1, 2, \ldots, k\}$) and β_{ij} is the estimated parameter (i.e., coefficient) for x_{ij}. ε_i is the error terms. The GWR calibration for the coefficients at each location *i* can be written in matrix form:

$$\widehat{\boldsymbol{\beta}_i} = (\boldsymbol{X}^T \boldsymbol{W}_i \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{W}_i \boldsymbol{y}, i \in \{1, 2, \dots, n\} \qquad (2)$$

where $\boldsymbol{X}$ is the n*k matrix of independent variables (including the intercept), and $\boldsymbol{y}$ is the n*1 response variable vector, and $\boldsymbol{W}_i$ is the n*n spatial weighting matrix for location *i* in which the spatial weights are calculated based on a specified kernel function and bandwidth. The bandwidth is assumed to be constant across all independent variables, indicating that the spatial process generating the observed data is at the same spatial scale for all independent variables.[15]

The major difference between MGWR and GWR is that MGWR relaxes the constant bandwidth assumption by allowing for variable-specific optimized bandwidths.[14,16] A MGWR model can be regarded as a generalized additive model, which can be expressed as[14]:

$$\boldsymbol{y} = \sum_{j=1}^{k} \boldsymbol{f}_j + \boldsymbol{\varepsilon} \qquad (3)$$

where $\boldsymbol{f}_j$ is a smooth function applied to the *j*th independent variable[17] and in MGWR, each smooth function is a spatial GWR parameter surface calculated with a specific bandwidth that is calibrated using a back-fitting algorithm.[14] As such, MGWR is more generalized than GWR and the spatial process generating the observed values is permitted to vary by spatial scale (i.e., the bandwidth for each independent variable). It should be noted that MGWR standardizes all variables in the back-fitting algorithm, which facilitates the comparison of estimated coefficients. The technical details of GWR and MGWR can be found elsewhere.[18]

The analytic strategy consists of 3 phases: (1) conducting descriptive analysis, (2) implementing the ordinary least squares (OLS) regression, a baseline model also referred to as a “global” model (in contrast to the “local” models generated by GWR/MGWR), and (3) using MGWR to obtain the local estimates for each county. The MGWR results are presented via summary statistics and maps[19] and the Monte Carlo method[20] is used to formally test whether spatial non-stationarity exists, which indicates that the direction and/or the magnitude of a relationship between an independent and dependent varies by location.[15]

The strengths of MGWR allows users to identify 3 dimensions of multiscale spatial process for each independent variable. Specifically, the first dimension refers to the level of influence, which is defined as the percentage of population affected by a certain independent variable across the entire study area. If a factor is found to influence more than 50% of the entire population, this factor is defined as a primary influence; otherwise (i.e., ≤50%), it is a secondary influencer.

The second dimension is scalability that can be drawn from the calibrated bandwidth of a factor. Scalability is categorized into 3 groups, namely global, regional, and local. Explicitly, if the bandwidth of a factor is greater than 75% of the global bandwidth (i.e., the total number of observations across the entire study area), it approximates a “global” determinant. If the bandwidth of a factor is between 75% and 25% of the global bandwidth, it is regarded as a “regional” determinant. When the bandwidth of a variable is smaller than 25% of the global bandwidth, it is a “local” determinant.

Specificity is the third dimension. For each unit of analysis (i.e., county in this study), the coefficient of each covariate can be compared directly (because of standardization of variables in MGWR) and identify the independent variable that has the "strongest" impact (regardless of direction) on the dependent variable. These variables across the entire study area can be visualized to show the uniqueness of a certain variable in space (calibrated for each focal county/local model). In conventional OLS regression covariates with larger variances tend to have larger standardized coefficients, making coefficient comparisons problematic.[21] However, under the MGWR framework, each variable has its own bandwidth and the comparison is specific to the population of a given county. As such, the concern about coefficient comparison is not directly applicable to this specificity measure.

RESULTS

Due to the space constraint, the discussion about the descriptive statistics of the variables is presented in Appendix Table 1 and the regression results in Table 1 are explained below. Column (a) of Table 1 presents the OLS (i.e., global) standardized coefficient estimates, and the variance inflation factors (VIF) among the independent variables are included in column (b). Columns (c)-(g) are the summary statistics of the MGWR local estimates and column (h) shows the Monte Carlo test results. Several findings can be drawn from columns (a) and (b). First, the OLS standardized coefficients suggest that the percentage of Republican votes has the strongest and negative association (β= –0.71) with the fully vaccinated rate, net of other covariates. Second, in the global model, the demographic covariates seem to play a more important role than the socioeconomic variables in explaining the fully vaccinated rate. For example, among socioeconomic variables, only median household income shows a positive relationship with the fully vaccinated rate (with marginal statistical significance). By contrast, except for the percentages of male population and older adults, all other demographic variables are significantly associated with the fully vaccinated rate. Finally, all the VIFs are smaller than 6, suggesting that multicollinearity among the independent variables is not a concern.

Regarding the MGWR results (columns (c)-(h)), it is important to note that the Monte Carlo test results suggest that 4 variables demonstrate spatial non-stationarity, namely percentage of Republican votes, percentage of older adults, percentage of non-Hispanic Blacks, and median household income. Among these variables, the local estimates vary in both direction and magnitude. For example, the estimated association between the percentage of Republican votes and fully vaccinated rate ranges between –1.82 and 0.50 and a wider range is observed for the percentage of non-Hispanic Blacks (min= –1.15; max=2.81). Although these 4 variables all show spatial non-stationarity, their calibrated bandwidths vary from 46 to 602, suggesting that the spatial process underlying these variables is different. The MGWR model fits the data better than the OLS model, as reflected in a smaller corrected Akaike Information Criterion (5,302.26 vs 7,076.37) and a higher adjusted r-squared (0.74 vs 0.43).

Figure 1 illustrates the spatial non-stationarity with the MGWR local estimates for the percentage of Republican votes (1a), percentage of non-Hispanic Blacks (1b), and median household income (1c). Two findings are worth noting. First, while percent of Republican votes and percent of non-Hispanic Blacks have comparable calibrated bandwidths (i.e., 48 and 46), their local estimates show different spatial coverages and patterns. Specifically, almost all counties show a significant and negative local association between the percentage of Republican votes and fully vaccinated rate, except for the northeastern region, parts of the Mid-Atlantic, and parts of Nevada/California. The strongest negative associations can be found in Georgia and Florida. By contrast, the local associations between non-Hispanic Blacks and fully vaccinated rate are positive (red/green areas) in West Virginia and California/Nevada, and these associations are negative (blue areas) in Maine, parts of the Plains and Mid-West, and eastern Black Belt states. Second, even though median household income has a relatively large bandwidth (i.e., 602), the local estimates are almost all positive and the strongest and significant estimates are regional in scale concentrated in the North East and Great Lakes regions. The local estimates in the Pacific region are also significant but the magnitude of the association is weaker.

The MGWR results help to identify the 3 dimensions of each independent variable in Table 2. Following the definitions discussed previously, variables were dichotomized into primary and secondary influencers based on the total population within the counties affected. Five variables are primary (e.g., percentage of Republican votes) and 6 are secondary (e.g., unemployment rate). For example, 68.6% of population in lower 48 states live in counties where poverty is a significant determinant, which is a primary influencer. Furthermore, compared with the global bandwidth (i.e., 3,106), 6 independent variables are classified as local factors because their bandwidths are smaller than 777 (i.e., 3,106*0.25). Three are global determinants as their bandwidths are larger than 2,330 (i.e., 3,106*0.75). Another 2 variables have regional influence given the bandwidths falling between the 25% and 75% of the global bandwidth. Finally, the percentage of Republican votes is the strongest determinant of fully vaccinated rate in most counties and the specificity dimension was visualized in Figure 2. Across the lower 48 states, the percentage of Republican votes has the strongest association with fully vaccinated rate in 82%

(2,556/3,106*100%) of counties. The percentage of non-Hispanic Blacks is the most dominant factor in almost 10% (302/3,106*100%) of the counties and these counties are found in clusters including, southern California/Nevada, central Appalachia, and Maine. Median household income has the strongest relationship with fully vaccinated rate in 122 counties, many of which are in New England and Virginia.

DISCUSSION

This study aims to investigate whether the spatial process underlying the fully vaccinated rate is universal across a range of social determinants in U.S. counties. By exploiting the recently developed MGWR method,[20] this study argues that the local estimates and calibrated bandwidth for each independent variable provide details about the spatial process that generates the observed patterns. Specifically, this study defines and operationalizes 3 dimensions of spatial process for each social determinant (i.e., level of influence, scalability, and specificity) and then demonstrates how these dimensions shed new light on how social determinants are associated with fully vaccinated rates in U.S. counties. While some studies have applied the MGWR to COVID-19 research,[6,10] no prior research has proposed and investigated the 3 dimensions of spatial process. The findings indicate that not all social determinants share the same spatial process. For example, only 13.9% of population live in counties where unemployment rate is associated with fully vaccinated rate (Table 2) but more than 80% of population reside in counties in which percentage of Republican votes is related to the fully vaccinated rate. That is, the influence of a factor on fully vaccination rates varies across U.S. counties.

The multiscale perspective allows users to classify social determinants into local, regional, and global factors based on bandwidths. Accordingly, percentage of Hispanics, percentage of married population, and percentage of population having at least a bachelor's or professional degree are globally/universally important. In Table 1, the variations in local estimates of these factors are low. Moreover, the percentage of Republican votes was found to have the strongest relationship with the fully vaccinated rate in 2,556 of 3,106 total counties, making this indicator the most dominant factor across space.

The different levels of spatial heterogeneity (e.g., local/regional) echo the argument that social processes appear to be non-stationary[18] as spatial variation in norms and preferences (or different administrative, political, or other contextual factors) produces different responses to the same stimuli. Specific to this study, the county is an appropriate level of analysis as it functions administratively as a coherent unit of local government and in many parts of the county corresponds to aggregate level daily routines and social interactions. More importantly, counties are embedded within larger governmental and administrative units such as metropolitan areas and in particular, states. States serve as a decision-making entity that has been prominent in guidelines and mandates regarding area based COVID-19 policy and action.

Several additional tests were conducted to examine whether the findings and conclusions are sensitive to unattended covariates or measurements. For example, considering other covariates, such as COVID-19 case rate, in the analysis does not alter the findings and conclusions. As the causality between fully vaccinated rate and COVID-19 case rate may be reciprocal, it is not include in this study. Furthermore, regarding the potential non-linear relationships between key independent variables and fully vaccinated rates, the analysis found that only percentage of Republican votes has a small quadratic effect and the vertex does not exist within the range of the percentage of Republican votes (results available upon request). Also, a composite social disadvantage index was created with the socioeconomic conditions,[22] and this index yields similar substantive MGWR results. Finally, using more stringent criteria to define scalability (e.g., 10% and 90% thresholds) does not change the conclusions.

Limitations

This study is subject to several limitations. First, given the cross-sectional research design, the findings cannot make any causal inferences between fully vaccinated rates and the independent variables. Second, this ecological study is subject to the modifiable areal unit problem,[23,24] and the results could not be generalized to the individual level. Third, as the unit of analysis is the county, the analysis may mask spatial heterogeneity at a finer geographic unit, such as ZIP code or census tracts. Finally, as the pandemic is ongoing, the analysis focuses on the early phase of vaccination rollout in the U.S. As such, boosters and other vaccination recommendations are not considered here, and using data from an extended (or shorter) time period may alter the conclusions.

Some scholars have used social media data to predict COVID-19 hospitalization and case rate[25] and conducted sensitivity analysis with environmental modeling approaches.[26] Such approaches are valuable but they do not examine different levels of spatial process with correlated data. Future research should incorporate these perspectives into county-level analysis, including vaccination rates.

CONCLUSIONS

Situating this study in the emerging COVID-19 ecological research,[27] this study has advanced the literature in 2 ways. Substantively, previous studies largely focus on COVID-19 cases and deaths[9,28–30] and as yet little attention has been paid to COVID-19 vaccination rates.[6,31] The MGWR results offer robust evidence identifying bipartisanship as playing a significant role in the differences observed in county-level fully vaccinated rates, net of other potential demographic and socioeconomic conditions. The specificity dimension further highlights the spatially varying patterns and offers insight into place-based policies aiming to increase fully vaccinated rates. Methodologically, this study introduces the three dimensions of spatial process to the literature and suggests that these dimensions improve the understanding of how ecological social determinants shape the spatial patterns of population health outcomes, such as COVID-19 fully vaccinated rates. Without the detailed information about the spatial process of individual ecological factors, it is difficult to assess their impacts on health.

ACKNOWLEDGMENTS

The authors acknowledge the support from the National Institute on Aging (NIA)-funded Interdisciplinary Network on Rural Population Health and Aging (INRPHA) group (R24-AG065159) and the Population Research Institute at Penn State University, which is supported by an infrastructure grant by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (P2CHD041025).

No financial disclosures were reported by the authors of this paper.

CRediT

Tse-Chuan Yang: Conceptualization; Methodology; Formal analysis; Visualization; Writing - Original Draft. Stephen A. Matthews: Conceptualization; Methodology; Validation; Writing - Review & Editing. Feinuo Sun: Data curation; Formal analysis; Validation.

REFERENCES

1. Macintyre S, Ellaway A. Ecological approaches: rediscovering the role of the physical and social environment. *Social Epidemiology*. 2000;9(5):332-348.
2. Richardson DB, Volkow ND, Kwan M-P, Kaplan RM, Goodchild MF, Croyle RT. Spatial turn in health research. *Science*. 2013;339(6126):1390-1392. https://doi.org/10.1126/science.1232257.
3. Franch-Pardo I, Napoletano BM, Rosete-Verges F, Billa L. Spatial analysis and GIS in the study of COVID-19. A review. *Sci Total Environ*. 2020;739:140033. https://doi.org/10.1016/j.scitotenv.2020.140033.
4. Allen WE, Altae-Tran H, Briggs J, et al. Population-scale longitudinal mapping of COVID-19 symptoms, behaviour and testing. *Nature Human Behaviour*. 2020;4(9):972-982. https://doi.org/10.1038/s41562-020-00944-2.
5. Sun F, Matthews SA, Yang T-C, Hu M-H. A spatial analysis of the COVID-19 period prevalence in US counties through June 28, 2020: where geography matters? *Ann Epidemiol*. 2020;52:54-59.e1. https://doi.org/10.1016/j.annepidem.2020.07.014.
6. Mollalo A, Tatar M. Spatial modeling of COVID-19 vaccine hesitancy in the United States. *Int J Environ Res Public Health*. 2021;18(18):9488. https://doi.org/10.3390/ijerph18189488.
7. Yan Y, Malik AA, Bayham J, Fenichel EP, Couzens C, Omer SB. Measuring voluntary and policy-induced social distancing behavior during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*. 2021;118(16). https://doi.org/10.1073/pnas.2008814118.

8. Strully K, Yang T-C, Liu H. Regional variation in COVID-19 disparities: connections with immigrant and Latinx communities in US counties. *Ann Epidemiol*. 2021;53:56-62.e2. https://doi.org/10.1016/j.annepidem.2020.08.016.
9. Maiti A, Zhang Q, Sannigrahi S, et al. Exploring spatiotemporal effects of the driving factors on COVID-19 incidences in the contiguous United States. *Sustain Cities Soc*. 2021;68:102784. https://doi.org/10.1016/j.scs.2021.102784.
10. Mollalo A, Vahedi B, Rivera KM. GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Sci Total Environ*. 2020;728:138884. https://doi.org/10.1016/j.scitotenv.2020.138884.
11. CDC. COVID-19 Vaccinations in the United States, County. https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh/data. Accessed November 30, 2021.
12. McGovern T. Data from: US County Level Election Results 08-20. 2020.
13. U.S. Census Bureau. Data from: American Community Survey, IPUMS. 2021.
14. Fotheringham AS, Yang W, Kang W. Multiscale geographically weighted regression (MGWR). *Ann Am Assoc Geogr*. 2017;107(6):1247-1265. https://doi.org/10.1080/24694452.2017.1352480.
15. Brunsdon C, Fotheringham S, Charlton M. Geographically weighted regression. *J Roy Stat Soc Ser D*. 1998;47(3):431-443. https://doi.org/10.1111/1467-9884.00145.
16. Yu H, Fotheringham AS, Li Z, Oshan T, Kang W, Wolf LJ. Inference in multiscale geographically weighted regression. *Geographical Analysis*. 2020;52(1):87-106. https://doi.org/10.1111/gean.12189.

17. Wood SN. *Generalized additive models: an introduction with R*. CRC press; 2017. https://doi.org/10.1201/9781315370279.
18. Fotheringham AS, Brunsdon C, Charlton M. *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons; 2003.
19. Matthews SA, Yang T-C. Mapping the results of local statistics: Using geographically weighted regression. *Demographic Res*. 2012;26:151. https://doi.org/10.4054/DemRes.2012.26.6.
20. Oshan TM, Li Z, Kang W, Wolf LJ, Fotheringham AS. mgwr: A Python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS Int J Geoinf*. 2019;8(6):269. https://doi.org/10.3390/ijgi8060269.
21. Greenland S, Maclure M, Schlesselman JJ, Poole C, Morgenstern H. Standardized regression coefficients: a further critique and review of some alternatives. *Epidemiol*. 1991:387-392. https://doi.org/10.1097/00001648-199109000-00015.
22. Singh GK. Area deprivation and widening inequalities in US mortality, 1969-1998. *Am J Public Health*. 2003;93(7):1137-1143. https://doi.org/10.2105/AJPH.93.7.1137.
23. Fotheringham AS, Wong DW. The modifiable areal unit problem in multivariate statistical analysis. *Environ Plan A*. 1991;23(7):1025-1044. https://doi.org/10.1068/a231025.
24. Openshaw S. *The modifiable areal unit problem*. Geo Books; 1983.
25. Galbraith E, Li J, Rio-Vilas VJD, Convertino M. In. To. COVID-19 socio-epidemiological co-causality. *Sci Rep*. 2022;12(1):1-25. https://doi.org/10.1038/s41598-022-09656-1.

26. Pianosi F, Beven K, Freer J, et al. Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environ Model Softw*. 2016;79:214-232. https://doi.org/10.1016/j.envsoft.2016.02.008.

27. Shakil MH, Munim ZH, Tasnia M, Sarowar S. COVID-19 and the environment: A critical review and research agenda. *Sci Total Environ*. 2020:141022. https://doi.org/10.1016/j.scitotenv.2020.141022.

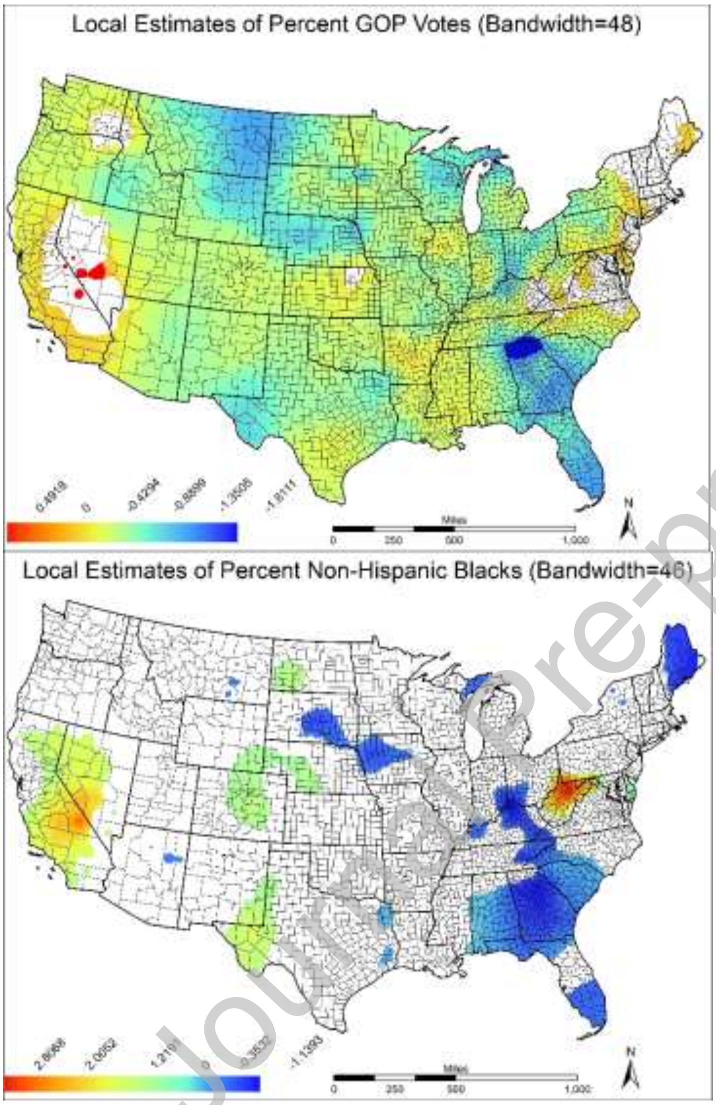
28. Fielding-Miller RK, Sundaram ME, Brouwer K. Social determinants of COVID-19 mortality at the county level. *PLoS One*. 2020;15(10):e0240151. https://doi.org/10.1371/journal.pone.0240151.

29. Karaye IM, Horney JA. The impact of social vulnerability on COVID-19 in the US: an analysis of spatially varying relationships. *Am J Prev Med*. 2020;59(3):317-325. https://doi.org/10.1016/j.amepre.2020.06.006.

30. Basu A. Estimating The Infection Fatality Rate Among Symptomatic COVID-19 Cases In The United States: Study estimates the COVID-19 infection fatality rate at the US county level. *Health Aff*. 2020;39(7):1229-1236. https://doi.org/10.1377/hlthaff.2020.00455.

31. Ye X. Exploring the relationship between political partisanship and COVID-19 vaccination rate. *J Public Health*. 2021: 1-8. https://doi.org/10.1093/pubmed/fdab364.

LIST OF FIGURES



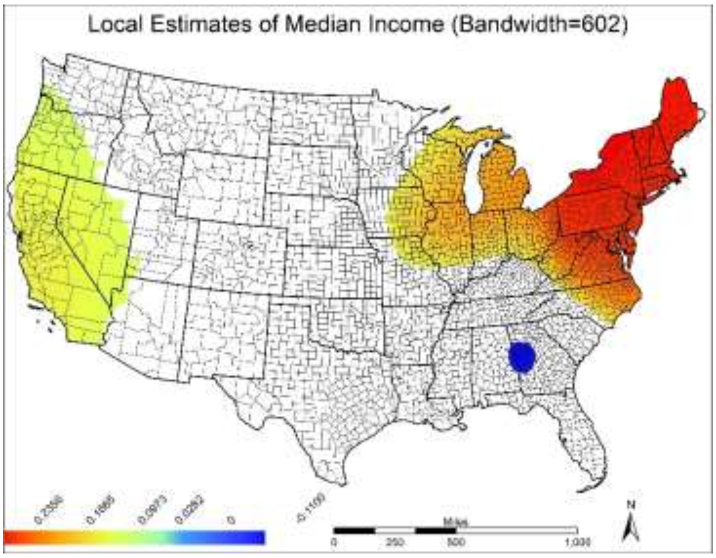


Figure 1. (1a) Local estimated relationship between percentage of Republican votes and fully vaccinated rate. (1b) Local estimated relationship between percentage of non-Hispanic Blacks and fully vaccinated rate. (1c) Local estimated relationship between median household income and fully vaccinated rate.

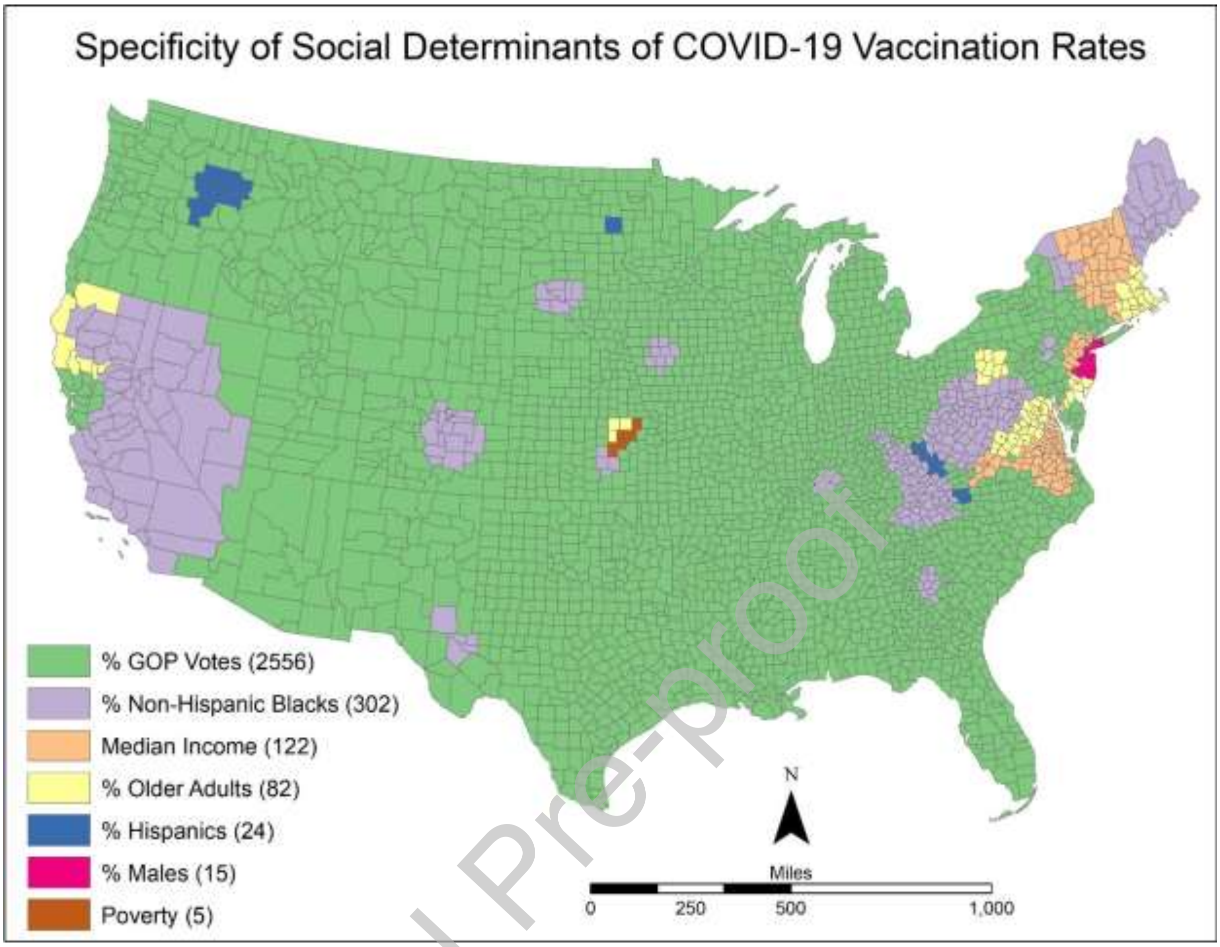


Figure 2. Specificity dimension of multiscale spatial process of fully vaccinated rate in U.S. counties.

Table 1. OLS and MGWR Results of COVID-19 Fully Vaccinated Rate

Variable	**Global estimates (a)**	**VIF[a] (b)**	**Mean (c)**	**SD (d)**	**Min (e)**	**Median (f)**	**Max (g)**	**Monte Carlo *p*-value (h)**	**MGWR Bandwidth (i)**
Percent Republican votes	**–0.71***	3.46	–0.63	0.28	–1.82	–0.62	0.50	**<0.001**	48
Percent aged ≥65 years	0.01	1.66	0.07	0.21	–1.22	0.08	0.84	**<0.001**	52
Percent males	–0.01	1.19	–0.04	0.09	–0.32	–0.03	0.20	0.35	144
Percent non-Hispanic Blacks	**–0.31***	2.13	–0.06	0.48	–1.15	–0.08	2.81	**<0.001**	46
Percent Hispanics	**0.07***	1.25	0.13	0.00	0.13	0.13	0.13	0.86	3104
Percent married	**0.14***	3.24	0.05	0.00	0.05	0.05	0.06	0.90	3104
Percent bachelor's degree and above	0.01	3.40	0.03	0.00	0.03	0.03	0.03	0.91	3104
Poverty rate	–0.01	4.26	–0.09	0.03	–0.15	–0.09	–0.02	0.18	1210
Unemployment rate	–0.00	2.00	–0.02	0.09	–0.35	–0.03	0.32	0.09	203
Percent on public assistance	–0.02	1.35	–0.04	0.03	–0.08	–0.03	0.00	0.28	1556
Median household income	0.06	5.65	0.08	0.09	–0.11	0.07	0.24	**<0.001**	602
Intercept	0.00	–	0.12	0.17	–0.43	0.15	0.57	**<0.001**	225
AIC_C	7076.37						5302.26		
Adjusted R^2	0.43						0.74		

Note: Boldface indicates statistical significance (*** p<0.001).
[a]The variance inflation factors (VIF) among the independent variables are all smaller than 6, indicating that multicollinearity is not a concern.

OLS, ordinary least squares; MGWR, multiscale geographically weighted regression; AIC, Akaike Information Criterion.

Table 2. Three Dimensions of Multiscale Spatial Process for Each Independent Variable Based on the MGWR Models

Variable (bandwidth)	**Level of influence**[a]	**Scalability**[b]	**Specificity**[c]
Percent Republican votes (48)	primary (83.6%)	local	2,556 (82.3%)
Percent aged ≥65 years (52)	secondary (42.5%)	local	82 (2.6%)
Percent males (144)	secondary (33.7%)	local	15 (0.5%)
Percent non-Hispanic Blacks (46)	secondary (29.8%)	local	302 (9.7%)
Percent Hispanics (3104)	primary (100.0%)	global	24 (0.8%)
Percent married (3104)	primary (100.0%)	global	0
Percent bachelor's degree and above (3104)	secondary (0.0%)	global	0
Poverty rate (1210)	primary (68.6%)	regional	5 (0.2%)
Unemployment rate (203)	secondary (13.9%)	local	0
Percent on public assistance (1556)	secondary (37.9%)	regional	0
Median household income (602)	primary (55.6%)	local	122 (3.9%)

[a]If the variable affects more than 50% of the total population, it is a primary influencer; otherwise (i.e., ≤50%), it is a secondary influencer. The percentage of population affected by a factor is included in the parentheses.
[b]If the bandwidth of a variable is larger than 75% of the global bandwidth (i.e., 2330), it is a global determinant; if the bandwidth is smaller than 25% of the global bandwidth (i.e., 777), it is a local determinant; if the bandwidth is between 75% and 25% of the global bandwidth, it is a regional determinant.
[c]The number and percentage of counties that the focal variable has the strongest significant impact on the dependent variable (i.e., the largest absolute value of the coefficients that are statistically significant).

MGWR, multiscale geographically weighted regression.