

# Documenting Social, Geographic, and Economic Inequities in COVID-19 Mortality at the County Level in the US Using Generalized Additive Models

Christian Testa<sup>1,\*</sup>

November 23, 2022

## Abstract

We present three types of applications of generalized additive models (GAMs) to COVID-19 mortality rates in the US for the purpose of advancing methods to document inequities with respect to which communities suffered disproportionate COVID-19 mortality rates. First, GAMs can be used to describe the changing relationship between COVID-19 mortality and county-level covariates (sociodemographic, economic, and political metrics) over time. Second, GAMs can be used to perform spatiotemporal smoothing that pools information over time and space so that county level mortality rates that tend to be noisy due to small population counts or stochasticity can be summarized by a smooth, dynamic latent surface describing the mortality risk associated with geographic locations over time. Third, estimation of the associations between county-level covariates and COVID-19 mortality after controlling for spatiotemporal effects allow for the distinguishing of what patterns in COVID-19 mortality were more plausibly due to geography than variation in county-level covariates. Each of these approaches provides a valuable approach and perspective to documenting inequities in COVID-19 mortality by addressing the question of which populations have suffered the worst burden of COVID-19 mortality taking into account the nonlinear spatial, temporal, and social patterning of disease.

Abbreviations used: United States (US), Coronavirus Disease 2019 (COVID-19), Generalized Additive Model (GAM), Centers for Disease Control and Prevention (CDC)

<sup>1</sup> Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health

\* Correspondence: [Christian Testa <catesta@hsph.harvard.edu>](mailto:catesta@hsph.harvard.edu)

## Introduction

As we enter the third winter with the novel coronavirus disease COVID-19 in the United States, evidence documenting the intense disparities in COVID-19 mortality rates comparing socially advantaged and disadvantaged populations continues to mount. Eliminating inequities in health outcomes has been stated as a major policy goal of the Biden administration (The White House, 2021, 2022), representing a revitalized focus on health justice and underlining the importance of adequate data reporting systems that measure the prevalent health inequities. To this end, we use generalized additive models (GAMs) as a flexible regression framework to illustrate the evolving roles and relationships sociodemographic, geographic, and economic conditions have with respect to trends in COVID-19 mortality. The code, data, and documentation necessary to reproduce the analyses contained in this paper are online and free to access at <https://github.com/catesta01/covid.gradient.estimation>.

## Background

Having passed over 1 million COVID-19 deaths in the United States earlier this year (Donovan, 2022), and facing uncertain prospects for the third COVID-19 winter looming even as new iterations on the COVID-19 vaccines become available, it remains critical that inequities in COVID-19 outcomes are documented and

analyzed to reckon with the unjust and unfair burden of preventable illness. Even though the first vaccines were granted emergency use authorization by the U.S. Food and Drug Administration in 2020 (Mayo Clinic, 2022), with the first shots going in arms in December 2020, COVID-19 is still continuing to cause hundreds of deaths a day in the US in the fall of 2022 (“United States COVID - Coronavirus Statistics - Worldometer,” 2022). The new bivalent vaccines released at the end of August 2022 contain mRNA sequences from both the original strain as well as the recently emergent BA.4 and BA.5 lineages in an effort to make the nation’s immunity more up-to-date and robust against the myriad of phylogenetic directions the COVID-19 virus is evolving to explore (Office of the Commissioner, 2022). Despite the updated bivalent boosters representing a significant step forward in prevention strategy, less than 4% of eligible Americans had taken the booster in the first month after it became available (Bendix, 2022; Lambert, 2022). As such, and with an enduring history of inequities in health care access in the US (Bailey et al., 2021; Blendon et al., 2002; Carpenter, 2021; Chrisler et al., 2016; Feldman et al., 2021; Okonkwo et al., 2021; Ortega & Roby, 2021; Rapp et al., 2022; Whitehead et al., 2016), it is clear that without further intervention not all communities will be equally able to benefit from the updated vaccines and inequities in COVID-19 illness and mortality may persist despite the technological innovations in vaccine technology.

## The Role of Geography in COVID-19

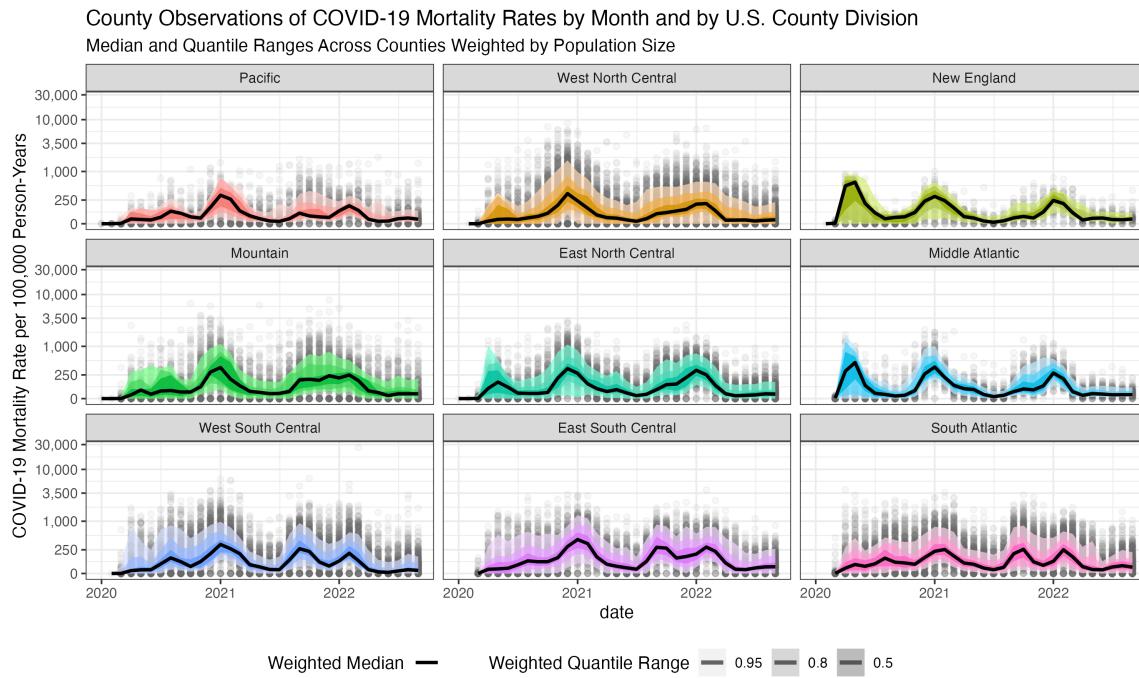


Figure 1: Estimates of monthly COVID-19 mortality rates per 100,000 person-years by county organized by Census Division. For each division, the median trendline and quantile ranges are shown weighted by county population size.

Prior literature has demonstrated that geography has played a significant role in the spread and impact of COVID-19. Methods employed to highlight the role geography plays have included quantile regression (Sigler et al., 2021), Besag-York-Mollie's mixed models (Whittle & Diaz-Artiles, 2020), spatial cluster analysis (Sugg et al., 2021), geographically weighted regression (Mollalo et al., 2020; Park et al., 2021), and others.

Figure 1 shows the monthly COVID-19 mortality rates for the counties grouped within each of the nine U.S. Census Divisions (U.S. Department of Commerce Economics and Statistics Administration & U.S. Census Bureau, 2000). The figure summarizes each division's median mortality rates weighted by county population size. Quantile ranges are included to illustrate the range of mortality rates observed, again weighted by county population size. Weighting based on county population size is used as opposed to equal weighting for

calculating the median and quantile ranges so that in divisions where low population counties predominate, mortality rate estimates that are noisy due to small population counts are not weighted equally to more reliable mortality rate estimates from larger population counties. Notably, the mortality associated with the early surge of cases starting in NYC and spreading through New York, New Jersey, and Massachusetts is visible in the Middle Atlantic and New England division figures. The figure also illustrates how the first peak in the mortality time-series for states in the Midwest (West North Central, East North Central) occurred later, in late 2020 and going into early 2021.

In the US context, one of the key aspects to the geographic story of COVID-19's spread and diffusion was the early surge of cases and epicenter in New York City during March 2020 (Thompson, 2020) followed by subsequent waves of cases in the South and Midwest (Glenza, 2020; Scott, 2020; Shumaker & Wu, 2020). As Park et al. stated summarized trends in the US from March 2020 to May 2021, "hot spots have shifted from densely populated cities and the states with a high percentage of socially vulnerable individuals to the states with relatively relaxed social distancing requirements, and then to the states with low vaccination rates" (2021).

When considering the drivers of the COVID-19 pandemic, it's necessary to note that geography and social conditions are inextricably linked. In July 2021, the CDC reported that "the COVID-19 cumulative death rate in non-metropolitan areas has exceeded that of metropolitan areas since December 2020," noting that of the approximately 1/5th of Americans who live in rural areas, many "are considered highly vulnerable according to CDC's Social Vulnerability Index (SVI), which includes factors such as housing, transportation, socioeconomic status, race, and ethnicity" (CDC, 2021). Moreover, rural communities often have lower health insurance rates, higher disability rates, older populations, and limited access to health care. One of CDC's Morbidity and Mortality Weekly Reports found that vaccination against COVID-19 was lower in rural communities than in urban communities between December 2020 and April 2021 (Murthy, 2021).

### The Social Determinants of COVID-19 Mortality

Even since the beginning of the COVID-19 outbreak in the US, data reflected sharp inequities in mortality rates. During January 22nd to May 5th 2020, county COVID-19 mortality rates were 4.94 (95% CI 4.78, 5.09) times higher in counties in the highest quintile of percent People of Color (61%-100%) compared to counties in the lowest quintile of percent People of Color (0%-17.2%) (Chen & Krieger, 2021). This was not wholly unanticipated: as COVID-19 was beginning to take off in the US, some were already calling attention to the fact that COVID-19 threatened to exacerbate existing disparities (Kim et al., 2020). Kim, Marrast, and Conigliaro noted three structural barriers in COVID-19 prevention and care: 1) originally requiring New York residents to have a doctor's prescription for a COVID-19 test reduced the opportunity for healthcare for People of Color as they are less likely to have a primary care provider; 2) drive-through testing made testing disadvantaged those who rely on public transportation; 3) quarantining at home while waiting the 7-10 days originally required for COVID-19 test results to come back posed an economic and social challenge that many in already financially difficult situations may not have been able to take on (2020). Others noted yet more reasons why COVID-19 threatened to worsen an already inequitable healthcare landscape in the US: those who reside in prisons and jails, immigrants and undocumented people, people with disabilities, and people experiencing homelessness all face additional challenges in seeking and getting the healthcare they deserve (Okonkwo et al., 2021). It's not just a hypothetical that those with more structurally enfranchised privileges have been more able to mitigate their risk of negative health outcomes associated with COVID-19 — during February 1 to April 1, 2020, New York City residents who lived in more affluent neighborhoods were more likely to have left the city, while New York City residents from more Black and Hispanic neighborhoods were more likely to continue working (Coven & Gupta, 2020).

Contrary to the oft used phrase that the 'virus does not discriminate', the data presented here suggest that this virus, as many other infectious diseases, has the greatest implications for the most vulnerable people. The intersections between health and human rights are clear—the health of a society and vulnerability to a pandemic are directly related to its human rights track record for those who are marginalised. (Okonkwo et al., 2021)

When vaccines became available, vaccine appointments were often only available to be scheduled through

online web-portals contributing to the inequities between those who had internet access and technological literacy and those who didn't (Press et al., 2021). Vaccination sites have not been equally distributed and areas determined to be vaccine deserts have been found to have disproportionately more Black and Hispanic residents (Rader et al., 2021). In fact, healthcare facilities in counties with higher Black composition had 32% (95% CI 14%-47%) lower odds of serving as vaccine sites (Hernandez et al., 2022).

What vaccination has been administered hasn't suddenly erased the unequal burden of COVID-19 either; in August 2022 the New York Times was reporting "Black death rates at this winter's peak were greater than those of white people by 34 percent in rural areas, 40 percent in small or medium cities and 57 percent in big cities and their suburbs" (Goldstein, 2022). As COVID-19 case and mortality rates have waxed and waned, the inequities have widen and shrunken, often with racial/ethnic inequities growing during times when COVID-19 rates have surged (Hill & Artiga, 2022). During 2022, the age-standardized COVID-19 mortality rates for white people have, at times, been slightly higher than those of Black and Hispanic people, predominantly because the mortality rate among white people has increased (Johnson & Keating, 2022). It's important to note that white COVID-19 mortality rates overtaking the Black mortality rates does not imply that equity has been established: neither does this undo the cumulative impact of mortality (which has been twice as high for Black people compared to white people (Hill & Artiga, 2022)), nor does it imply the underlying systemic barriers to equity have been overturned (Del Rios et al., 2022). As Del Rios, Chomilo, and Lewis note, instead, COVID-19 leaves in its wake more years of life expectancy lost, wages lost, and degradation of mental health in Communities of Color.

## Methods

### Data Sources

The following variables were retrieved at the county level:

- Counts of COVID-19 deaths (The New York Times, 2021).
- Population size estimates for 2020 from the U.S. Census (Redistricting File Public-Law 94-171 Dataset) (US Census Bureau, 2021).
- Median age, median household income, racial/ethnic composition, population density, percent below the federal poverty line, and number of households with high (\$100k+)/low (<\$25k) household income by racial/ethnic group from the 2014-2019 5-year American Community Survey (US Census Bureau, 2020) through the `tidyverse` R package (Walker & Herman, 2022).
- Votes cast in the 2020 presidential election (MIT Election Data and Science Lab, 2022)

### Generalized Additive Models

Traditional generalized linear models fit a regression model using the functional relation  $g(\mu_i) = \mathbf{X}_i\beta$  where Examples of common link functions include the identity function for linear regression, log for Poisson regression, and the logit function for logistic regression.

In contrast, generalized additive models (GAMs) improve upon generalized linear models by allowing for the fitting of smooth functions that transform the  $x$  variables. This is a convenient means to include nonlinear relationships between the outcome and predictor variables. Whereas a generalized linear model may have looked like

$$g(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} \dots$$

a generalized additive model could look like

$$g(\mu_i) = \mathbf{A}_i\theta + f_1(x_{i1}) + f_2(x_{i2}) + f_3(x_{i3}, x_{i4}) + \dots$$

where  $\mu_i \equiv \mathbb{E}(Y_i)$ ,  $\mathbf{A}_i$  is a row of the model matrix for any strictly non-parametric model components,  $\theta$  is the corresponding parameter vector, and the  $f_j$  are smooth functions of the covariates  $x_k$  (S. N. Wood, 2017).

The smooth functions estimated as part of fitting a GAM are constructed using spline basis functions. These spline basis functions allow for the smooth interpolation of trends in the data allowing for the incorporation

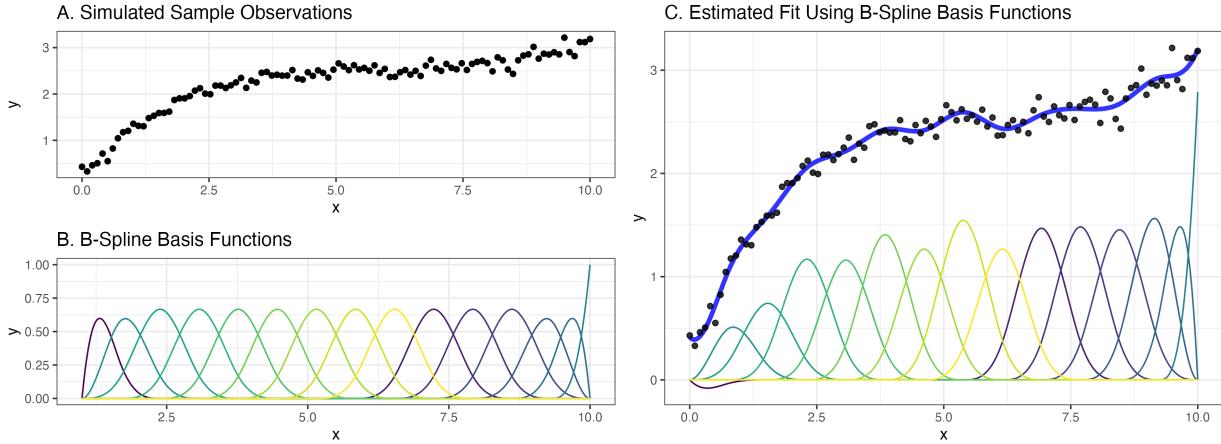


Figure 2: A demonstration of fitting regression models using B-spline basis functions. Part A. shows the simulated dataset. Part B. shows the B-spline basis functions. Part C. shows the weighting of the B-spline basis functions and their weighted sum with an intercept added to create the regression model fit to the data.

of nonlinearity. To avoid overfitting the data, a penalty term is introduced that controls the degree of “wigginess” or smoothness and this penalty term is fit using generalized cross validation. B-splines are one kind of spline basis function that is commonly used and are especially popular because of their property that they are non-negative on only a finite interval.

While a variety of varying kinds of spline bases exist (cubic splines, B- and P-splines, thin-plate splines, etc.), our particular application setting warrants the use of tensor-product smooths because we require anisotropic penalization. That is to say, when we seek to create spatiotemporally smoothed model estimates, it’s inappropriate to assume that the amount of smoothing across space should be the same as the amount of smoothing across time because they’re in fundamentally different units (space being measured in Cartesian coordinates and time being measured in years, months, days, etc.). More details about tensor-product smooths are available in Simon Wood’s *Generalized Additive Models - An Introduction with R* (2017).

## Overdispersion and Event Modeling

Above and beyond using the GAM framework to allow for flexible, nonlinear relationships between our observed county-level variables and COVID-19 mortality, we must have a model specification that agrees with the data generating process. In our case, the data generated are counts of deaths per population, and the class of models most suited to represent counts of events are Poisson, quasi-Poisson, and negative binomial models. Here we’ve chosen to use the negative binomial model as it accounts for overdispersion and has the intuitive interpretation of a Poisson model with gamma distributed underlying rate parameter (Gelman et al., 2013).

A negative binomial model in the context of modeling count data can be written

$$y_i \sim \text{NegativeBinomial}(u_i \theta_i, \phi),$$

where  $y_i$  are the counts observed,  $u_i$  is the exposure,  $\theta_i$  are the rates, and  $\phi$  determines the amount of overdispersion. The rates are modeled as  $\theta_i = e^{X_i \beta}$  where  $X_i$  are the observed covariates and  $\beta$  are the coefficients on the covariates corresponding to log rate ratios. The logarithm of the exposure,  $\log(u_i)$  is often called (and later herein referenced as) the offset. In epidemiological contexts, the offset is often representative of the amount of person-time during which observations were recorded. Whereas the Poisson model holds that  $\text{var}(y) = \mu$  where  $\mu$  is the average rate, the negative binomial model instead assumes that  $\text{var}(y) = \mu + \mu^2/\phi$ . The Poisson model is a special case of the negative binomial model when  $\phi \rightarrow \infty$ .

An alternative and equivalent formulation of the negative binomial that is commonly used makes the

connection to the Poisson model even more clear: instead of  $\theta_i$  and  $\phi$ , using  $\alpha$  and  $\beta$ ,

$$y \sim \text{NegativeBinomial}(\alpha, \beta), \text{ and}$$

$$\text{NegativeBinomial}(y|\alpha, \beta) = \int \text{Poisson}(y|\theta)\text{Gamma}(\theta|\alpha, \beta)d\theta.$$

Introductions to and additional exposition on the negative binomial model, especially in the context of modeling counts of events outcome data, are available in *Bayesian Data Analysis* and *Regression and Other Stories* (Gelman et al., 2020, 2013).

## Variables of Interest

Besides age, which is the greatest risk factor for COVID-19 mortality, we have included the following variables as covariates of interest:

- Population Density per Square Mile
- Median Household Income
- Proportion in Poverty
- the Index of Concentration at the Extremes for Racialized Economic Segregation
- Political Lean in the 2020 Election (1 = 100% Republican votes, -1 = 100% Democratic Votes)

**Index of Concentration at the Extremes for Racialized Economic Segregation** The Index of Concentration at the Extremes (ICE) is a measure which describes how concentrated a given area's population is in terms of the extreme ends of privilege and marginalization (Krieger et al., 2016). In general, the ICE measure is formulated as

$$\text{ICE} = \frac{\text{Number of People in Most Privileged Category} - \text{Number of People in Least Privileged Category}}{\text{Total Population}}$$

Applying the ICE approach to a specific context involves defining the axes of privilege of interest. In this case, data on racialized economic segregation are used from the US Census American Community Survey on White households earning more than \$100,000 a year (the most privileged group) or households of People of Color earning less than \$25,000 a year (the least privileged group). This variable is referred to throughout as the ICE for racialized economic segregation, or `ICEraceinc` in the code.

Compared with the Gini coefficient which is one of the most popular methods for summarizing area-level rates of inequities, the ICE the major advantage that it is more suitable for describing inequities at smaller area levels (Krieger et al., 2016). While the Gini coefficient measures within-area dissimilarity (as in, for example, how unequal wealth is distributed within a county), the ICE measure establishes where on a spectrum a given county's population falls allowing for comparison across counties. The Gini coefficient suffers from the fact that areas which are made up of relatively homogeneous populations will appear as having low within-area inequality (and therefore low Gini coefficient). Instead, the ICE measures how much of the population is either privileged or not. The Gini coefficient remains useful for reporting on the degree of inequity in larger areal units (like countries, states, and regions), but at smaller areal units (like counties, ZIP codes, census tracts) can be more difficult to interpret and compare. Maps of ICE measure can elucidate what spatial social segregation and polarization exist, and the ICE for racialized economic segregation has been repeatedly and significantly associated with COVID-19 outcomes (Brown et al., 2021; Chen & Krieger, 2021; Eichenbaum & Tate, 2022; Hanage et al., 2020; Krieger et al., 2022; Saha & Feldman, 2020; Sonderlund et al., 2022).

**Political Lean** Political lean has been associated with COVID-19 mortality in numerous studies, with plausible mechanisms explaining the association including differences in non-pharmaceutical intervention uptake (mask usage, social distancing, quarantining), differences in rates of vaccination, differences in political leadership's messaging, resource allocation, and the adoption of policy interventions (Gonzalez et al., 2021; Grossman et al., 2020; Kaashoek et al., 2021; Krieger et al., 2022; Leonhardt, 2021).

# Results

## Application 1: Non-Spatial Covariate Effects Over Time

The GAM models are fit with the following structure using the `gam` function from the `mgcv` package (S. Wood, 2022):

```
gam(  
  formula = deaths ~ s(median_age) + te(covariate, date, d=c(1,1)), # regression formula  
  offset = log(popsiz/1e5/12), # our offset represents the person-time  
  data = df,      # our data-set of county-level observations  
  family = nb(link='log') # indicates negative binomial family and a log-link function  
)
```

The formula used puts a one-dimensional smoothing spline to represent a nonlinear age-effect and a two-dimensional tensor-product smooth on the interaction between the given covariate and the date. The `d=c(1,1)` argument provides the instruction necessary to consider covariate and date as being on separate scales and therefore to fit the tensor-product smooth with anisotropy — that is, to allow for independent amounts of scaling in the dimensions of the covariate and time. The offset used structures the regression to estimate rates in units of person-time per 100,000 person-years. Since the death counts are aggregated to the monthly level, the person-time in units of 100,000 person-years are calculated by taking each county's population size, dividing by 100,000, and dividing by 12 (for the 12 months in a year).

The above structure is used to estimate models for our different `covariate` variables of interest: median income, percent in poverty, the ICE for racialized economic segregation, political lean. Whereas the other models treat median age as a variable to adjust for, the model presenting the effect of median age over time treats median age as the main covariate of interest including it as the `te(covariate, date, d=c(1,1))` and dropping the `s(median_age)` term which is redundant. Results of these models are summarized in 3 where the additional COVID-19 mortality associated with each covariate is visualized.

Likelihood ratio tests confirmed that models with covariates interacted with time had significantly lower residual deviance ( $p \leq 2.2e^{-16}$  for all models) compared to models only including a spline term on median age and the given covariate not interacted with time. The model interacting median age and time was compared to a model with a spline for median age alone. Akaike Information Criteria values were also lower for all models compared to models that did not interact the covariates with time.

Another non-spatial application of GAMs to describe the distribution of COVID-19 mortality in the US over time is to consider the associations of mortality with three-way interactions of time and two covariates taken together. In the following example, the formula used is `deaths ~ te(date, ICEraceinc, median_age, d=c(1,1,1))`. Again, the `d` argument specifying the marginal basis dimensions is used indicate that each of the `date`, `ICEraceinc`, and `median_age` measures are in different units and should not be smoothed assuming that a one unit difference in one variable is comparable to a one unit difference in another variable.

## Evaluating Spatial and Temporal Autocorrelation

## Regression Modeling Results

We used GAMs to describe the following:

- the changing associations between individual area-based measures and COVID-19 mortality over time;
- the changing associations between bivariate area-based measures and COVID-19 mortality over time;
- associations with area-based measures and COVID-19 mortality after controlling for spatio-temporal autocorrelation.

Additionally, as sensitivity analyses, random forest models were fit to compare with the coefficient estimates from the GAMs.

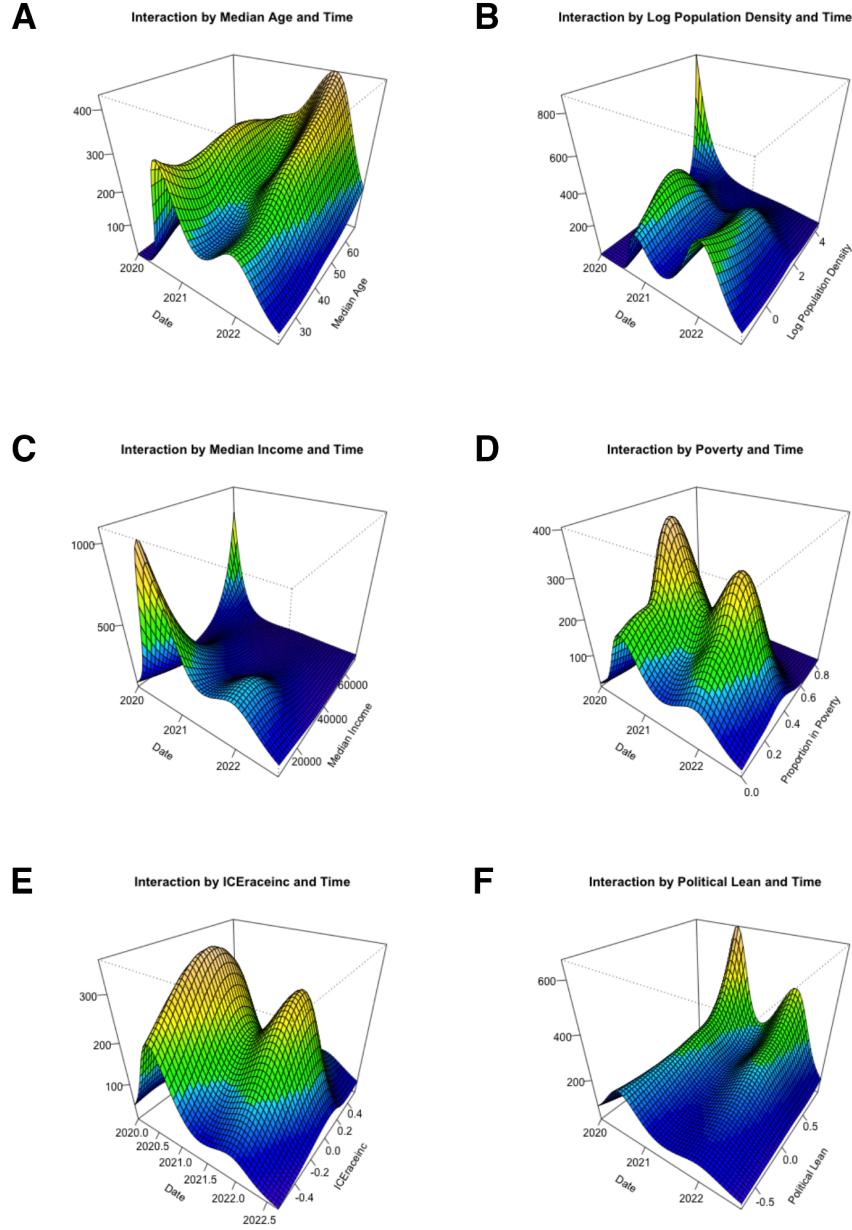


Figure 3: Additional COVID-19 mortality associated with covariates over time. A) age and time, B) log (base 10) population density and time, C) median income and time, D) proportion in poverty and time, E) ICEraceinc and time, F) political lean and time. In panels B-F the effect of age is marginalized out using the median age in the US, 38.8 (US Census Bureau, n.d.)

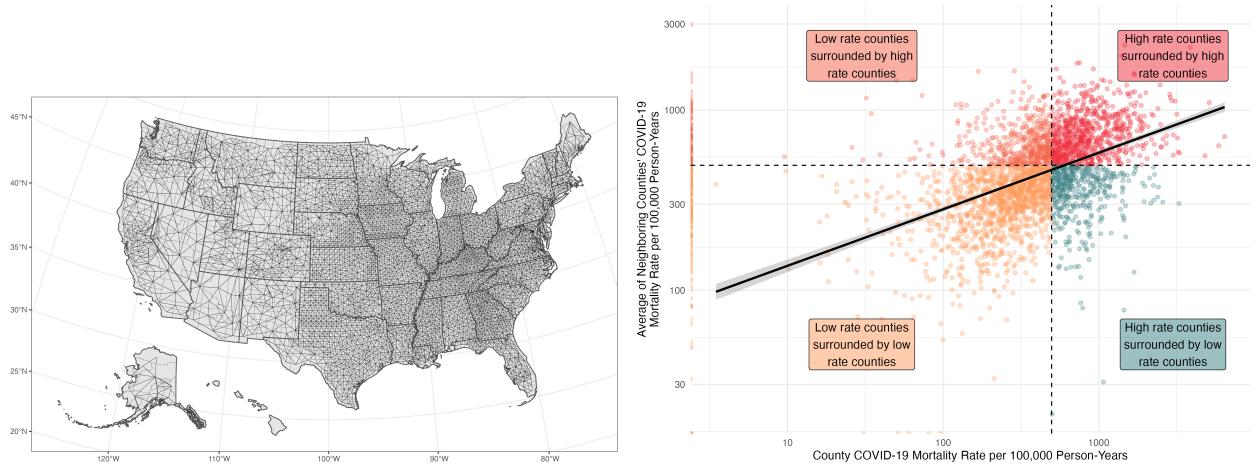
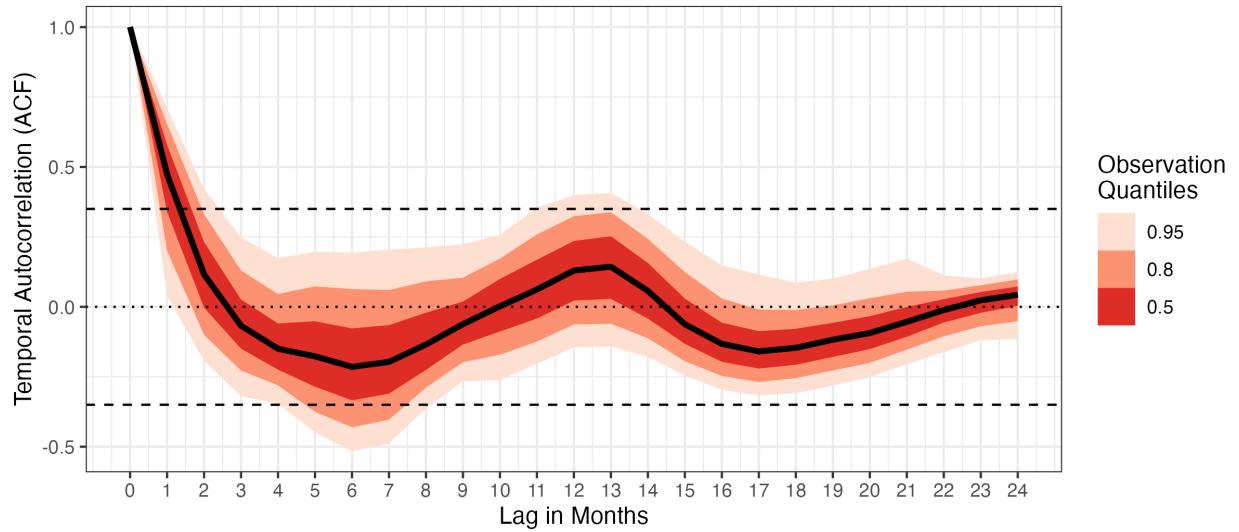
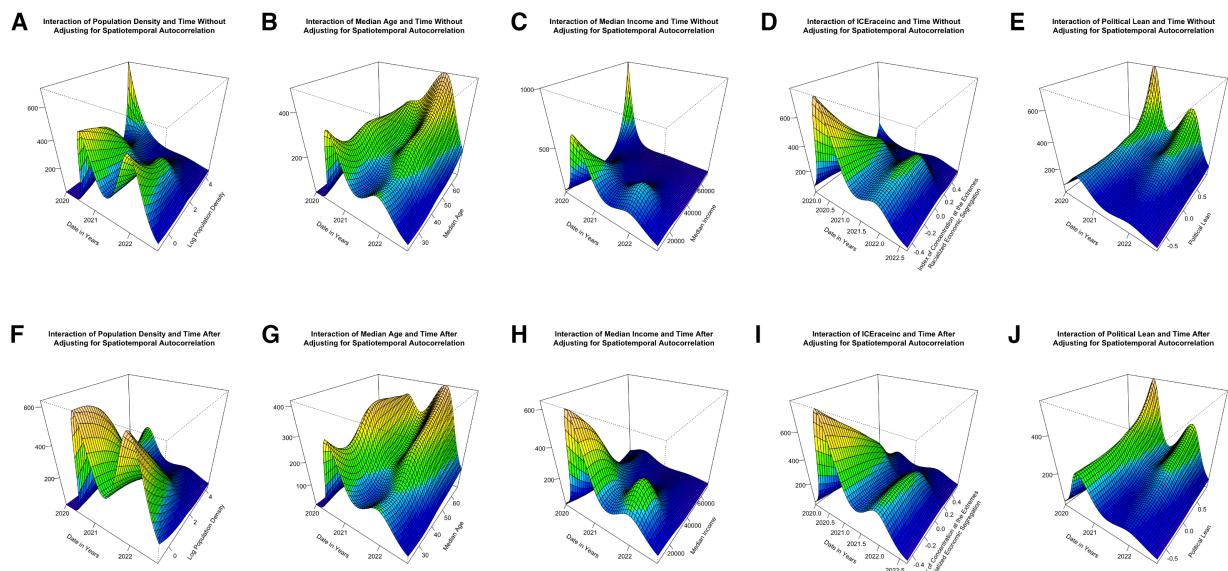
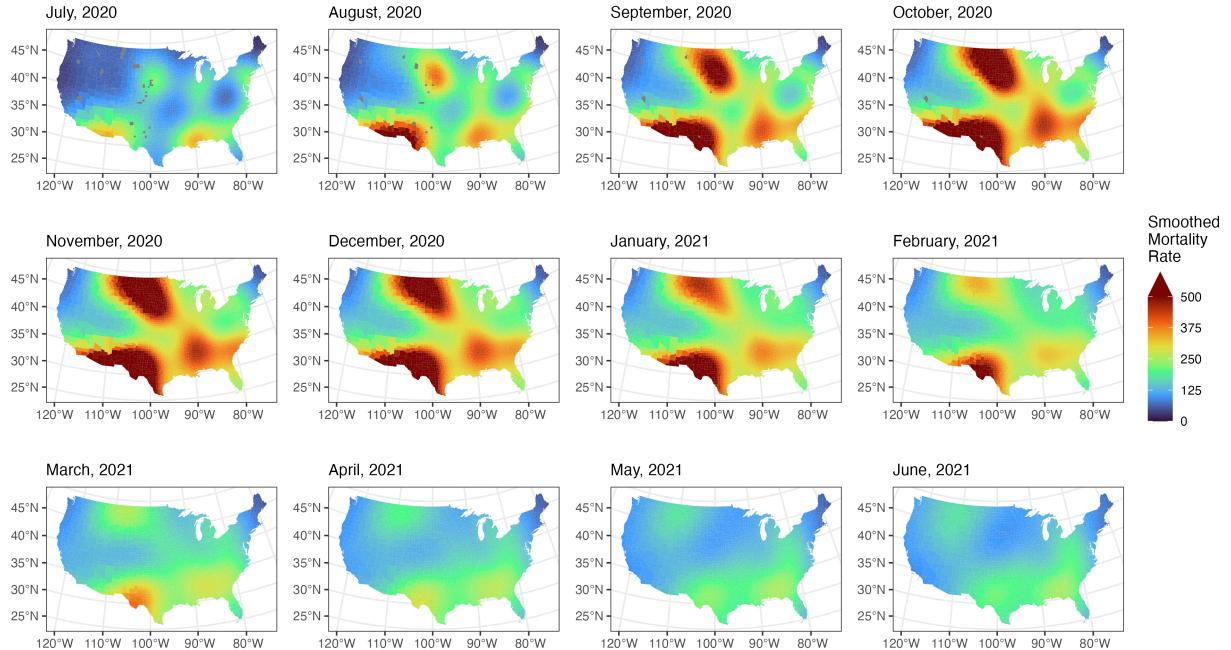


Figure 4: A) The spatial neighbor graph for counties in the US. In this graph, county centroids are connected if the two counties are adjacent to one another (i.e., share a boundary). This and similar graphs that represent neighbor relationships or geographic proximity are often used in modeling contexts to create spatial weights that account for which regions help to explain a given region's observed data.  
 (#fig:figure county neighboring graph)





# Discussion

## Limitations

- Ecological fallacy
- Age effects
- Fixed population denominators -> real-time denominator estimates
- Vaccination and mobility effects
- Equal Smoothing Across All Areas
- Soap Film Smoothers

# Conclusions

## References

- Bailey, Z. D., Feldman, J. M., & Bassett, M. T. (2021). How Structural Racism Works — Racist Policies as a Root Cause of U.S. Racial Health Inequities. *New England Journal of Medicine*, 384(8), 768–773. <https://doi.org/10.1056/NEJMms2025396>
- Bendix, A. (2022). Less than 4% of eligible people have gotten updated Covid booster shots, one month into the rollout. In *NBC News*. <https://www.nbcnews.com/health/health-news/updated-covid-booster-shots-doses-administered-cdc-rcna48960>
- Blendon, R. J., Schoen, C., DesRoches, C. M., Osborn, R., Scoles, K. L., & Zapert, K. (2002). Inequities In Health Care: A Five-Country Survey. *Health Affairs*, 21(3), 182–191. <https://doi.org/10.1377/hlthaff.21.3.182>
- Brown, K. M., Lewis, J. Y., & Davis, S. K. (2021). An ecological study of the association between neighborhood racial and economic residential segregation with COVID-19 vulnerability in the United States' capital city. *Annals of Epidemiology*, 59, 33–36. <https://doi.org/10.1016/j.anepidem.2021.04.003>
- Carpenter, E. (2021). “The Health System Just Wasn’t Built for Us”: Queer Cisgender Women and Gender Expansive Individuals’ Strategies for Navigating Reproductive Health Care. *Women’s Health Issues*, 31(5), 478–484. <https://doi.org/10.1016/j.whi.2021.06.004>
- CDC. (2021). Location, location, location – COVID Data Tracker Weekly Review. In *Centers for Disease Control and Prevention*. <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/past-reports/07162021.html>
- Chen, J. T., & Krieger, N. (2021). Revealing the Unequal Burden of COVID-19 by Income, Race/Ethnicity, and Household Crowding: US County Versus Zip Code Analyses. *Journal of Public Health Management and Practice*, 27(1), S43–S56. <https://doi.org/10.1097/PHH.0000000000001263>
- Chrisler, J. C., Barney, A., & Palatino, B. (2016). Ageism can be Hazardous to Women’s Health: Ageism, Sexism, and Stereotypes of Older Women in the Healthcare System. *Journal of Social Issues*, 72(1), 86–104. <https://doi.org/10.1111/josi.12157>
- Coven, J., & Gupta, A. (2020). *Disparities in Mobility Responses to COVID-19*. <https://arpitgupta.info/s/DemographicCovid.pdf>
- Del Rios, M., Chomilo, N. T., & Lewis, N. A. (2022). Covid-19 is an inverse equity story, not a racial equity success story. In *STAT*. <https://www.statnews.com/2022/10/25/covid-19-inverse-equity-story-not-racial-equity-success-story/>
- Donovan, D. (2022). U.S. Officially Surpasses 1 Million COVID-19 Deaths. In *Johns Hopkins Coronavirus Resource Center*. <https://coronavirus.jhu.edu/from-our-experts/u-s-officially-surpasses-1-million-covid-19-deaths>
- Eichenbaum, A., & Tate, A. D. (2022). Health Inequity in Georgia During the COVID-19 Pandemic: An Ecological Analysis Assessing the Relationship Between County-Level Racial/Ethnic and Economic Polarization Using the ICE and SARS-CoV-2 Cases, Hospitalizations, and Deaths in Georgia as of October 2020. *Health Equity*, 6(1), 230–239. <https://doi.org/10.1089/heq.2021.0118>
- Feldman, J. L., Luhur, W. E., Herman, J. L., Poteat, T., & Meyer, I. H. (2021). Health and health care access in the US transgender population health (TransPop) survey. *Andrology*, 9(6), 1707–1718. <https://doi.org/10.1111/andr.13052>

- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). Bayesian Data Analysis. In *Routledge & CRC Press*. <https://www.routledge.com/Bayesian-Data-Analysis/Gelman-Carlin-Stern-Dunson-Vehtari-Rubin/p/book/9781439840955>
- Gelman, A., Hill, J., & Vehtari, A. (2020). *Regression and Other Stories* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781139161879>
- Glenza, J. (2020). Covid cases increase across US as upper midwest sees rapid rise. *The Guardian*. <https://www.theguardian.com/world/2020/oct/22/covid-cases-coronavirus-us-midwest>
- Goldstein, J. (2022). During the Omicron surge, Black New Yorkers were hospitalized at a rate more than twice that of white residents. *The New York Times*. <https://www.nytimes.com/2022/03/03/nyregion/omicron-hospitalizations-new-york-city.html>
- Gonzalez, K. E., James, R., Bjorklund, E. T., & Hill, T. D. (2021). Conservatism and infrequent mask usage: A study of US counties during the novel coronavirus (COVID-19) pandemic. *Social Science Quarterly*, 102(5), 2368–2382. <https://doi.org/10.1111/ssqu.13025>
- Grossman, G., Kim, S., Rexer, J. M., & Thirumurthy, H. (2020). Political partisanship influences behavioral responses to governors' recommendations for COVID-19 prevention in the United States. *Proceedings of the National Academy of Sciences*, 117(39), 24144–24153. <https://doi.org/10.1073/pnas.2007835117>
- Hanage, W. P., Testa, C., Chen, J. T., Davis, L., Pechter, E., Seminario, P., Santillana, M., & Krieger, N. (2020). COVID-19: US federal accountability for entry, spread, and inequities—lessons for the future. *European Journal of Epidemiology*, 35(11), 995–1006. <https://doi.org/10.1007/s10654-020-00689-2>
- Hernandez, I., Dickson, S., Tang, S., Gabriel, N., Berenbrok, L. A., & Guo, J. (2022). Disparities in distribution of COVID-19 vaccines across US counties: A geographic information system-based cross-sectional study. *PLOS Medicine*, 19(7), e1004069. <https://doi.org/10.1371/journal.pmed.1004069>
- Hill, L., & Artiga, S. (2022). COVID-19 Cases and Deaths by Race/Ethnicity: Current Data and Changes Over Time. In KFF. <https://www.kff.org/coronavirus-covid-19/issue-brief/covid-19-cases-and-deaths-by-race-ethnicity-current-data-and-changes-over-time/>
- Johnson, A., & Keating, D. (2022). Whites now more likely to die from covid than Blacks: Why the pandemic shifted. In *Washington Post*. <https://www.washingtonpost.com/health/2022/10/19/covid-deaths-us-race/>
- Kaashoek, J., Testa, C., Chen, J., Stolerman, L., Krieger, N., Hanage, W. P., & Santillana, M. (2021). *The Evolving Roles of US Political Partisanship and Social Vulnerability in the COVID-19 Pandemic from February 2020 - February 2021* [{SSRN} {Scholarly} {Paper}]. <https://doi.org/10.2139/ssrn.3933453>
- Kim, E. J., Marrast, L., & Conigliaro, J. (2020). COVID-19: Magnifying the Effect of Health Disparities. *Journal of General Internal Medicine*, 35(8), 2441–2442. <https://doi.org/10.1007/s11606-020-05881-4>
- Krieger, N., Testa, C., Chen, J. T., Hanage, W. P., & McGregor, A. J. (2022). Relationship of political ideology of US federal and state elected officials and key COVID pandemic outcomes following vaccine rollout to adults: April 2021–March 2022. *The Lancet Regional Health – Americas*, 16. <https://doi.org/10.1016/j.lana.2022.100384>
- Krieger, N., Waterman, P. D., Spasojevic, J., Li, W., Maduro, G., & Van Wye, G. (2016). Public Health Monitoring of Privilege and Deprivation With the Index of Concentration at the Extremes. *American Journal of Public Health*, 106(2), 256–263. <https://doi.org/10.2105/AJPH.2015.302955>
- Lambert, J. (2022). Most Americans Don't Know About the Omicron Covid Boosters. In *Grid News*. <https://www.grid.news/story/science/2022/10/04/most-americans-dont-know-about-the-omicron-covid-boosters-that-spells-trouble-for-the-coming-winter/>
- Leonhardt, D. (2021). Red Covid. *The New York Times*. <https://www.nytimes.com/2021/09/27/briefing/covid-red-states-vaccinations.html>
- Mayo Clinic. (2022). History of COVID-19: Outbreaks and Vaccine Timeline. In *MayoClinic.org*. <https://www.mayoclinic.org/coronavirus-covid-19/history-disease-outbreaks-vaccine-timeline/covid-19>
- MIT Election Data and Science Lab. (2022). *County Presidential Election Returns 2000-2020*. Harvard Dataverse. <https://doi.org/10.7910/DVN/VOQCHQ>
- Mollalo, A., Vahedi, B., & Rivera, K. M. (2020). GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Science of The Total Environment*, 728, 138884. <https://doi.org/10.1016/j.scitotenv.2020.138884>
- Murthy, B. P. (2021). Disparities in COVID-19 Vaccination Coverage Between Urban and Rural Counties — United States, December 14, 2020–April 10, 2021. *MMWR. Morbidity and Mortality Weekly Report*, 70. <https://doi.org/10.15585/mmwr.mm7020e3>

- Office of the Commissioner. (2022). Coronavirus (COVID-19) Update: FDA Authorizes Moderna, Pfizer-BioNTech Bivalent COVID-19 Vaccines for Use as a Booster Dose. In *FDA*. <https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-moderna-pfizer-biontech-bivalent-covid-19-vaccines-use>
- Okonkwo, N. E., Aguwa, U. T., Jang, M., Barré, I. A., Page, K. R., Sullivan, P. S., Beyrer, C., & Baral, S. (2021). COVID-19 and the US response: Accelerating health inequities. *BMJ Evidence-Based Medicine*, 26(4), 176–179. <https://doi.org/10.1136/bmjebm-2020-111426>
- Ortega, A. N., & Roby, D. H. (2021). Ending Structural Racism in the US Health Care System to Eliminate Health Care Inequities. *JAMA*, 326(7), 613–615. <https://doi.org/10.1001/jama.2021.11160>
- Park, Y. M., Kearney, G. D., Wall, B., Jones, K., Howard, R. J., & Hylock, R. H. (2021). COVID-19 Deaths in the United States: Shifts in Hot Spots over the Three Phases of the Pandemic and the Spatiotemporally Varying Impact of Pandemic Vulnerability. *International Journal of Environmental Research and Public Health*, 18(17), 8987. <https://doi.org/10.3390/ijerph18178987>
- Press, V. G., Huisings-Scheetz, M., & Arora, V. M. (2021). Inequities in Technology Contribute to Disparities in COVID-19 Vaccine Distribution. *JAMA Health Forum*, 2(3), e210264. <https://doi.org/10.1001/jamahealthforum.2021.0264>
- Rader, B., Astley, C. M., Sewalk, K., Delamater, P. L., Cordiano, K., Wronski, L., Rivera, J. M., Hallberg, K., Pera, M. F., Cantor, J., Whaley, C. M., Bravata, D. M., & Brownstein, J. S. (2021). *Spatial Accessibility Modeling of Vaccine Deserts as Barriers to Controlling SARS-CoV-2*. medRxiv. <https://doi.org/10.1101/2021.06.09.21252858>
- Rapp, K. S., Volpe, V. V., Hale, T. L., & Quartararo, D. F. (2022). State-Level Sexism and Gender Disparities in Health Care Access and Quality in the United States. *Journal of Health and Social Behavior*, 63(1), 2–18. <https://doi.org/10.1177/00221465211058153>
- Saha, S., & Feldman, J. M. (2020). *Neighborhood-level Racial/Ethnic and Economic Inequities in COVID-19 Burden Within Urban Areas in the US and Canada*. medRxiv. <https://doi.org/10.1101/2020.12.07.20241018>
- Scott, D. (2020). These 4 Midwestern states are seeing worrying Covid-19 spikes. In *Vox*. <https://www.vox.com/2020/9/2/21418812/covid-19-coronavirus-us-cases-midwest-surge>
- Shumaker, L., & Wu, T. (2020). COVID-19 cases surge in U.S. Midwest and Northeast. *Reuters*. <https://www.reuters.com/article/us-health-coronavirus-usa-trends-idUSKBN26Q30B>
- Sigler, T., Mahmuda, S., Kimpton, A., Loginova, J., Wohland, P., Charles-Edwards, E., & Corcoran, J. (2021). The socio-spatial determinants of COVID-19 diffusion: The impact of globalisation, settlement characteristics and population. *Globalization and Health*, 17(1), 56. <https://doi.org/10.1186/s12992-021-00707-2>
- Sonderlund, A. L., Charifson, M., Schoenthaler, A., Carson, T., & Williams, N. J. (2022). Racialized economic segregation and health outcomes: A systematic review of studies that use the Index of Concentration at the Extremes for race, income, and their interaction. *PLOS ONE*, 17(1), e0262962. <https://doi.org/10.1371/journal.pone.0262962>
- Sugg, M. M., Spaulding, T. J., Lane, S. J., Runkle, J. D., Harden, S. R., Hege, A., & Iyer, L. S. (2021). Mapping community-level determinants of COVID-19 transmission in nursing homes: A multi-scale approach. *Science of The Total Environment*, 752, 141946. <https://doi.org/10.1016/j.scitotenv.2020.141946>
- The New York Times. (2021). *Coronavirus (Covid-19) Data in the United States*. The New York Times. <https://github.com/nytimes/covid-19-data>
- The White House. (2021). Executive Order On Advancing Racial Equity and Support for Underserved Communities Through the Federal Government. In *The White House*. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/>
- The White House. (2022). Advancing Equity and Racial Justice Through the Federal Government. In *The White House*. <https://www.whitehouse.gov/equity/>
- Thompson, C. N. (2020). COVID-19 Outbreak — New York City, February 29–June 1, 2020. *MMWR. Morbidity and Mortality Weekly Report*, 69. <https://doi.org/10.15585/mmwr.mm6946a2>
- United States COVID - Coronavirus Statistics - Worldometer. (2022). In *Worldometers*. <https://www.worldometers.info/coronavirus/country/us/>
- U.S. Department of Commerce Economics and Statistics Administration, & U.S. Census Bureau. (2000).

- Census Regions and Divisions of the United States.* [https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us\\_regdiv.pdf](https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf)
- US Census Bureau. (n.d.). Nation Continues to Age as It Becomes More Diverse. In *Census.gov*. Retrieved November 23, 2022, from <https://www.census.gov/newsroom/press-releases/2022/population-estimates-characteristics.html>
- US Census Bureau. (2021). 2020 Census: Redistricting File (Public Law 94-171) Dataset. In *Census.gov*. <https://www.census.gov/data/datasets/2020/dec/2020-census-redistricting-summary-file-dataset.html>
- US Census Bureau. (2020). American Community Survey 2015-2019 5-Year Data Release. In *Census.gov*. <https://www.census.gov/newsroom/press-kits/2020/acs-5-year.html>
- Walker, K., & Herman, M. (2022). *Tidycensus: Load US Census Boundary and Attribute Data as 'tidyverse' and 'sf'-Ready Data Frames*.
- Whitehead, J., Shaver, J., & Stephenson, R. (2016). Outness, Stigma, and Primary Health Care Utilization among Rural LGBT Populations. *PLOS ONE*, 11(1), e0146139. <https://doi.org/10.1371/journal.pone.0146139>
- Whittle, R. S., & Diaz-Artilles, A. (2020). An ecological study of socioeconomic predictors in detection of COVID-19 cases across neighborhoods in New York City. *BMC Medicine*, 18(1), 271. <https://doi.org/10.1186/s12916-020-01731-6>
- Wood, S. (2022). *Mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation*. <https://CRAN.R-project.org/package=mgcv>
- Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R* (2nd ed.). Chapman & Hall/CRC Texts in Statistical Science. <https://www.routledge.com/Generalized-Additive-Models-An-Introduction-with-R-Second-Edition/Wood/p/book/9781498728331>