(2024) 54, 27–52

The Review of Regional Studies

The Official Journal of the Southern Regional Science Association

Social Vulnerability and COVID-19 Pandemic Outcomes: Evidence from Spatial Quantile Regression*

Jim Lee[a] and Yuxia Huang[b]

[a]College of Business, Texas A&M University-Corpus Christi, USA

[b]College of Engineering and Computer Science, Texas A&M University-Corpus Christi, USA

Abstract: This paper evaluates the empirical relevance of a prominent social vulnerability measure to local economic and public health outcomes at two distinct stages of the COVID-19 pandemic. In contrast to traditional least-squares regression, conditional quantile regression offers stronger evidence in support of social vulnerability for predicting the pandemic's uneven economic devastation and death tolls across the United States. An extension of the spatial autoregressive geographically weighted regression model to conditional quantile regression reveals the varying role of social vulnerability within and across broad regions that were exposed to different levels of pandemic impacts.

Keywords: social vulnerability, COVID-19 pandemic, conditional quantile regression, spatial dependence, geographically weighted autoregressive model

JEL Codes: R1, C5, E2, I1

1. INTRODUCTION

Social vulnerability has long been considered a key factor that affects how diverse communities respond to disruptive events like natural disasters and disease outbreaks. Social vulnerability commonly refers to the inability of certain sociodemographic groups to withstand or adapt to an external shock's adverse impacts (Bakkensen et al., 2017). The voluminous literature on social vulnerability and its related concept of community resilience has spawned a proliferation of composite, quantitative indicators for individual communities as summary measures of the high-dimensional nature of inherent sociodemographic attributes (Fatemi et al., 2017). Yet it remains unclear whether those social vulnerability models align with post-disaster outcomes (Bakkensen et al., 2017; Rufat et al., 2019; Spielman et al., 2020). This paper aims to fill this knowledge gap.

*Jim Lee is Regents Professor of Economics at Texas A&M University-Corpus Christi, USA. Yuxia Huang is an Associate Professor of Spatial Science at Texas A&M University-Corpus Christi, USA. *Corresponding Author:* Jim Lee, E-mail: jim.lee@tamucc.edu

ISSN 1553-0892, 0048-49X (online)
www.srsa.org/rrs

Our study focuses on the Social Vulnerability Index (SVI) constructed by the Centers for Disease Control and Prevention (CDC) and the Agency for Toxic Substances and Disease Registry (ATSDR). Most social vulnerability or resilience indices are geared toward a wide range of disasters, such as tropical storms and wildfires. The CDC/ATSDR instead constructed its SVI and the constituent indicators with a focus on providing "socially and spatially relevant information to help public health officials and local planners better prepare communities to respond to emergency events such as ... disease outbreaks." (ATSDR, 2022).

The COVID-19 pandemic provided a natural testing ground for exploring the empirical validity of social vulnerability constructs. Disease outbreaks were widespread across the United States. In early 2020, socially vulnerable communities, particularly those with disproportionately large ethnic-minority populations, experienced high COVID-19 incidence and death rates (Neelon et al., 2022). Over the next two years, the nation continued to undergo several waves of outbreaks while its overall economy steadily returned to its pre-pandemic conditions. Like the historic economic recession during the early months of the pandemic, the pace of subsequent recovery varied widely across local communities and broad regions. Case studies have found that natural disasters disproportionately impacted socially vulnerable populations (Bakkensen et al., 2017; Park and Xu, 2020). In the context of the COVID-19 pandemic, the question we seek to address is: How well do pre-existing social vulnerability characteristics, as captured by the CDC SVI, explain disparities in local economic and public health outcomes?

To evaluate the empirical validity of social vulnerability measures, the typical approach is to perform regression analysis that estimates the extent to which social vulnerability predicts disaster outcomes (e.g., Bakkensen et al., 2017; Rufat et al., 2019). However, standard regression methods, such as ordinary least-squares (OLS), suffer two shortcomings that are addressed in this study. First, regression results ignore spatial effects over different geographic locations. Spatial analysis in the context of spatial dependence and spatial heterogeneity is common in modeling social vulnerability characteristics (e.g., Cutter et al., 2014; Cutter and Derakhsan, 2020). Drawing from the first law of geography, spatial dependence, or autocorrelation, occurs when the observations of nearby locations tend to be similar. Interactions between the labor market and the economy of a community and those of its neighbors tend to intensify in the wake of a disaster (Lee, 2021). For instance, Belasen and Polachek (2009) described how counties in the state of Florida hit directly by hurricanes affected the labor supply and wages in their surrounding counties. As people commuted to work during the pandemic, neighboring areas tended to share the impacts of COVID-19 outbreaks.

Another aspect of spatial nonstationarity is unobserved spatial heterogeneity, which describes the uneven distribution of model relationships over different geographic locations. Spatial heterogeneity is particularly relevant to social vulnerability conditions, which exhibit remarkable disparities across U.S. regions (Cutter et al., 2014, 2016; Park and Xu, 2020). Local regulations and public services, such as school closures and online class delivery, also vary among U.S. communities. Omitting the presence of spatial effects would lead to biased and inconsistent model estimates and spurious inferences (LeSage and Pace, 2009).

The second shortcoming of OLS arises from a lack of flexibility in estimating model relationships beyond the mean of the dependent or outcome variable. The literature has

documented the nonlinear nature of disaster damage in the sense that physical and economic losses tend to be disproportionately larger at the top end of the disaster intensity than its lower end (Schumacher and Strobl, 2011). This may also bring about outliers in disaster outcomes that potentially distort OLS estimation results. As shown in this study, the association between sociodemographic attributes and disaster outcomes might differ between communities severely impacted by the pandemic and those less exposed to COVID-19 outbreaks, such as rural regions in the U.S. Midwest. As such, the OLS estimator that is restricted to the conditional mean is not informative about model relationships at other points of the conditional distribution of the dependent variable. This might explain the mixed findings in studies (e.g., Bakkensen et al., 2017; Rufat et al., 2019; Lee, 2023) that empirically evaluate social vulnerability indices, including the CDC SVI, in the context of historical disasters.

We overcome the above two drawbacks by incorporating spatial dependence and spatial heterogeneity simultaneously with the consideration of varying model relationships across the entire conditional distribution of the outcome variable. The approach essentially extends the spatial autoregressive geographically weighted regression (SAR-GWR) model (Geniaux and Martinetti, 2018) to a conditional quantile regression (QR) setting. A spatial autoregression model incorporates spatial interactions in the dependent variable, whereas geographically weighted regression relaxes the assumption of "global" or constant model relationships for the entire study area and estimates those relationships "locally" at different geographic locations. The spatial autoregressive geographically weighted quantile regression (SAR-GWQR) helps us explore the uneven geography of the COVID-19 pandemic impacts across the United States by allowing model parameters to vary over the conditional distribution of disaster outcomes in addition to over geographic space (Tomal and Helbich, 2023).

2. DATA AND METHODS

2.1. Data Description

2.1.1. Outcome Variables

This study evaluates the role of social vulnerability in explaining disparities in local economic and public health outcomes across the United States during the COVID-19 pandemic. Motivated by the finding of the pandemic's evolving impact on communities of different social vulnerability characteristics (Neelon et al., 2021), we consider outcomes in two periods. The first period is April 2020, at the depths of a nationwide recession induced largely by stay-at-home orders and business lockdowns during the first wave of COVID-19 outbreaks. The second period is June 2022, when outbreaks subsided across the U.S., and the overall national economy returned to its pre-pandemic level.

We employ county-level data to compare local communities. Following Rose (2021) and Walmsley et al. (2021), we characterize local economic losses and subsequent recovery alternatively by employment and output. The employment levels of the nation's 3,142 counties and county equivalents capture local labor market conditions. The data are sourced from the Bureau of Labor Statistics' Local Area Unemployment Statistics (LAUS) program. Seasonal

factors are removed using the U.S. Bureau of Census X-11 procedure.

The total size of economic activity, or output, is another metric indicative of how well or poorly an economy performs, but such a measure is not typically available at the county level. As a proxy, we leverage the County Economic Impact Index (CEII) developed by the Department of Energy's Argonne National Laboratory (Smith et al., 2021). The CEII is equivalent to the relative measure of the total value added of all industries, or Gross Domestic Product, within a county. The monthly CEII data are seasonally adjusted at the source.

Data for the two economic outcomes are expressed as percentage deviations from the January 2020 levels as the pre-pandemic baselines. For instance, a positive value for employment indicates the percentage of a given month's employment level that was below the January 2020 baseline, and a negative value indicates the percentage that was above the baseline. Essentially, the employment and output data in April 2020 and June 2022 depict COVID-19's adverse impact on local economies, respectively, at the early and late stages of the pandemic. In April 2020, employment losses across U.S. counties averaged about 11% with a range between 7% and 41%. By June 2022, a typical county had restored most of the employment lost during the early months of the pandemic (i.e., 0% median), but the employment levels of 1,372 counties (44%) remained below their pre-pandemic levels. The overall geographic distribution of the CEII data across the nation is comparable to that of the employment data.

For public health outcomes, we consider the number of confirmed coronavirus cases and deaths per 1,000 persons in each county. The number of confirmed cases tracks the spread of the virus within a county, whereas the number of reported COVID-19 deaths represents the ultimate impact on public health, especially in populations at higher health risk. The data are monthly cumulative COVID-19 incidence and death data obtained from the New York Times database. The cumulative public health data for April 2020 and June 2022 allow us to compare how individual counties were susceptible to the public health impacts of COVID-19 outbreaks at the onset of the pandemic with their susceptibility over nearly the entire pandemic.

To facilitate comparisons of data in different scales, all dependent and explanatory variables in model regressions are standardized into z-scores. Table 1 displays their descriptive statistics. All variables have a zero mean and a standard deviation equal to one due to standardization. However, specifically for the dependent variables, the 0.25 ($\tau = 0.25$) and 0.75 ($\tau = 0.75$) quantiles, along with the minimum and maximum values, reveal remarkable variability around the median. Cross-sectional variation, as captured by the ranges between the 0.25 and 0.75 quantiles and between extreme values, reduces for the employment variable from April 2020 to June 2022. The output variable also becomes less dispersed during the latter period, but its maximum value is much higher, reflecting exceptionally strong economic recovery among a few counties. By contrast, the two quantiles of the COVID case and death variables remain in relatively narrow ranges by the end of the pandemic.

The county-level outcome data are further presented as choropleth maps in Figure A1 of the Appendix. In April 2020, local employment and output were more exposed to the COVID-19 pandemic in the Great Lakes and Northeastern region, the state of Florida, and on the West Coast (blue color). Economic losses were relatively modest among counties

Table 1: Descriptive Statistics for Regression Variables

	Mean	**Std.Dev.**	**Min.**	**τ=0.25**	**Median**	**τ=0.75**	**Max.**
SVI	0.000	1.000	-2.819	-0.724	0.019	0.723	2.607
Pop. Density	0.000	1.000	-3.958	-0.556	0.005	0.551	4.133
				April 2020 Regressions			
Employment	0.000	1.000	-3.082	-0.677	-0.010	0.592	5.261
Output	0.000	1.000	-4.258	-0.596	-0.011	0.593	8.975
COVID Cases	0.000	1.000	-5.488	-0.261	0.072	0.589	2.502
COVID Deaths	0.000	1.000	-2.94	-0.873	-0.148	0.877	2.464
Policy Stringency	0.000	1.000	-1.770	-0.580	-0.207	0.238	2.951
Policy Support	0.000	1.000	-1.882	-0.630	-0.213	0.738	2.639
				June 2022 Regressions			
Employment	0.000	1.000	-26.502	-0.438	0.045	0.473	3.895
Output	0.000	1.000	-3.575	-0.574	0.053	0.577	18.093
COVID Cases	0.000	1.000	-2.854	-0.101	0.176	0.824	2.760
COVID Deaths	0.000	1.000	-1.602	-0.237	0.165	0.952	2.601
Policy Stringency	0.000	1.000	-2.449	-0.648	0.019	0.505	3.825
Policy Support	0.000	1.000	-2.159	-0.641	-0.055	0.277	2.530

Notes: The sample consists of observations for 3,142 U.S. counties and county equivalents. The original employment and output data are expressed relative to the January 2020 levels. The SVI data are obtained from the CDC/ATSDR 2020 database. The original data of population density is the log level of residents per square mile of county area. The original data for COVID cases and deaths are log levels of cumulative numbers per 1,000 persons through the specific regression month. The original data for policy stringency and policy support are individual states' OxCGRT policy indices that are averaged over the period between January 2020 and the specific regression month. The data of all variables are standardized into z-scores.

in the Great Plains (e.g., Montana, Wyoming, Colorado, North Dakota, South Dakota, Nebraska, and Kansas) and Midwestern regions (red color). For economic outcomes in June 2022, the data appear more scattered than those in the earlier period, especially in the Great Plains region. In addition, most counties in the state of Florida performed relatively better in mid-2022 than they did about two years earlier.

The geographic distributions of the four sets of public health data across broad U.S. regions appear to align with the economic data in some regions but not others. The COVID-19 case rates and death rates tended to be higher along much of the East Coast and parts of the Southwestern region. Except for a few regions, such as the New England area in the Northeast, the cumulative death rates differed remarkably between the two periods. For instance, the death rates in parts of the Midwest were relatively higher over the entire course of the pandemic than during the first wave of COVID-19 outbreaks.

2.1.2. Explanatory Variables

Given the COVID-19 pandemic as our event study, the measure of social vulnerability draws from the CDC county-level database released in 2020. The CDC SVI is a composite measure of the summed percentile rankings of 15 variables based on Census data (Flanagan et al., 2011). The constituent components represent a broad range of sociodemographic factors, which are grouped into four themes: a) socioeconomic status (population living below poverty; unemployed workers; per capita income; population without a high school diploma), b) household composition and disability (population aged 65 and older; population aged 17 and younger; population living with a disability; single-parent households), c) race/ethnicity and language proficiency (racial/ethnic minority population; people with limited English proficiency), and d) housing type and mobility (multi-unit housing; mobile homes; crowed housing; households without a vehicle; population living in group quarters).

The role of sociodemographic conditions in determining a community's vulnerability to disasters is well documented (e.g., Cutter et al., 2014). Some recent studies question the empirical validity of their quantitative measures (e.g., Rufat et al., 2019). In the context of the COVID-19 pandemic, the incidence of virus caseload and excess mortality was found to be associated with the presence of certain demographic groups, such as elderly people (Rodriguez-Pose and Burlina, 2021; Ramirez et al., 2022). To supplement the descriptive statistics in Table 1, Figure A2 in the Appendix shows that socially vulnerable counties are more concentrated in the South and Southwest, especially rural areas in those regions (blue color). They are less likely to be found in the Northeast and Midwest.

Other than the CDC SVI as the primary predictor variable of interest, our regression analyses account for confounding factors that might have also affected local economic and public health outcomes over the course of the pandemic.[1] The first control variable is population density (the number of residents per square mile of county area). During the early months of the pandemic, COVID-19 outbreaks tended to intensify in more populous city centers than in suburbs or rural areas (Carozzi et al., 2024; Ramirez et al., 2022). Figure A2 in the Appendix indicates relatively lower population density in the western half of the United States, except the West Coast.

The empirical model also accounts for government policy measures and interventions in response to COVID-19 outbreaks and their impact on the local economy. The diversity of government responses makes it difficult to measure policy outcomes. We, nevertheless, draw on the monthly data of two composite policy measures of Oxford COVID-19 Government Response Tracker (OxCGRT) for U.S. individual states (Hallas et al., 2021). State governments' responses to COVID-19 were likely to be exogenous to county-level conditions. The first is the policy stringency index, which measures the strictness of closure and COVID-19 containment measures, such as stay-at-home orders, lockdown restrictions, and closures for schools and workplaces. A higher value represents stricter policy measures. Famiglietti and Leibovici (2021) showed that states with higher policy stringency scores during the early months of the pandemic experienced relatively higher unemployment. Dergiades et al.

[1]We have also considered other controlling factors, such as remote work (Althoff et al., 2022) and economic diversification (Coulson et al., 2020). These variables are not statistically significant in preliminary analysis and so are excluded from final regression results.

(2022) reported the effectiveness of government interventions for reducing coronavirus-related deaths.

The second policy measure captures the extent of economic support, such as financial aid to people who lost their jobs and debt relief for households. Government financial aid could arguably alleviate the devastation of the pandemic on communities and their economies (Edelberg and Sheiner, 2021). A higher index value represents more economic support. The OxCGRT policy data series begins in January 2020. To account for lagged policy effects, we take the monthly average of the indices between January and April 2020 in the model for outcomes in April 2020. Similarly, for analyzing outcomes in June 2022, we take the average of the 30 months between January 2020 and June 2022.

Figure A2 in the Appendix displays the state policy data of the two periods. State policy measures that aimed to contain COVID-19 outbreaks were the strictest throughout the pandemic in the state of California on the West Coast, the Great Lakes, and the Northeastern regions (blue color). Containment measures were especially lax among states in the Southeast, such as Florida and Alabama, and the Great Plains, such as North Dakota, South Dakota, and Kansas (red color). By comparison, the relative extent of economic support across states was less consistent over time than policy stringency. Only the states along the West Coast (i.e., California, Oregon, and Washington) and in the lower portion of the Northeast (e.g., New York, New Jersey, and Pennsylvania) continued to provide the most financial relief to households within their states. Changes in state-level policy measures during the pandemic affect the distributions of the two policy variables, as shown in Table 1.

2.2. Empirical Models

This section outlines the workhorse of our empirical study of the extent to which the CDC's social vulnerability index explains disaster outcomes in the COVID-19 pandemic. A standard regression model for a cross-section of n geographic locations, or counties in our case, can be written as:

$$y_i = \sum_k \beta_k x_{ik} + \varepsilon_i \qquad i= 1, \ldots, n;\ k = 1, \ldots, p{+}1 \tag{1}$$

where y_i is the dependent variable, β_k is a vector of parameters, x_{ik} denotes an explanatory or predictor variable in the covariate set $X_i{=}(1,\ x_{i1},\ \ x_{i2}, \ldots,\ x_{ip})^{\mathrm{T}}$, and ε_i is an error term that is assumed to be normally distributed in OLS regression. If spatial dependence exists among data observations with a spatial dimension, then the OLS assumption about the error term is violated. A spatial autoregressive (SAR) model that accounts for the presence of spatial autocorrelation, a form of spatial dependence, can be expressed as:

$$y_i = \rho \sum_j w_{ij} y_j + \sum_k \beta_k x_{ik} + \varepsilon_i \tag{2}$$

where ρ is the spatial autoregressive parameter, and w_{ij} is an element of a spatial weight matrix W, which quantifies interactions between nearby locations with the first-order queen

contiguity-based spatial weights. The matrix W is row-standardized so that $\sum_j w_{ij} = 1$. The key feature of SAR is the spatially lagged dependent variable, Wy, which is a weighted average of all neighbors of location i.

LeSage and Pace (2009) indicated that the SAR can conceptually provide coefficient estimates for covariates in the model at every location as $(I_n - \rho W)^{-1}\beta_k$, which is a full $n \times n$ matrix. However, to simplify interpretations, it is customary in the SAR literature to apply a scalar specification, which assumes that, like OLS, model parameters are the same for all geographic locations. The "total" effect of x_{ik} on y_i, which includes the effects from neighboring observations, becomes $\beta_k/(1-\rho)$. To relax this assumption, geographically weighted regression (GWR) accounts for unobserved spatial heterogeneity by allowing model parameters to vary across data observations of different locations. The spatially lagged dependent variable can supplement standard GWR to yield an encompassing spatial model SAR-GWR, which accounts for both spatial dependence and spatial heterogeneity (Geniaux and Martinetti, 2018):

$$y_i = \rho(u_i, v_i) \sum_j w_{ij} y_j + \sum_k \beta_k(u_i, v_i)\, x_{ik} + \varepsilon_i \tag{3}$$

where (u_i, v_i) denotes the geographic coordinates of location i, and so the parameters $\rho(u_i, v_i)$ and $\beta_k(u_i, v_i)$ reflect values at location i. Compared with its counterpart ρ in equation (2), $\rho(u_i, v_i)$ is a "local" measure of spatial autocorrelation.

In line with the first law of geography, the spatial weight matrix $\sum_j w_{ij} = W$ assumes that data observations closer to location i receive a higher weight than observations located farther from that location (Fotheringham et al., 2002). In this study, the specific weighting scheme follows the "adaptive" bandwidth approach, in which the distance for capturing the same number of neighboring locations varies across different locations. This is more attune to our study region in which the density of observations associated with individual counties varies across the United States. For location i, the weight of data observations of location j is given by an adaptive Gaussian distance decay-based weighting function:

$$w_{ij} = \exp\left(-\frac{d_{ij}^2}{h_i^2}\right) \tag{4}$$

where $d_{ij}^2 = [(u_i - u_j)^2 + (v_i - v_j)^2]$ is the Euclidean distance between location i and location j, and h_i is the spatial bandwidth (or neighborhood size) for location i. The optimal spatial bandwidth is identified using cross-validation (CV), which minimizes the sum of squared error (Fotheringham et al., 2002):

$$CV(h) = \sum_i \left[y_i - \widehat{y}^{(\neq i)}(X_i, u_i, v_i)\right]^2 \tag{5}$$

where h is a given bandwidth and $\widehat{y}^{(\neq i)}$ is the predicted value from model regression conditional on h without the ith location.

As for OLS, all the above spatial models (SAR, GWR, and SAR-GWR) are limited to the characterization of conditional-mean relationships between the dependent variable and

covariates. These conditional-mean regression models may not be informative over the entire distribution of the dependent variable. This is particularly crucial for understanding the impacts of social vulnerability across U.S. communities with potential outliers at both tails of the distribution. Instead, conditional quantile regression allows for modeling spatially varying relationships at different points of the conditional distribution of the dependent variable (Koenker and Bassett, 1978). This framework contrasts unconditional quantile regression, which estimates the impact on the dependent variable at a specific quantile for all observations in a sample unconditional on the explanatory variable. The SAR-GWQR model that extends SAR-GWR to a QR setting can be expressed as follows:

$$y_i = \rho^{\tau}(u_i, v_i) \sum_j w_{ij} y_j + \sum_k \beta_k^{\tau}(u_i, v_i)\, x_{ik} + \varepsilon_i^{\tau} \tag{6}$$

where ρ^{τ} and β_k^{τ} are parameters at quantile τ $(0 < \tau < 1)$. Following Chen et al. (2012), equation (6) can be estimated by solving a linear programming problem that minimizes the following weighted loss function for a given location (u_0, v_0):[2]

$$\sum_i \gamma_\tau [y_i - \rho^{\tau}(u_0, v_0) \sum_j w_{ij} y_j - \sum_k \beta_k^{\tau}(u_0, v_0)\, x_{ik}] K\left(\frac{d_{i0}}{h}\right) \tag{7}$$

where $\gamma_\tau(z) = z[\tau - I(z < 0)]$ is a V-shaped piecewise linear "check" function at quantile τ and $I(\cdot)$ is the indicator function. The term $K = diag(\alpha_{i1}, \ldots, \alpha_{in})$ is a vector of weights, α_{ij}, depending on a pre-specified kernel function with the bandwidth h and the distance d_{i0} between the location (u_0, v_0) and the ith location (u_i, v_i). The bandwidth essentially controls the smoothness and efficiency of parameter estimates. We adopt the Gaussian kernel function analogous to equation (4) and an adaptive weighting scheme for the bandwidth. Instead of the CV score defined by equation (5) for conditional-mean regression, the CV value for determining the optimal bandwidth in QR is as follows:

$$CV(h) = \sum_i \gamma_\tau [y_i - \widehat{y}_\tau^{(\neq i)}(X_i, u_i, v_i)] \tag{8}$$

where $\widehat{y}_\tau^{(\neq i)}$ is the predicted value from SAR-GWQR with the ith location being removed. The optimal bandwidth is determined by the model specification that yields the lowest CV value.

For SAR-GWQR model estimation, we arrange the data of dependent variables into quantiles and then apply a two-stage least-squares (2SLS) procedure suggested by Tomal and Helbich (2023). First, we run conditional QR for the spatial lag term, Wy, as a dependent variable using the variables X and WX as instruments (Anselin, 2003). This step is taken to avoid an endogeneity bias from estimating Wy that is correlated with the error term. Next, the fitted values from the first step are used to estimate the SAR-GWR with the

[2]An alternative is a local linear estimator (Chen et al., 2012; Wang et al., 2018), which is computationally more demanding. The performance of the two methods is comparable (Yu and Jones, 1997), and so we chose the local constant estimator in this study.

local-constant QR method detailed by Chen et al. (2020). Essentially, the SAR-GWQR method yields "local" quantile parameter estimates of the spatial lag and other covariates at each geographic location across the conditional distribution of the dependent variable.

3. EMPIRICAL RESULTS

3.1. OLS Results

This section presents results from model regressions with the CDC SVI as the primary variable for explaining local economic and public health outcomes in the COVID-19 pandemic. As described in Section 2 above, the empirical models include population density, state-level OxCGRT COVID-19-related policy stringency, and economic support measures in addition to SVI. The dependent variable is measured alternatively by the relative employment and economic activity levels as economic outcomes, as well as the COVID-19 case and death rates. To facilitate interpretations among variables of different scales, especially the indices, the data of all variables are converted to standard normal values. Regressions are run with cross-section data of 3,142 U.S. counties separately in April 2020 and June 2022, which correspond to the early and late stages of the pandemic, respectively. The expected sign for SVI is positive, meaning more adverse impact on economic and public health outcomes among more socially vulnerable counties.

Table 2 (panel A) first presents the results of OLS regressions as the baselines. Because the four outcome variables have a standard normal distribution, the estimates for all intercept terms are close to zero and, thus, are not reported. For the outcomes in April 2020, the SVI parameter estimates are not statistically meaningful in the models for employment and output, but statistically significant with the expected positive sign for the COVID-19 case and death rates. For the outcomes in June 2022, the SVI parameter estimate is positive for all outcome variables except the statistically insignificant estimate for employment. The positive estimates confirm that more socially vulnerable counties experienced more pandemic-related economic or public health impacts.

Overall, estimation results for the control variables are relatively more consistent across the outcome variables than those for SVI. For instance, the results for population density confirm that more populated counties were more exposed to both economic and public health impacts at the onset of the pandemic. The negative parameter estimates for the June 2022 data indicate that cities or urban areas had recovered faster than rural areas by the end of the pandemic. As a result, the economies of more populated counties performed relatively better and had lower cumulative human tolls.

For the two pandemic-related state policy variables, most estimates are also statistically significant. Higher policy stringency, or tighter containment measures, are associated with more severe economic setbacks in the wake of the nationwide lockdowns in 2020. The estimation results also support the efficacy of state pandemic interventions for curbing local COVID-19 outbreaks and death tolls. The results for the economic support variable ("support") are mixed, however. In April 2020, state support was negatively associated with economic outcomes and positively associated with the two public health outcomes. Their corresponding parameter estimates switch signs for outcomes about two years later: State

Table 2: "Global" Model Regression Results

A) OLS Models

	April 2020								June 2022							
	Employment	s.e.	Output	s.e.	Cases	s.e.	Deaths	s.e.	Employment	s.e.	Output	s.e.	Cases	s.e.	Deaths	s.e.
SVI	0.005	0.016	-0.003	0.017	0.126	0.016***	0.071	0.017***	0.001	0.019	0.087	0.019***	0.030	0.018*	0.217	0.018***
Density	0.412	0.016*	0.357	0.017***	0.440	0.016***	0.282	0.017***	-0.039	0.018**	-0.055	0.018***	-0.020	0.018	-0.058	0.017***
Stringency	0.206	0.017***	0.187	0.018***	-0.101	0.017***	-0.067	0.018***	0.010	0.020	0.002	0.020	-0.084	0.020***	-0.190	0.019***
Support	-0.023	0.017	-0.048	0.018***	0.042	0.017***	0.068	0.019***	0.129	0.020***	0.072	0.020***	-0.075	0.020***	-0.088	0.019***
Adjusted R^2	0.224		0.165		0.224		0.094		0.017		0.123		0.018		0.106	
Likelihood	-4056.12		-4171.37		-4057.14		-4300.42		-4429.56		-4439.54		-4426.06		-4279.43	
Moran's I	0.558***		0.486***		0.317***		0.249***		0.21***		0.198***		0.394***		0.373***	
Robust LM	157.467***		262.703***		32.104***		19.801***		412.168***		348.109***		1.944***		0.723***	
KB Test	26.650***		49.112***		295.046***		169.553***		38.475***		47.835***		423.49***		438.892***	
QR Pseudo R^2	0.390		0.350		0.401		0.187		0.269		0.266		0.073		0.333	

B) SAR Models

	April 2020								June 2022							
	Employment	s.e.	Output	s.e.	Cases	s.e.	Deaths	s.e.	Employment	s.e.	Output	s.e.	Cases	s.e.	Deaths	s.e.
SVI	-0.021	0.016	-0.007	0.012	0.095	0.014***	0.043	0.016**	0.025	0.018	0.078	0.018***	0.040	0.015***	0.143	0.015***
Density	0.136	0.012*	0.102	0.013***	0.258	0.016***	0.177	0.017*	-0.028	0.017*	-0.051	0.017***	0.017	0.014	-0.024	0.014
Stringency	0.036	0.011***	0.033	0.013***	-0.054	0.015***	-0.037	0.017***	0.009	0.019	0.001	0.019	-0.06	0.016***	-0.112	0.016***
Support	-0.012	0.011	-0.016	0.013	0.015	0.015	0.032	0.017	0.081	0.019***	0.05	0.019***	-0.015	0.016	-0.013	0.016
Spatial Lag	0.775	0.013***	0.741	0.015***	0.541	0.020***	0.479	0.022***	0.435	0.024***	0.410	0.024***	0.704	0.017***	0.634	0.018***
Adjusted R^2	0.657		0.571		0.409		0.246		0.143		0.127		0.379		0.384	
Likelihood	-3003.50		-3329.16		-3724.06		-4084.61		-4273.41		-4305.32		-3883.82		-3832.87	
QR Pseudo R^2	0.705		0.658		0.666		0.809		0.351		0.298		0.671		0.550	

Notes: s.e. denotes standard error. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

support was positively associated with economic outcomes and negatively associated with public health outcomes.

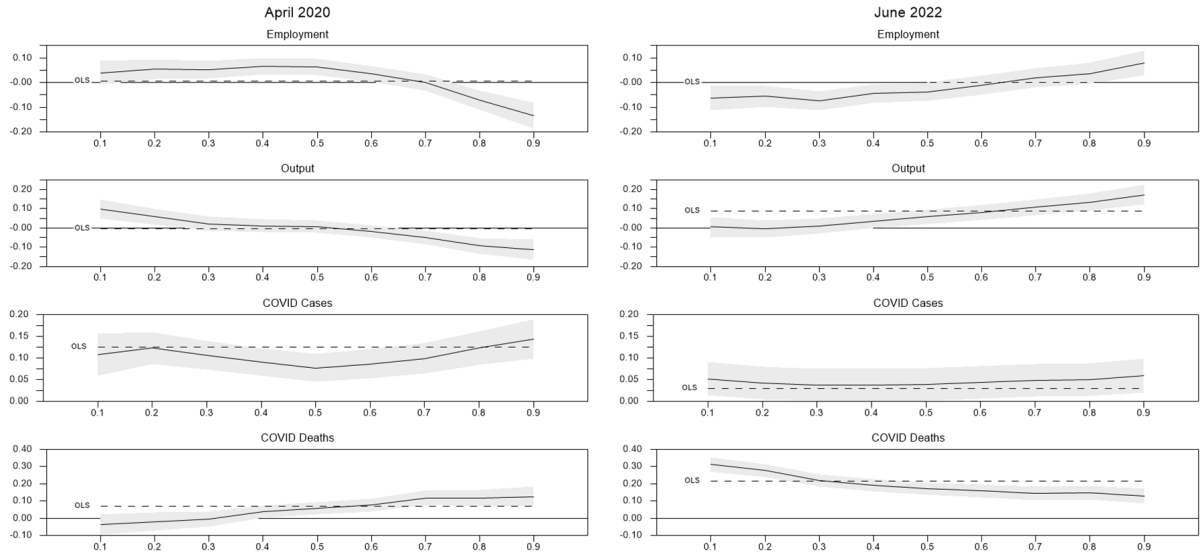
Below the OLS parameter estimates are the models' overall diagnostic statistics. The Moran's I statistics test for spatial autocorrelation in residuals. Along with the corresponding Lagrange Multiplier (LM) statistics, all Moran's I test statistics reject the null hypothesis of no spatial correlation in OLS residuals. Such statistical evidence supports our application of the SAR model below. The KB test results, below the Moran's I statistics, are Koenker and Bassett (KB, 1978) statistics for detecting heteroskedasticity in regression residuals. All test results are statistically significant, suggesting that our disaster outcome data do not follow a normal distribution and plausibly contain extreme values.

3.2. QR Results

Unlike OLS, quantile regression is robust to outliers. For illustration, we ran standard, "global" conditional quantile regressions with the procedure developed by Koenker and Bassett (1978). The bottom row of panel A in Table 2 shows the pseudo-adjusted R^2's, which are appreciably higher than their OLS counterparts. Figure 1 shows the QR parameter estimates for SVI at each decile (10th percentile) between the 0.1 and 0.9 quantiles of the dependent variables. The shaded bands delineate the 95% confidence levels of the point estimates. Keep in mind that all parameter estimates are expected to be positive to capture the role of social vulnerability. In each case, the OLS parameter estimate (horizontal dashed line) lies near the mean of the parameter estimates across the entire quantile range.

For employment in April 2020, the QR parameter estimate hovers around 0.05 up to the median ($\tau = 0.5$) and then becomes negative above the 0.75 quantile of employment losses. The QR for output data of the same period yields a similar downtrend, turning negative at the upper half of the output distribution. The patterns of the QR estimates for both employment and output data contrast their corresponding OLS estimates, which are close

Figure 1: OLS and Quantile Regression Estimates for SVI by Outcome Quantile



Notes: Horizontal dashed lines indicate OLS parameter estimates, and shaded bands indicate 95% confidence intervals of quantile regression.

to the means of the conditional quantile estimates but not statistically different from zero (see Table 2).

For the two economic outcomes in June 2022, however, the SVI parameter estimates trend up instead of down along the outcome distributions. During that period, most of the employment and output data above their medians are positive (i.e., employment and output losses) as opposed to negative below the medians (i.e., higher employment and output levels than the baselines). As such, the plots reveal that near the end of the pandemic, the expected positive relationship between social vulnerability and pandemic-related economic tolls holds up mostly among counties whose economies remained behind their pre-pandemic levels. Those counties are mostly in states in the Eastern region. In the West, most counties in the state of Nevada also lagged the rest of the nation due to the pandemic's disproportionate impact on tourism.

Unlike the results for economic outcomes, conditional quantile parameter estimates vary modestly across the conditional distribution of COVID-19 case rates. Their corresponding OLS parameter estimates are also mostly encapsulated by the 95% confidence bands of the QR parameter estimates. This is not the case for local death tolls, however. The bottom two plots of Figure 1 indicate that the relationship between social vulnerability and the death rate was stronger among counties with higher death rates in April 2020, but this relationship reversed in June 2022.

3.3. SAR Model Results

Like OLS, standard "global" quantile regressions ignore spatial effects. To illustrate the impacts of spatial interactions on economic and public health outcomes, panel B of Table

2 displays the results of "global" SAR model regressions. The SAR model supplements the OLS model with a spatially lagged dependent variable, Wy, in equation (2) above. In line with the test statistics for spatial dependence in panel A, the parameter estimates for the spatial lags between 0.410 and 0.775 suggest strong, positive spillovers of economic and public health outcomes of a county with its neighboring counties. The presence of the spatial lag term does not alter the qualitative results of most covariates including SVI.

The SAR estimation results, nonetheless, are limited to the conditional means of the dependent variable. To highlight this deficiency, the bottom row of Table 2 lists the pseudo-adjusted R^2's of the SAR models estimated with QR. For all eight cases, QR further improves the conventional SAR specification's overall explanatory power, although the extent of improvement varies across different outcome data.

3.4. SAR-GWR Results

All results presented in Table 2 are generated from "global" models, which assume the same parameter estimates for all U.S. counties. To explore varying model relationships across counties in the context of spatial heterogeneity, we extend those "global" models to geographically weighted regression. Table 2 displays estimation results for the SAR-GWR in the form of equation (3) above. In addition to the models' overall diagnostic statistics, the table lists summary statistics for the "local" estimated parameters of the primary explanatory variable, SVI.

Table 3 first lists the optimal bandwidths selected for model calibrations. For the April 2020 outcome data, the bandwidths between 243 and 568 correspond to broad U.S. regions. The bandwidths for outcomes in June 2022 are considerably smaller, reflecting less spatial interaction in the economic and public health outcomes over the course of the pandemic than in early 2020. The next column lists the adjusted R^2's of the SAR-GWR models, which are all higher than their SAR counterparts in Table 2. The better performance of SAR-GWR over SAR is further supported by the statistically significant likelihood ratio (LR) test statistics for comparing the two models.[3]

Instead of parameter estimates that capture "direct" effects, Table 3 lists the "total" effects of the parameter estimates, which also include "indirect" spatial effects of neighboring counties, are larger by a factor of $1/(1\text{-}\widehat{\rho})$ where $\widehat{\rho}$ is the estimated spatial autoregressive parameter. For each of the eight outcomes, the median of "local" parameter estimates for SVI is comparable to their corresponding "global" parameter estimates. However, the spreads between the minimum and maximum values, as well as their standard deviations, indicate large dispersions in local parameter estimates. Chen et al. (2012) proposed a test for spatial heterogeneity in a model parameter, which compares the interquartile range (IQR, between 25th percentile and 75th percentile) of the local parameter estimates against twice the standard error of the "global" model parameter estimate. In all cases, there is evidence in support of spatial heterogeneity in the SVI parameters as the IQR exceeds twice the standard error (s.e.) of the corresponding "global" SVI parameter estimate.

[3]The LR test statistics are two times the difference between the two models' log-likelihood values with one degree of freedom.

Table 3: SAR-GWR Results

	Bandwidth	Adjusted R^2	Log Likelihood	LR (SAR)	SVI Local Parameter (Total Effects)					
					Minimum	Median	Maximum	Std. Dev.	IQR	2×s.e. (SAR)
A) April 2020										
Employment	412	0.685	-2605.304	796.392***	-0.377	-0.036	0.094	0.056	0.112 >	0.048
Output	243	0.618	-2880.102	898.116***	-0.338	-0.029	0.25	0.091	0.182 >	0.047
COVID Cases	357	0.459	-3447.903	552.314***	-0.309	0.101	0.352	0.096	0.192 >	0.046
COVID Deaths	568	0.295	-3879.926	409.368***	-0.164	0.021	0.254	0.100	0.200 >	0.031
B) June 2022										
Employment	58	0.440	-3271.042	2004.736***	-0.732	0.076	1.648	0.167	0.334 >	0.032
Output	89	0.372	-3726.607	1157.426***	-0.681	0.098	1.406	0.175	0.350 >	0.030
COVID Cases	276	0.462	-3407.371	952.898***	-0.244	0.132	0.896	0.035	0.070 >	0.050
COVID Deaths	217	0.685	-2605.304	2455.132***	-0.377	-0.036	0.094	0.125	0.250 >	0.042

Notes: IQR denotes the interquartile range. s.e. denotes standard error. *** denotes statistical significance at the 1% level.

3.5. SAR-GWQR Results

Despite evidence of spatial heterogeneity in model relationships other than spatial autocorrelation, it is important to observe that the adjusted R^2's of the SAR-GWR models (Table 3) are lower than their corresponding statistics for models that extend SAR to QR (bottom row of Table 2. This caveat motivates the adoption of a framework that nests the SAR-GWR model with QR, as captured by equation (6) above.

For illustration, Table 4 reports key results of the SAR-GWQR models at three conditional quantiles of the dependent variable, namely the 0.25, 0.5 (median), and 0.75 quantiles. Standard errors are obtained with a bootstrap method with 500 replications (Chen et al., 2020). The optimal bandwidth for model calibration varies as much among the three quantiles as different outcomes. The bandwidths tend to be smaller for employment and COVID-19 death data than output and COVID-19 case data.

According to the pseudo-adjusted R^2's, the model goodness-of-fit differs considerably across different quantile levels. Overall, the empirical model fits the economic and public health data better at the middle quantile (τ=0.5) than the upper (τ=0.75) and lower (τ=0.25) quantiles. The R^2 values at the middle quantile are also most comparable with the corresponding statistics for the SAR-GWR model in Table 3. Despite a dearth of decisive improvement in the overall explanatory power, the primary advantage of the QR extension stems from the estimation results for quantiles away from the median. This is illustrated by the summary statistics of the local parameter estimates for SVI in Table 4. Overall, the median tends to increase with a higher quantile. The standard deviations are also comparable across quantiles and between the two periods. However, the ranges between the maximum and minimum values are substantially larger for the earlier than the later period, suggesting the presence of more extreme outcomes in early 2020. At all quantiles, spatial heterogeneity is confirmed by Chen et al.'s (2012) test, in which the IQR of the local parameter estimates exceeds two standard errors of the corresponding "global" SAR model estimated with QR.

To visualize the results of QR in comparison with the conditional-mean models, we constructed choropleth maps of "local" parameter estimates for SVI in the GWR, SAR-GWR, and SAR-GWQR. Estimates from the latter two models with a spatial autoregressive

Table 4: Results of SAR-GWR with Quantile Regression (SAR-GWQR)

		Pseudo	SVI Local Parameter (Total Effects)						
	Bandwidth	Adj. R^2	Minimum	Median	Maximum	Std. Dev.	IQR		2 × s.e. (SAR-QR)
A) τ = 0.25									
April 2020									
Employment	134	0.574	-1.764	-0.258	0.750	0.360	0.448	>	0.042
Output	1659	0.461	-1.566	-0.913	-0.017	0.364	0.540	>	0.226
COVID Cases	243	0.295	-1.495	0.781	3.369	0.835	1.139	>	0.087
COVID Deaths	82	0.275	-1.601	-0.031	1.115	0.264	0.252	>	0.033
June 2022									
Employment	478	0.267	-0.692	0.161	0.789	0.346	0.563	>	0.052
Output	248	0.238	-0.566	0.228	1.351	0.351	0.435	>	0.045
COVID Cases	1643	0.384	-0.037	0.269	0.604	0.166	0.266	>	0.073
COVID Deaths	356	0.369	-0.011	0.464	1.426	0.209	0.273	>	0.028
B) τ = 0.50									
April 2020									
Employment	258	0.681	-1.993	-0.263	1.013	0.527	0.785	>	0.073
Output	1093	0.594	-1.133	-0.009	0.685	0.430	0.623	>	0.145
COVID Cases	284	0.428	-0.181	0.241	0.804	0.180	0.257	>	0.029
COVID Deaths	121	0.285	-0.821	0.021	1.116	0.330	0.432	>	0.033
June 2022									
Employment	392	0.349	-0.487	0.223	1.506	0.316	0.349	>	0.043
Output	265	0.325	-0.307	0.279	1.302	0.289	0.333	>	0.037
COVID Cases	1242	0.503	0.122	0.236	0.544	0.094	0.112	>	0.026
COVID Deaths	433	0.395	0.078	0.330	0.741	0.121	0.194	>	0.021
C) τ = 0.75									
April 2020									
Employment	377	0.603	-8.313	0.526	9.571	4.642	5.860	>	0.869
Output	1026	0.490	-1.519	0.732	3.273	1.218	2.119	>	0.419
COVID Cases	217	0.183	-0.382	0.201	0.795	0.183	0.240	>	0.021
COVID Deaths	437	0.237	-0.513	0.208	0.852	0.266	0.323	>	0.046
June 2022									
Employment	294	0.226	-0.725	0.578	2.201	0.471	0.619	>	0.064
Output	308	0.277	-0.478	0.357	1.273	0.290	0.371	>	0.044
COVID Cases	176	0.190	-0.165	0.173	0.605	0.128	0.144	>	0.012
COVID Deaths	237	0.297	-0.109	0.245	0.856	0.156	0.218	>	0.018

Notes: IQR denotes the interquartile range. s.e. denotes standard error obtained through 500 bootstrap replications.

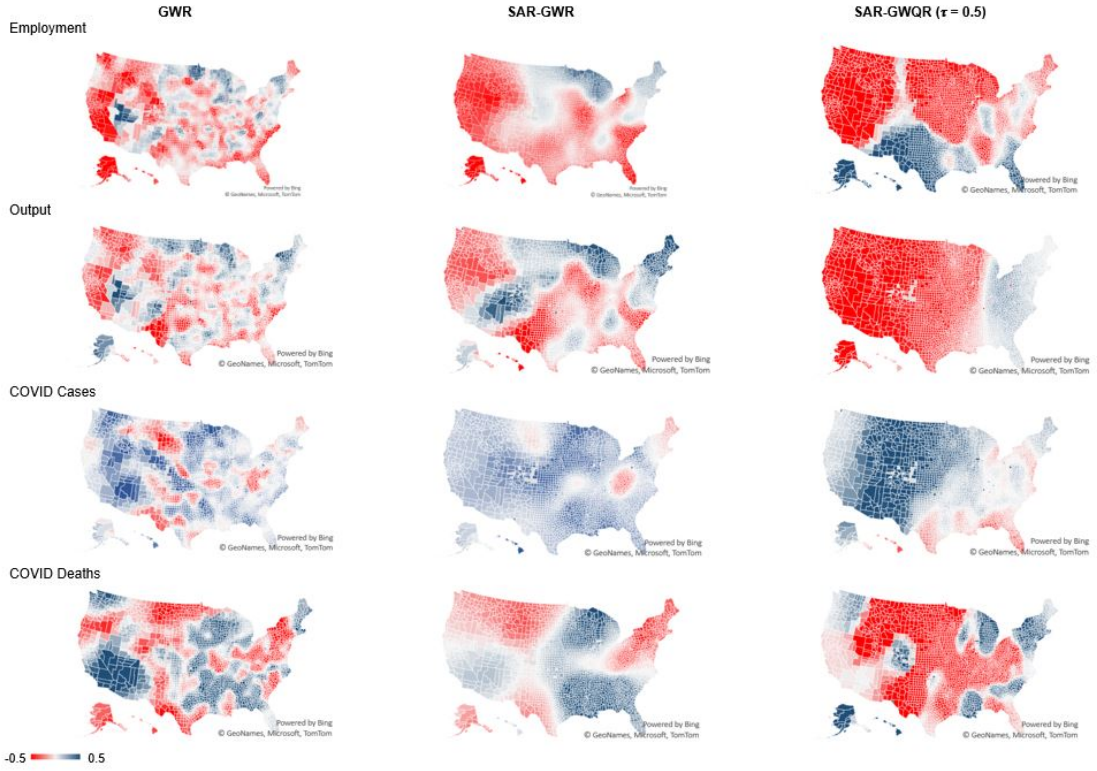
term represent "total" effects, which include "indirect" spatial effects captured by the spatial autoregressive term. Figure 2 displays the results for outcomes in April 2020 and Figure 3 displays the results for outcomes in June 2022. For comparison purposes, we present the conditional quantile regression results for $\tau = 0.5$ (median), and only estimates that are statistically significant at the 0.01 level are visible.[4]

The maps in Figure 2 convey two general observations. First, the spatial patterns of the SAR-GWQR parameter estimates (third column) for the two economic outcomes in April 2020 appear to correspond to the relative sizes of pandemic-related economic losses. For employment, positive SVI parameter estimates (blue color) are found largely in the South-

[4]A vast majority of local SVI parameter estimates are negative at the 0.25 quantile and positive at the 0.75 quantile. To conserve space, these results are not reported here but are available upon request.

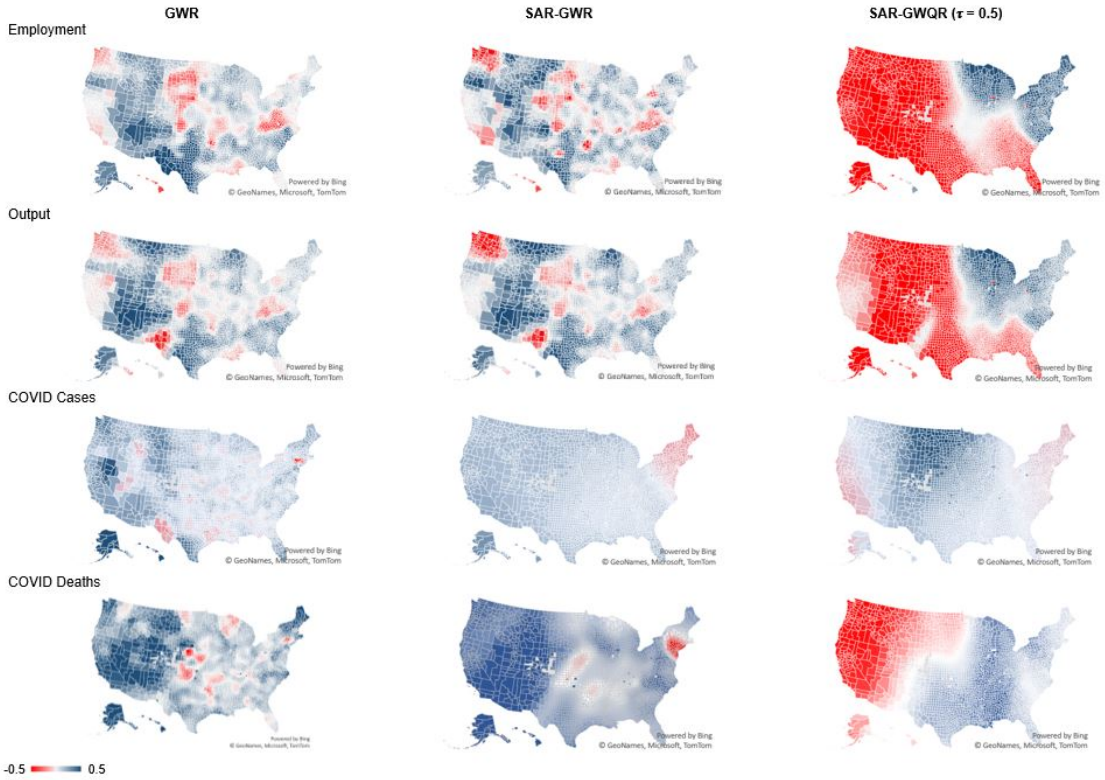
eastern U.S., such as the states of Florida and Texas. For output, positive SVI parameter estimates are found along the Atlantic Coast. Those regions also tended to experience disproportionately more economic losses in the wake of business lockdowns in early 2020 (Figure A1 in the Appendix). Conversely, negative SVI parameter estimates (red color) coincide with relatively lower economic losses in much of the Midwest and Western regions. The map for the relationship between the SVI and the death rate (bottom right map in Figure 2) displays a similar overall pattern in association with the distribution of death rates across the United States. For the June 2022 outcomes in Figure 3, negative SVI parameter estimates for employment, output, and deaths also cluster in the Western United States, particularly counties less exposed to the pandemic's adverse impacts.

Figure 2: Geographically Weighted Regression Local Parameter Estimates for SVI (April 2020 Outcomes)



The second observation is that the SAR models, which augment GWR with a spatial lag term, smooth out the "local" parameter estimates for SVI over geographic space as the estimated impact on each county's disaster outcomes includes spillover effects from its neighboring counties. As evident in the first column of Figure 2, the standard GWR generates rather scattered parameter estimates both in their signs and magnitudes across the United States. Dispersion is much less evident among local parameter estimates from SAR-GWR (second column). An extension of SAR-GWR to QR (third column) yields even smoother local parameter estimates due in part to larger bandwidths (Table 4). The smoothing effects of large bandwidths are striking for the output data in April 2020 (bandwidth of 1,093 counties) and the COVID-19 incidence data in June 2022 (bandwidth of 1,242 counties).

Figure 3: Geographically Weighted Regression Local Parameter Estimates for SVI (June 2022 Outcomes)



The overall pattern of local SVI parameter estimates of SAR-GWQR at the median of the outcome distribution ($\tau = 0.5$) is drastically different from the corresponding results in conditional-mean regressions (GWR and SAR-GWR). In particular, parameter estimates are negative for economic outcomes in both periods for most counties in the Western region. For the mid-2022 period, positive parameter estimates for economic outcomes can be found mostly in the Northeast.

Between the two public health outcomes, the spatial patterns of the results for the COVID-19 death data, as opposed to incidence data, align with the results for economic outcomes. The SVI parameter estimates for COVID-19 cases are predominantly positive. Those comparative results highlight the extent to which socially vulnerable populations are exposed to the disease's most severe health impact. The bottom right map in Figure 3 shows that, at the median of the conditional distribution, the parameter estimates for the relationship between SVI and the death rate are positive among most counties in the eastern half of the United States and negative in the western half. This contrasts the results of the conditional-mean regressions, which generate mostly positive local parameter estimates.

4. DISCUSSION

Despite a proliferation of social vulnerability constructs to guide policymakers in efforts to mitigate and respond to the adverse effects of disasters, some studies have questioned the empirical validity of those composite measures, especially their reliability in predicting different disaster outcomes (Bakkensen et al., 2017; Rufat et al., 2019). One reason for the weak or conflicting findings stems from disparities in a variety of sociodemographic conditions as well as disaster outcomes across local communities and broad U.S. regions (Cutter et al., 2014, 2016). From this perspective, geospatial dynamics play a key role in modeling community-level disaster outcomes with pre-disaster social vulnerability attributes (e.g., Rufat et al., 2019; Park and Xu, 2020).

Our empirical model estimated with OLS generates mixed results for the CDC SVI as the key predictor for local economic performance at two distinct stages of the COVID-19 pandemic. By comparison, the OLS results for predicting public health outcomes are more decisive. In particular, more socially vulnerable counties are found to witness relatively more coronavirus-related death tolls. This finding accords with results on the role of sociodemographic factors in determining excess mortality during the pandemic (Rodriguez-Pose and Burlina, 2021; Ramirez et al., 2022).

We attribute the distinct results for estimating the role of social vulnerability in economic outcomes as opposed to public health outcomes to the varying model relationships over the distribution of economic outcomes, particularly outliers at the tails of the distributions. As a result, the OLS estimator, which is limited to the conditional mean of the dependent variable, conceals a great deal of variation in model relationships over the entire distribution of economic outcomes. An application of QR reveals positive parameter estimates for SVI up to the median of local employment and output losses during the onset of the pandemic in April 2020. For counties most exposed to economic disruptions (i.e., top quartile), the parameter estimates for SVI turn negative. This might explain the near-zero OLS parameter estimates, which are close to the medians of the QR parameter estimates. In other words, as a result of some counties that were hit exceptionally hard in early 2020, standard conditional-mean regressions underestimate the role of social vulnerability for others.

For economic outcomes near the end of the pandemic in June 2022, the QR results for SVI infer patterns opposite to those based on data about two years earlier. The finding of an increasing trend along the outcome conditional distribution complements recent studies on disaster determinants. For example, Kim and Marcouiller (2019) reported that the determinants of flood losses become relevant only when the damage reaches a high level. Sociodemographic factors were less relevant to local economies, largely immune to the pandemic's devastation, especially the best-performing U.S. counties. Conditional quantile regressions also yield different trends of parameter estimates in models for public health outcomes between the two periods. For COVID-19 death tolls up to April 2020, the SVI plays a larger role among counties with higher death rates, but the opposite pattern occurs over the pandemic through June 2022. Neelon et al. (2021) found that more socially vulnerable counties began to exhibit *lower* instead of higher COVID-19 death rates in late 2020, plausibly due to government policy responses and interventions. Ramirez et al. (2022) showed that while the coronavirus spread more widely over time, local death tolls continued to be affected by

government responses along with sociodemographic and institutional factors.

Other than social vulnerability's varying role at different levels of local disaster outcomes, spatial dependence and unobserved spatial heterogeneity in its relationships with disaster outcomes prevail across the United States. Spatial spillovers of disaster outcomes are well documented in the literature (Belasen and Polacheck, 2009; Lee, 2021). Estimates of the spatial lag term in the SAR models highlight more spatial interactions among counties at the early stage than at the late stage of the pandemic. Local parameter estimates from incorporating GWR into the standard SAR model add insight into variation in model relationships among counties in different regions. The overall findings accord with studies (Cutter et al., 2014, 2016, 2020; Park and Xu, 2020) that highlight the extent of spatial variation in social vulnerability across U.S. regions and its implication for understanding disaster outcomes.

Still, inferences from the integrated SAR-GWR model are limited to the conditional means of the economic and public health outcomes. Fotheringham et al. (2002) asserted that local GWR estimates are highly sensitive to outlying observations, and Mur and Lauridsen (2007) showed how outliers affect the reliability of the tests for social dependence, such as Moran's *I*. To understand spatially varying model relationships at different points of the disaster outcome distribution, we have further integrated QR into the SAR-GWR framework.

Overall, local economic and public health outcomes were associated with their pre-existing sociodemographic characteristics, as captured by the CDC SVI, but their relationships varied not only spatially across counties and broad regions but also among communities exposed to the pandemic at different degrees. Our integrated SAR-GWQR framework highlights the varying role of social vulnerability in predicting economic and human losses of different counties in different regions, depending on their relative exposure to COVID-19 outbreaks. At the typical levels, or medians, of pandemic-related economic and human losses, for instance, a divide emerged between the eastern and western United States.

5. CONCLUSION

We have explored the empirical validity of the CDC's composite measure of social vulnerability in the context of the COVID-19 pandemic. The OLS method, the standard workhorse in econometric analysis, yields weak support for the association of the CDC SVI with economic outcomes at two stages of the pandemic. To shed light on model relationships beyond the conditional mean of the dependent variable, we have applied quantile regression to estimate different effects of the SVI at different points of the disaster outcome distribution.

Compared to OLS, "global" conditional quantile regression offers more decisive evidence in support of social vulnerability in predicting the pandemic's uneven economic devastation across the United States. In the wake of the nationwide lockdowns in early 2020, more socially vulnerable counties experienced disproportionately more employment and output losses, except those hardest-hit county economies (top quartile). About two years into the pandemic, the nationwide economy had restored most of its employment and output losses, but recovery was uneven across the nation. For local economies remaining behind their pre-pandemic levels (upper quantiles), social vulnerability was directly associated with their relative performance. Local GWR estimation results further infer that many of those slowly

recovering counties were clustered in the Northeastern region. The SAR-GWQR framework, which incorporates conditional quantile regression into spatial models, distinguishes the role of social vulnerability in disaster outcomes between the eastern and western United States at the median of the conditional outcome distribution.

In contrast to the results for economic outcomes, OLS results support the role of social vulnerability in explaining different counties' exposure to COVID-19 health risks and their death tolls. Conditional quantile regressions further indicate that, during the first wave of COVID-19 outbreaks, social vulnerability was positively associated with death tolls among most counties except those with little exposure to COVID-19 health risks (lower quartiles). Those outliers were mostly rural counties in the Great Lakes and Midwest regions. Over the course of the pandemic through June 2022, more socially vulnerable counties saw more cumulative death tolls, although the role of social vulnerability diminished as the exposure to the public health impact increased (upper quartiles).

We attribute the shortfall of OLS as a conditional-mean estimator to the presence of extreme local outcomes in the wake of major developments in the COVID-19 pandemic. Still, quantile regression is a "global" model that seems to mask substantial variation in model relationships within and across regions. The SAR-GWQR framework allows for spatial autocorrelation and spatial heterogeneity in model relationships at different points of the outcome distribution. Applying this framework to the economic and public health outcomes in the pandemic highlights the distinct role of social vulnerability in the western versus eastern United States at the middle quantile of the conditional distribution.

Taken together, our regression results reinforce pre-existing social vulnerability conditions as key determinants for a *typical* U.S. community's susceptibility to the pandemic's woes as well as a hindrance to its eventual recovery. Yet our study also highlights the complexity, or nonlinearity, of the relationships between *local* disaster outcomes and their determinants. This helps explain the fragile findings on the predictive power of social vulnerability in the literature (e.g., Rufat et al., 2019; Spielman et al., 2020). In the case of the COVID-19 pandemic, regions within the U.S. differ widely in their vulnerability to pandemic-related economic and public health outcomes. From this perspective, it would be fruitful in future research to incorporate the spatial and quantile dimensions into the study of various social vulnerability outcomes. In addition, despite much overlap between the CDC and other popular social vulnerability measurements (Derakhshan et al., 2022), the multidimensional nature of community-level sociodemographic conditions calls for event studies with other social vulnerability determinants.

By comparing disaster outcomes in two time periods as snapshots, we shed light on the evolving role of social vulnerability over time. More specifically, we have identified changes in the relationships between social vulnerability and some disaster outcomes near the beginning and the end of the pandemic. These findings underscore the cyclical, or transitory, nature of socioeconomic conditions, particularly at the local level, in the wake of a major disruptive event. For economic outcomes, U.S. counties hit hardest in the wake of the nationwide lockdown were not particularly socially vulnerable. In the longer run, however, the role of social vulnerability became the strongest among worst-performing counties.

The opposite was true for public health outcomes. More socially vulnerable counties witnessed more death tolls during the first wave of COVID-19 outbreaks. Over time, gov-

ernment interventions might have ameliorated the public health outcomes of socially vulnerable communities. To guide disaster mitigation and responses, policymakers would benefit from a better understanding of the spatial and temporal dynamics uncovered in this paper, especially among socially vulnerable communities.

REFERENCES

ATSDR. (2022). CDC/ATSDR SVI 2020 Documentation. Assessed online: https://www.atsdr.cdc.gov/placeandhealth/svi/documentation/SVI_documentation_2020.html.

Anselin, L. (2003). Spatial Econometrics. In Baltagi, B.H. (ed.), *A Companion to Theoretical Econometrics*, Blackwell Publishing: Malden, USA, 310-330. https://doi.org/10.1002/9780470996249.ch15.

Althoff, L., F. Eckert, S. Ganapati, and C. Walsh. (2022). The Geography of Remote Work. *Regional Science and Urban Economics*, 93, 103770. https://doi.org/10.1016/j.regsciurbeco.2022.103770.

Bakkensen, L.A., C. Fox-Lent, L.K. Read, and I. Linkov. (2017). Validating Resilience and Vulnerability Indices in the Context of Natural Disasters. *Risk Analysis*, 37(5), 982-1004. https://doi.org/10.1111/risa.12677.

Belasen, A.R., and S.W. Polachek. (2009). How Disasters Affect Local Labor Markets: The Effects of Hurricanes in Florida. *Journal of Human Resources*, 44(1), 251-276. https://doi.org/10.3368/jhr.44.1.251.

Carozzi F., S. Provenzano, and S. Roth. (2024). Urban Density and COVID-19: Understanding the US Experience. *The Annals of Regional Science*, 72, 163-194. https://doi.org/10.1007/s00168-022-01193-z.

Chen, V.Y.J., W.C. Deng, T.C. Yang, and S.A. Matthews. (2012). Geographically Weighted Quantile Regression (GWQR): An Application to U.S. Mortality Data. *Geographical Analysis*, 44(2), 134-150. https://doi.org/10.1111/j.1538-4632.2012.00841.x.

Chen, V.Y.J., T.C. Yang, and S.A. Matthews. (2020). Geographically Weighted Quantile Regression: An Enhancement Based on the Bootstrap Approach. *Geographical Analysis*, 52(4), 642-661. https://doi.org/10.1111/gean.12229.

Cutter, S.L., K.D. Ash, and C.T. Emrich. (2014). The Geographies of Community Disaster Resilience. *Global Environmental Change*, 29, 65-77. https://doi.org/10.1016/j.gloenvcha.2014.08.005.

Cutter, S.L., K.D. Ash, and C.T. Emrich. (2016). Urban-Rural Differences in Disaster Resilience. *Annals of American Association of Geographers*, 106(6), 1236-1252. https://doi.org/10.1080/24694452.2016.1194740.

Cutter, S., and S. Derakhshan. (2020). Temporal and Spatial Change in Disaster Resilience in US Counties, 2010-2015. *Environmental Hazards*. 19(1), 10-29. https://doi.org/10.1080/17477891.2018.1511405.

Coulson, N.E., S.J. McCoy, and I.K. McDonough. (2020). Economic Diversification and the Resiliency Hypothesis: Evidence from the Impact of Natural Disasters on Regional Housing Values. *Regional Science and Urban Economics*, 85, 103581. https://doi.org/10.

1016/j.regsciurbeco.2020.103581.

Dergiades, T., C. Milas, E. Mossialos, and T. Panagiotidis. (2022). Effectiveness of Government Policies In Response to the First COVID-19 Outbreak. *PLoS Global Public Health*, 2(4), e0000242. https://doi.org/10.1371/journal.pgph.0000242.

Derakhshan S., C.T. Emrich, and S.L. Cutter. (2022). Degree and Direction of Overlap Between Social Vulnerability and Community Resilience Measurements. *PLoS ONE*, 17(10), e0275975. https://doi.org/10.1371/journal.pone.0275975.

Edelberg, W., and L. Sheiner. (2021). The Macroeconomic Implications of Biden's $1.9 Trillion Fiscal Package. The Brookings Institution. Accessed online: https://www.brookings.edu/articles/the-macroeconomic-implications-of-bidens-1-9-trillion-fiscal-package/.

Famiglietti, M., and F. Leibovici. (2021). COVID-19 Containment Measures, Health and the Economy. Federal Reserve Bank of St. Louis, Regional Economist, February 18, 2021. Accessed online: https://www.stlouisfed.org/publications/regional-economist/first-quarter-2021/covid-19-containment-measures-health-economy.

Fatemi, F., A. Ardalan, B. Aguirre, N. Mansouri, and I. Mohammadfam. (2017). Social Vulnerability Indicators in Disasters: Findings from a Systematic Review. *International Journal of Disaster Risk Reduction*, 22, 219-227. https://doi.org/10.1016/j.ijdrr.2016.09.006.

Flanagan, B.E., E.W. Gregory, E.J. Hallisey, J.L. Heitgerd, and B. Lewis. (2011). A Social Vulnerability Index of Disaster Management. *Journal of Homeland Security and Emergency Management*, 8(1), Article 3. https://doi.org/10.2202/1547-7355.1792.

Fotheringham, A.S., C. Brunsdon, and M. Charlton. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons: Chichester, UK.

Geniaux, G., and D. Martinetti. (2018). A New Method for Dealing Simultaneously With Spatial Autocorrelation and Spatial Heterogeneity in Regression Models. *Regional Science and Urban Economics*, 72, 74-85. https://doi.org/10.1016/j.regsciurbeco.2017.04.001.

Hallas, L., A. Hatibie, R. Koch, S. Majumdar, M. Pyarali, A. Wood, and T. Hale. (2021) Variation in US States' COVID-19 Policy Responses, University of Oxford, Blavatnik School of Government Working Paper Series, BSG-WP-2020/034.

Kim, H., and D.W. Marcouiller. (2018). Mitigating Flood Risk and Enhancing Community Resilience to Natural Disasters: Plan Quality Matters. *Environmental Hazards*, 17(5), 397-417. https://doi.org/10.1080/17477891.2017.1407743.

Koenker, R., and G. Bassett. (1978). Regression Quantiles. *Econometrics: Journal of the Econometric Society*, 46, 33-50. https://doi.org/10.2307/1913643.

LeSage, J., and K. Pace. (2009). *Introduction to Spatial Econometrics*. SRC Press: Boca Raton, USA. https://doi.org/10.1201/9781420064254.

Lee, J. (2021). The Economic Aftermath of Hurricane Harvey and Irma: The of Federal Aid. *International Journal of Disaster Risk Reduction*, 61, 102301. https://doi.org/10.1016/j.ijdrr.2021.102301.

Lee, J. (2023). Social Vulnerability and Local Economic Outcomes During the COVID-19 Pandemic. *Regional Studies, Regional Science*, 10(1), 845-869. https://doi.org/10.1080/

21681376.2023.2274097.

Lehnert, E.A., G. Wilt, B. Flanagan, and E. Hallisey. (2020). Spatial Exploration of the CDC's Social Vulnerability Index and Heat-Related Health Outcomes in Georgia. *International Journal of Disaster Risk Reduction*, 46, 101717. https://doi.org/10.1016/j.ijdrr.2020.101517.

Mur, J., and J. Lauridsen. (2007). Outliers and Spatial Dependence in Cross-Sectional Regressions. *Environment and Planning A: Economy and Space*, 39(7), 1752-1769. https://doi.org/10.1068/a38207.

Neelon, B., F. Mutiso, N.T. Mueller, J.L. Pearce, and S.E. Benjamin-Neelon. (2021). Spatial and Temporal Trends in Social Vulnerability and COVID-19 Incidence and Death Rates in the United States. *PLoS ONE*, 16(3), e0248702. https://doi.org/10.1371/journal.pone.0248702.

Park, G., and Z. Xu. (2020). Spatial and Temporal Dynamics of Social Vulnerability in the United States from 1970 and 2010: A County Trajectory Analysis. *International Journal of Applied Geospatial Research*, 11(1), 36-54. http://dx.doi.org/10.4018/IJAGR.2020010103.

Ramirez, M.D., P. Veneri, and A.C. Lembcke. (2022). Where Did It Hit Harder? Understanding the Geography of Excess Mortality During the COVID-19 Pandemic. *Journal of Regional Science*, 62(3), 889-908. https://doi.org/10.1111/jors.12595.

Rodriguez-Pose, A., and C. Burlina. (2021). Institutions and the Uneven Geography of the First Wave of the COVID-19 Pandemic. *Journal of Regional Science*, 61(4), 728-752. https://doi.org/10.1111/jors.12541.

Rose, A. (2021). COVID-19 Economic Impacts in Perspective: A Comparison to Recent U.S. Disasters. *International Journal of Disaster Risk Reduction*, 60, 102317. https://doi.org/10.1016/j.ijdrr.2021.102317.

Rufat, S., E. Tate, C. Emrich, and F. Antolini. (2019). How Valid Are Social Vulnerability Models? *Annals of the American Association of Geographers*, 109(4), 1131-1153. https://doi.org/10.1080/24694452.2018.1535887.

Schumacher, I., and E. Strobl. (2011). Economic Development and Losses due to Natural Disasters: The Role of Hazard Exposure. *Ecological Economics*, 72, 97-105. https://doi.org/10.1016/j.ecolecon.2011.09.002.

Spielman, S.E., J. Tuccillo, D.C. Folch, A. Schweikert, R. Davies, N. Wood, and E. Tate. (2020). Evaluating Social Vulnerability Indicators: Criteria and Their Application to the Social Vulnerability Index. *Natural Hazards*, 100(1), 417-436. https://doi.org/10.1007/s11069-019-03820-z.

Smith, B., M. Riddle, A. Wagner, L. Edgemon, C. Burdi, and I. Hyde. (2021). County Economic Impact Index: Measuring the Ongoing Economic Effects of COVID-19. Argonne National Laboratory. Accessed online: https://publications.anl.gov/anlpubs/2021/09/169295.pdf.

Tomal, M., and M. Helbich. (2023). A Spatial Autoregressive Geographically Weighted Quantile Regression to Explore Housing Rent Determinants in Amsterdam and Warsaw. *Environment and Planning B: Urban Analytics and City Science*, 50(3), 579-599. https:

//doi.org/10.1177/23998083221122790.

Walmsley, T., A. Rose, and D. Wei. (2021). The Impacts of the Coronavirus on the Economy of the United States. *Economics of Disasters and Climate Change*, 5(1), 1-52. https://doi.org/10.1007/s41885-020-00080-1.

Wang, W., S. Xu, and T. Yan. (2018). Structure Identification and Model Selection in Geographically Weighted Quantile Regression Models. *Spatial Statistics*, 26, 21-37. https://doi.org/10.1016/j.spasta.2018.05.003.

Yu, K., and M.C. Jones. (1997). A Comparison of Local Constant and Local Linear Regression Quantile Estimators. *Computational Statistics & Data Analysis*, 25(2), 159-166. https://doi.org/10.1016/S0167-9473(97)00006-6.

APPENDIX

Figure A1: Data of COVID-19 Pandemic Outcomes

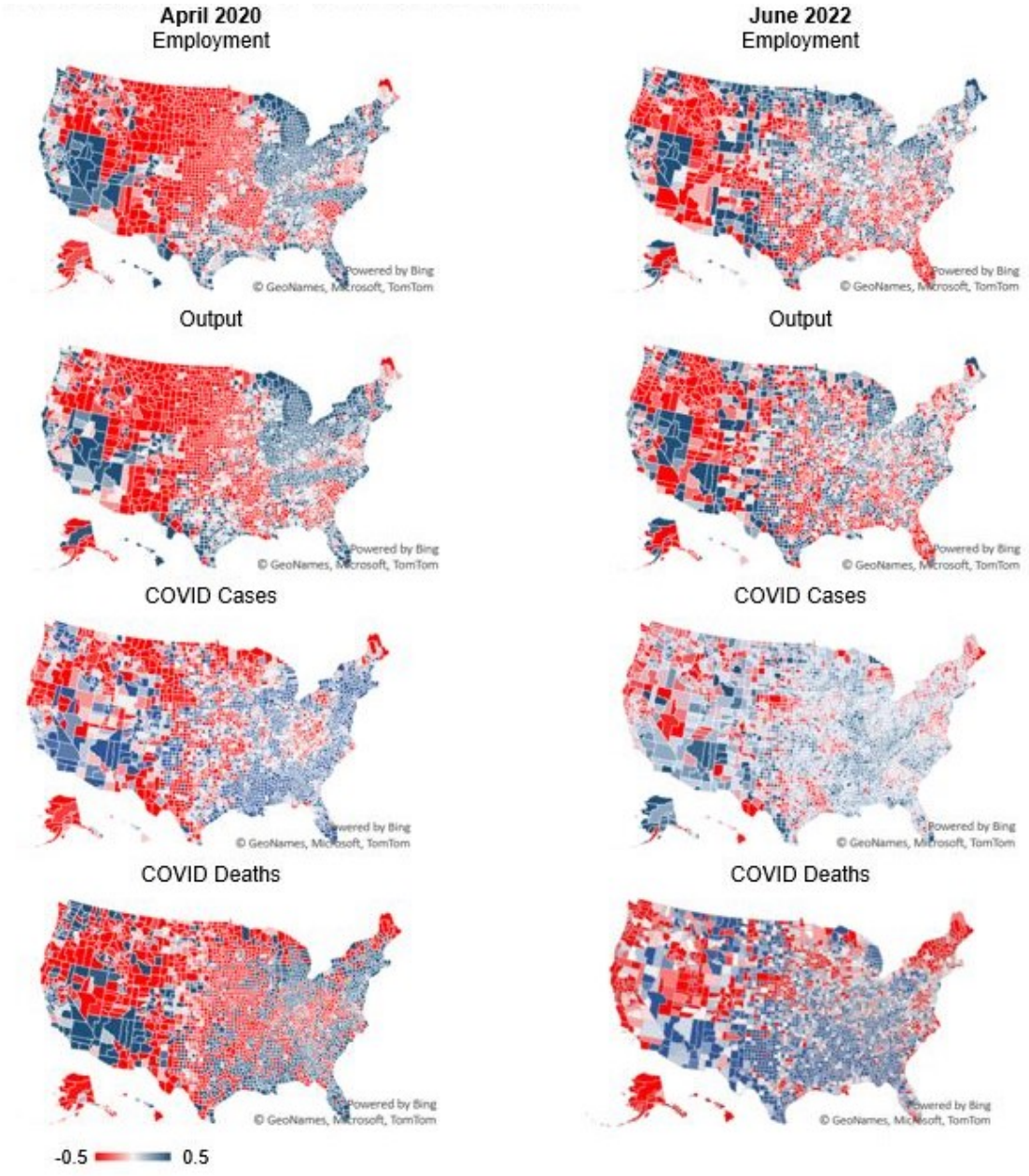


Figure A2: Data of Explanatory Variables