ANNALS OF THE AMERICAN ASSOCIATION OF GEOGRAPHERS
https://doi.org/10.1080/24694452.2023.2292807

Governance Policy Evaluation in the United States during the Pandemic: Nonpharmaceutical Interventions or Else?

Danlin Yu,[a] Shenyang Guo,[b] Yuanyuan Yang,[b] Linyun Fu,[c] Timothy McBride,[b] and Ruopeng An[b]

[a]*Department of Earth and Environmental Studies, Montclair State University, USA;* [b]*Brown School, Washington University in St. Louis, USA;* [c]*Crown Family School of Social Work, Policy and Practice, The University of Chicago, USA*

Scientific evidence suggests that nonpharmaceutical interventions (NPIs) effectively curb the spread of COVID-19 before a pharmaceutical solution. Implementing these interventions also significantly affects regular socioeconomic activities and practices of social, racial, and political justice. Local governments often face conflicting goals during policymaking. Striking a balance among competing goals during a global pandemic is a fine science of governance. How well state governments consume the scientific evidence and maintain such a balance remains less understood. This study employs a set of Bayesian hierarchical models to evaluate how state governments in the United States use scientific evidence to balance the fighting against the spread of COVID-19 disease and socioeconomic, racial, social justice, and other demands. We modeled the relationships between five NPI strategies and COVID-19 caseload information and used the modeled result to perform a balanced governance evaluation. The results suggest that governmental attitude and guidance effectively guide the public to fight back against a global pandemic. The more detailed spatiotemporally varying coefficient process model produces 612,000 spatiotemporally varying coefficients, suggesting all measures sometimes work somewhere. Summarized results indicate that states emphasizing NPIs fared well in curbing the spread of COVID-19. With over 1 million deaths due to COVID-19 in the United States, we feel the balance scale likely needs to tip toward preserving human lives. Our evaluation of governance policies is hence based on such an argument. This study aims to provide decision support for policymaking during a national emergency. *Key Words: balanced governance, Bayesian spatial models, COVID-19, policy evaluation.*

For policymakers, having a solid understanding of what nonpharmaceutical interventions (NPIs) work and to which extent these interventions curb the spread of COVID-19 is a matter of urgency. The right decisions to act against the spread of a contagious disease could reduce mortality rates for individuals during this global pandemic, especially when the different variants of the virus are rampaging through the country while people's attitude toward vaccination remains mixed (Kreps et al. 2020; Largent et al. 2020).

Few studies have used advanced geospatial analytical methods to account for the timing and geographic dispersion simultaneously of such interventions, to investigate the entire pandemic period before a solid vaccination solution in the United States to distill implications for policymaking, to learn lessons to prepare for similar pandemics in the future, and more important, to balance among economic demands, public health imperatives, and political considerations. This is because of the complex and entwined nature of the policymaking and governance practices during a pandemic caused by a new virus. Responses and actions that significantly affect people's daily lives must be made on a short-term, sometimes even daily basis, because of the volatility of the caseload information and because we still have limited knowledge about the virus. Few studies, though, account for daily policy changes. In addition, most traditional methods often lack the capacity to deal with the highly skewed, often zero-inflated daily caseload data required for dynamic

ARTICLE HISTORY
Initial submission, July 2023; revised submissions, July and October 2023; final acceptance, October 2023

CORRESPONDING AUTHORS Danlin Yu yud@montclair.edu; Shenyang Guo s.guo@wustl.edu

policymaking and daily governance practices (Yu et al. 2023). Moreover, in the United States, people's attitudes toward the pandemic vary dramatically from place to place (Puspitasari et al. 2020; Kerr, Panagopoulos, and van der Linden 2021). Whereas governments and people in some places regarded COVID-19 as an immediately life-threatening disease, in others, COVID-19 was treated no more seriously than seasonal influenza (Kerr, Panagopoulos, and van der Linden 2021). Such highly complex spatial and temporal variations of the disease and governmental and individual attitudes toward the disease lead to a highly complex and uneven landscape of NPI implementation strategies across the United States. This complexity further convoluted the effectiveness of policies designed at the federal and local levels to curb the spread of COVID-19.

Although many scholars have attempted to model and predict the trend and patterns of COVID-19 (Enserink and Kupferschmidt 2020; Dowd et al. 2020; Nazia, Law, and Butt 2022), or what social distancing (or physical distancing) strategies should be taken (Kraemer et al. 2020; Auger et al. 2020), these conventional research models often fall short of delving deeply into the complex spatiotemporal dynamics among disease caseload, NPI implementation, individual behaviors, and governmental decisions during the pandemic. NPIs have been shown on many occasions to be effective in preventing the spread of COVID-19 (Baker et al. 2020; Kraemer et al. 2020; Tian et al. 2020). The efficacy of NPIs, however, is closely tied to individual attitudes and behaviors and governments' decisions. Although personal attitudes and behaviors are generally hard to identify, they are shown to be closely related to individual knowledge levels and governments' political orientation (which affects governments' attitude and decisions) during a public health event like the COVID-19 outbreak (Puspitasari et al. 2020; Kerr, Panagopoulos, and van der Linden 2021; Salomon et al. 2021). In addition, in a global pandemic that was caused by a new virus, the situation is constantly evolving and struggles between different decision options (implementing NPI vs. maintaining economic performance, social justice vs. economic justice vs. health threat, among others) are omnipresent. Many such decisions (both individual and governmental) must be made daily. For this matter, investigating governments' decisions and evaluating how such decisions generate the current complex NPI implementing landscape, especially at a daily temporal resolution, provide a much-needed scientifically oriented policymaking practice support, which could lead to better policy guidance for state governments in future emergent public health events. Recently developed data processing (e.g., indirect inverse-normal transformation) and advanced spatiotemporal methods (e.g., Bayesian hierarchical modeling) can allow us to undertake a thorough investigation of these complex dynamics so that a more complete, more detailed picture of the policymaking and governance during a pandemic can be depicted to guide better future responses and actions.

The main objectives of this study were to (1) investigate the impacts of the NPIs on curbing the spread of COVID-19 before a pharmaceutical solution and (2) evaluate how the U.S. state governments use this information for balanced policymaking. The evaluation intends to provide governments with a clear starting point to wade through a myriad of choices in a chaotic public health crisis like this so that they will clearly understand the consequences of their decisions and provide their citizens with the best possible balanced governance. We aim to address two research questions. First, we aim to analyze how NPIs influence the caseload of COVID-19 daily with spatiotemporally varying coefficients models, which intends to provide detailed policy-related insights into the effectiveness of these interventions in controlling the spread of the virus. Second, we devise an evaluation scheme of the U.S. state governments' policies in fighting the spread of COVID-19 by comparing the NPI implementation decisions made by the U.S. state governments and the modeled NPI efficacy results.

Considering the loss of more than 1 million lives in the United States, our evaluation framework is predicated on prioritizing the containment of the disease spread above other considerations. The study aims to draw lessons that shed light on policymaking that enhances racial, social, and political justice.

Literature Review

Scientific evidence suggests that NPIs are effective against COVID-19 to various degrees (Dehning et al. 2020; Guo et al. 2020; Brauner et al. 2021). How such scientific evidence was used in the United

States to develop policies, however, has been a mixed story. There were considerations of a delicate balance between interventions against COVID-19, tensions with individual "freedoms," and reducing economic costs. There was also the complex situation of the 2020 presidential election, during which politicians often held to their own agenda and policies more tightly than in other times (Bartels 1993; Jacobs and Shapiro 2000; Masket, Winburn, and Wright 2012; Desmarais, La Raja, and Kowal 2015). In addition, the death of George Floyd and the ensuing antiracism movement further convoluted the enactment and implementation of the NPIs during the pandemic, especially given strong evidence of racial disparities in the impact of the pandemic in the United States (Boserup, McKenney, and Elkbuli 2020; Agarwal et al. 2021). We briefly review some highly dynamic and controversial arguments during the pandemic to provide a solid background for the Bayesian spatial and spatiotemporal analysis and later governance evaluation practices.

The Vulnerable Populations Suffered from the Pandemic Most Severely

Vulnerable populations (i.e., racial and ethnic minority people, economically deprived individuals, and unemployed people) were significantly affected by the pandemic (Guo et al. 2020; Cheung 2022). Studies documented that COVID-19 has had a disproportionate impact on minority communities in terms of higher infection rates, higher mortality rates, and less access to health care (Laurencin and McClinton 2020; Tai et al. 2020; Webb Hooper, Nápoles, and Pérez-Stable 2020). During the initial stages of the outbreak, the infection rate was over three times greater in Black communities when compared to predominantly White communities (Thebault, Tran, and Williams 2020). A recent study systematically reviewed literature published from 3 June 2020 to 31 August 2020. Results revealed that African-American and Latinx populations experienced disproportionately higher COVID-19 infection and death rates (Mackey et al. 2021).

The COVID-19 pandemic has exacerbated the susceptibility of socioeconomically disadvantaged individuals to health risks and daily life challenges. An ecological study that collected data from the seven most severely affected states during the early outbreak of the pandemic indicated that counties with higher poverty rates were associated with higher mortality rates (Abedi et al. 2021). Similar patterns were found in other observational studies (Raifman and Raifman 2020). Another investigation based on hospital data examined the influence of patients' socioeconomic status on COVID-19 outcomes. The study revealed that individuals from low-income communities faced an elevated mortality risk, medical ventilation, and admission to the intensive care unit due to COVID-19 (Quan et al. 2021). Moreover, despite the income transfers allocated by the Coronavirus Aid, Relief, and Economic Security (CARES) Act, the poverty rate continued to escalate from 15.0 percent to 16.7 percent between February and September 2020, with a particularly pronounced trend among Black and Hispanic individuals and children (Parolin et al. 2020). Evidence showed that households below 200 percent of the federal poverty line experienced a notable increase in energy insecurity and utility disconnection during the COVID-19 pandemic (Memmott et al. 2021).

Similarly, unemployed people were more severely affected by the pandemic. Han, Meyer, and Sullivan (2020) reported that the COVID-19 pandemic took a heavy toll on the U.S. labor market. The employment rate fell significantly in April 2020 by over eight percentage points (14 percent), marking recorded history's most prominent single-month decline. Simultaneously, earnings decreased by more than 10 percent at the same time. Although employment and incomes had a slight uptick in May and June, they were still far behind where they were at the beginning of 2020. Studies conducted in the past decades consistently demonstrated a strong association between unemployment and a wide range of adverse health outcomes, including mortality, cardiovascular disease, suicide deaths, and elevated rates of mental distress, substance abuse, depression, and anxiety (Moser, Fox, and Jones 1984; Virgolino et al. 2022).

These observations pose an immediate and complex conundrum for local governance. The first issue to grapple with is whether implementing NPIs would further intensify the difficulties faced by the most vulnerable segments of the population, thereby widening the gaps of social and economic inequality. This issue raises an urgent question: Can NPIs be employed to both mitigate the spread of disease and prevent the exacerbation of existing inequalities?

Conversely, the second issue is equally as challenging: If policies were to prioritize economic performance over public health measures such as NPIs, what would be the potential fallout? One must consider the possible repercussions on public health, social cohesion, and overall quality of life. Would such a strategy yield short-term economic gains at the expense of long-term societal well-being? Would it, ironically, lead to deeper economic crises in the future due to prolonged health impacts?

Thus, it is clear that any approach to this predicament must consider the delicate balance between protecting public health and safeguarding economic stability while also ensuring the fair and equitable treatment of all societal groups. This is the basis and fundamental reason for us to devise a government evaluation scheme so that a clear understanding of the public health consequences of governments' decisions is present while balanced governance is sought.

The Impact of Politics on the NPI Policies

Studies have documented that early NPIs implemented by state governors were significant predictors of decreasing cumulative cases, new cases, and death rates (Guo et al. 2020; White and Hébert-Dufresne 2020). Existing literature reported the crucial impact of partisanship on the effectiveness of implementing social distancing interventions in the United States at the individual level. Using SafeGraph social distancing data, Allcott et al. (2020) demonstrated that within the counties that supported Donald Trump in the 2016 U.S. presidential election, people were less likely to keep social distancing than people in counties that supported Hillary Clinton. Moreover, Painter and Qiu (2020) further revealed that Democrats were less likely to comply with a state-level order issued by a Republican governor than a Democratic one. This intricacy of politics and governance significantly affects how public health interventions are implemented and received. It underscores the intersection of political alignment and compliance with health directives, implying a complex interplay between political, social, and health factors that cannot be overlooked. The implications are manifold, suggesting that the success of such health measures is inextricably tied to the political leanings of the populations they aim to protect.

Indeed, it brings to the fore the crucial role of political leaders in shaping public responses to crises. Notably, their ability to guide their constituents toward behaviors beneficial for public health depends on political affiliations, belief systems, and trust. The evidence just presented indicates that the effectiveness of crucial health interventions during a pandemic, such as social distancing, can be significantly compromised by partisan differences. Ultimately, this intricate entanglement of politics and governance has far-reaching implications for public health, necessitating an approach to pandemic management that is scientifically sound, politically aware, and socially sensitive.

Few studies have ventured this far to untangle the intricacy of the interplay between politics, socioeconomic justice, NPI strategies against a pandemic, and other demands that governing bodies face every day. Adolph et al. (2021) examined the early response of five state-level mitigation intervention policies (i.e., large gathering ban, stay-at-home order, school closure, restaurant restrictions, and nonessential business closure) across all fifty states, with a study window from 26 February 2020 to 23 March 2020. Results indicated that governors' party affiliation was the most crucial predictor of when states issued mitigation interventions; other things being equal, Republican-led states were slower to enact such policies during the early outbreak of COVID-19. These findings were aligned with another study that showed that Democratic governors adopted stay-at-home orders more quickly than their Republican counterparts (Patterson 2022). Still, these studies aimed at understanding the timing of enacting early mitigation policies and state political interaction with such implementation until April 2020, without examining these policies over a more extended period, including issuing, lifting, and reissuing social distancing policies. This study aims to fill in this gap.

Studies of Ranking the Fifty States and Washington, DC, in Terms of Governance Performance

We reviewed the existing literature on rating the efficiency of the state's responses to the COVID-19 pandemic and found a dearth of research conducted in the United States. Xu, Park, and Park (2021) adopted data envelopment analysis and four distinct machine-learning approaches to evaluate response performance to COVID-19 at the state level in the United States. Results showed that twenty-three states were efficient regarding the number of tested, public funding, health care workers, and hospital beds. In addition, Radley, Baumgartner, and Collins (2023) integrated fifty-six measures, including health care access, health inequalities, and health outcomes, to assess fifty states' overall health system performance and readiness in 2020. They ranked Hawaii and Massachusetts at the top and Mississippi, Oklahoma, and West Virginia as the low-performance participants. These studies centered on the evaluation of public health infrastructure in response to the pandemic, rather than understanding the interaction between the governments' efforts and the effectiveness of mitigation policies.

Although these studies provide valuable insights into health infrastructure and readiness, there remains a gap in understanding how state governance practices influenced the effectiveness of pandemic mitigation strategies. We posit that it is essential to evaluate the static conditions of public health infrastructure and how dynamic governmental decision-making can shape outcomes. Our aim in this study is to bridge this gap by developing a ranking system that prioritizes the value of human lives above other societal demands, such as economic development. We integrate this human-centric focus with a robust analysis of state-level governance practices, specifically their enaction or lifting of the NPI strategies during the pandemic, coupled with the modeled efficacy of those NPI strategies. Our study intends to create a more nuanced and comprehensive understanding of how state governments are doing when facing complex and conflicting goals during unprecedented public health emergencies like COVID-19. The study aims to provide an angle that might be of essential importance to guide future governments' balanced practices during emergencies like COVID-19, be it public health, natural disasters, or coping with the increasing crises related to climate change.

Data Collection

COVID-19 Caseload Data

In this study, for a period of 304 days (13 March 2020–10 January 2021), for each state and the District of Columbia, we collected data from eight outcome variables to represent from a broad perspective the spread of COVID-19. They are the cumulative cases, cumulative deaths, new cases, and new deaths per 1,000 people of these four variables (all collected daily). The starting point was chosen because 11 March 2020 marked the enactment of the National Emergency in the United States to fight the pandemic. The endpoint was chosen because the current investigation attempts to isolate the impact of COVID-19 vaccination and instead focuses on the effectiveness and relevant policymaking of the NPI strategies. At the specified time of 10 January 2021, there were only a few health workers who were fully vaccinated in the country. Centers for Disease Control and Prevention (CDC) data show that on 10 January 2021, only 0.35 percent of the U.S. population were fully vaccinated (CDC 2021). This long period provides an excellent opportunity to study how NPIs prevent the spread of COVID-19 and how states use the information for their own policymaking during the pandemic on a daily basis. We use all eight caseload variables to represent the spread of COVID-19 because the spread of the novel coronavirus is a highly complex process and how the data are generated also varies from place to place and time to time. A composite index based on the eight caseload variables might be tempting, but the different recording and reporting mechanisms in different states suggest that a multiple-outcome representation strategy works best to capture the full image of the spread of the virus (Guo et al. 2020). It is critical to study how different NPIs curb the spread of the virus from the broadest possible perspective. In addition, to remove the potential daily fluctuation of the reported data, we applied a five-day-moving-averaging process on the raw data to produce a smooth daily change pattern of the caseloads, which renders the final analysis ranging from 15 March 2020 to 8 January 2021, for a total of 300 days.

NPI Data

Data for the five NPI strategies during the 300-day study period for each state are also collected, including the exact dates and hours of adopting, lifting, and readopting each strategy. These mitigation strategies include the stay-at-home order or advisory, restaurant and bar limit (closure or outside dining only), large gathering ban (no more than ten people), nonessential business closure, and mask-wearing mandate.

A total of 3,703 executive orders pertinent to the NPI orders from the state government Web sites for all fifty states and Washington, DC, were extracted. Figure 1 shows the data extraction process using New Mexico as an example (all other states and NPI measures follow the same procedure and are not repeated here). This is followed by content coding and analysis. Review and coding guidelines were developed in advance. Two researchers from our team separately examined all documents and resolved discrepancies in coding by consensus in a discussion. Considering variation across states in the frequency of use of alternative social distancing strategies and prior literature on social distancing, we extracted information on five types of mitigation interventions. The study window was defined from the declaration of national emergency (13 March 2020) to the moment when major vaccination efforts from the Biden administration had not yet been initiated (10 January 2021).

Social Determinants of Disease Control and Epidemic or Pandemic Prevention

In addition to these daily varying COVID-19 caseload and NPI data, we have also used daily invariant background information for each state. From a disease control perspective, the effectiveness of NPIs against COVID-19 is sensitive to each state's socioeconomic, demographic, and cultural-political backgrounds. Background information was assembled under the guidance of the National Academies of Science, Engineering, and Medicine's report on adapting the World Health Organization's (WHO) "social determinants of health" (SDOH) framework for the United States. The SDOH framework emphasizes the strong relationships between health outcomes and socioeconomic, cultural, ethnic, political, and infrastructural factors. While the changes of COVID-19 spread and the implementation of the mitigation measures for each day are volatile, these changes are related to the comparatively invariant background factors for each state. Based on the National Academies' report under the SDOH framework, this research identified five blocks of state-level factors that might influence the spread of the COVID-19 and the implementation of various NPI mitigation measures. These include (1) demographic factors, which include the percentage of people sixty-five and older in 2018, and the percentage of African and Latino Americans; (2) economic factors, which include per-capita gross domestic product in 2019, the poverty rate in 2018, and the unemployment rate in February 2020; (3) public health infrastructure factors, which include the number of hospital beds per 1,000 population in 2018, the percentage of noninsured individuals in 2018, and per-capita public health budget in 2019; (4) political factors, which include whether the governor was a Democrat in 2020 and whether the state had Democratic senators in 2020; and (5) infrastructure factors, which include the number of international airports in 2020. All data were obtained from the U.S. Census, Bureau of Labor Statistics, Kaiser Family Foundation, and other public data sources, including articles published in the *New York Times*, *Washington Post*, and other relevant news agencies.

Methods

The Inverse-Normal Transformation and the Indirect INT Method

This study uses both daily variant caseload and intervention data and daily invariant state background information. To consider the state background information and control for the relative volatility of the daily varying details, and to avoid potential modeling misspecification with the highly right-skewed and zero-inflated caseload variables (the histograms for all variables are not reported here due to space constraints), we adopt the indirect inverse-normal transformation (INT) and mixed modeling approach proposed in McCaw et al. (2020) to preprocess the data.

Denoting y_{it} the caseload variables for the ith state at time t, $\in \{1, \ldots, T\}$, the analysis followed these steps.

1. Separately for each time point $t \in \{1, \ldots, T\}$, regress each y_{it} on the time-invariant background information w_j ($j = 1, \ldots, k$, k is the number of time-invariant background covariates) to obtain the residuals ε_{it} :

$$\varepsilon_{it} = y_{it} - \sum \beta_j^0 w_j$$

where β_j^0 is the coefficients (not reported) of the jth time-invariant covariate, including the intercept. After this step, the influence of the background SDOH information is removed.

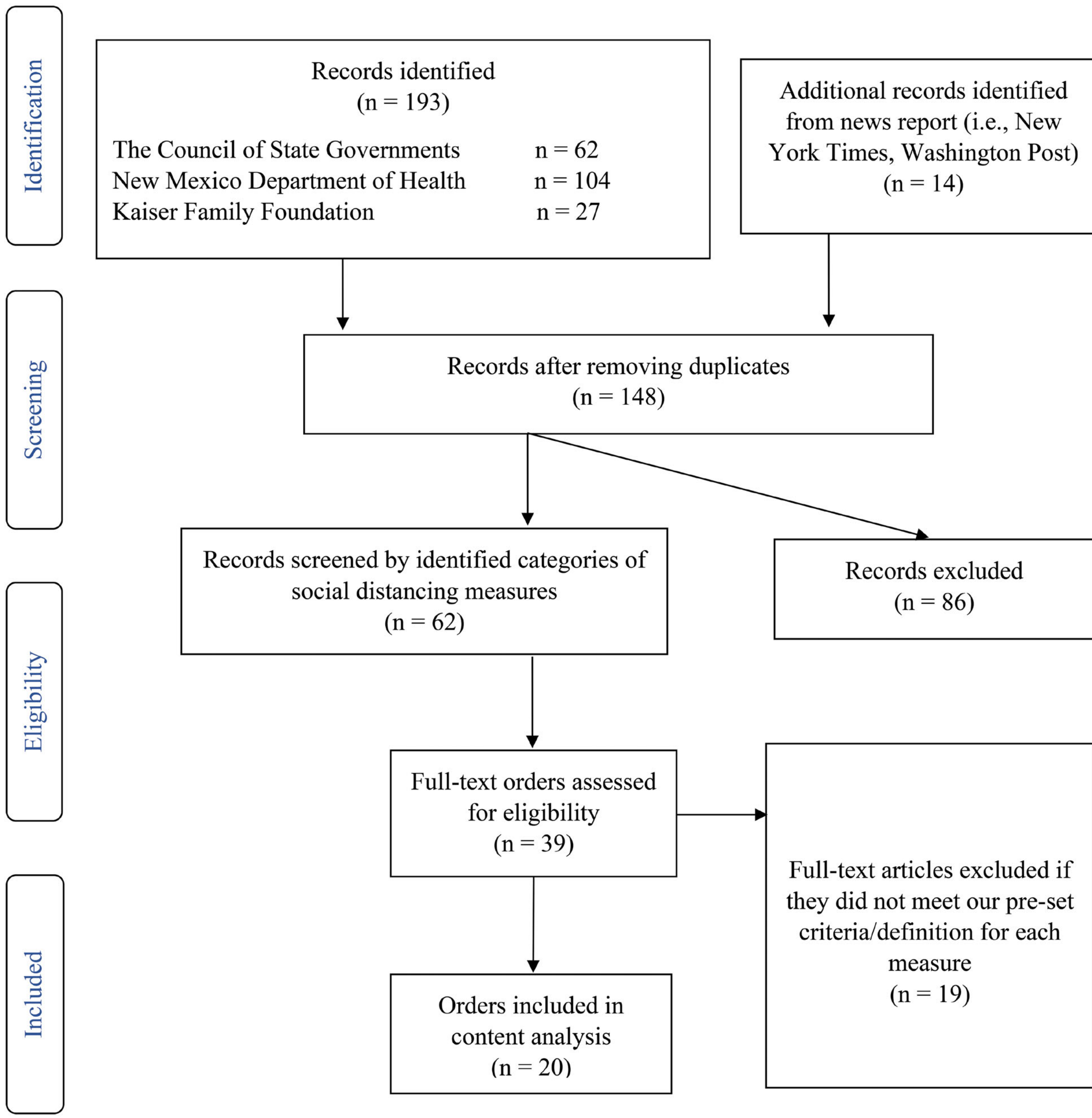

Figure 1. Illustration: The content analysis data search flowchart for New Mexico's social distancing measures.

2. Conduct INT on the residuals to obtain the Z scores $z_{it} \equiv INT(\varepsilon_{it})$, again separately for each time point t. INT procedures follow: Suppose u is a zero-inflated skewed-distributed variable. Let $rank(u_i)$ denote the sample rank of u_i when the measurements are placed in ascending order. The rank-based INT is defined as:

$$INT(u_i) = \phi^{-1}\left[\frac{rank(u_i) - k}{n - 2k + 1}\right].$$

Here ϕ^{-1} is the normal density function or the inverse cumulative distribution function of a normal distribution from which the name of INT comes, $k \in (0, 1/2)$ is an adjustable offset, and n is the sample size. By default, the Blom offset of $k = 3/8$ is adopted (McCaw et al. 2020). This transformation will both eliminate the zero-inflation problem in the data and make the distribution normal. This procedure is performed for both the discrete caseload data and the continuous proportion data as both data sets are heavily zero-inflated.

3. Because the INT transformation is not linear, the transformed z_{it} is again correlated with the background information w_j. To remove this redundant correlation (so that their influences can be fully considered), we regress each of the five time-variant NPI variables, x_{ipt}, $p \in \{1, \ldots, 5\}$, on the time-invariant covariates (background information) w_i to obtain the residuals ε_{ipt} (McCaw et al. 2020):

$$\varepsilon_{ipt} = x_{ipt} - \sum \beta_j^p w_j$$

where β_j^p is the coefficients (not reported) of the pth mitigation measures on the jth time-invariant covariate, including the intercepts.
4. The newly obtained z_{it} and ε_{ipt} represent the outcome variable and mitigation measures that will be used for understanding the impact of NPI mitigation measures on the spread of COVID-19 after controlling for the socioeconomic, cultural, and infrastructure background information for each state (because now the background SDOH variables' influences are fully removed after this procedure). The final model takes the form:

$$z_{it} = \beta_0 + \sum_{i=1}^{p} \beta_i \varepsilon_{ipt} + \varepsilon^*$$

where β_0 is the constant term, β_i is the coefficient of the ith mitigation measure, and ε^* is the model residual.

Method of the Impact Analysis Considering Spatial and Temporal Structures and Scoring Evaluation for State Governments

Bayesian Hierarchical Models to Assess NPI Strategies' Efficacy

To understand how our system of ranking states' responses to the pandemic works, we have used the Bayesian hierarchical model. This model allows us to analyze data collected across different regions and time frames, commonly referred to as geopanel data. Such data often exhibit correlations between nearby geographical regions or similar time points (autocorrelation), a factor we must account for in our models for unbiased and efficient modeling results.

Unlike typical spatial autoregressive models as detailed in Elhorst (2014), or geographically weighted panel regression models as detailed in Yu (2010) and Yu et al. (2023), the Bayesian hierarchical model is uniquely suited to incorporating both spatially and temporally varying coefficient processes explicitly in the modeling framework. This is because the hierarchical structure employed in this model allows the spatial autoregressive structures to be explicitly expressed using certain statistical distributions (Besag 1974; Besag and Green 1993), such as the Matérn kernel for points (Genton 2001), or the Gaussian Markov random field for areal units (Rue and Held 2005). The temporal autoregressive structure is often modeled with a first-order autoregressive process as is typical in time series analysis (Hamilton 2020).

More important, the interactions between spatial and temporal structures that are often treated as latent in other models can also be expressed by an expanded Gaussian Markov random field distribution (Rue and Held 2005). The model hence provides an excellent opportunity for examining the intricate dynamics and potentially varying relationships between daily COVID-19 caseload information and the enaction and lifting of the NPI strategies in each state. After controlling for the temporal, spatial, and spatiotemporal autoregressive structures explicitly using the Bayesian hierarchical structure, we can produce the posteriors for the coefficients at each state and on each day, hence the spatiotemporally varying coefficients. These potentially daily and state-wise varying relationships are the foundation for detailed policy evaluation in our proposed methodological framework.

We have applied three versions of this model, each more complex than the last. The first model incorporates a spatial random effect and a temporally autoregressive prior to accounting for the correlation structures (this corresponds to a nonspatial model). The second model goes a step further and considers the spatial random effect as a Gaussian Markov random field, a technique to model spatially correlated data (this corresponds to a spatial but nonvarying coefficient model). Our third and most advanced model accounts for spatial and temporal varying influences and interactions on both the model's intercept and its slopes. It is called the Bayesian hierarchical spatiotemporally varying coefficient process (BHSTVCP) model. The BHSTVCP model allows these elements to change over time and location, yielding a richer understanding of how NPIs affect COVID-19 caseloads. Detailed methodological elaboration is provided in the Supplemental Material for the interested audience.

The versatility of these models allows us to paint a more precise picture of the NPI measures' daily effectiveness. This picture is then compared with the state governments' actions regarding the NPI measures, creating an evaluation of government performance in terms of battling COVID-19 versus other demands, hence a basis for balanced governance practices.

We have performed these calculations using a technique called integrated nested Laplace approximations (INLA), a deterministic algorithm proposed by Rue, Martino, and Chopin (2009) and implemented in the R platform for efficient and accurate Bayesian simulation. The availability of the INLA algorithm enables the complex BHSTVCP model to be estimated with reasonable computational cost (less than ten minutes) on a Microsoft Windows workstation with a Xeon Platinum 8280 CPU, fifty-six cores, and 512 GB of RAM. This complex analysis offers us a nuanced view of each state's COVID-19 responses while considering the multiple demands they face.

Applying the spatiotemporally varying coefficients model provides a channel to produce a daily effectiveness impression of the five NPI measures (whether or not the NPI measures are significant on each day). When comparing this effectiveness impression with the state government's actions in terms of the five NPI measures, it serves the purpose of evaluating balanced government performance between fighting against the spread of COVID-19 and other urgent needs. We detail the evaluation procedure with a scoring system in what follows.

Scoring and Ranking of States' Efforts

To evaluate how governments strike the balance, we devise a 4-point Likert scale to score the intervention performance for each state on each day of the study period from 15 March 2020 to 8 January 2021. The scoring system was devised by applying a counterfactual framework for program evaluation. The framework was originally developed by Neyman (1923) for evaluations of programs using a randomized experiment, and later expanded by Rubin (1974, 1986) to evaluations of programs with a nonrandomized quasi-experiment.

A counterfactual is a potential outcome, or the state of affairs that would have happened in the absence of the cause. The counterfactual framework emphasizes that program participants selected into either treatment or nontreatment conditions have potential outcomes in both states: the one in which they are observed and the one in which they are not observed. The difference between the observed and potential outcomes is the treatment effect. Under the current context, a state actually implementing an intervention is the outcome of treatment; likewise, a state not implementing an intervention is the outcome of nontreatment. Using the BHSTVCP model-predicted coefficients, we define the model-predicted significance of implementing a particular NPI as potential outcome and whether a state actually implemented the intervention as an observed outcome. A state government might or might not take a treatment because of various factors affecting the decision-making, notably economic, health, and political factors. States that should have taken certain NPIs to fight against the spread of COVID-19 might not choose to do so. Hence, whether the decisions made by the states align with the epidemiological requirement—defined here as the model-predicted significance of implementing a certain NPI—represents the state's performance in fighting the spread of the disease. The scoring system employed a 4-point Likert scale to measure the treatment effects, or various levels of the difference between the counterfactual (i.e., what would have happened had the state "followed" the action suggested by the BHSTVCP model) and the actual outcome, with 4 being excellent, 3 being very good, 2 being fair, and 1 being poor. We gave the scores based on the following comparisons between whether a state implemented the intervention and whether the BHSTVCP model showed that the implementation would significantly reduce COVID-19 caseloads (based on the null hypothesis that there was no difference between pre- and postgovernment intervention, and a one-tailed credibility level of 95 percent). It is also to be noted, however, that policy-making in the face of a new pandemic is fundamentally a choice under (very great) uncertainty. Our evaluation metrics assess but one aspect (fighting against the deadly virus) of the full spectrum of governance.

1. If a state implemented an intervention, and the BHSTVCP model shows the intervention was not significant on that day, we gave the state a score of 4, to acknowledge the government's precautionary efforts. This is because the actual implementation of

any of the interventions was a decision made under very great uncertainty. If an intervention would not be effective as reported by the model, the governor still made the decision, which means that the decision-maker was more cautious. The governor's extra effort to prevent the spread of COVID-19 warrants the highest points.

2. If a state implemented an intervention, and the BHSVCP model shows the intervention was significant on that day, we gave the state a score of 3. If the governor enacted a measure that turned out to be necessary for that day, it means that the governor made a correct decision. In this context, the governor is awarded the next highest point, as opposed to the previous situation in which the governor made a more prudent decision.
3. If a state did not implement an intervention, and the BHSVCP model shows the intervention was not necessary (statistically not significant) on that day, we gave the state a score of 2. In this case, the governors were just lucky.
4. If a state did not implement an intervention, but the BHSVCP model shows the intervention was necessary on that day, we gave the state a score of 1. This is the scenario when the balanced scale tipped toward other priorities than preventing the spread. The daily scores are then added up over the 300 days to tell the story of the governments' efforts to consume scientific evidence in preventing the spread of COVID-19.

Results

The raw data were first transformed through the indirect INT approach and then fed to the Bayesian nonspatial, spatial, and BHSTVCP models. The models produce sixteen sets of global (eight Bayesian hierarchical nonspatial and spatial models, respectively) and 122,400 sets of locally varying relationships between the five NPI strategies and the spread of COVID-19 in the United States (in total, there are $8 * 5 = 40$ regression coefficients each for the Bayesian hierarchical nonspatial and spatial models, and $122{,}400 * 5 = 612{,}000$ spatiotemporally varying regression coefficients for the BHSTVCP models). The global coefficients are reported in Table 1. The saturated deviance information criterion (DIC) that are used to compare the Bayesian hierarchical models are also produced and reported in Table 2. These results provide solid evidence to access the effectiveness of the COVID-19 intervention and evaluate how such information is consumed during the policymaking process in different states in the United States.

Results from the Bayesian Nonspatial, Spatial, and BHSTVCP models

Following the specifications of the global models, we calibrated eight nonspatial Bayesian models (spatial random effect assumed to be i.i.d with temporal effect), eight spatial (but not spatiotemporally varying) Bayesian models (spatial random effect assumed to be "besag" with temporal effect) and eight BHSTVCP models for the eight outcome variables regressed on the five NPI measures after the indirect INT transformation.

Quite interestingly, for the Bayesian hierarchical models, the nonspatial model suggests that none of the NPIs work against the spread of COVID-19 (as all the one-tailed credibility interval contains the value zero), but when spatial information is modeled with the prior "besag," "stay-at-home order" and "no large gathering" are shown to be statistically significant to prevent the spread of COVID-19 (Table 1), but all other NPI mitigation strategies remain nonsignificant. This practice resonates with many geographical studies that when data were collected over geographical spaces (or geo-temporal spaces), considering the autocorrelation structure of the data makes quite a difference in modeled results and modeling performance (Anselin 1988; LeSage and Pace 2009). Still, these models are global in nature and detailed local patterns require the varying coefficient process model.

From the BHSTVCP models, for each of the five NPI measures, the models produce 8 (the number of outcome variables) * 51 (the number of states) * 300 (the number of days) = 122,400 spatiotemporally varying coefficients. There are in total $5 * 122{,}400 = 612{,}000$ spatiotemporally varying coefficients produced. In addition, for these 612,000 varying coefficients, we conducted the one-tailed significance (95 percent confidence level) test for each individual estimated coefficient to see if the NPI measure effectively prevents the spread of COVID-19 on the specified day and state.

The amount of information produced by the BHSTVCP models is rich. It will not be practical to visualize all 612,000 varying coefficients, nor will it be informative. In addition, the actual values of the

Table 1. Results of nonspatial and spatial nonvarying coefficient Bayesian hierarchical models

		Intercept	Stay-at-home order	Restaurant closure	No large gathering	Nonessential business closure	Mask wearing
Nonspatial							
Cumulative cases	Estimate	0.000	−0.007	0.008	−0.002	0.001	−0.002
	95% CI	<0.722	<0.001	<0.016	<0.005	<0.009	<0.009
Cumulative deaths	Estimate	0.000	0.001	−0.005	−0.001	0.034	0.005
	95% CI	<0.182	<0.009	<0.003	<0.006	<0.042	<0.016
New cases	Estimate	0.000	−0.012	0.004	0.022	−0.010	0.024
	95% CI	<0.116	<0.027	<0.040	<0.056	<0.029	<0.076
New deaths	Estimate	0.000	−0.027	0.026	0.040	0.037	0.012
	95% CI	<0.093	<0.029	<0.080	<0.089	<0.094	<0.087
Cumulative cases per 1,000 people	Estimate	0.000	0.014	−0.006	0.002	0.026	0.002
	95% CI	<0.189	<0.024	<0.004	<0.010	<0.036	<0.015
Cumulative deaths per 1,000 people	Estimate	0.000	0.014	−0.006	0.002	0.026	0.002
	95% CI	<0.192	<0.024	<0.004	<0.010	<0.036	<0.015
New cases per 1,000 people	Estimate	0.000	−0.058	0.003	0.020	0.061	−0.039
	95% CI	<0.072	<0.013	<0.072	<0.084	<0.135	<0.056
New deaths per 1,000 people	Estimate	0.000	−0.030	0.048	0.096	0.080	−0.091
	95% CI	<0.052	<0.057	<0.137	<0.176	<0.175	<0.019
Spatial nonvarying							
Cumulative cases	Estimate	0.000	−0.009	0.003	−0.004	−0.003	0.001
	95% CI	<1.638	<−0.001	<0.010	<0.003	<0.005	<0.011
Cumulative deaths	Estimate	0.000	−0.001	0.004	−0.003	0.026	0.009
	95% CI	<1.637	<0.007	<0.012	<0.004	<0.034	<0.020
New cases	Estimate	0.000	−0.030	0.025	0.017	0.001	0.011
	95% CI	<1.635	<0.009	<0.061	<0.050	<0.039	<0.062
New deaths	Estimate	0.000	−0.024	0.009	0.051	0.032	0.009
	95% CI	<1.637	<0.032	<0.061	<0.099	<0.088	<0.081
Cumulative cases per 1,000 people	Estimate	0.000	−0.004	−0.004	−0.011	0.018	0.008
	95% CI	<1.642	<0.006	<0.006	<−0.003	<0.028	<0.021
Cumulative deaths per 1,000 people	Estimate	0.000	0.017	0.003	0.001	0.023	0.000
	95% CI	<1.642	<0.027	<0.013	<0.009	<0.033	<0.013
New cases per 1,000 people	Estimate	0.000	−0.059	0.035	0.015	0.080	0.012
	95% CI	<1.642	<0.010	<0.101	<0.074	<0.151	<0.101
New deaths per 1,000 people	Estimate	0.000	0.019	0.053	0.055	0.065	−0.072
	95% CI	<1.638	<0.108	<0.140	<0.132	<0.158	<0.040

coefficients are likely of less interest under the spatiotemporally varying modeling scheme. The spatiotemporally varying pattern is of more interest instead. Moreover, it is necessary to summarize these coefficients along with their varying z scores that indicate whether they are statistically significant or not.

To create the spatiotemporally varying pattern, the first thing we did was pick out the coefficients with z scores that suggest their 95 percent credibility intervals do not contain the value zero. We then combine the coefficients over the eight COVID-19 caseload variables because we regard all eight variables as different aspects of the spread of COVID-19. During the combination, because the values of the coefficients are of less interest, for each NPI mitigation strategy, on each state and each day, if one of the eight coefficients is statistically significant (95 percent credibility interval does not contain the value zero), the first such coefficient will be retained for that NPI mitigation strategy on the specific state and day. After these two steps, the total number of coefficients that we use for visualization is reduced from 612,000 to 25,000 (10,229 for stay at home, 2,850 for restaurant and bar limits, 9,521 for large gathering ban, 1,658 for nonessential business closure, and 742 for mask-wearing mandate). Using time as the horizontal axis, and state and the varying

Table 2. Saturated deviance information criterion (DIC) across the three Bayesian hierarchical models

Caseload variables	Nonspatial model	Spatial model	Varying coefficient model
Cumulative cases	30,500.18	30,466.40	−114,971.330
Cumulative deaths	30,535.86	30,488.59	−120,081.633
New cases	30,422.73	24,175.28	−18,805.373
New deaths	27,790.51	23,684.59	−6,311.103
Cumulative cases per 1,000 people	30,418.24	29,196.28	−142,402.369
Cumulative deaths per 1,000 people	30,514.70	30,395.58	−120,547.353
New cases per 1,000 people	28,564.83	23,872.18	−1,868.621
New deaths per 1,000 people	29,465.73	23,457.20	−128,070.319

coefficients as vertical axes, we can produce the five summarized charts shown in Figures 2 through 6. All charts are produced in R (R Core Team 2022).

The immediate impression from the five charts is that the effectiveness (significant NPI mitigation strategies' coefficients) varies widely in both geography and time. Although there is information loss during the summarization and combination procedure, it is clear that contrary to the nonvarying model, all NPI mitigation strategies work to prevent the spread of COVID-19 at some places and sometimes, although the statistical significance varies in different places and at different times. The result agrees in general with many previous studies (Abel and McQueen 2020; Anderson et al. 2020; Auger et al. 2020; Güner, Hasanoğlu, and Aktaş 2020; Brauner et al. 2021) and common practices in epidemiological scenarios. More important, the saturated DIC values reported in Table 2 are evidence that the varying coefficient models perform far better than the global spatial models, which perform slightly better than the global nonspatial models. It is worth noting here, though, the great change of the saturated DIC value from the global to local model might be because of the added flexibility of the local model that considers not only the spatial and temporal structures of the data, but also their interactions. We do invite the readers to be cautious when interpreting the fitness of the models to the data. It might be more prudent to situate different models to different application scenarios instead of comparing the models and celebrating the local models' seemingly superiority. In addition, the results from the local models also suggest that the effectiveness of the NPI mitigation strategies might be better evaluated at the local (individual state) level instead of the global (collective state) level. In countries like the United States in which each local administrative unit (state) has its own unique policymaking and implementation practices and priorities, a global model that summarizes over the vast details might not provide specific enough information to evaluate the effectiveness of some policies. Local models, on the other hand, although specifically focusing on the details of the individualized differences, could produce more reasonable results for policy evaluation and hence guide policymaking and governance in the future.

Evaluation of Governments' Performance with the BHSTVCP Model

To use this rich amount of information to evaluate the state governments' performance during the pandemic, we summarize the information by making the following calculations, ranking, and mapping the results in Table 3 and the five maps (Figure 7).

First, for any specific state and day, if one of the five NPI strategies is significantly effective on any one of the eight outcome variables, the NPI measure will be marked as significantly effective to prevent the spread of COVID-19 for that state on that day. The number of days that NPI measure is significantly effective for each state is then recorded. This information is visualized in Figures 2 through 6 with the five charts.

Second, the number of days a specific NPI measure is significantly effective for each state is compared with the number of days that the NPI measure is implemented for the corresponding state. The comparison of all fifty-one states is shown as the bar charts in the maps (Figure 7).

Third, a 4-point Likert scale is used to score the intervention performance for each state on each day of the study period from 15 March 2020 to 8 January

2021. The scale uses the following values to represent the performance of a state government against the spread of COVID-19: 4 = excellent, 3 = very good, 2 = fair, 1 = poor (details given earlier). After each state is scored for all 300 days, the scores are summed to form the basis of ranking and evaluating how states consume the scientific evidence that NPIs are effective against COVID-19. The summed scores for all the states are presented in Table 3. The ranking was the ranking order of the score that further considered tied states. The ranking is also gray-scaled for each state in the five maps in Figure 7.

Discussion and Conclusion

The Overall Lesson

To our first research objective and questions, our results clearly suggest that implementation of NPI strategies has a significant impact on the spread of COVID-19. This is especially true when the spatial and temporal structure of our data is incorporated within the model.

The research team combed through thousands of governmental Websites, policy-related documents, and other sources to build a full set of COVID-19 caseload and NPI strategies data sets. We investigated the policymaking and governance practices at the state level during the pandemic by unraveling how policymakers acted in such difficult times through advanced spatiotemporal modeling. Global and local models are employed to fully use the very rich spatial and temporal information in the collected data. Both the global and local models' results clearly suggest NPI mitigation measures work to curb the spread of the disease, if not all the time and at all places, at least in some places (most of the places) and at some times (most of the time). From a general, global perspective, Table 1 suggests the stay-at-home order or advisory and no large gatherings work to curb the spread under the Bayesian hierarchical modeling framework across the spectrum of the eight COVID-19 spread outcome variables when spatial structure is explicitly modeled. This result agrees with previous studies that government enforcement and social distancing play a significant role in curbing the spread of COVID-19, at least from an overall perspective, as witnessed in many other countries (Dehning et al. 2020; Guo et al. 2020; Brauner et al. 2021). The global models provide evidence that governmental attitude and guidance effectively guide the public to fight back a global pandemic.

The spread of COVID-19 and the fight against it in the United States for the years 2020 and 2021 were a struggle for both the people and the governments. Preventing the spread of the disease through NPI measures (e.g., stay-at-home orders and advisories, social distancing mandates, mask-wearing mandates, and business closures, among many others) to mitigate the most severe effects of the pandemic on the people and the economy and to maintain social justice were constantly in conflict during the pandemic. Whether the government should enact NPI strategies, whether individuals should get vaccinated, and whether travel bans (both domestic and international) should be enforced are questions that attracted vigorous debates that did not always rely on scientific evidence. Balancing between NPI implementation and the perceived impacts on economic performance, social justice, and many other day-to-day businesses was often the issue causing the most conflict as state governors made decisions. Although this implied the weighing of the value of a human life against other factors such as economic costs and freedoms, governors and other policymakers made these decisions without using explicit value-of-life studies or considering the complex ethics of such decisions. Considering governmental attitudes and particularly policies could have far-reaching effects on individuals' behaviors (Patrick and Cormier 2020; Brauner et al. 2021; Kerr, Panagopoulos, and van der Linden 2021), combing through the day-to-day chaos of COVID-19 cases, day-to-day government-issued orders, advisories, and mandates that attempt to battle the spread of COVID-19, and struggle to maintain a feasible socioeconomic performance and justice, is of critical importance for science-based policy studies.

The global models, however, assume the regressed relationships between the outcomes and NPIs stay the same in places and times. The geographic (individual) and temporal variation effects are usually not of concern. The results we see here only represent an aggregated overall picture of the relationships. Although the results are useful for a quick evaluation of overall policies, we feel a more detailed understanding of the intertwined relationships between implementation of NPI strategies and the spread of COVID-19 is essential for governance

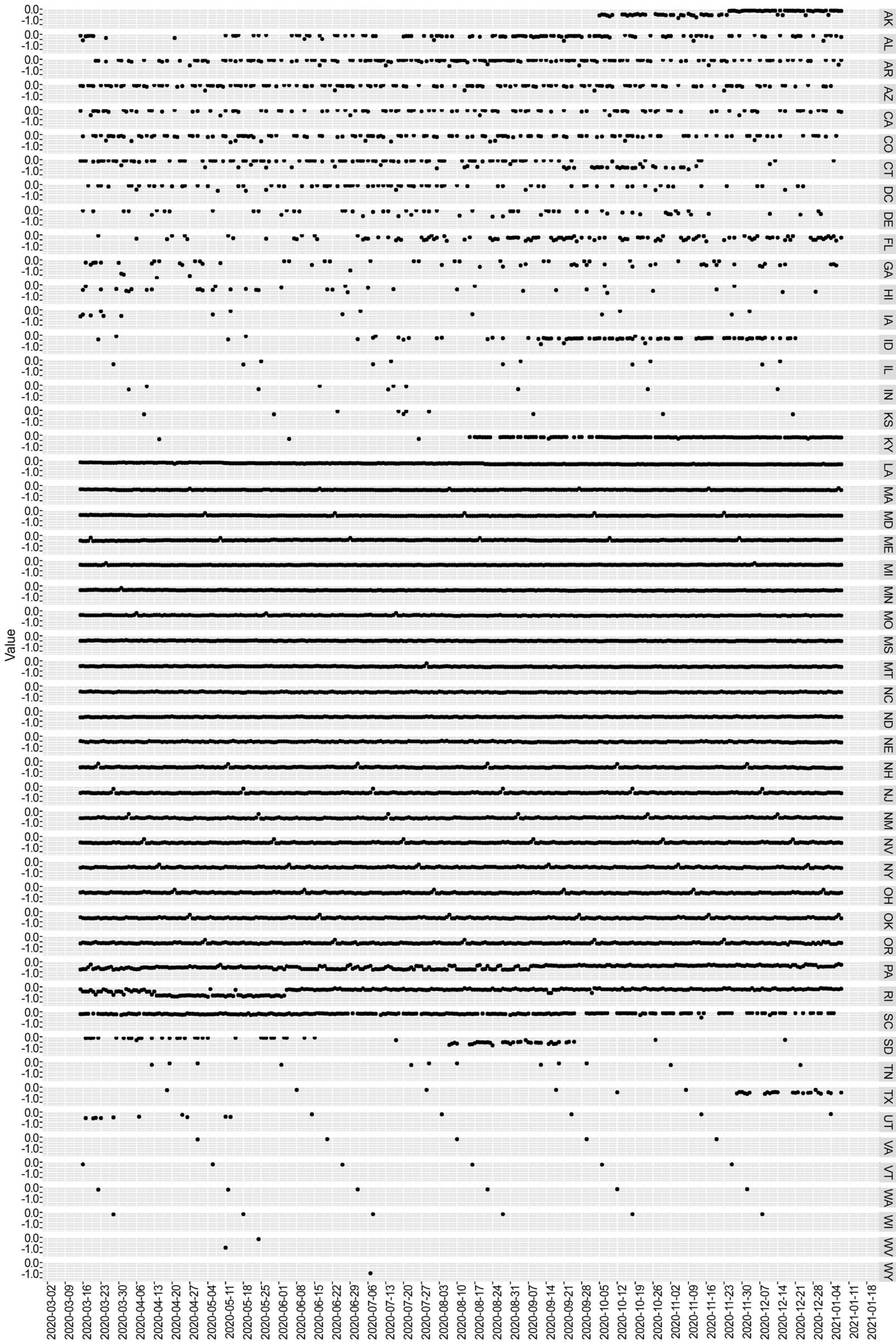

Figure 2. Spatiotemporally varying significant coefficients for stay-at-home orders.

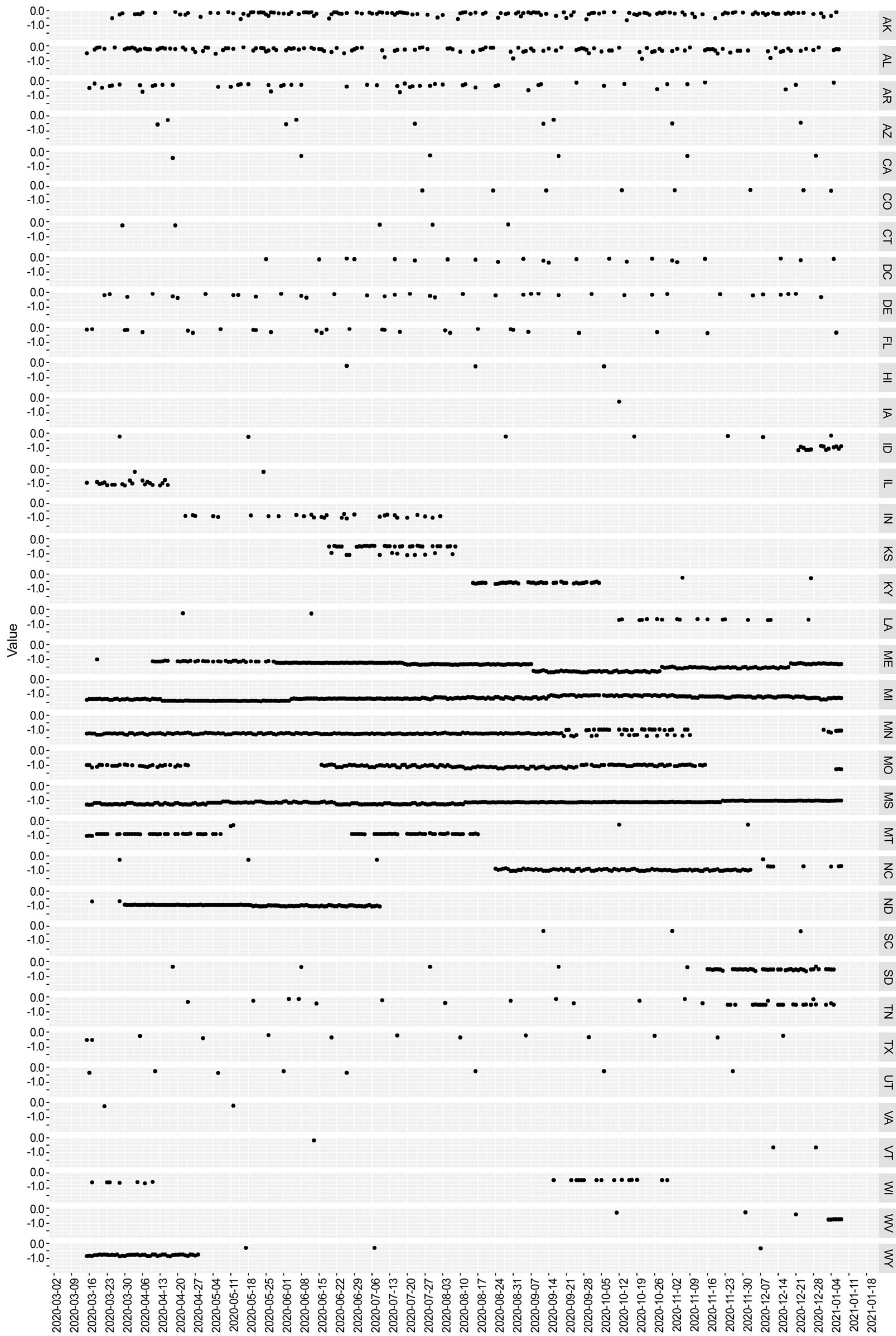

Figure 3. Spatiotemporally varying significant coefficients for restaurant and bar limits.

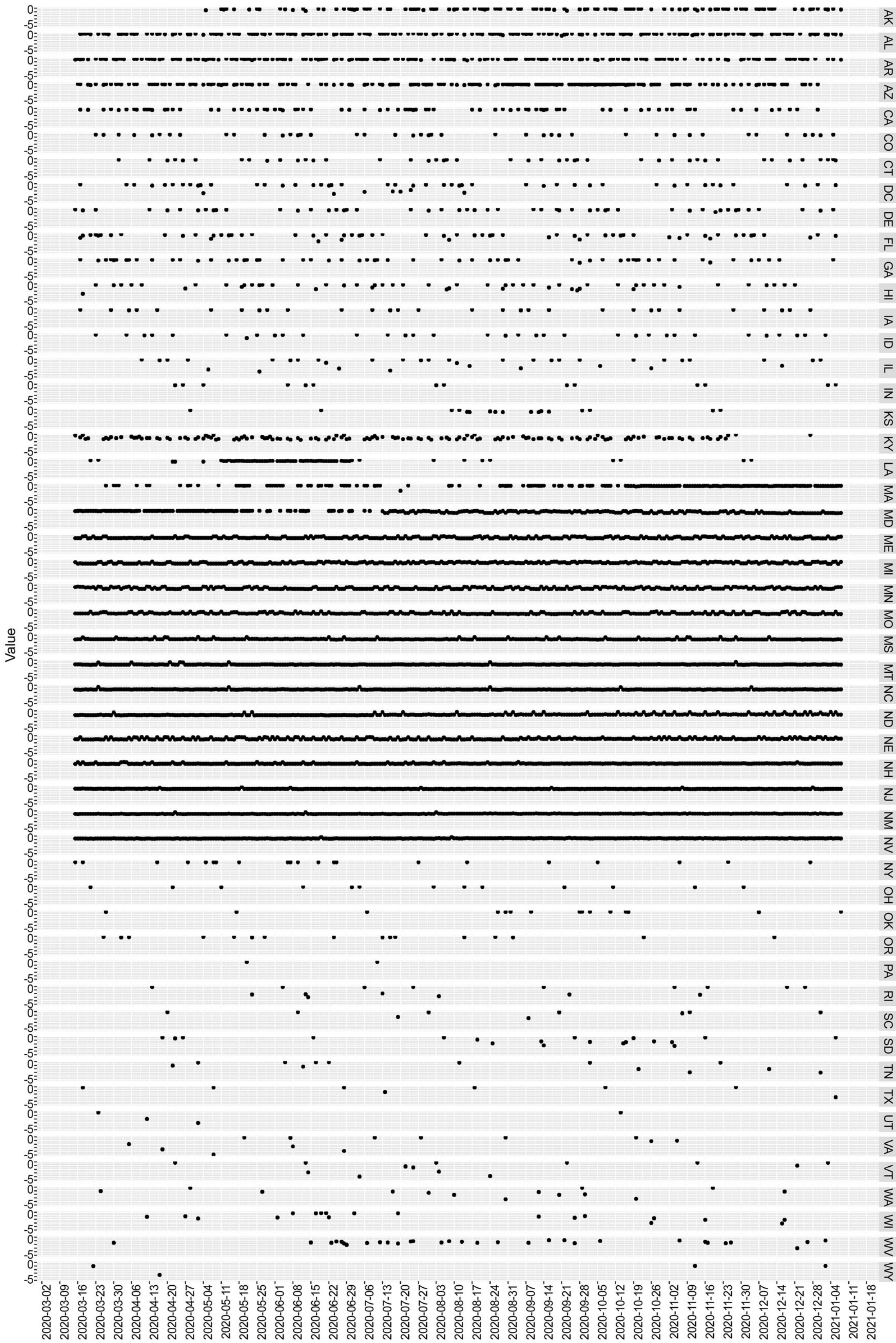

Figure 4. Spatiotemporally varying significant coefficients for large gathering bans.

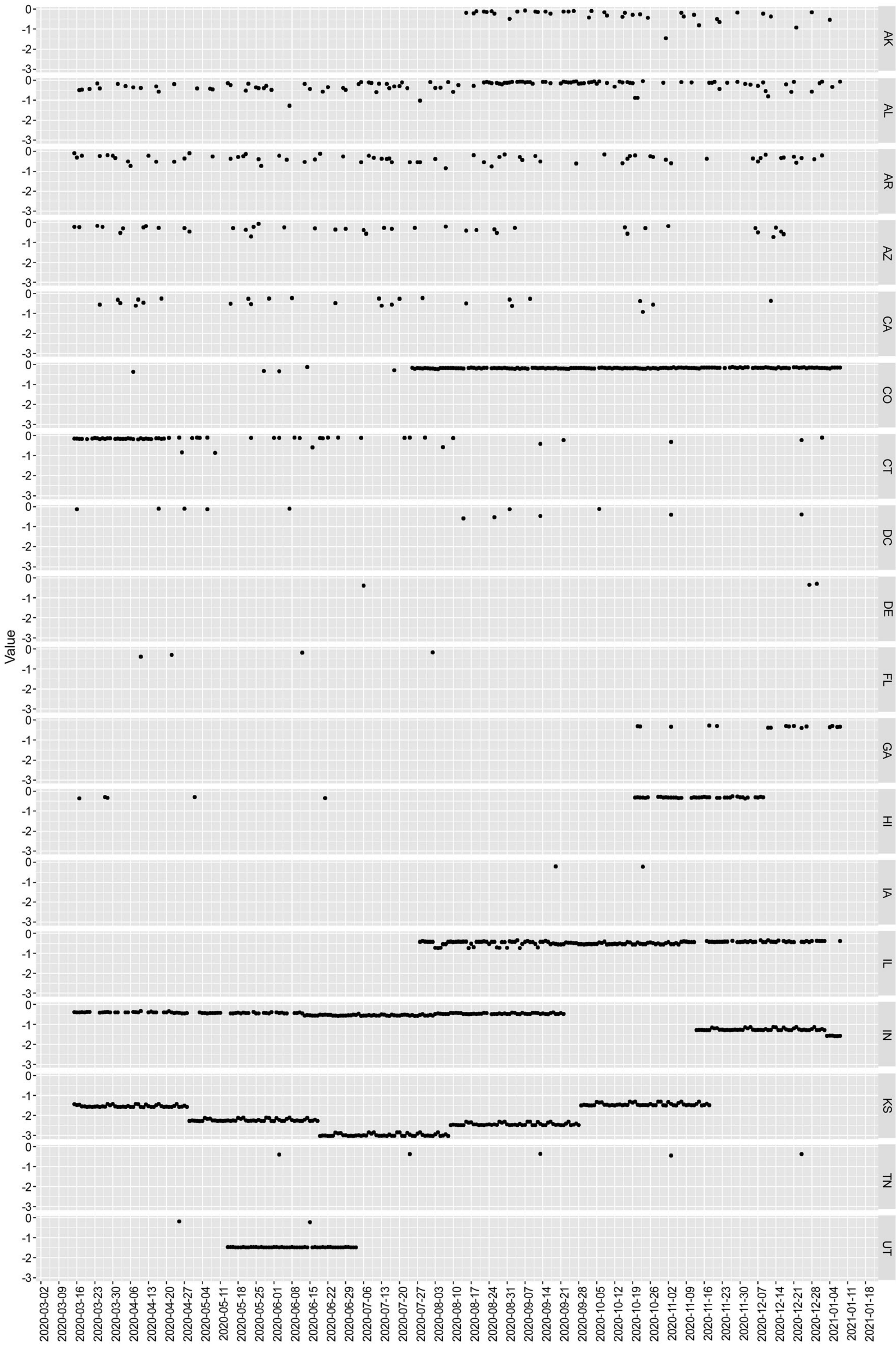

Figure 5. Spatiotemporally varying significant coefficients for nonessential business closures.

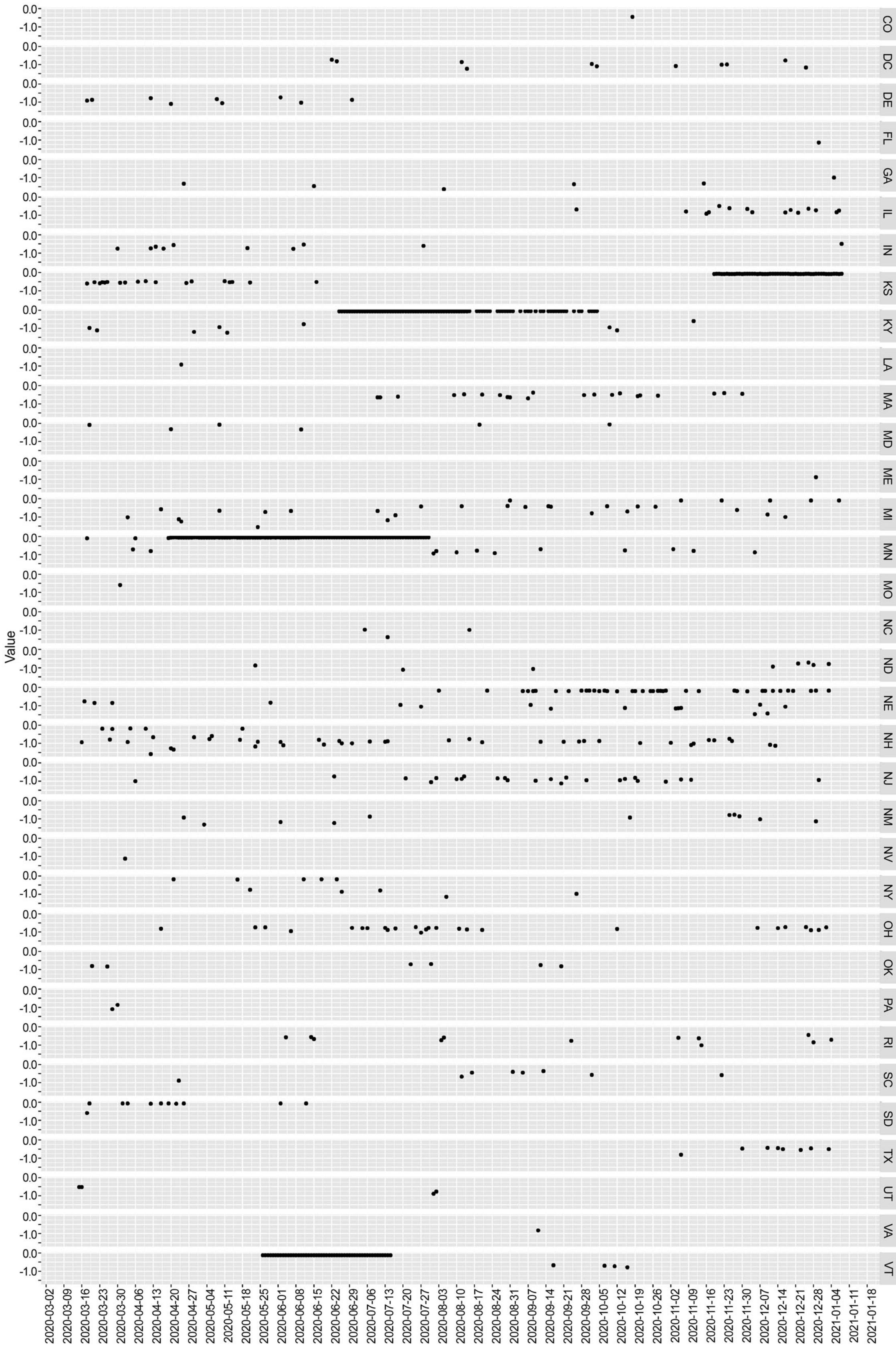

Figure 6. Spatiotemporally varying significant coefficients for mask-wearing mandates.

Table 3. The scores and ranking of the five intervention measures for each state.

Stay-at-home order			Restaurant/bar limits			Large gathering bans			Nonessential business bans			Mask-wearing mandate		
State	Score	Rank	State	Score	Rank	State	Score	Rank	State	Score	Rank	State	Score	Rank
West Virginia	580	1	New Jersey	344	1	Ohio	293	1	New Jersey	186	1	Connecticut	528	1
Colorado	447	2	California	330	2	New Mexico	292	2	New Mexico	168	2	New York	524	2
Alabama	441	3	New Mexico	300	3	Wisconsin	281	3	Michigan	166	3	Maine	505	3
California	412	4	Washington, DC	208	4	Rhode Island	268	4.5	Maine	158	4	Delaware	503	4
New Mexico	282	5	Oklahoma	194	5.5	Colorado	268	4.5	Minnesota	150	5	Illinois	491	5
Illinois	227	6	Massachusetts	194	5.5	New York	263	6	Pennsylvania	142	6	Rhode Island	480	6
Wisconsin	212	7	Pennsylvania	192	7	Illinois	258	7	New York	136	7	Massachusetts	475	7
Virginia	195	8	Washington	188	8	Texas	248	8	Vermont	134	8	New Mexico	465	8
Washington	134	9	New York	182	9.5	Vermont	245	9	New Hampshire	130	9	Virginia	449	9
Delaware	130	10	New Hampshire	182	9.5	Washington	222	10	Rhode Island	126	10	California	410	10
Hawaii	99	11	Illinois	181	11	Washington, DC	217	11	Iowa	124	11.5	Washington	394	11
Vermont	96	12	Connecticut	179	12	Idaho	210	12	Massachusetts	124	11.5	Nevada	393	12
Washington, DC	69	13	Maryland	176	13	Virginia	183	13	Kentucky	120	13	North Carolina	391	13
Indiana	66	14	Oregon	168	14	Wyoming	170	14	Washington, DC	118	14	Pennsylvania	382	14
Kansas	60	15	Vermont	157	15	Oklahoma	165	15	California	112	15	West Virginia	372	15.5
Tennessee	48	16	Rhode Island	152	16	Tennessee	145	16	Maryland	106	17	Texas	372	15.5
Texas	28	17	Hawaii	149	17	Iowa	122	18	Louisiana	106	17	Louisiana	359	17
Wyoming	−1	18	Virginia	142	18	Delaware	122	18	North Carolina	106	17	Oregon	356	18
Idaho	−6	19	Kentucky	139	19	Connecticut	122	18	Oregon	104	20	Alabama	354	19
Georgia	−9	20	Ohio	134	20.5	South Carolina	121	20	Ohio	104	20	Colorado	351	20
Utah	−15	21.5	Colorado	134	20.5	Oregon	120	21	Idaho	104	20	New Jersey	346	21.5
Iowa	−15	21.5	West Virginia	119	22	Pennsylvania	116	22	Virginia	102	22	Arkansas	346	21.5
Connecticut	−18	23	Delaware	117	23	Kentucky	108	23	Nevada	98	23	Michigan	335	23
Alaska	−29	24	Wisconsin	114	24	Kansas	99	24	South Carolina	94	24.5	Washington, DC	331	24
Arizona	−50	25	South Carolina	105	25	Hawaii	98	25	Wisconsin	94	24.5	Indiana	322	25.5
Florida	−52	26	Louisiana	99	26	Florida	84	26	Delaware	87	26	Wisconsin	322	25.5
Massachusetts	−62	27	Nevada	98	27	Georgia	79	27	Washington	86	27	Maryland	318	27
Kentucky	−70	28.5	Arizona	95	28	Massachusetts	75	28	Hawaii	84	28	Ohio	314	28
South Dakota	−70	28.5	Iowa	89	29	Indiana	66	29	West Virginia	82	29	Kansas	311	29
Ohio	−84	30	Indiana	83	30.5	California	61	30	Connecticut	80	30	Kentucky	270	30
Rhode Island	−92	31	Idaho	83	30.5	South Dakota	53	31	Wyoming	74	31	Vermont	267	31
Oregon	−124	32	Utah	78	32	West Virginia	51	32	Florida	72	32.5	Minnesota	219	32
New York	−132	33	Texas	69	33	Louisiana	43	33	Texas	72	32.5	Mississippi	154	33
New Hampshire	−140	34.5	Georgia	68	34	Arizona	38	34	Mississippi	70	34	Utah	118	34
New Jersey	−140	34.5	Wyoming	66	35	Utah	26	35	Tennessee	65	35	Hawaii	108	35
Arkansas	−156	36	Florida	62	36.5	Minnesota	−38	36	Montana	60	36	Iowa	106	36
Michigan	−162	37	Nebraska	62	36.5	Alaska	−49	37	Oklahoma	58	37	North Dakota	104	37
North Carolina	−164	38	Arkansas	51	38	Maryland	−55	38	Missouri	56	38	Montana	100	38
Mississippi	−182	39.5	Tennessee	34	39	Alabama	−91	39	Alaska	43	39	Wyoming	62	39
Maine	−182	39.5	South Dakota	26	40	Arkansas	−96	40	Arizona	27	40	New Hampshire	54	40
South Carolina	−191	41	North Carolina	19	41	New Hampshire	−130	41	North Dakota	0	42	Alaska	0	42.5
Louisiana	−196	42	Montana	9	42	New Jersey	−140	42	South Dakota	0	42	Idaho	0	42.5
Minnesota	−198	43	North Dakota	−20	43	Nebraska	−146	43	Nebraska	0	42	Tennessee	0	42.5
Maryland	−208	44	Alabama	−32	44	Maine	−150	44	Illinois	−11	44	Arizona	0	42.5
Nevada	−210	45	Kansas	−41	45	Michigan	−162	45.5	Georgia	−16	45	Missouri	−1	45.5
Pennsylvania	−226	46	Alaska	−46	46	Montana	−162	45.5	Alabama	−32	46	Florida	−1	45.5
Montana	−242	47	Minnesota	−66	47	Nevada	−168	47	Utah	−52	47	Oklahoma	−6	47.5
Missouri	−244	48	Missouri	−97	48	Mississippi	−182	48	Arkansas	−70	48	Georgia	−6	47.5
Nebraska	−258	49	Maine	−140	49	North Carolina	−194	49	Colorado	−93	49	South Carolina	−8	49
Oklahoma	−300	50.5	Michigan	−165	50	Missouri	−216	50	Indiana	−131	50	South Dakota	−11	50
North Dakota	−300	50.5	Mississippi	−212	51	North Dakota	−300	51	Kansas	−151	51	Nebraska	−55	51

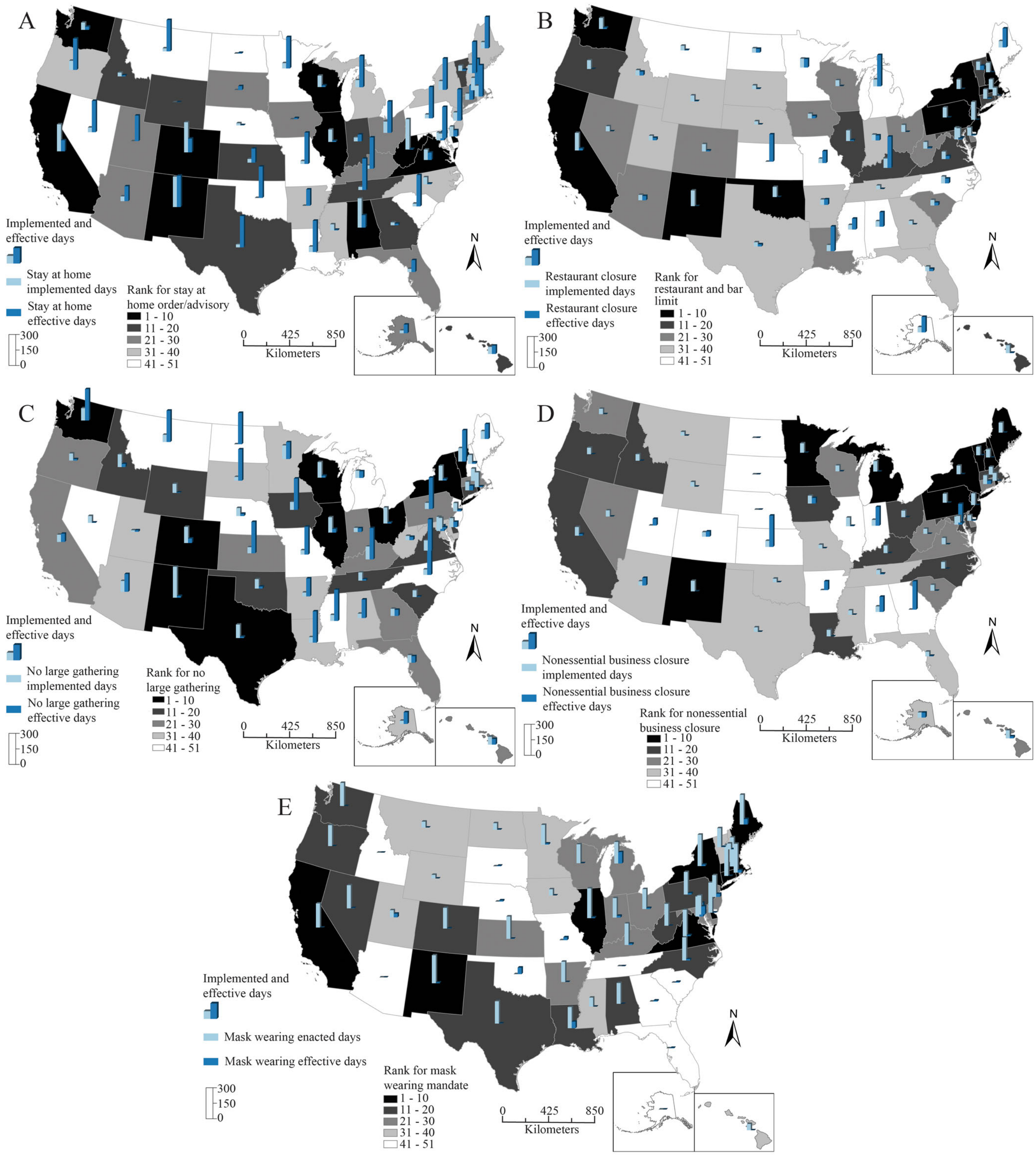

Figure 7. Contrast of implemented and effective days of the five mitigation measures and the rank for the fifty states and Washington, DC. (A) Implementation, effectiveness, and rank of stay-at-home orders; (B) Implementation, effectiveness, and rank of restaurant and bar limits; (7C) Implementation, effectiveness, and rank of no large gatherings; (D) Implementation, effectiveness, and rank of nonessential business closures; and (E) Implementation, effectiveness, and rank of mask-wearing mandates.

policy evaluation. The results from the global model invite further investigation. We do, however, realize that there are many factors other than the NPI strategies that are at work to curb the spread of COVID-19. We also realize that the administrative units that implement NPI strategies (municipalities) and the data collection units (states) are not always the same, so this modifiable areal unit problem (MAUP) might also mean this analytical result needs to be taken with caution and merits further investigation.

The Detailed Pattern

To our second research objective and question, with the Likert scale system and results from the spatiotemporally varying coefficient model, we are able to generate a governance policy evaluation scheme. The scheme compares the modeled spatiotemporally varying coefficients and the implementation of NPI strategies at each state on each day.

First, we found that under the Bayesian analytical framework, the Bayesian hierarchical spatiotemporally varying coefficient model poses the capability to unravel the seemingly chaotic relationships caused by the MAUP. The model combines the data and our prior experiences regarding how policies will interact over geographic units and temporal periods and allows the coefficient to vary in both space and time. This model is able to examine the relationship at a more detailed local level.

Second, supplementary to the global model, the local model suggests that all NPIs work in some place and on some days during the study period. Not surprisingly, stay-at-home orders or advisories and large gathering bans are the two factors that are most effective because they work on most days and in most places.

Based on the results, we contend that individuals' decisions in staying at home or participating in large gatherings regardless of the order or advisory might differ from place to place. Whereas the global models are not designed to detect the locational difference in individuals' behaviors, the local models provide more details that these measures are effective against COVID-19 in some places on some days, especially when and where the individuals' behaviors agree with the order or advisory.

More important, the local model depicts a political divide in consuming the scientific evidence that enacting NPIs facilitates curbing the spread of COVID-19. The states that actively enacted NPIs, which happen to be primarily pro-Democrat states, often rank higher when considering both NPI strategies' implementations and their effectiveness (Table 3). States where economic development took higher priority and NPIs and public health were considered less important in their policies, as advocated by most Republican governors, often ranked lower in dealing with the spreading of the virus (Table 3) in our ranking scheme.

For instance, for the top quintile states in the ranking (Table 3), more than 80 percent of them either have a Democratic governor or voted for Joe Biden in 2020. New Mexico ranks the highest (always among the top ten), followed by New York (four times) and California, Rhode Island, and Washington (three times). For the bottom quintile states, more than 60 percent of them voted for Trump in 2020. Missouri and Nebraska rank the lowest (four times among the bottom ten), followed by North Dakota and Montana (three times; Table 3 and Figures 2–6). When governmental views are in direct conflict with the scientific evidence, especially when leadership in curbing the spread of COVID-19 is lacking, chaos often ensues. Although preventing the spread of a global pandemic is supposed to be a scientific exercise, the reality is that science and politics are often mixed in devising local policies. The different attitudes toward the scientific evidence led to different strategies for handling the implementation of the various NPIs, which had direct consequences on the daily caseloads of COVID-19.

The Take-Home Message

Conflicts between the importance of human lives and economic performance are inherently convoluted. It was rare that this explicit trade-off was made explicit empirically during the pandemic as policymakers debated policy changes. Yet troubled economies could also lead to a tragic aftermath for both individuals and governments. Striking a balance between science and politics is a delicate science of policy and governance. We do provide a ranking of governance in this study, but we also acknowledge that the ranking is strictly applied to curbing the spread of COVID-19. Whether or not curbing the spread of COVID-19 should be taken as the highest priority, however, is a policy decision that all governance bodies must consider carefully, and the consequences are felt by all citizens under governance. The delicacy of such balance is beyond the scope of this study to fully explore, although with more than 1 million deaths due to COVID-19 since 2020 in the United States, we contend the scale of the balance needs to tip toward saving human lives. With advanced spatiotemporal modeling, this study facilitates achieving a possible such balance. The elegant Bayesian analytical framework that allows scholars to combine the observed (the data) and the experienced (the priors) to produce more reliable posteriors renders a potential scientifically based compromise among different priorities, at different times, and in different places. This is especially true under the context of the continued new variants of the virus and the government's difficulty in persuading more individuals to get vaccinated. Even with vaccination, NPIs are still effective strategies that fight against the spread of COVID-19 (Baker et al. 2020). We hope the results from this study will not only provide practical policymaking support for state governments on the ongoing fight against COVID-19, but also create a platform for discussion regarding the delicate balance between fighting a global pandemic and maintaining socioeconomic normalcy.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Supplemental Material

Supplemental data for this article can be accessed on the publisher's site at: https://doi.org/10.1080/24694452.2023.2292807

ORCID

Danlin Yu http://orcid.org/0000-0002-4470-7011

References

Abedi, V., O. Olulana, V. Avula, D. Chaudhary, A. Khan, S. Shahjouei, J. Li, and R. Zand. 2021. Racial, economic, and health inequality and COVID-19 infection in the United States. *Journal of Racial and Ethnic Health Disparities* 8 (3):732–42. doi: 10.1007/s40615-020-00833-4.

Abel, T., and D. McQueen. 2020. The COVID-19 pandemic calls for spatial distancing and social closeness: Not for social distancing! *International Journal of Public Health* 65 (3):231. doi: 10.1007/s00038-020-01366-7.

Adolph, C., K. Amano, B. Bang-Jensen, N. Fullman, and J. Wilkerson. 2021. Pandemic politics: Timing state-level social distancing responses to COVID-19. *Journal of Health Politics, Policy and Law* 46 (2):211–33. doi: 10.1215/03616878-8802162.

Agarwal, R., M. Dugas, J. Ramaprasad, J. J. Luo, G. J. Li, and G. D. Gao. 2021. Socioeconomic privilege and political ideology are associated with racial disparity in COVID-19 vaccination. *Proceedings of the National Academy of Sciences of the United States of America* 118 (33):3. doi: 10.1073/pnas.2107873118.

Allcott, H., L. Boxell, J. Conway, M. Gentzkow, M. Thaler, and D. Yang. 2020. Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *Journal of Public Economics* 191:104254. doi: 10.1016/j.jpubeco.2020.104254.

Anderson, R. M., H. Heesterbeek, D. Klinkenberg, and T. D. Hollingsworth. 2020. How will country-based mitigation measures influence the course of the COVID-19 epidemic? *Lancet* 395 (10228):931–34. doi: 10.1016/S0140-6736(20)30567-5.

Anselin, L. 1988. *Spatial econometrics: Methods and models*. Dordrecht, The Netherlands: Kluwer Academic.

Auger, K. A., S. S. Shah, T. Richardson, D. Hartley, M. Hall, A. Warniment, K. Timmons, D. Bosse, S. A. Ferris, P. W. Brady, et al. 2020. Association between statewide school closure and COVID-19 incidence and mortality in the US. *JAMA* 324 (9):859–70. doi: 10.1001/jama.2020.14348.

Baker, R. E., S. W. Park, W. C. Yang, G. A. Vecchi, C. J. E. Metcalf, and B. T. Grenfell. 2020. The impact of COVID-19 nonpharmaceutical interventions on the future dynamics of endemic infections. *Proceedings of the National Academy of Sciences of the United States of America* 117 (48):30547–53. doi: 10.1073/pnas.2013182117.

Bartels, L. M. 1993. Messages received: The political impact of media exposure. *American Political Science Review* 87 (2):267–85. doi: 10.2307/2939040.

Besag, J. 1974. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society: Series B (Methodological)* 36 (2):192–225.

Besag, J., and P. J. Green. 1993. Spatial statistics and Bayesian computation. *Journal of the Royal Statistical Society: Series B (Methodological)* 55 (1):25–37.

Boserup, B., M. McKenney, and A. Elkbuli. 2020. Disproportionate impact of COVID-19 pandemic on racial and ethnic minorities. *The American Surgeon* 86 (12):1615–22. doi: 10.1177/0003134820973356.

Brauner, J. M., S. Mindermann, M. Sharma, D. Johnston, J. Salvatier, T. Gavenčiak, A. B. Stephenson, G. Leech, G. Altman, V. Mikulik, et al. 2021. Inferring the effectiveness of government interventions against COVID-19. *Science* 371 (6531):802. doi: 10.1126/science.abd9338.

Centers for Disease Control and Prevention (CDC). 2021. Trends in number of COVID-19 vaccinations in the US 2021. Accessed September 28, 2021. https://covid.cdc.gov/covid-data-tracker/#vaccination-trends_vacctrends-fully-daily.

Cheung, J. C.-S. 2022. Responses to COVID-19 in major social work journals: A systematic review of empirical studies, comments, and editorials. *Research on Social Work Practice* 32 (2):168–85. doi: 10.1177/10497315211046846.

Dehning, J., J. Zierenberg, F. P. Spitzner, M. Wibral, J. P. Neto, M. Wilczek, and V. Priesemann. 2020. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* 369 (6500):eabb9789. doi: 10.1126/science.abb9789.

Desmarais, B. A., R. J. La Raja, and M. S. Kowal. 2015. The fates of challengers in U.S. House elections: The role of extended party networks in supporting candidates and shaping electoral outcomes. *American Journal of Political Science* 59 (1):194–211. doi: 10.1111/ajps.12106.

Dowd, J. B., L. Andriano, D. M. Brazel, V. Rotondi, P. Block, X. J. Ding, Y. Liu, and M. C. Mills. 2020. Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proceedings of the National Academy of Sciences of the United States of America* 117 (18):9696–98. doi: 10.1073/pnas.2004911117.

Elhorst, J. P. 2014. *Spatial econometrics: From cross-sectional data to spatial panels*. New York: Springer.

Enserink, M., and K. Kupferschmidt. 2020. With COVID-19, modeling takes on life and death importance. *Science* 367 (6485):1414–15. doi: 10.1126/science.367.6485.1414-b.

Genton, M. G. 2001. Classes of kernels for machine learning: A statistics perspective. *Journal of Machine Learning Research* 2 (December):299–312.

Güner, R., I. Hasanoğlu, and F. Aktaş. 2020. COVID-19: Prevention and control measures in community. *Turkish Journal of Medical Sciences* 50 (SI-1):571–577. doi: 10.3906/sag-2004-146.

Guo, S., R. An, T. D. McBride, D. Yu, L. Fu, and Y. Yang. 2020. Mitigation interventions in the United States: An exploratory investigation of determinants and impacts. *Research on Social Work Practice* 31 (1):26–41. doi: 10.1177/1049731520957415.

Hamilton, J. D. 2020. *Time series analysis*. Princeton, NJ: Princeton University Press.

Han, J., B. D. Meyer, and J. X. Sullivan. 2020. *Income and poverty in the COVID-19 pandemic*. Cambridge, MA: National Bureau of Economic Research.

Jacobs, L. R., and R. Y. Shapiro. 2000. *Politicians don't pander: Political manipulation and the loss of democratic responsiveness*. Chicago: University of Chicago Press.

Kerr, J., C. Panagopoulos, and S. van der Linden. 2021. Political polarization on COVID-19 pandemic response in the United States. *Personality and Individual Differences* 179 (9):110892. doi: 10.1016/j.paid.2021.110892.

Kraemer, M. U. G., C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D. M. Pigott, L. Du Plessis, N. R. Faria, R. Li, W. P. Hanage, et al. 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 368 (6490):493–97. doi: 10.1126/science.abb4218.

Kreps, S., S. Prasad, J. S. Brownstein, Y. Hswen, B. T. Garibaldi, B. B. Zhang, and D. L. Kriner. 2020. Factors associated with US adults' likelihood of accepting COVID-19 vaccination. *JAMA Network Open* 3 (10):e2025594. doi: 10.1001/jamanetworkopen.2020.25594.

Largent, E. A., G. Persad, S. Sangenito, A. Glickman, C. Boyle, and E. J. Emanuel. 2020. US public attitudes toward COVID-19 vaccine mandates. *JAMA Network Open* 3 (12):e2033324. doi: 10.1001/jamanetworkopen.2020.33324.

Laurencin, C. T., and A. McClinton. 2020. The COVID-19 pandemic: A call to action to identify and address racial and ethnic disparities. *Journal of Racial and Ethnic Health Disparities* 7 (3):398–402. doi: 10.1007/s40615-020-00756-0.

LeSage, J. P., and R. K. Pace. 2009. *Introduction to spatial econometrics*. Boca Raton, FL: Taylor & Francis/CRC Press.

Mackey, K., C. K. Ayers, K. K. Kondo, S. Saha, S. M. Advani, S. Young, H. Spencer, M. Rusek, J. Anderson, S. Veazie, et al. 2021. Racial and ethnic disparities in COVID-19-related infections, hospitalizations, and deaths a systematic review. *Annals of Internal Medicine* 174 (3):362–73. doi: 10.7326/M20-6306.

Masket, S. E., J. Winburn, and G. C. Wright. 2012. The gerrymanderers are coming! Legislative redistricting won't affect competition or polarization much, no matter who does it. *PS: Political Science and Politics* 45 (1):39–43.

McCaw, Z. R., J. M. Lane, R. Saxena, S. Redline, and X. H. Lin. 2020. Operating characteristics of the rank-based inverse normal transformation for quantitative trait analysis in genome-wide association studies. *Biometrics* 76 (4):1262–72.

Memmott, T., S. Carley, M. Graff, and D. M. Konisky. 2021. Sociodemographic disparities in energy insecurity among low-income households before and during the COVID-19 pandemic. *Nature Energy* 6 (2):186–93. doi: 10.1038/s41560-020-00763-9.

Moser, K. A., A. J. Fox, and D. R. Jones. 1984. Unemployment and mortality in the OPCS longitudinal study. *Lancet* 2 (8415):1324–29. doi: 10.1016/s0140-6736(84)90832-8.

Nazia, N., J. Law, and Z. A. Butt. 2022. Identifying spatiotemporal patterns of COVID-19 transmissions and the drivers of the patterns in Toronto: A Bayesian hierarchical spatiotemporal modelling. *Scientific Reports* 12 (1):9369. doi: 10.1038/s41598-022-13403-x.

Neyman, J. 1923. On the application of probability theory to agricultural experiments: Essay on principles. *Statistical Science* 5:465–80.

Painter, M., and T. Qiu. 2020. Political beliefs affect compliance with COVID-19 social distancing orders. *Covid Economics* 4:103–23.

Parolin, Z., M. Curran, J. Matsudaira, J. Waldfogel, and C. Wimer. 2020. Monthly poverty rates in the United States during the COVID-19 pandemic.

Poverty and Social Policy Working Paper, Center on Poverty & Social Policy, Columbia University, New York.

Patrick, S. L., and H. C. Cormier. 2020. Are our lives the experiment? COVID-19 lessons during a chaotic natural experiment—A commentary. *Health Behavior and Policy Review* 7 (2):165–69.

Patterson, S. 2022. The politics of pandemics: The effect of stay-at-home orders on COVID-19 mitigation. *State Politics & Policy Quarterly* 22 (1):1–23. doi: 10.1017/spq.2021.14.

Puspitasari, I. M., L. Yusuf, R. K. Sinuraya, R. Abdulah, and H. Koyama. 2020. Knowledge, attitude, and practice during the COVID-19 pandemic: A review. *Journal of Multidisciplinary Healthcare* 13:727–33. doi: 10.2147/JMDH.S265527.

Quan, D., L. Luna Wong, A. Shallal, R. Madan, A. Hamdan, H. Ahdi, A. Daneshvar, M. Mahajan, M. Nasereldin, M. Van Harn, et al. 2021. Impact of race and socioeconomic status on outcomes in patients hospitalized with COVID-19. *Journal of General Internal Medicine* 36 (5):1302–1309. doi: 10.1007/s11606-020-06527-1.

R Core Team. 2022. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Radley, D. C., J. C. Baumgartner, and S. R. Collins. 2023. 2022 Scorecard on state health system performance: How did states do during the COVID-19 pandemic? The Commonwealth Fund. Accessed March 6, 2023. https://www.commonwealthfund.org/publications/scorecard/2022/jun/2022-scorecard-state-health-system-performance.

Raifman, M. A., and J. R. Raifman. 2020. Disparities in the population at risk of severe illness from COVID-19 by race/ethnicity and income. *American Journal of Preventive Medicine* 59 (1):137–39. doi: 10.1016/j.amepre.2020.04.003.

Rubin, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66 (5):688–701. doi: 10.1037/h0037350.

Rubin, D. B. 1986. Comment. *Journal of the American Statistical Association* 81 (396):961–62.

Rue, H., and L. Held. 2005. *Gaussian Markov random fields: Theory and applications*. Boca Raton, FL: CRC Press.

Rue, H., S. Martino, and N. Chopin. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 71 (2):319–92. doi: 10.1111/j.1467-9868.2008.00700.x.

Salomon, J. A., A. Reinhart, A. Bilinski, E. J. Chua, W. La Motte-Kerr, M. M. Rönn, M. B. Reitsma, K. A. Morris, S. LaRocca, T. H. Farag, et al. 2021. The US COVID-19 trends and impact survey: Continuous real-time measurement of COVID-19 symptoms, risks, protective behaviors, testing, and vaccination. *Proceedings of the National Academy of Sciences of the United States of America* 118 (51):9. doi: 10.1073/pnas.2111454118.

Tai, D. B. G., A. Shah, C. A. Doubeni, I. G. Sia, and M. L. Wieland. 2020. The disproportionate impact of COVID-19 on racial and ethnic minorities in the United States. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 72 (4):703–706. doi: 10.1093/cid/ciaa815.

Thebault, R., A. B. Tran, and V. Williams. 2020. The coronavirus is infecting and killing Black Americans at an alarmingly high rate. *Washington Post*, April 7.

Tian, H. Y., Y. H. Liu, Y. D. Li, C. H. Wu, B. Chen, M. U. G. Kraemer, B. Y. Li, J. Cai, B. Xu, Q. Q. Yang, et al. 2020. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* 368 (6491):638–42. doi: 10.1126/science.abb6105.

Virgolino, A., J. Costa, O. Santos, M. E. Pereira, R. Antunes, S. Ambrósio, M. J. Heitor, and A. Vaz Carneiro. 2022. Lost in transition: A systematic review of the association between unemployment and mental health. *Journal of Mental Health* 31 (3):432–44. doi: 10.1080/09638237.2021.2022615.

Webb Hooper, M., A. M. Nápoles, and E. J. Pérez-Stable. 2020. COVID-19 and racial/ethnic disparities. *JAMA* 323 (24):2466–77. doi: 10.1001/jama.2020.8598.

White, E. R., and L. Hébert-Dufresne. 2020. State-level variation of initial COVID-19 dynamics in the United States. *PLoS ONE* 15 (10):e0240648. doi: 10.1371/journal.pone.0240648.

Xu, Y., Y. S. Park, and J. D. Park. 2021. Measuring the response performance of U.S. states against COVID-19 using an integrated DEA, CART, and logistic regression approach. *Healthcare* 9 (3):268. doi: 10.3390/healthcare9030268.

Yu, D. L. 2010. Exploring spatiotemporally varying regressed relationships: The geographically weighted panel regression analysis. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 38:134–39.

Yu, D., Y. Zhang, J. Meng, X. Wang, L. He, M. Jia, J. Ouyang, Y. Han, G. Zhang, and Y. Lu. 2023. Seeing the forest and the trees: Holistic view of social distancing on the spread of COVID-19 in China. *Applied Geography* 154:102941. doi: 10.1016/j.apgeog.2023.102941.

DANLIN YU is Professor of Urban Geography and GIS in the Department of Earth and Environmental Studies, Montclair State University, Montclair, NJ 07043. E-mail: yud@montclair.edu. His research interests include spatial and spatiotemporal data analysis, urban geography, and complex system reconstruction and simulation.

SHENYANG GUO is Frank J. Bruno Distinguished Professor at the Brown School of Washington University in St Louis, St. Louis, MO 63130.

E-mail: s.guo@wustl.edu. His research interests include quantitative research methods, child welfare, child mental health, and intervention research.

YUANYUAN YANG is a PhD Student in the Brown School of Social Work at Washington University in St. Louis, St. Louis, MO 63130. E-mail: yuanyuan.yang@wustl.edu. Her research interests include asset-based poverty policy and social determinants of children's mental health.

LINYUN FU is a PhD Student in the Crown Family School of Social Work, Policy, and Practice at the University of Chicago, Chicago, IL 60637. E-mail: fulinyun@uchicago.edu. Her research interests include program and policy evaluation, intervention research, and child mental health.

TIMOTHY McBRIDE is a Professor at the Brown School, Washington University in St. Louis, St. Louis, MO 63130. E-mail: tmcbride@wustl.edu. His research interests include health insurance policy, Medicaid, Medicare, health economics, rural health, and long-term care.

RUOPENG AN is an Associate Professor in the Brown School and the Division of Computational and Data Sciences at Washington University in St. Louis, St. Louis, MO 63124. E-mail: ruopeng@wustl.edu. His research assesses population-level policies, local food and built environment, and socioeconomic determinants that affect individuals' dietary behavior, physical activity, sedentary lifestyle, and adiposity in children, adults of all ages, and people with disabilities.