BMC Public Health

Spatial spillover and COVID-19 spread in the U.S.

John Ulimwengu[1*] and Aziza Kibonge[2]

Abstract

Background: This research estimates the effects of vulnerability on the spread of COVID-19 cases across U.S. counties. Vulnerability factors (Socioeconomic Status, Minority Status & Language, Housing type, Transportation, Household Composition & Disability, Epidemiological Factors, Healthcare system Factors, High-risk Environments, and Population density) do not only influence an individual's likelihood of getting infected but also influence the likelihood of his/her neighbors getting infected. Thus, spatial interactions occurring among individuals are likely to lead to spillover effects which may cause further virus transmission.

Methods: This research uses the COVID-19 community index (CCVI), which defines communities likely vulnerable to the impact of the pandemic and captures the multi-dimensionality of vulnerability. The spatial Durbin model was used to estimate the spillover effects of vulnerability to COVID-19 in U.S. counties, from May 1 to December 15, 2020.

Results: The findings confirm the existence of spatial spillover effects; with indirect effects (from neighboring counties) dominating the direct effects (from county-own vulnerability level). This not only validates social distancing as a strategy to contain the spread of the pandemic but also calls for comprehensive and coordinated approach to fight its effects. By keeping vulnerability factors constant but varying the number of reported infected cases every 2 weeks, we found that marginal effects of vulnerability vary significantly across counties. This might be the reflection of both the changing intensity of the pandemic itself but also the lack of consistency in the measures implemented to combat it.

Conclusion: Overall, the results indicate that high vulnerability in Minority, Epidemiological factors, Healthcare System Factors, and High-Risk Environments in each county and adjacent counties leads to an increase in COVID-19 confirmed cases.

Keywords: COVID-19, Vulnerability index, Spatial spillovers, Minorities, Pre-existing conditions, Transportation, Housing, Healthcare

Background

As of May 4, 2021, the Centers for Disease Control and Prevention (CDC) reported effects of the COVID-19 pandemic included 32,267,958 cases, and 574,679 deaths in the U.S. [1]. At the same time, there is increasing evidence that rates of infections and mortality from COVID-19 are not evenly distributed across locations and socio-economic groups. For example, the percentage of hospitalizations, intensive care unit admissions, and deaths were highest among persons aged 70 years or higher, regardless of underlying conditions [2]. Moreover, the presence of underlying health conditions increases vulnerability to severe complications from COVID-19 [3]. Furthermore, communities with high levels of poverty, high concentration of less-educated adults, and minorities are more likely to face devastating

* Correspondence: j.ulimwengu@cgiar.org
[1]International Food Policy Research Institute (IFPRI), Eye Street, 1201 I St NW, Washington, DC 20005, USA
Full list of author information is available at the end of the article

effects of the virus [4, 5]. As evidence suggests, lower socioeconomic status, homelessness, poverty, racial and ethnic disparities [6], and health-related factors [7, 8] may significantly contribute to the COVID-19 transmission and incidence rates, thereby accounting for some of the spatial variation.

These factors do not only influence an individual's likelihood of getting infected but do also influence the likelihood of his/her neighbors getting infected. Spatial interactions occurring among individuals are likely to lead to spillover effects which may cause further virus transmission. Hence, analyzing the spread of COVID-19 without accounting for spatial disparities and spillover effects will lead to biased estimates and erroneous policy recommendations. This study aims at estimating the effects of vulnerability factors on COVID-19 cases while explicitly accounting for spatial spillover effects.

Although several studies have looked at disparities, vulnerability, and COVID-19 spread [9–13], none has examined the related spillover effects. This is particularly important as people travel across counties and states. McLaren (2020) examined the relationship between COVID-19 severity and racial composition and found that minority's population share is strongly correlated with overall COVID-19 deaths [14]; neighborhoods that are predominantly comprised of Black and Hispanic persons are disproportionately at risk to COVID-19 infections and mortality [15]. Unemployment has also been a contributing factor as the share of people not working and thus not commuting is correlated with higher death rates [16]. Likewise, age structure matters as the proportion of people living in nursing homes were shown to be significant and persistent predictors of number of deaths [17].

Several indices have been developed to measure vulnerability and identify populations at risk of contracting the virus. The CDC's Social Vulnerability Index (SVI) in particular, is widely used in health research and was recently applied to COVID-19 analysis in U.S. counties ([9, 11, 18]). Moreover, spatial models have been used to capture the dynamics of COVID-19 prevalence in U.S. counties as they partially explain the geographic disparities in COVID-19 period prevalence [19]. Indeed, substantial spatial heterogeneity is observed across U.S. counties, and the association between case counts of COVID-19 infection and social vulnerability in U.S. counties allowed the most vulnerable counties to be identified. These counties had on average higher incidence rates compared to least vulnerable counties [9, 20, 21]. Neelon et al. (2021) reported that COVID-19 disproportionately affected less vulnerable counties at the outset of the pandemic, before spreading to more vulnerable counties over time [11].

Methods
Social interactions models
Social interaction models study how interaction among individuals or entities can lead to collective behavior and aggregate patterns [22]. These models incorporate endogenous effects, contextual effects, and unobserved correlation effects. As pointed out by Lewnard et al. (2020) in the health literature, symptoms of COVID-19 vary – some infected people seem to experience no symptoms, while others suffer from many symptoms which typically take from two to fourteen days to appear after infection [23]. The rate at which a susceptible person in a given subpopulation acquires the infection is determined by his/her interactions with infectious persons, either in the home subpopulation or in its neighboring subpopulations [18].

In general, latent persons will transit to asymptomatic infectious stage with high probability of becoming symptomatic infectious (Fig. 1). Infectious persons with symptoms can further be divided between those who can travel and those who are travel restricted. This paper focuses on those who can transmit the disease through social interactions, which further sparks disease spread across locations.

Infections occur mainly though exposure to respiratory droplets when a person is in close contact with someone who has COVID-19 (six feet of one another) [24]. Thus, there is potential for spillover effects because people who are physically near a person with COVID-19 or have direct contact with that person are at greater risk of infection [25]. Qu et al. (2012) argue that ignoring the fact that spatial or social interactions among individuals simultaneously determine the outcomes can cause biased estimates for the effects of exogenous variables and may also underestimate their influences [26]. Indeed, location i 's exogenous attributes will not only influence its own rate of pandemic spread but, due to spatial interactions, its neighbors' pandemic rate may also be influenced, which will in turn affect location i 's pandemic level and so on as described in Fig. 2.

Given the presence of spatial interactions, the number of infections (y) and its determinants (X), we implement a spatial Durbin model (SDM) of the following form [27]:

$$y = \rho Wy + X\beta + WX\gamma + \varepsilon \qquad (1)$$

where y is an $n \times 1$ vector of observations on the dependent variable, X is an $n \times K$ matrix of observations on the explanatory variables, ρ is $n \times 1$ vector of spillover parameter, β and γ are $K \times 1$ vectors of parameters, W is $n \times n$ spatial weight matrix that capture the structure spatial interactions, and $\varepsilon \sim N(0, I_n)$ is $n \times 1$ vector of the error term.
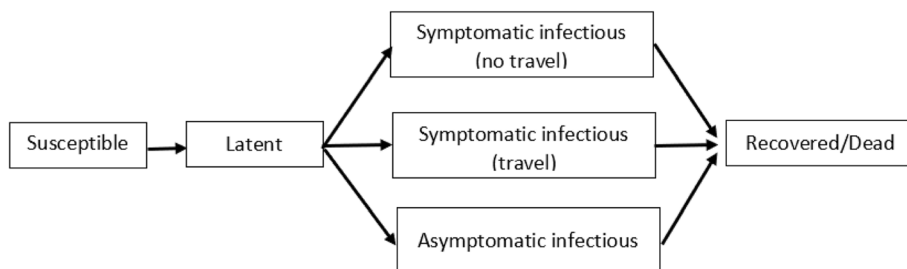
Fig. 1 Compartmental structure of the epidemic model within each subpopulation, adapted from [18]

Where

Y1 = number of infections in location 1 (number of COVID-19 confirmed cases in county 1).

Y2 = number of infections in location 2 (number of COVID-19 confirmed cases in county 2).

X1 = determinant of Y1 (explanatory variable in county 1).

X2 = determinant of Y2 (explanatory variable in county 2).

ε1 = error term in county 1.

ε2 = error term in county 2.

As pointed out by Lesage et al. (2008), the SDM includes most models used in applied spatial econometrics literature: i) if $\gamma = 0$, eq. (1) becomes a spatial autoregressive (SAR) model that includes a spatial lag of infected cases from related counties, but excludes these counties' vulnerability factors; ii) if $\gamma = -\rho\beta$, it becomes a spatial error model (SEM); iii) if $\gamma = 0$ and $\rho = 0$, it is a non-spatial least-squares model that assumes locations' dependent variables are independent [28].

For K locations, Lesage and Fisher (2008) show that eq. (1) can be rewritten as

$$y = \sum_{r=1}^{k} K_r(W)X_r + A(W)^{-1}\varepsilon, \qquad (2)$$

where $K_r(W) = A(W)^{-1}(I_n\beta_r + W\theta_r)$ and $A(W)^{-1} = (I_n - \rho W)^{-1}$, and
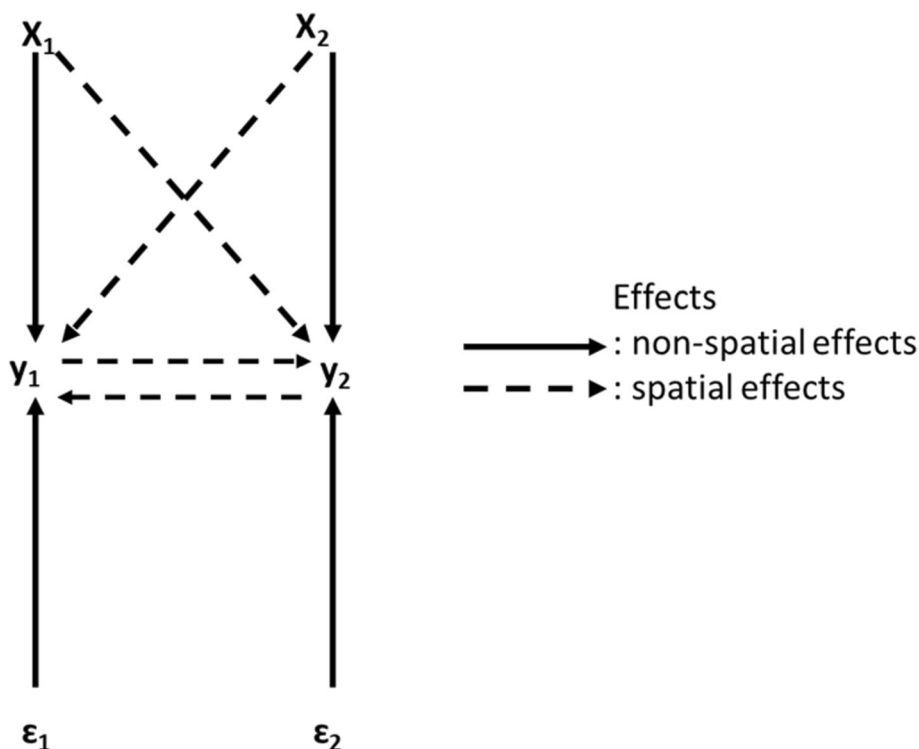

Fig. 2 SMD for 2 locations

$$X_r = \begin{pmatrix} x_{r1} \\ x_{r2} \\ \cdot \\ \cdot \\ \cdot \\ x_{rn} \end{pmatrix}$$

Lesage et al. (2011) propose a series of metrics to interpret estimates from eq. (2); direct, indirect, and total marginal effects of the determinants of y [29]. Direct marginal effects show how each determinant (Xs) affects y while indirect effects show how each determinant affects y in neighboring counties through social interactions. The total marginal effects provide a complete picture of impacts of each variable, considering the spillover effects. The expansion of the expected value of eq. (2) is given by

$$\begin{pmatrix} E(y_1) \\ E(y_2) \\ \vdots \\ E(y_n) \end{pmatrix} = \begin{pmatrix} K_r(W)_{11} & K_r(W)_{12} & \dots & K_r(W)_{1n} \\ K_r(W)_{21} & K_r(W)_{21} & \dots & K_r(W)_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ K_r(W)_{n1} & K_r(W)_{n1} & \dots & K_r(W)_{nn} \end{pmatrix} \begin{pmatrix} x_{1r} \\ x_{2r} \\ \vdots \\ x_{nr} \end{pmatrix}$$

(3)

for variable i, the expected value of (3) takes the following form:

$$E(y_i) = \sum_{r=1}^{k} \left[K_r(W)_{i1} x_{1r} + \dots + K_r(W)_{n1} x_{nr} \right] \quad (4)$$

Spatial direct and indirect effects are derived as follows:

i)   Indirect effects: the impact on the expected value of location i given a change in the explanatory variable x_r in location j is

$$\frac{\partial E(y_i)}{\partial x_{jr}} = K_r(W)_{ij} \quad (5)$$

where $K_r(W)_{ij}$ represents the i, j th element of matrix $K_r(W)$.

ii) Direct effects: the impact on the expected value of location i, given a change in certain variable for the same location is given by

$$\frac{\partial E(y_i)}{\partial x_{ir}} = K_r(W)_{ii} \quad (6)$$

It follows that for K locations, there will be $K \times n^2$ marginal effects which might be burdensome to report on.

Data

This study uses the COVID-19 Community Vulnerability Index (CCVI) developed by the Surgo Foundation which combines the Social Vulnerability Index (SVI) with other vulnerability factors specific to COVID-19 (Epidemiological factors, public health system capacity, and variables capturing high risk environments).

Developed by the CDC, the SVI has been applied in disaster management to help identify vulnerable people and areas. It specifies the relative vulnerability of every U.S. Census tract, and is built by ranking each of the 15 census variables from highest to lowest across all census tracts in the U.S. The tracts are then grouped into four themes (Socioeconomic Status, Household Composition and Disability, Minority Status & Language, and Housing type & Transportation).

A percentile rank was then calculated for each census tract and each theme, and percentile values range from 0 to 1, with higher values indicating greater vulnerability. For each of the four themes, the percentiles are summed up for the variables comprising each theme. Then the summed percentiles are ordered for each theme to determine percentile ranking that are specific to each theme [30].

The CCVI builds on the SVI and identifies communities that are vulnerable by aggregating dozens of indicators across six themes (citation: Surgo ventures).

The Surgo Foundation[1] defines "more vulnerable" communities as those that "have a limited ability to mitigate, treat, and delay transmission of a pandemic disease, and to reduce its economic and social impacts" [31]. More recently, a new version of the CCVI was developed and focuses on 40 variables covering seven core themes to account for additional factors that make a community or individual susceptible to the COVID-19 pandemic. This new and composite CCVI metric ranks each geography (state, county, or census tract) relative to one another on a 0–1 scale (0 = least vulnerable, 1 = most vulnerable), and the data is available in the Surgo Foundation website [31], Appendix D.

Dependent and independent variables

The dependent variable is the COVID-19 confirmed cumulative cases from May 1 to December 15, 2020. Fifteen sub-periods are considered: May 1–May 15, May 16–May 31, June 1–June 15, June 16–June 30, July 1–15,

---

[1]The Surgo Foundation collected data from the CDC, Centers for Medicare, & Medicaid Services (CMS), the Harvard Global Health Institute, PolicyMap, the U.S Bureau of Labor Statistics (BLS), the U.S. Census Bureau (USCB), the National Cancer Institute (NCI), the National Center for HIV, STD and TB Prevention (NCHSTP), Kaiser Health News, Centers for Medicare & Medicaid Services (CMS), Census Economic Annual Surveys, the Vera institute for Justice, and the Association of Public Health Laboratories.

July 16–31, August 1–15, August 16–31, September 1–15, September 16–30, October 1–15, October 16–31, November 1–15, November 16–30, December 1–15. For each county and each sub-period, the cumulative counts of confirmed cases by county is divided by the total population multiplied by 100,000. Data on the number of people tested for COVID-19 was retrieved from the CDC COVID Data Tracker (https://covidtracking.com/about-data/cdc-comparison).

The independent variables are the CCVI for each of the 7 themes that account for the following types of vulnerability: Socioeconomic status, Minority status & Language, Transportation, Household composition & Disability, Epidemiological factors, Healthcare system, High-risk environment, and Population density. It is a percentile-based index and ranks each county relative to one another on a 0–1 scale, with 0 indicating the least vulnerable, and 1 indicating the most vulnerable.

CCVI Theme 1: Socioeconomic Status is composed of percentile rank data for the percentage below poverty, percentage unemployed, per capita income, and percentage with no high school diploma [31].

CCVI Theme 2: Minority Status & Language include percentage of minority and percentage of those who speaks English "less than well" [31].

CCVI Theme 3: Transportation, Household Composition & Disability include persons aged 17 or younger, older than 5 with a disability, single-parent households, multiunit structures, mobile homes, crowding, no vehicle, group quarters, and access to indoor plumbing [31].

CCVI Theme 4: Epidemiological factors was selected based on the CDC guidelines and identify high risk populations as elderly adults and individuals with underlying conditions (respiratory conditions, cardiovascular conditions, obesity, diabetes, immuno-compromised, and persons aged 65 or Older) [31].

CCVI Theme 5: Healthcare System factors provide a measure of the capacity, strength, accessibility, and preparedness of the healthcare system in the context of COVID-19. Health System Capacity includes hospital beds per 100,000, intensive care unit beds per 100,000, and epidemiologist per 100,000. Healthcare accessibility includes the percentage of the population with a primary care physician. Health system strength is measured by total health expenditure and quality of care and includes the Agency for Healthcare Research and Quality's (AHRQ) prevention quality indicator (PQI) composite, Health spending per capita, and Aggregate cost of medical care. Health system preparedness includes public health emergency preparedness funding available, health labs per 100,000, and emergency services per 100,000 [31].

CCVI Theme 6: High- Risk Environments measure the percentage of population working or living in environments with high infection risk. The variables are long-term care residents per 100,000, prisons population per 100,000, and the percentage of populations employed in high-risk industry. The CCVI considers industries where workers are in enclosed spaces for extended time, in proximity of or interacting with others, and/or exposed to diseases. Rural counties for example were particularly affected by outbreaks due to their employment dependence on industries with high-risk work conditions like the meatpacking industry [32].

CCVI Theme 7: Population Density is measured as the total number of people per unit area and represent an important factor in health outcomes.

Table 1 presents the descriptive statistics of variables used for the contiguous U.S. counties.

Results and discussion
Spatial dependence
The Ordinary Least Squares (OLS) model does not allow for spillovers as it implicitly assumes that outcomes for different units (counties) are independent of each other. To assess the presence of spatial interactions between counties, several statistics are used including the Moran's Indices which are positive and significant at 1% suggesting that the COVID-19 confirmed cases exhibit spatial correlation (see Appendix A). The Lagrange Multiplier test (LM) and the Robust Lagrange Multiplier test (Robust LM) were conducted and confirmed the presence of spatial lag or error terms in the model. As expected, OLS estimates are biased and inconsistent as the spillover parameters (ρ and γ) in eq. (1) are significantly different from zero for all sub-periods.

The SDM was then used to estimate the interaction effects and exogenous interaction effects. The results from non-spatial (OLS) and spatial (SDM) specifications models for all the sub-periods are presented below (see Appendix C).

Furthermore, the average impacts (direct, indirect, and total effects) are estimated to determine the magnitude of the change in confirmed COVID-19 cases resulting from changes in the selected vulnerability factors. The direct effects measure the impact of change in vulnerability factors in one county on the confirmed COVID-19 cases within that county, while the indirect effects (spillover effects), measure the impact of same change on confirmed COVID-19 cases from the neighboring county. The total effect is the sum of both direct and indirect effects. To compare the change in spillover effects over time, the average impacts were calculated for fifteen sub-periods (May 1 – December 15, 2020). As shown in Appendix C, the increase in some of the vulnerability factors will increase the possibility of COVID-19 cases in the target county.

Table 1 Descriptive statistics

Variables	Observations	Mean	Standard deviation	Minimum	Maximum
Index (0 to 1)					
Socioeconomic Status	3142	0.499953	0.288811	0	1
Minority Status & Language	3142	0.499835	0.288924	0	1
Transportation, Household Composition & Disability	3142	0.499962	0.288809	0	1
Epidemiological factors	3142	0.499979	0.288811	0	1
Healthcare System factors	3142	0.499963	0.288809	0	1
High Risk- Environments	3142	0.499931	0.288816	0	1
Population Density	3142	0.500000	0.288813	0	1

Overall, the results indicate significant differences between direct and indirect effects across time and across vulnerability factors. This suggests that failing to account for spatial correlation would lead to biased estimates.

Socioeconomic factors

As expected, marginal effects of socioeconomic vulnerability on the number of confirmed cases are positive and significant; except for the last four periods indicating that COVID-19 confirmed cases increase with higher socioeconomic vulnerability (Fig. 3 and Appendix C). Similarly, Nayak found that socioeconomic status was associated with higher incidence after adjusting for age and comorbidities, and other studies found that most vulnerable counties based on socioeconomic status had greater risk of COVID-19 diagnosis and death [9, 11, 20].

Except for the last four sub-periods, the average direct effect of socioeconomic vulnerability indicates that a marginal point increase in vulnerability index is likely to increase COVID-19 cases by 0.7%. With respect to indirect effects and total effects, the socioeconomic factors appear to have negative spillover effects on the number of confirmed cases, especially early in the pandemic (May 1–June 15) and later (Sept 1 - August 31).

Minorities & Language

The average direct effect suggest that a marginal point increase in vulnerability index is likely to increase confirmed cases in a given county by 1.7% (Fig. 4 and Appendix C).

The indirect effects of Minorities & Languages are only significant early in the pandemic and later. Additionally, the indirect effects from neighboring counties in some cases is ten times the magnitude of the direct impacts, suggesting large spillover effects. On average, the total impact is positive, with about two third comprised of the spillover effects from neighboring counties. Similar findings indicate that people in the most vulnerable counties by Minority & Language domain had greater risks of COVID-19 diagnosis [20], and increased case counts [19].

Housing type, Transportation, Household Composition & Disability

The results indicate that Housing & Transportation vulnerability exert a positive direct impact on the number
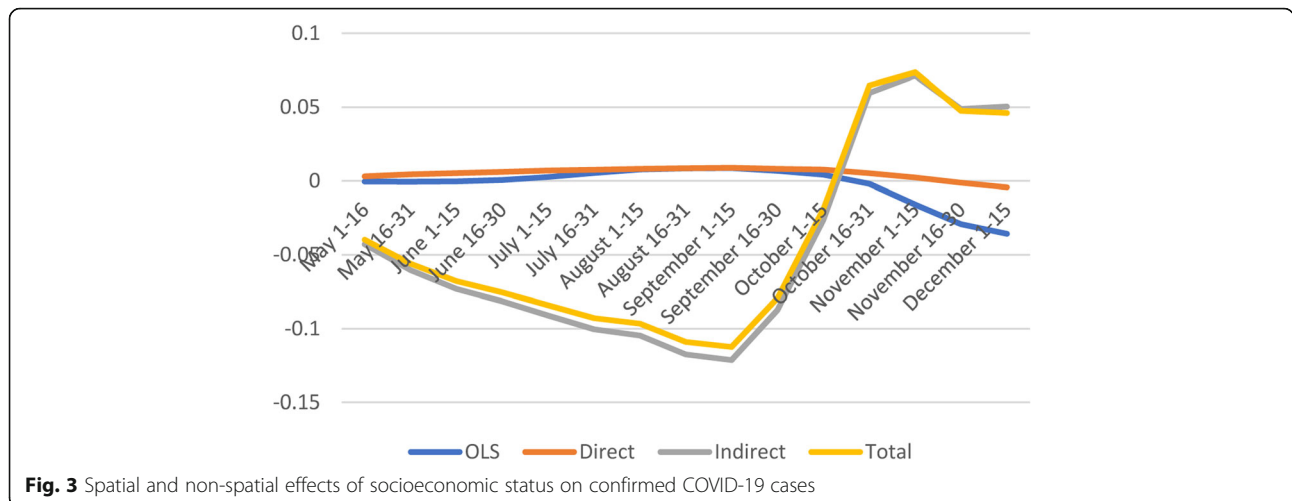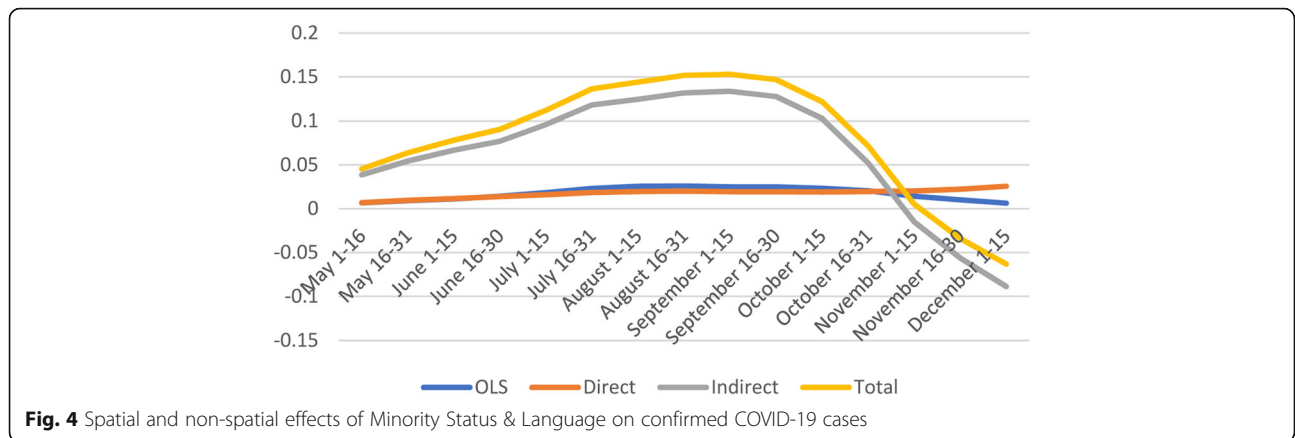

Fig. 3 Spatial and non-spatial effects of socioeconomic status on confirmed COVID-19 cases

Fig. 4 Spatial and non-spatial effects of Minority Status & Language on confirmed COVID-19 cases

of confirmed cases. Counties with higher vulnerability in Housing & Transportation also experienced higher confirmed COVID-19 cases (Fig. 5). Less urban counties with high-density housing structures, and crowded housing units had a higher probability of being identified as a COVID-19 hotspot [2].

However, the indirect effects and total effects from nearby counties are negative and significant (October 1–December 15) suggesting a negative relationship between Housing & Transportation vulnerability and confirmed COVID-19 cases. Although counterintuitive, these findings are similar to those of Karaye et al. (2020) who pointed out that the direction and magnitude of social vulnerability varied among U.S. states, and perhaps could be better explained by local spatial variations between counties [19].
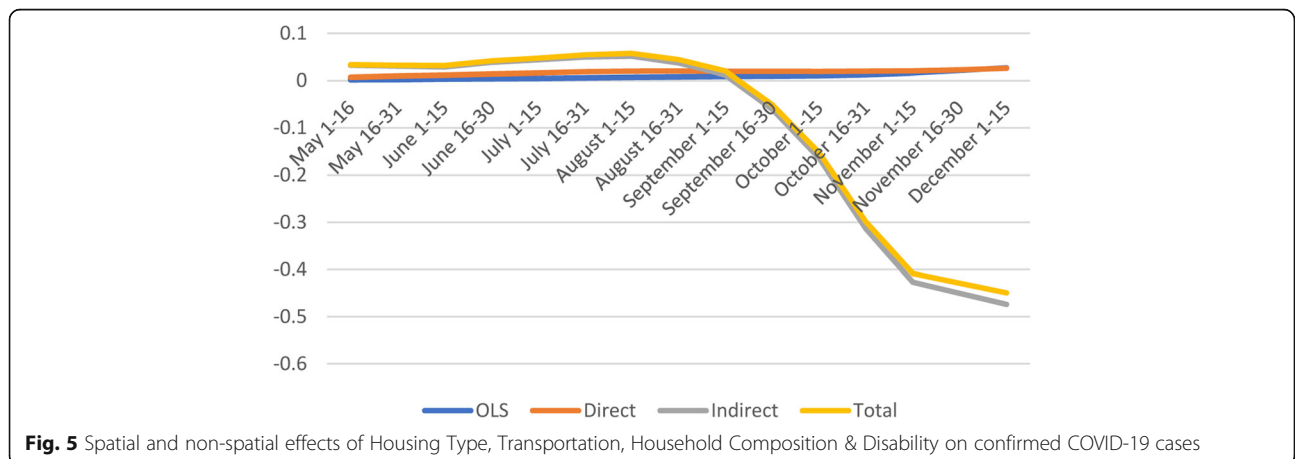
Epidemiological factors
From August 16 to December 15, the direct effects of epidemiological factors are negative and significant
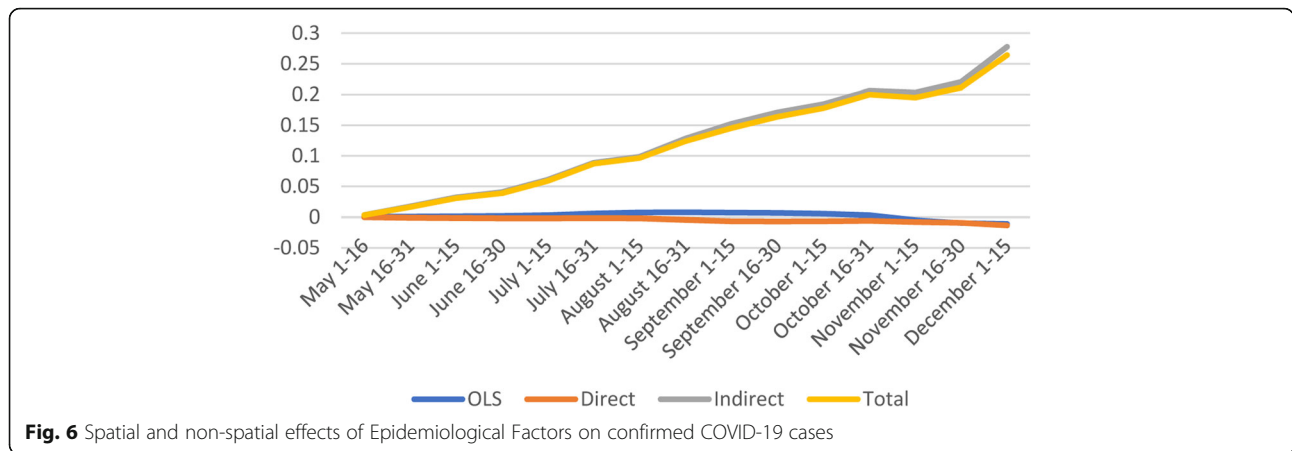
whereas indirect effects show positive effects during the same period (Fig. 6). However, total effects are positive and significant during the same sub-periods which is partially consistent with Neelon et al. (2020) who found that the percentage in fair or poor health tested was positively associated with more COVID-19 cases; patients with comorbidities had a higher risk of mortality than patient without [11, 33].

Healthcare System
Healthcare System vulnerability factors do contribute to higher number of confirmed cases as shown by the positive and significant effects in most subperiods except during the later months of 2020 (Fig. 7). Interestingly, the indirect and total effects show significant negative coefficients early in the pandemic (May 1–June 15) and positive coefficients for the last quarter of 2020 (September 1–November 30).

The U.S health care system experienced tremendous pressure during the pandemic, and certain measures

Fig. 5 Spatial and non-spatial effects of Housing Type, Transportation, Household Composition & Disability on confirmed COVID-19 cases

Fig. 6 Spatial and non-spatial effects of Epidemiological Factors on confirmed COVID-19 cases

have been implemented to reduce the burden such as elective procedures, inpatient/outpatient surgery, procedures cancelled [34]. The evolution in treatment protocols has decreased dependence on ventilators, lowered overall hospital length of stay, and decreased clinical and demographic-adjusted mortality rates [35]. These measures allowed hospitals to increase inpatient bed capacity, equipment, the availability of health care providers, and decrease the risk of spread of the virus.

High-Risk Environments
High-Risk environment vulnerability has negative and significant direct effects coefficients early in the pandemic and turn positive later. Indirect and total effects are positive and significant from August to December (Fig. 8).
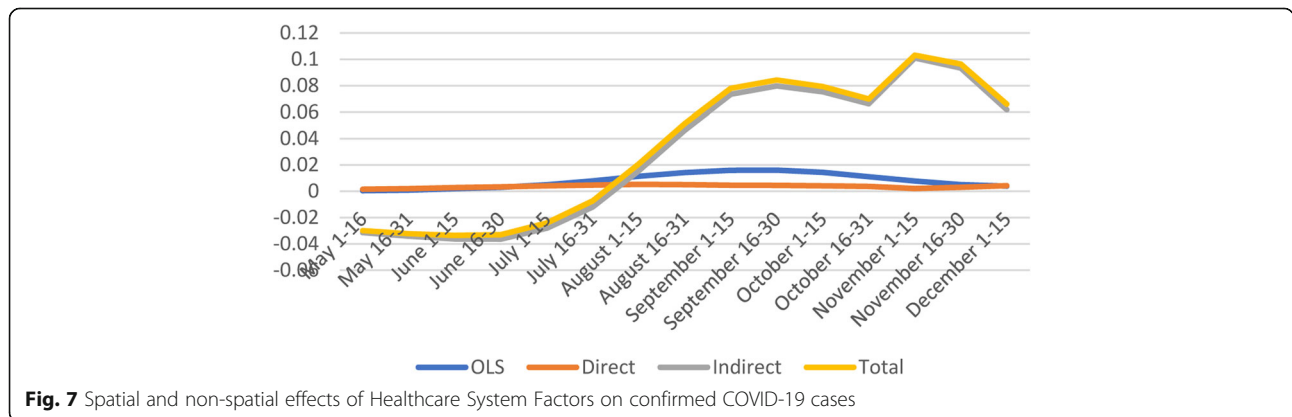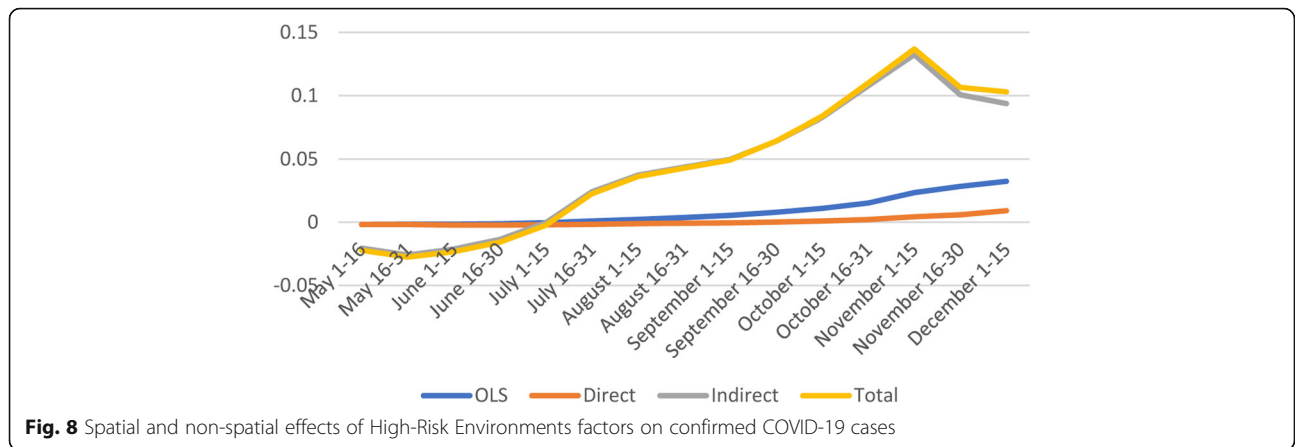
Population density
The direct effects are positive and significant only during a short period (July 1–August 15), and insignificant for the remaining sub-periods. Along with other studies [36, 37], Kadi et al. (2020) used Algerian data and found a strong correlation between the population density and

the number of COVID-19 infections [38]. The indirect and total effects start off positive (May 1–June 15) and turned negative during the last quarter of 2020 (Fig. 9). Studies found that population density is not the sole factor contributing to the adverse COVID-related impacts in cities [39].

Conclusions
The purpose of this study was to assess the effects of pre-existing vulnerability factors on COVID-19 confirmed cases while accounting for spatial spillover effects. The findings show that the effect on COVID-19 cases differs widely across vulnerability factors and across counties. Consistent with previous studies, vulnerability factors such as Minorities & Language, Epidemiological factors, Transport & Housing, and Population Density are positively associated with COVID-19 confirmed cases, suggesting that higher vulnerability is associated with higher number of COVID-19 confirmed cases. The results also indicate that spatial spillover effects are important determinants of COVID-19 confirmed cases, and in some cases, these effects are higher than the direct non-spatial effects. On the other

Fig. 7 Spatial and non-spatial effects of Healthcare System Factors on confirmed COVID-19 cases

Fig. 8 Spatial and non-spatial effects of High-Risk Environments factors on confirmed COVID-19 cases

hand, direct effects estimates are on average significant and carry the expected positive sign, except for some subperiods.

A major difference with previous studies is that COVID-19 confirmed cases are examined at different points in time (March 1 – December 15) which provides a better understanding of the rapidly evolving relationship between vulnerability factors and COVID-19 confirmed cases. Although Neelon et al. (2020) used longitudinal data to examine social determinants of health and disparities in COVID-19 incidence, spatial spillovers effects were not accounted for [11]. The magnitude and direction of the relationship between vulnerability factors and COVID-19 cases vary widely among U.S. counties. Karaye et al. (2020) found that in the states of Washington and Oregon, minority status & Language as well as household composition and disability were more predictive of COVID-19 cases than housing and transportation. However, vulnerability to COVID-19 was better explained by housing and transportation than by minority status and language

in other states (Gulf Coast states, Southern Arkansas, and Western Tennessee) [19].

Because of significant spillover effects, the first step in the fight against the pandemic should be a federal commitment to containing the spread of coronavirus and related fears. Using Italy's experience, Pisano et al. (2020) notes that the most effective time to take strong action is extremely early, when the threat appears to be small — or even before there are any cases [40].

If anything, the COVID-19 crisis has taught us that pandemics precipitate much more than a global health catastrophe, countries or states cannot fix it alone, and even though pandemics have been anticipated for many years, countries were profoundly unprepared [41]. Because of potential negative spillover effects of individual actions on nearby communities, there is a need for global coalition against COVID-19 pandemic.

Moreover, the more people continue to move from one location to the other, containing infection would be challenging and would largely depend on the coordination across borders. An effective approach to

Fig. 9 Spatial and non-spatial effects of Population Density on confirmed COVID-19 cases
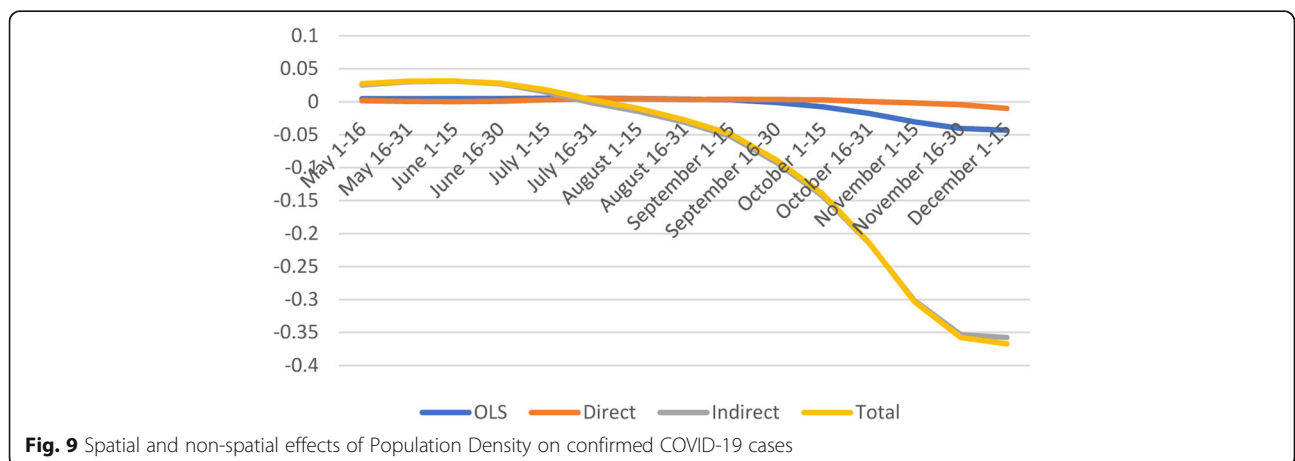
controlling a pandemic such as COVID-19 should involve national/federal coordination.

One limitation of this study is that rural and urban counties were not differentiated. Indeed, counties with the highest relative pressure on healthcare system are in more rural areas of the country and have also relatively fewer resources nearby to support patients as opposed to major metropolitan areas.

Abbreviations
CDC: Center for Disease Control; SVI: Social Vulnerability Index; CCVI: COVID-19 Community Vulnerability Index; AHRQ: Agency for Healthcare Research and Quality; PQI: Prevention Quality Indicator (PQI) composite; USDA: The United States Department of Agriculture

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12889-021-11809-2.

Additional file 1.

Acknowledgements
Not applicable.

Authors' information (optional)
Not applicable'.

Authors' contributions
JU: acquisition of data, analysis (modelling and estimation), results and discussion, conclusion. AK: literature review, description of data, results and discussion, conclusion. All authors read and approved the final manuscript.".

Funding
Not applicable.

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
"The authors declare that they have no competing interests."

Author details
[1]International Food Policy Research Institute (IFPRI), Eye Street, 1201 I St NW, Washington, DC 20005, USA. [2]Institution: Ohio University, 1 Ohio University, Athens, OH 45701, USA.

References
1. COVID Data Tracker [Internet]. Centers for Disease Control and Prevention. 2021 [cited 4 May 2021]. Available from: https://covid.cdc.gov/covid-data-tracker/#datatracker-home
2. Dasgupta S, Bowen VB, Leidner A, Fletcher K, Musial T, Rose C, et al. Association between social vulnerability and a County's risk for becoming a COVID-19 hotspot—United States, June 1–July 25, 2020. Morb Mortal Wkly Rep. 2020;69(42):1535–41. https://doi.org/10.15585/mmwr.mm6942a3.
3. Clark A, Jit M, Warren-Gash C, Guthrie B, Wang HH, Mercer SW, et al. Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: a modelling study. Lancet Glob Health. 2020;8(8):e1003–17. https://doi.org/10.1016/S2214-109X(20)30264-3.
4. Raifman MA, Raifman JR. Disparities in the population at risk of severe illness from COVID-19 by race/ethnicity and income. Am J Prev Med. 2020;59(1):137–9. https://doi.org/10.1016/j.amepre.2020.04.003.
5. Wiemers EE, Abrahams S, AlFakhri M, Hotz VJ, Schoeni RF, Seltzer JA. Disparities in vulnerability to severe complications from COVID-19 in the United States. National Bureau of Economic Research; 2020.
6. Laurencin CT, McClinton A. The COVID-19 Pandemic: a Call to Action to Identify and Address Racial and Ethnic Disparities. J Racial Ethn Health Disparities. 2020;7(3):398–402. (https://doi.org/10.1007/s40615-020-00756-0).
7. Huang R, Zhu L, Xue L, Liu L, Yan X, Wang J, et al. Clinical findings of patients with coronavirus disease 2019 in Jiangsu province, China: a retrospective, multi-center study. PLoS Negl Trop Dis. 2020;14(5):e0008280. https://doi.org/10.1371/journal.pntd.0008280.
8. Livingston E, Bucher K. Coronavirus disease 2019 (COVID-19) in Italy. JAMA. 2020;323(14):1335.
9. Nayak A, Islam SJ, Mehta A, Ko YA, Patel SA, Goyal A, et al. Impact of social vulnerability on COVID-19 incidence and outcomes in the United States. medRxiv. [Preprint] 2020. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7217093/. Accessed 9 Oct 2020.
10. Kiaghadi A, Rifai HS, Liaw W. Assessing COVID-19 risk, vulnerability and infection prevalence in communities. PLoS One. 2020;15(10):e0241166. https://doi.org/10.1371/journal.pone.0241166.
11. Neelon B, Mutiso F, Mueller NT, Pearce JL, Benjamin-Neelon SE. Spatial and temporal trends in social vulnerability and COVID-19 incidence and death rates in the United States. PLoS One. 2021;16(3):e0248702. https://doi.org/10.1371/journal.pone.0248702.
12. Macharia PM, Joseph NK, Okiro EA. A vulnerability index for COVID-19: spatial analysis at the subnational level in Kenya. BMJ Glob Health. 2020;5(8):e003014. https://doi.org/10.1136/bmjgh-2020-003014.
13. Nicodemo C, Barzin S, Lasserson D, Moscone F, Redding S, Shaikh M. Measuring geographical disparities in England at the time of COVID-19: results using a composite indicator of population vulnerability. BMJ Open. 2020;10(9):e039749. https://doi.org/10.1136/bmjopen-2020-039749.
14. McLaren J. Racial disparity in COVID-19 deaths: seeking economic roots with census data. National Bureau of Economic Research; 2020.
15. Benitez J, Courtemanche C, Yelowitz A. Racial and ethnic disparities in Covid-19: evidence from six large cities. J Econ Race Policy. 2020 Dec;3(4):243–61. https://doi.org/10.1007/s41996-020-00068-9.
16. Knittel CR, Ozaltun B. What does and does not correlate with COVID-19 death rates. National Bureau of Economic Research; 2020.
17. Desmet K, Wacziarg R. JUE insight: understanding spatial variation in COVID-19 across the United States. J Urban Econ. 2021:103332. https://doi.org/10.1016/j.jue.2021.103332.
18. Balcan D, Gonçalves B, Hu H, Ramasco J, Colizza V, Vespignani A. Modeling the spatial spread of infectious diseases: the GLobal epidemic and mobility computational model. J Comput Sci. 2010;1(3):132–45. https://doi.org/10.1016/j.jocs.2010.07.002.
19. Sun F, Matthews SA, Yang TC, Hu MH. A spatial analysis of the COVID-19 period prevalence in US counties through June 28, 2020: where geography matters? Ann Epidemiol. 2020;52:54–9. https://doi.org/10.1016/j.annepidem.2020.07.014.
20. Karaye IM, Horney JA. The impact of social vulnerability on COVID-19 in the US: an analysis of spatially varying relationships. Am J Prev Med. 2020;59(3):317–25. https://doi.org/10.1016/j.amepre.2020.06.006.
21. Khazanchi R, Beiter ER, Gondi S, Beckman AL, Bilinski A, Ganguli I. County-level association of social vulnerability with COVID-19 cases and deaths in the USA. J Gen Intern Med. 2020;35(9):2784–7. https://doi.org/10.1007/s11606-020-05882-3.
22. Anselin L, Florax RJ, Rey SJ. Econometrics for spatial models: recent advances. Adv Spat Econometrics. 2004:1–25. https://doi.org/10.1007/978-3-662-05617-2_1.
23. Lewnard JA, Lo NC. Scientific and ethical basis for social-distancing interventions against COVID-19. Lancet Infect Dis. 2020;20(6):631–3. https://doi.org/10.1016/S1473-3099(20)30190-0.

24. Social Media at CDC [online]. Centers for Disease Control and Prevention. 2021. [Accessed 7 August 2021]. Available from https://www.cdc.gov/socialmedia/syndication/405380/403327.html.
25. COVID-19 and Your Health [Internet]. Centers for Disease Control and Prevention. 2021 [cited 8 March 2021]. Available from: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html
26. Qu X, Lee LF. LM tests for spatial correlation in spatial models with limited dependent variables. Reg Sci Urban Econ. 2012;42(3):430–45. https://doi.org/10.1016/j.regsciurbeco.2011.11.001.
27. Anselin L. Spatial econometrics: methods and models. Kluwer Academic, Dordrecht: Springer Netherlands; 1988. https://doi.org/10.1007/978-94-015-7799-1.
28. LeSage JP, Fischer MM. Spatial growth regressions: model specification, estimation and interpretation. Spat Econ Anal. 2008;3(3):275–304. https://doi.org/10.1080/17421770802353758.
29. LeSage JP, Kelley Pace R, Lam N, Campanella R, Liu X. New Orleans business recovery in the aftermath of hurricane Katrina. J R Stat Soc A (Statistics in Society). 2011 Oct;174(4):1007–27. https://doi.org/10.1111/j.1467-985X.2011.00712.x.
30. CDC/ATSDR's Social Vulnerability Index (SVI) [Internet]. Atsdr.cdc.gov. 2021 [cited 5 May 2021]. Available from: https://www.atsdr.cdc.gov/placeandhealth/svi/index.html
31. Bringing Greater Precision to the COVID-19 Response [Internet]. Precisionforcovid.org. 2021 [cited 9 September 2020]. Available from: https://precisionforcovid.org/ccvi
32. USDA ERS - The Meatpacking Industry in Rural America During the COVID-19 Pandemic [Internet]. Ers.usda.gov. 2021 [cited 5 March 2021]. Available from: https://www.ers.usda.gov/covid-19/rural-america/meatpacking-industry/
33. Ssentongo P, Ssentongo AE, Heilbrunn ES, Ba DM, Chinchilli VM. Association of cardiovascular disease and 10 other pre-existing comorbidities with COVID-19 mortality: a systematic review and meta-analysis. PLoS One. 2020;15(8):e0238215. https://doi.org/10.1371/journal.pone.0238215.
34. Barnett ML, Mehrotra A, Landon BE. Covid-19 and the Upcoming Financial Crisis in Health Care. NEJM Catalyst Innovations Care Delivery. 2020;1(2):1–6. (https://catalyst.nejm.org/doi/full/10.1056/CAT.20.0153).
35. Horwitz L, Jones SA, Cerfolio RJ, Francois F, Greco J, Rudy B, et al. Trends in COVID-19 Risk-Adjusted Mortality Rates. J Hosp Med. 2021;16(2):90–2. (https://doi.org/10.12788/jhm.3552).
36. Bouba-Olga, O. La densité favorise-t-elle l'épidémie ? (épisode 26) | [Internet]. Blogs.univ-poitiers.fr. 2020 [cited 3 November 2020]. Available from: https://blogs.univpoitiers.fr/o-bouba-olga/2020/04/29/la-densite-favorise-t-elle-lepidemie-episode-26/
37. Babbitt D, Garland P, Johnson O. Lived population density and the spread of COVID-19. arXiv preprint arXiv. 2005;01167:2020.
38. Kadi N, Khelfaoui M. Population density, a factor in the spread of COVID-19 in Algeria: statistic study. Bull Natl Res Centre. 2020;44(1):1–7. https://doi.org/10.1186/s42269-020-00393-x.
39. Hsu J. Population density does not doom cities to pandemic dangers. Sci Am. 2020. Available from: https://www.scientificamerican.com/article/population-density-does-not-doom-cities-to-pandemic-dangers/. Accessed 6 Jan 2021.
40. Pisano GP, Sadun R, Zanini M. Lessons from Italy's response to coronavirus. Harv Bus Rev. 2020. Available from: https://hbr.org/2020/03/lessons-from-italys-response-to-coronavirus. Accessed 2 Nov 2020.
41. R. Allen J, von Hippel K. Where is the global coalition to fight the coronavirus? [Internet]. Brookings. 2021; [cited 4 May 2021]. Available from: https://www.brookings.edu/president/where-is-the-global-coalition-to-fight-the-coronavirus/.

Publisher's Note