SPECIAL ISSUE PAPER

Regional poverty and infection disease: early exploratory evidence from the COVID-19 pandemic

Abu Bakkar Siddique[1] · Kingsley E. Haynes[1] · Rajendra Kulkarni[1] · Meng-Hao Li[1]

Abstract

This paper examines the role of regional poverty on the COVID-19 pandemic in the USA. It also explores how the effects differ with the concentration of ethnic minorities. We find that poverty is a significant and consistent determinant of higher COVID-19 infections and fatalities. Prevalent poverty areas experienced higher infections due to economic structure that require hypermobility (high mobility and interpersonal interaction)—more physical human to human contact resulting in higher deaths from limited access to health services. These are also regions where minority groups are concentrated. Disproportionate infections and fatalities occurred within the black, Hispanic, and Asian population. Our evidence is robust to state fixed effects that capture local COVID-19 mitigation policies, multi-level hierarchical modeling, spatial autoregressive assessment, and large sets of county-level health, social, and economic factors. This paper contributes to the literature on health and economic disparities and their resulting consequences for infectious diseases.

JEL Classification I30 · I14 · R12 · C31

1 Introduction

Over the last century, different types of pathogens have killed more people than any armed conflict (Adda 2016). COVID-19 is a disease caused by SARS-CoV-2, and it can infect people via direct and indirect human-to-human contact. The differences we observe in the infections and fatalities are due to built-in characteristics within the society that generate different exposure levels and different responses to the virus attack. All regions are not equally responsive to the diffusion and impact of the virus

---

✉ Abu Bakkar Siddique
    asiddi@gmu.edu

✉ Kingsley E. Haynes
    khaynes@gmu.edu

1   Schar School of Policy and Government, George Mason University, Arlington, VA 22201, USA

since most regions reflect different social, demographic, behavior, and economic characteristics and related responses. Therefore, although the attack of COVID-19 on human society was exogenous, the response to the virus was completely endogenous to society's established characteristics.

COVID-19 appears to be impacting people disproportionately across different socioeconomic groups and varies by ethnicity, income, age, gender, and social groups or communities. This is reflected in the relative concentration of these communities in different regions. However, it is not possible to simply and reliably put together the threads of evidence that clarify these discriminatory COVID-19 impacts. The evidence of differences linked to impact differentials needs to be better understood. What are the aggregated regional characters that are fundamental to these disparities in virus infections and fatalities? This paper explores these disproportionate impacts for COVID-19.

The poverty rate is an economic variable, and we evaluate how the poverty rate influences the rate of transmission and consequences for viral disease—here COVID-19. Earlier literature has investigated several economic variables and their role in health conditions related to economic activity in the context of infectious diseases (Adda 2016). These variables include income, GDP per capita, and unemployment rates (Adda et al. 2009; Ettner 1996; Ruhm 2000, 2003, 2005). Others have also looked at the impact of trade (imports and exports) and international exposure in a context of viruses and bacteria as they relate to human infections, including small-pox, whopping cough, scarlet fever, mumps, measles, polio, tuberculosis, and HIV among many others (Oster 2005).

In this paper, we test two main hypotheses fundamental to understanding the relationships between poverty and COVID-19. In theory, it is not clear whether poverty should worsen the transmission of infectious diseases, or help it contain the diseases. It is even more complex and obscure in the developed world context. If poverty prone areas have limited economic activities and that in turn limits human interaction, poverty may help to control infectious diseases like COVID-19. On the other hand, if poverty is related to physical interpersonal contacts, limited health care capacity, and certain interaction behaviors, it might dominate other poverty-related limited economic activity, which means we may experience higher infections in poverty prone areas. For example, financially poorer people are often essential workers that demand physical presence and are not permitted to work from home. The increase in physical interpersonal contacts and decline of compliance with social distancing in poverty prone areas would exacerbate the spread of the COVID-19 disease (Kim and Kwan 2021; Liu et al. 2020). Furthermore, economically disadvantaged neighborhoods may be clustered by houses with poor conditions and have less safety in outdoor spaces. Overcrowding in poverty prone areas will put people at higher risk of spreading the contagious disease (Rollston and Galea 2020). Moreover, poverty and public shame may shape how people experience being observed and judged by others during the pandemic. A strict disciplinary measure and shaming in any form (rather than harm-reduction strategy) can cause stigmatization and eventually undermine public health efforts by discouraging people, particularly poor people, from getting tested and disclosing their health status to contact-tracers or cooperating with other mitigation programs (Collins 2020; Gold 2020). Saxe (2020) presented how stigmatization and

criminalization deter testing and exacerbate problems of abusive policing of members of marginalized communities in the HIV epidemic. Similarly, a higher poverty rate can be responsible for higher death and fatality rates (fatality rate is the ratio of COVID-19 infections to deaths, converted into percentage) because poor regions may have limited capacity to provide medical access and treatment to infected (COVID-19) patients. Additionally, economically disadvantaged people may not be able to afford health insurance and hesitate to get the COVID-19 test which in turn limit their access to health resources and treatments in poverty prone areas (Cordes and Castro 2020).

The majority of this literature covering the characteristics and consequences of the pandemic[1] can be grouped as follows: (a) the measurement of the spread of COVID19 and the role of mitigation such as masks, crowding limitation, and social distancing, (b) the degrees of disease transmission, plus the effectiveness and compliance with social distancing, (c) the economic impacts of COVID-19 such as the impact on employment, (d) the socioeconomic consequences of extreme measures such as shelter-in-place or lockdowns, and e) the governmental response to the pandemic (Brodeur et al. 2020).

Some literature attempts to relate COVID-19 to key socioeconomic variables as causes for the difference in impacts. Most of these have yet to be peer-reviewed (Chin et al. 2020; Li et al. 2020). Other studies focus on spatial analysis (Sun et al. 2020), ethnic disparities (Li et al. 2020), development of online dashboards for tracking COVID-19 (Wissel et al. 2020), use of vulnerability indices (Mukherji 2020), and issues in equitable COVID-19 response (Chin et al. 2020). The major concern of these studies is that they do not adequately address the confounding factors. For example, counties in each state were subject to various COVID-19 mitigation policies. Therefore, simple county-level cross section analysis often suffers from selection biases by not controlling for state mitigation policies. This is the case for most pre-print unpublished articles. We address these problems using multi-level data with state fixed effects which control all COVID-19 policies and other factors across the states. We observe a significant difference in results with and without state fixed effect controls. We also estimate a multi-level mixed method for robustness checks which is a suitable and appropriate strategy given in our research design and data set. In this paper, we hypothesize that poverty worsens the COVID-19 impact, and we test it using US county-level data. However, this expectation is not obvious or not agreed upon.

The next section states the hypotheses about the relationship between COVID-19 and poverty. Section 3 reports methods, data, and descriptive statistics. Section 4 presents the results of cumulative COVID-19 data until the end of July 2020. The next section, Sect. 5, presents a robustness check by disaggregating results by monthly cumulative cases and deaths and by estimating a multi-level mixed method. Section 6 concludes the paper.

---

[1] See Brodeur et al. (2020) for a review of these literature.

2 Poverty and COVID-19

Earlier studies on similar infectious diseases demonstrated that poverty, poor health, and poor sanitary conditions make such health crises worse. This is reflected in studies in most developing countries including most recently Campos et al. (2018) for the Zika virus and Redding et al. (2019) for the Ebola virus. The Infectious Disease Vulnerability Index for countries in Africa developed by Moore et al. (2017) using Ebola to inform actions for preparedness and response to infectious disease outbreaks was dependent on national health systems worldwide. Such studies are rare for developed economies but are still quite relevant even for developed economies and are particularly important when it comes to COVID-19 in the USA. Except for Adda (2016) who offers an extensive analysis of the transmission of three viruses—influenza, gastroenteritis, and chickenpox using data across a century of national health systems in France; these national data systems have not been as widely used as they should have been. One of the questions that Adda (2016) asked was whether the virus spread more rapidly during periods of economic growth and if the spread followed a "gradient determined by economic factors." He found that the viruses studied propagated faster during times of economic boom due to increased economic activity and contact between people. Qiu et al. (2020) conducted a similar analysis for Wuhan, China, and also had a positive relationship between the spread of the virus and economic activity. These studies imply economic expansion, not poverty, helps spread of disease. Hence, a decrease in poverty rather than its increase may be related to growth in the disease.

When it comes to relating infectious diseases such as COVID-19 to poverty, the relationship is not straightforward. Whether poverty will precipitate the spread of the virus or will limit its spread may depend more on the micro-level characteristics that result from poverty. For example, if we assume that poorer regions have limited economic activities, they should have lower diffusion rates of infectious disease and lower aggregated deaths. Adda (2016) noted that higher economic activities cause the spread of infectious diseases while poverty-prone areas may have limited economic activities that contain the virus because of less mobility and traveling, which in turn reduces interpersonal contacts and reduces the diffusion (spread) of the diseases.

On the other hand, alternative evidence shows that economic downturns precipitate the transmission of infectious diseases due to limiting the capacity to control the disease (Suhrcke et al. 2011). If poverty-prone areas have more work that requires physical presence and person-to-person contact, this can increase the spread of the virus. Moreover, poverty makes people reluctant to take sick leave from work, fearing unemployment, while increasing the risk of disease transmission at the workplace (Barmby and Larguem 2009). During the 1990s, countries of the former Soviet Union (FSU) and Eastern Europe experienced a devastating economic crisis, as GDP fell by one-third on average, which markedly increased the incidence, prevalence, and mortality of tuberculosis, and worsening treatment. This, in turn, led to the emergence of drug-resistant strains (Shilova and Dye 2001). Similar experiences have also been recorded for HIV prevalence, outbreaks of diphtheria (Markina et al. 2000),

tick-borne encephalitis (Randolph 2008), and leptospirosis (Stoilova and Popivanova 1999).

Related literature suggests that infectious diseases disproportionately affect vulnerable groups. In a review of the European literature, this effect could be found in every single EU Member State (Semenza and Giesecke 2008). A separate study comparing wealth distribution and tuberculosis (TB) rates across the EU Member States demonstrated a strong correlation between income equality and lower TB rates (Suk et al. 2009).

Thus, it is unclear which dominates the overall impact of higher poverty and the spread of the virus. If limited economic activities resulting from poverty dominate, the impact we will see is the lowering of the spread of the COVID-19 in poverty-concentrated areas. However, if it is the latter, where poverty drives limited health facility capacity, hypermobility, interactive economic behavior, and job characteristics and it is dominant, then we should see a higher level of infections in poverty concentrated areas. While we test this hypothesis (H1) in the US context at the county level, we expect to see higher infections in poorer counties since we believe the latter scenario dominates in the USA. This is especially true when it is widely viewed that poverty is related to hypermobility (Schafft 2006).

H1 Higher poverty concentrated regions will have higher COVID-19 infections

Regarding deaths, the association is even more predictable. Poverty-prone areas have limited health sector capacity and higher budget constraints. A pandemic like COVID-19 can easily disrupt the limited health and treatment systems and infected people may die more in regions with higher poverty than in regions of less poverty. Moreover, poverty is expected to be associated with limited to no health insurance coverage, and if people cannot afford to cover the cost of treatment (out of pocket payment), the consequences can be dire.[2] Therefore, people will seek less health treatment and have less access to health services leading to higher deaths in poverty concentrated areas. We test the second hypothesis (H2) to understand the relationship between poverty and COVID-19-related deaths and fatalities.

H2 Higher poverty concentrated regions will have higher COVID-19 deaths and fatalities.

---

[2] These people can be covered through public services such as Medicaid. It is the people who have low income, certain immigrants, and those who do not meet special state determined characteristics who are not eligible for Medicaid. It is this population and those who are close to or right above the poverty line that have the most coverage problems.

3 Methods and data

3.1 Methods

The unexpected incidence of the pandemic created an opportunity for a quasi-natural experiment over time. We use the pre-pandemic poverty rate to see the relations between poverty and COVID-19 infectious disease which rule out the possibility of reverse causality and selection bias. We estimate the following Eq. (1) to identify the consequence of poverty on the COVID-19 pandemic.

$$Y_{cs} = \beta_0 + \beta_1 \text{Poverty}_{cs} + \beta_2 X_{cs} + \beta_3 \mu_s + \varepsilon_{cs} \tag{1}$$

where Y_{cs} is the cumulative COVID-19 cases, or deaths, or fatality rate in county c in the state s, and Poverty_{cs} is the poverty rate at the county-level in the state s. X_{cs} is the county-level covariates and ε_{cs} is the error term. The μ_s is the state fixed effect, which captures the state-level observable and unobservable factors that are specific to a state. Since the pandemic started, many states initiated various COVID-19 mitigation policies such as stay-at-home order, mandating masks, closing many businesses and schools, and others that are expected to have an important role on both the infections and death rates. Moreover, besides direct mitigation policies, other spatial factors differ across states including summer school holidays, regional time zones that characterize regions differently, and variation in regional transport systems—public and private. These are all potential confounding components impacting the spread of infectious disease (Adda 2016; Siddique 2021). State fixed effect should control all these variations not captured elsewhere as they impact COVID-19 diffusion. In this fixed effect approach, we are only considering heterogeneity within states and abandoning the heterogeneity across the states that are time-invariant, but we are considering heterogeneity both within and across counties. Our primary interest is β_1, which is the estimator for the average effect of the poverty rate on COVID-19 outcomes.

3.2 Data

We conducted this analysis using county-level data. Our COVID-19 data are the cumulative confirmed cases and deaths per million population. We also estimated the impact of poverty on the COVID-19 fatality rate. The COVID-19 data cover the period from late January to July 28, 2020 and are collected from www.usafacts.org, which provides an identical data set as The Johns Hopkins University COVID-19 dashboard data set. We acknowledge that the measures of COVID-19 data may suffer from undercounting problems due to limited testing, particularly during the early pandemic.[3] However, we believe this will not be a serious problem for testing our models for two reasons. Firstly, undercounting is likely to be more prevalent in

---

[3] An alternative measure of positivity rate could have solved this issue partially; however, we did not utilize it since positivity rate data at the county level were incomplete and could have generated additional problems.

poorer counties with limited capacity and stigmatization of poorer people. Since we are expecting a positive association between the COVID-19 pandemic and poverty, if there is real undercounting, it will underestimate the impact of poverty relative to the true size of the coefficient, not overestimate it. Secondly, we believe undercounting is more likely to happen early in the pandemic than later, so we disaggregated the cumulative data by months and check the robustness of our results that might be driven by potential errors in COVID-19 case and death counts.

Data for all independent variables are for the year 2019. We do this to avoid reverse causality, i.e., the impact of COVID-19 on independent variables such as poverty rate (Siddique 2016). For example, recently, Lima et al. (2020) reported that Brazil, which has suffered one of the world's worst pandemic tolls, has responded to the crisis by distributing a large amount of cash directly to citizens that poverty and inequality have approached national historic lows. Therefore, we applied a one-year lag value to all independent variables so that there is no reverse impact on right-hand side variables in Eq. (1). At this writing, it should also be noted that 2020 data are not yet available for most variables. Poverty data are collected from the U.S. Census Bureau, Small Area Income, and Poverty Estimates (SAIPE) for the year 2018. These are the latest data available at the county level. Income inequality data are also collected from the SAIPE for the year 2018.

Most other data are collected from the County Health Rankings & Roadmaps (CHR&R) program, which is derived from a collaboration between the Robert Wood Johnson Foundation and the University of Wisconsin Population Health Institute. To avoid multicollinearity and too many control variables in the models, we created three composite health indexes, health behavior, clinical care, and the physical environment. We then calculated z-scores to rank the counties. The health behavior ranking is composed of tobacco use, diet and exercise, alcohol and drug use, and sexual activity. Clinical care is composed of access to care and quality of care. Ranking of the physical environment includes air and water quality and housing and transit. To rank them, we followed guidance from the CHR&R.[4] Most of these data have been compiled from the American Community Survey (ACS), USDA Food Environment Atlas, National Center for Health Statistics, Behavioral Risk Factor Surveillance System, Small Area Health Insurance Estimates, Centers for Medicare and Medicaid Services (CMS), National Provider Identification and Environmental Justice Screening and Mapping Tool (EJSCREEN).

We also controlled for many economic and demographic variables that can potentially impact both the poverty rate and the COVID-19 pandemic. While economic variables are collected from ACS, the Bureau of Labor Statistics, and SAIPE, the demographic variables are from the Census Bureau's Population Estimates.

3.3 Descriptive statistics

Table 1 presents the descriptive statistics of all variables used in the models. The mean infection and death per million population are 9077.99 and 224.43, respectively,

---

[4] Details of the ranking process and weighting factors are available in the CHR&R website at https://www.countyhealthrankings.org/explore-health-rankings/measures-data-sources.

Table 1 Descriptive statistics

Variable	Obs	Mean	Std. Dev	Min	Max
Dependent variables					
Infections per million population[†]	3,142	9077.99	9724.26	0.00	155,211.70
Death per million population[†]	3,142	224.43	377.28	0.00	3971.50
Fatality rate[†]	3,107	2.22	2.90	0.00	25.00
Independent variables					
Poverty rate	3,192	15.13	6.09	2.60	54.00
Income inequality (top 20% / bottom 20%)	3,142	4.52	0.75	2.56	10.10
Health and physical					
Ranking of health behavior[®]	3,081	0.00	0.17	− 0.68	0.71
Ranking of clinical care[®]	3,081	0.00	0.12	− 0.48	0.41
Ranking of physical environment[®]	3,081	0.00	0.05	− 0.18	0.19
Economic					
Rural (%)	3,135	58.58	31.48	0.00	100.00
Median income	3,141	51,090.53	13,497.97	22,679	136,191
Unemployment rate	3,141	4.62	1.68	1.62	20.07
Some college (%)	3,142	57.61	11.70	16.76	93.67
Long commute—driving alone	3,142	30.96	12.40	0.00	84.50
Demographic					
Female (%)	3,142	49.89	2.24	26.57	57.00
Sixty-five plus (%)	3,142	18.81	4.62	4.77	56.94
Black (%)	3,142	8.97	14.31	0.00	85.33
Asian (%)	3,142	1.53	2.91	0.00	43.01
Hispanic (%)	3,142	9.48	13.75	0.52	96.32

[†]Indicates COVID-19 data are the cumulative data till July 28, 2020. Fatality rate ((case per million/death per million) * 100). [®]Indicates composite index and see at https://www.countyhealthrankings.org/ for the details of the indices and their measures. The ranking value is a standardized z-score and reversely sorted from lowest to highest within each state. The lowest score represents best health, and the highest score represents the worst health. Rural (%): The Census Bureau defines rural as any population, housing, or territory NOT in an urban area. Some College is the percentage of the population ages 25–44 with some post-secondary education, such as enrollment in vocational/technical schools, junior colleges, or four-year colleges. Long Commute—Driving Alone is the percentage of workers who drive alone with a commute longer than 30 min

while the mean fatality rate is 2.22. The mean poverty rate is 15.13. Figures 1, 2, 3 present the maps of the poverty rate and COVID-19 infections, deaths, and fatality rates. In the maps, the polygon layer shows the poverty rate. Color codes range from dark blue to yellow. The yellow color represents the highest poverty rate category. The 3D cones are for infections and deaths per million population and fatality rate in all three figures. These maps clearly show that more 3D cones with higher height are located in the higher poverty counties. The simple correlation test also shows that the correlation between poverty rate and infection per million population is 0.31 and between poverty rate and death per million population is 0.20. The next section,
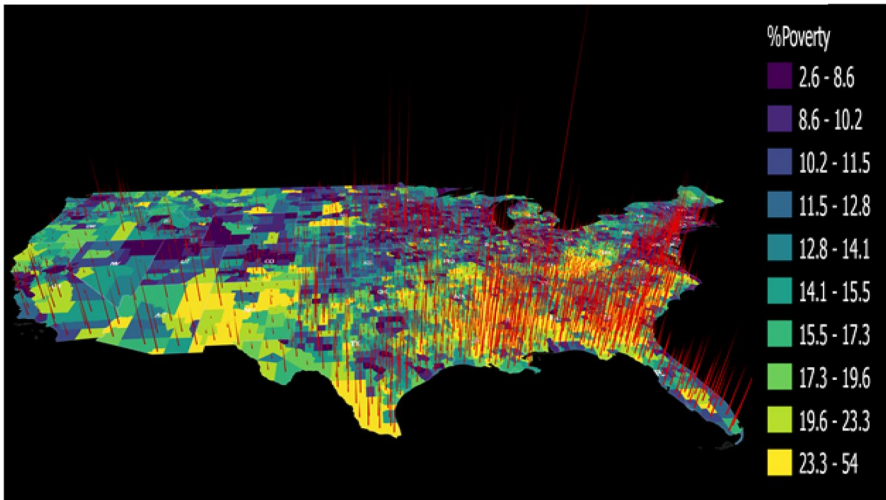
Fig. 1 Poverty rate and COVID-19 infections (poverty in color and infection in bar height). 3D cones represent infection data (higher the height higher the infections) and color codes represent the poverty rate at the county level
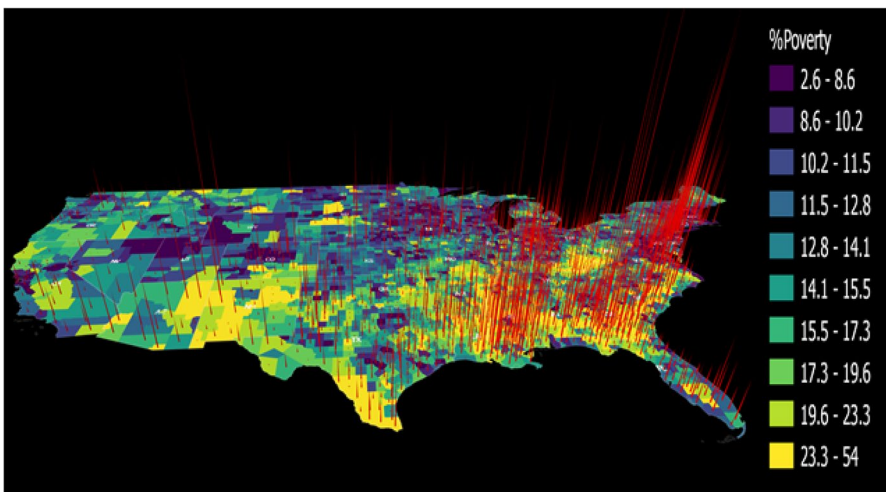

Fig. 2 Poverty rate and COVID-19 deaths**.** 3D cones represent the death (higher the height higher the deaths) and color codes represent the poverty rate at the county level

following the Figures and Descriptive Statistics, provides the output from the multivariate regression with state fixed effects and test our hypotheses.
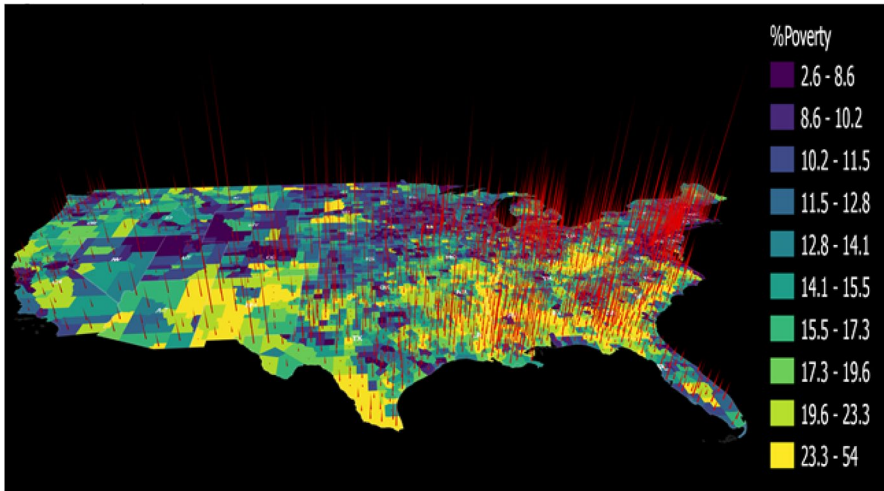
Fig. 3 Poverty rate and COVID-19 fatalities. 3D cones represent the fatality rate (higher the height higher the fatality rate) and color codes represent the poverty rate at the county level

4 Results analysis

Table 2 presents the results from an ordinary least square (OLS) estimate with state fixed effects using cumulative COVID-19 data from January 22 to the end of July 2020. Panel A to C in Fig. 4 presents the marginal effect of poverty on the COVID-19 infections, deaths, and fatalities, respectively. Results confirm our hypothesis that poverty worsened the COVID-19 impact both in terms of infections and deaths. This means poverty-related elements linked to health facility access and capacity, hypermobility, and direct/face to face/interactive jobs dominate over other remote economic activities in the USA while impacting the COVID-19 pandemic. We also estimate the impact of poverty on the fatality rate and we see that the higher the county poverty level, the greater the overall COVID-19 fatality rates. This also confirms our other hypothesis that poverty-driven limits on treating COVID-19 patients cause higher deaths in the USA. Note that, we did not test the direct effect of limited health facility access and capacity, hypermobility, and job characteristics (interactive behavior) on the infections and deaths and we do not claim that as a causal outcome. However, this might be a fair interpretation of the positive association between poverty and COVID-19 infections and deaths given the possible alternative scenarios.

Models 2.1 to 2.3 include some basic controls such as income inequality, the ranking of health behavior, clinical care, and physical environment, percentage of the rural population, the log of median household income, gender, population percentage

Table 2 Regression results on COVID-19 confirmed case, death, and fatality rate until the end of July

	Model 2.1 Confirmed cases per million	Model 2.2 Deaths per million	Model 2.3 Fatality rate	Model 2.4 Confirmed cases per million	Model 2.5 Deaths per million	Model 2.6 Fatality rate
Poverty rate	403.748***	17.814***	0.072***	207.651**	8.456***	0.047**
	(95.961)	(3.255)	(0.022)	(96.723)	(3.260)	(0.023)
Income inequality (top 20%/bottom 20%)	3.228	47.621***	0.084	334.367	49.819***	0.073
	(322.525)	(16.198)	(0.112)	(323.194)	(15.813)	(0.118)
Ranking of health behavior	− 3,494.850**	− 49.756	0.819**	− 2,504.923	− 131.273**	0.437
	(1,592.705)	(65.765)	(0.416)	(1,624.129)	(62.049)	(0.505)
Ranking of clinical care	15,083.133***	395.988***	0.014	6,030.034**	202.517***	− 0.249
	(2,250.001)	(94.619)	(0.540)	(2,563.789)	(88.066)	(0.612)
Ranking of physical environment	9,795.149***	453.042***	1.605	10,654.6***	179.832	0.414
	(2,842.045)	(128.374)	(1.218)	(2,704.545)	(116.804)	(1.305)
Rural (%)	− 56.018***	− 1.617***	− 0.004*	− 11.511	− 1.223***	− 0.006**
	(7.620)	(0.327)	(0.002)	(7.342)	(0.338)	(0.003)
Log of median income	4,172.761*	355.343***	2.225***	7,577.10***	230.270***	1.491***
	(2,199.564)	(69.933)	(0.531)	(2,123.528)	(74.752)	(0.557)
Female (%)	− 403.444***	10.457***	0.084***	− 201.327	14.902***	0.089***
	(154.253)	(4.046)	(0.029)	(166.634)	(3.782)	(0.030)
Sixty-five plus (%)	− 338.934***	3.851**	0.073***	− 197.411***	5.117**	0.068***
	(57.793)	(1.934)	(0.019)	(59.315)	(2.158)	(0.020)
Some college (%)				− 166.916***	− 5.082***	− 0.011
				(33.700)	(1.093)	(0.009)
Unemployment rate				− 108.081	− 7.432	− 0.006
				(184.805)	(6.106)	(0.046)
Black (%)				201.049***	8.962***	0.020***
				(184.805)	(6.106)	(0.046)
Asian (%)				195.336*	10.494***	0.034*
				(19.636)	(1.191)	(0.005)
Hispanic (%)				227.376***	3.210***	− 0.003
				(100.163)	(4.410)	(0.019)

Table 2 (continued)

	Model 2.1 Confirmed cases per million	Model 2.2 Deaths per million	Model 2.3 Fatality rate	Model 2.4 Confirmed cases per million	Model 2.5 Deaths per million	Model 2.6 Fatality rate
Long commute—driving alone				(28.097)	(1.018)	(0.006)
				− 45.236***	2.301***	0.013**
				(14.606)	(0.726)	(0.006)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Constant	− 2,572.39	− 4,426.41***	− 28.81***	− 48,341.59**	− 3,257.87***	− 21.01***
	(24,835.687)	(852.525)	(6.546)	(24,244.458)	(879.500)	(6.820)
R^2	0.41	0.36	0.21	0.48	0.41	0.22
N	3075	3075	3062	3075	3075	3062

Standard errors are robust and shown in parenthesis * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Fatality rate ((case per million/death per million) * 100). We conducted VIF: Variance Inflation Factor and Ridge regression and found no problem of multicollinearity
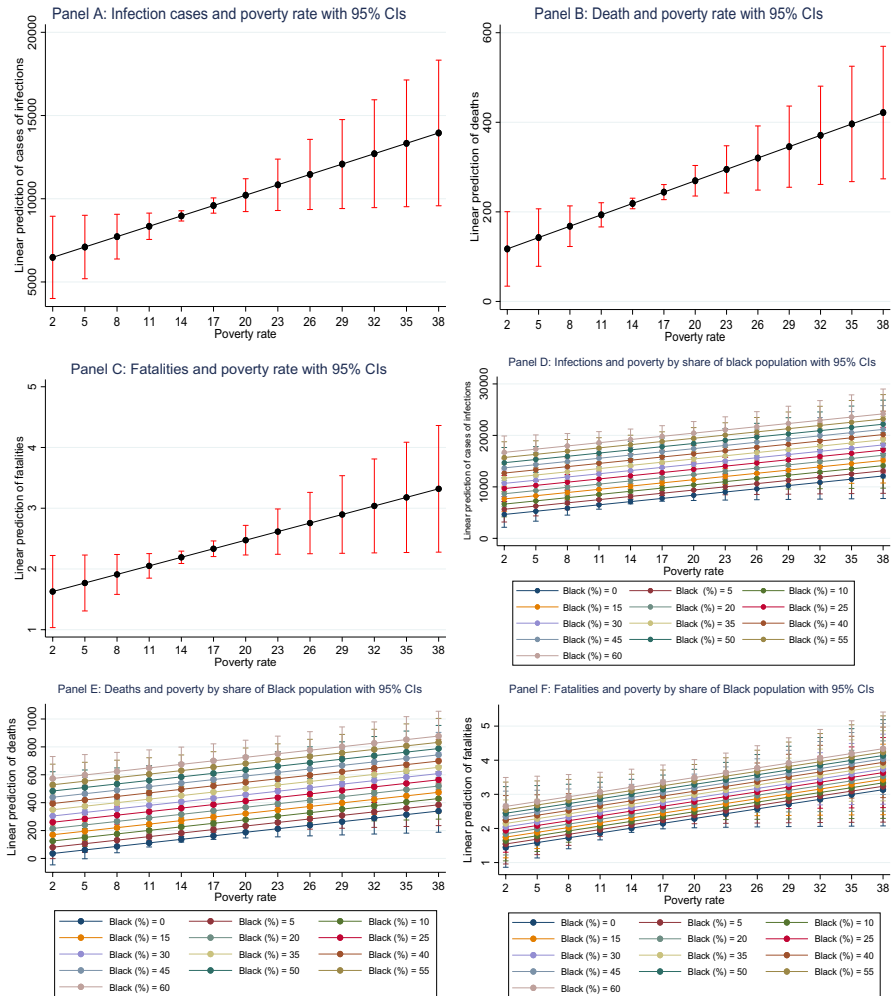
Fig. 4 Marginal effect of poverty rate on COVID-19. Panel A to C presents the marginal effect of poverty on the COVID-19 infections, deaths, and fatalities based on the models 2.4–2.6, respectively. Panel D to F presents the marginal effect of poverty on the COVID-19 infections, deaths, and fatalities by share of Black population based on the models 2.4–2.6, respectively. All these estimations have been adjusted for all covariates included in Table 2. They all show a positive impact of poverty on COVID-19 and it gets worse for counties with higher share of Black population. Figure 4 only shows the impact of poverty by Black population; however, they are equally true for other minorities like Asian and Hispanic

over sixty-five, and education. Models 2.4 to 2.6 add more controls for ethnicity, unemployment, and the local population's commuting habit.[5]

Model 2.1 reports the regression results on confirmed COVID-19 cases. It shows that a 1% increase in the poverty rate is correlated with an increase in COVID-19 confirmed cases by 403.75 per million population for a US county. This is an important and considerable number linked to the poverty rate. The impact is statistically significant at the 1% level. Poverty is also associated with a higher number of deaths per million population. Model 2.2 reports that a 1% increase in the poverty rate is related to 17.81 more deaths per million population for a county. We also estimate the impact of poverty on the fatality rate ((death per million/case per million) * 100) and find a statistically significant impact. Results in model 2.3 show that a 1% increase in poverty is associated with an increase in the fatality of 0.07, which is again statistically significant. This provides another piece of evidence suggesting that poverty is a key link to the consequences of this infectious disease.

After controlling for ethnicity, unemployment rate, and commuting behavior, the impact of poverty remains statistically significant although the magnitude of the impact drops by almost 50% for all infections, deaths, and fatality rates. This is understandable since the poverty rate is higher among minority groups which is well documented (Massey, Gross, and Shibuya 1994). It is particularly high for the black population. Panel D to F in Fig. 4 presents the marginal effect of poverty on COVID-19 infections, deaths, and fatalities by share of black population. The black population is more likely to be infected, and then to die from the virus relative to the population at large. Overall fatality rates are also high among the black population. The coefficient of the black population in models 2.4–2.6 indicates that a 1% increase in the black population is associated with additional 201 infection cases and 8.96 deaths per million and a 0.02 increase in the fatality rate.

The Hispanic population is more likely to be infected and die although their fatality rate is not statistically significant. A 1% increase in the Hispanic population is associated with 227 more infections and 3.2 deaths per million as shown in models 2.4 and 2.5. Their insignificant fatality rate suggests that the Hispanic population was more able to get medical treatment to save their lives if they are infected. Another reason that Hispanics could die less frequently when they are infected is health. They may die less frequently because more Hispanics do not have underlying health conditions and their population is younger than other races. However, their higher infections and deaths per million indicate that they live in an economic environment that requires more human to human interactions.

---

[5] Generally, if the VIF is more than 10, it is considered a matter of concern and none of our variables has a 10 or higher VIF score. However, we acknowledge that VIF for poverty (7.91) and median income (9.8). It is because poverty and median income are negatively correlated. To confirm whether this negative correlation is a problem for our estimates, we tried to verify in several ways, we are mentioning two of them here: First, we run Generalized Ridge Regression which is a technique for analyzing multiple regression data that may suffer from multicollinearity. The selected Ridge k Value = 0.008 which is a very small penalty to the efficiency of the model but corrects the multicollinearity problems. We confirm that this Ridge regression produced very identical results compared to what we reported Table 2. Since Ridge regression does not allow the State FE model, we stick to our original choice of the model. Second, we also rerun the model after dropping the median income and that gave us lower VIF for all variables (less than 5), but our results/conclusion of the paper remain unchanged.

Asian minorities are also more likely to be infected. The fatality rate is also positively associated with Asian minorities. The coefficients for the Asian population indicate that a 1% increase in the Asian population in a county is associated with 195 more infections and 10.49 more deaths per million population resulting in a 0.03 higher fatality rate. Interestingly, the marginal effect on death and fatality rate for Asian is higher than that for the Black and Hispanic populations. However, the average, income and wealth for Asians are higher than the other minority groups and they are believed to have access to better health care services. On the other hand, it is important to note that the Asian population is quite diverse, and the "average" number is driven by some outliers of very rich Asian groups. In contrast, some Asians have the lowest income and wealth of all minorities such as the recently immigrated Myanmar or Burmese Asians. As we can see the coefficient of Asian is less statistically significant than Black which means the effect is noisier reflecting their diversity. However, here, it is hard to make any decisive conclusion without further evidence.

If we look back to the coefficient of poverty after controlling for ethnic minority populations in models 2.4–2.6 and make a comparison, we see that the coefficient of poverty drops by almost 50% (drops from 403 to 207 infection, 17.81–8.46 deaths per million and 0.07–0.05 fatality rate). This suggests that the minority populations have absorbed part of the impact of the poverty rate on the COVID-19 catastrophe due to their intercorrelation with poverty.

Other control variables also reveal interesting results. The population share that is sixty-five plus has a lower infection rate, but higher deaths per million and hence higher fatality levels. This is expected because the retired and elderly have lower engagement with outside employment which is one of the main sources of infection. However, due to their poorer immunity levels, the infection often leads to death. Counties with a higher share of the population living in rural areas have lower infections and lower deaths per million and therefore a lower fatality rate. On the other hand, log of median income is positive and significant which seems not to support our argument that poverty worsens infectious disease. Note that, higher median income does not represent lower poverty in some counties. This further suggests that something else such as an industrialization of a county and its related urbanization is likely driving higher median income than in poor non-industrialized counties. For example, the national average median household income was $51,090, and the national average poverty rate was 15.13% (Table 1). In Suffolk County (including Boston city), the median household income was $65,999, but the poverty rate was 17.5%. Hence, median household income is not necessary to be negatively correlated with the poverty rate. Similarly, share of the population with some college degrees is negatively associated with infections and deaths per million but not with fatality rate. Among the health ranking variables, clinical care, which consists of access to care and quality of care, is more consistently significant than other health rankings. The interpretation will be worst rank in health clinical care is associated with higher infections and deaths per million population. Since a lower score represents a better rank in clinical care and value ranges from -0.48 to 0.41, one unit increase in the score means moving from best in the ranking to the worst in the ranking, leading to a rise in 6 to 15 thousand additional infections and up to 2 hundred to 4 hundred additional deaths. On the other hand, worst ranking of health behavior is associated with

less death although infections per million and fatality rate are not statistically significant in the full model in columns (2.4) and (2.6). The coefficient of health behavior ranking is difficult to interpret since it is likely confounded by some other behavioral factors such as better healthy behavior may be correlated with more social engagement which is not favorable in the case of infectious diseases, with higher age groups who have inferior immune systems, and with lower motivation to seek medical care. All these can lead to more deaths, however, confirming such conclusion is beyond the scope of this paper and needs more relevant measures for such behavioral factors.

5 Robustness checks

5.1 Data disaggregation to check impact overtime

Throughout the pandemic, we observed a large variation in coronavirus surges across regions. Early in the pandemic, the northeastern region of the USA experienced high impact levels, while later, southern regions experienced high impact levels (Haynes et al. 2020; Haynes and Kulkarni 2021). These differences in temporal disease diffusion generated some questions concerning the role of poverty as a significant determinant of COVID-19 infections and deaths. The concern is whether our results hold for both earlier and later periods in the pandemic? To account for the variations in COVID-19 cases and deaths across regions and over time, we conducted a robustness check by disaggregating the analysis and dividing cumulative COVID-19 data into four monthly intervals: April, May, June, and July. Note that, results on COVID-19 cumulative data till the end of July are presented in Table 2. Table 3 presents three separate analyses for cumulative data for April, May, and June for both infections and deaths. We find that our results hold for both early and later in the pandemic. The impact of poverty on COVID-19 infections and deaths were always significant. Figure 5 plots the coefficient of the poverty rate for COVID-19 infections and death per million population. It reveals that an increase in the poverty rate by 1% resulted in an increase of 121.30 infections and 7.99 deaths per million population in April. This coefficient increased to 403.75 infections and 17.81 deaths per million population in July. The size of the coefficient of poverty for May and June remains within the lowest and highest level, meaning the increased amount of cumulative death and infections per million that are linked to a higher level of poverty and this keeps rising over time. These results provide important evidence that poverty is a strong continuing correlate for an infectious disease such as COVID-19. Note that, in these disaggregated results, the unemployment rate shows a significant negative relationship to both infections and death except for infection for July. The impact of the unemployment rate supports the argument posed by Adda (2016). Unemployment is a marker of economic cycles, which reduces human to human interaction when it is high. Alternatively, it also captures additional effects such as different socialization patterns for those out of the labor force. If unemployment encourages people to engage in more socialization without any precaution/mitigation measures, it can increase infections and thereby more death. We do not see the evidence for this alternative interpretation but the possibility of less infection and death when

Table 3 Disaggregated results by monthly COVID-19 confirmed cases and deaths per million population

	Model 3.1 Confirmed cases per million (April)	Model 3.2 Confirmed cases per million (May)	Model 3.3 Confirmed cases per million (June)	Model 3.4 Confirmed cases per million (July)	Model 3.5 Death per million (April)	Model 3.6 Death per million (May)	Model 3.7 Death per million (June)	Model 3.8 Death per million (July)
Poverty rate	121.301***	215.627***	342.710***	410.415***	7.994***	12.153***	15.024***	17.619***
	(34.209)	(54.063)	(86.555)	(99.254)	(1.835)	(2.346)	(2.972)	(3.314)
Income inequality (top 20%/bottom 20%)	488.330***	696.127***	602.336**	488.079	21.262**	42.902***	53.313***	59.184***
	(124.407)	(215.047)	(300.547)	(329.954)	(9.273)	(11.905)	(14.912)	(16.257)
Ranking of health behavior	− 1490.127**	− 2585.192**	− 4264.21***	− 6,413.88***	− 96.686***	− 88.281*	− 90.650	− 111.469*
	(669.342)	(1119.664)	(1468.484)	(1639.343)	(36.245)	(50.281)	(59.118)	(64.482)
Ranking of clinical care	2153.669***	6294.251***	8197.776***	9474.772***	100.843*	126.632	176.103*	188.559*
	(699.200)	(1930.937)	(2240.832)	(2454.481)	(55.919)	(78.301)	(90.256)	(97.606)
Ranking of physical environment	1573.812	3064.322	6667.131***	11149.690***	70.934	167.877*	200.815*	237.658*
	(1123.540)	(1965.798)	(2539.666)	(2919.682)	(60.971)	(92.138)	(110.222)	(121.801)
Rural (%)	− 21.122***	− 33.797***	− 40.905***	− 53.906***	− 1.296***	− 1.896***	− 2.168***	− 2.346***
	(3.122)	(5.596)	(6.795)	(7.782)	(0.223)	(0.307)	(0.364)	(0.391)
Log of median income	3730.284***	6974.365***	8495.211***	8243.350***	116.974***	230.315***	285.173***	321.438***
	(904.641)	(1344.147)	(1927.881)	(2171.370)	(38.770)	(55.271)	(67.005)	(73.356)
Female (%)	− 32.329	− 144.676	− 192.262	− 157.568	4.292**	8.375***	13.065***	16.124***
	(60.288)	(117.623)	(156.976)	(167.129)	(1.984)	(2.880)	(3.537)	(3.960)
Sixty-five plus (%)	11.876	− 85.847**	− 200.640***	− 376.260***	3.411***	5.003***	4.463***	3.079
	(17.730)	(34.752)	(51.530)	(59.702)	(0.868)	(1.351)	(1.687)	(1.961)
Some college (%)	− 58.218***	− 142.918***	− 184.884***	− 229.845***	− 0.927**	− 2.687***	− 3.925***	− 5.582***
	(12.944)	(30.783)	(35.116)	(36.823)	(0.436)	(0.771)	(0.974)	(1.062)
Unemployment rate	− 152.566***	− 247.976***	− 143.890	69.242	− 9.513***	− 12.435***	− 10.835**	− 5.594
	(39.900)	(81.411)	(170.265)	(209.397)	(2.600)	(3.927)	(4.992)	(6.354)
Long commute—driving alone	20.758***	6.442	− 18.929	− 36.133**	2.046***	3.140***	3.218***	3.098***

Table 3 (continued)

	Model 3.1 Confirmed cases per million (April)	Model 3.2 Confirmed cases per million (May)	Model 3.3 Confirmed cases per million (June)	Model 3.4 Confirmed cases per million (July)	Model 3.5 Death per million (April)	Model 3.6 Death per million (May)	Model 3.7 Death per million (June)	Model 3.8 Death per million (July)
	(5.995)	(10.542)	(13.789)	(15.290)	(0.459)	(0.669)	(0.767)	(0.817)
State FE	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Constant	− 36,984.58***	− 57,164.25***	− 62,741.67***	− 47,724.24*	− 1,631.65***	− 3,035.66***	− 3,778.95***	− 4,149.48***
	(9,301.906)	(14,171.890)	(21,452.565)	(24,613.726)	(466.472)	(651.779)	(796.588)	(875.303)
R^2	0.24	0.22	0.27	0.43	0.30	0.33	0.35	0.37
N	3075	3075	3075	3075	3075	3075	3075	3075

Robust standard errors are in parenthesis *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$. The VIF test result is 2.32 in all models, so there is no issue of multicollinearity problem
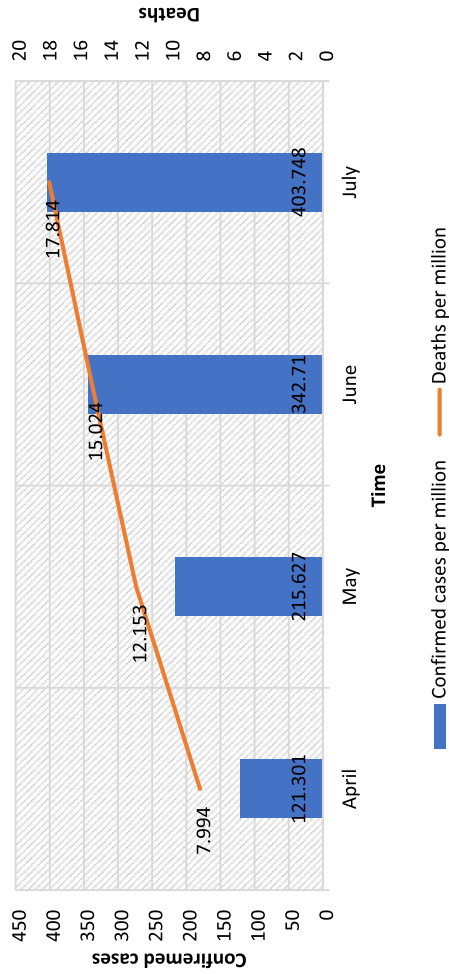
Fig. 5 Plotting the coefficients of poverty rate for infections and deaths

they are not in the workplace is real. This is in line with our original argument that hypermobility or physical person to person interaction is the key sources of infections.

5.2 Applying multi-level mixed model estimation

Now we apply a more complicated model for the same hypothesis tests as part of our robustness checks. Our models incorporate two levels of factors—level 1—the county-level factors and—level 2—the state-level factors which practice a spectrum of policies and operate under the unique political and economic system. We take advantage of this opportunity to examine the impact of factors at various levels on the COVID-19 pandemic. This is a hierarchical modeling environment where the combining process occurs at level 1; however, both level 1 and level 2 influence this combining process (Charnes et al. 1975; Otani et al. 2019). The utility of using hierarchical modeling is relevant since what is happening at the county level may not be completely independent and a multi-level model can account for both lower and higher levels as distinct levels but simultaneously and counties are clustered within the state.[6] We replicate the estimates presented in Table 2 by estimating Eq. (2) for a multi-level mixed effect (MLME) model. Here, $Poverty_{cs}$ is the county level poverty within a state and X_{cs} is the vector of county level characteristics. In μ_s, we gather state level factors. U_{cs} is a random effect in the model that accounts for specific variations in COVID-19 within a state, and U_c operates at the state level to account for variations specific to each state. The ϵ_{cs} is the idiosyncratic residual that captures anything that is not in the model.

$$Y_{cs} = \beta_0 + \beta_1 \text{Poverty}_{cs} + \beta_2 X_{cs} + \beta_3 \mu_s + U_{cs} + U_s + \in_{cs} \tag{2}$$

We present this MLME result in Table 4. We see that the impact of poverty on both COVID-19 infections and deaths is larger than what we have seen in the case of state fixed effect estimation (i.e., coefficient of poverty rate in model 2.1 is 403.75 versus coefficient of 421.88 in model 4.1 for the same specification). So, the main results presented in Table 2 are even more conservative than this alternative MLME. The MLME provides us additional information that it has two intercepts: The variance due to the heterogeneity across states is $U_s/\epsilon_{cs} = 2.13/(2.13 + (-25.57) = 49.93$, which about 50% of the variance is due to the variations at the state level (column 4.6) and the overall constant identical to a constant from non-MLME model.

5.3 Applying spatial autoregressive model estimation

Public health scholars have called for attention on COVID-19 infection patterns between spatially contiguous entities or interconnections between counties. As a result, most state governments imposed lockdown policies to reduce the frequency of human interactions for the purpose of curbing the spread of the COVID-19 disease. Some local governments even implemented stricter restrictions to constrain human mobility (Goolsbee et al. 2020). That is, spatial dependences between COVID-19 infections and counties may impact our proposed models in Table 2. To assess the

---

[6] See Otani et al. (2019) for details of this modeling.

Table 4 Multi-level mixed model estimates on COVID-19 confirmed case, death, and fatality rate until the end of July

	Model 4.1	Model 4.2	Model 4.3	Model 4.4	Model 4.5	Model 4.6
	Confirmed cases per million	Deaths per million	Fatality rate	Confirmed cases per million	Deaths per million	Fatality rate
Poverty rate	421.876*** (59.352)	18.304*** (2.407)	0.074*** (0.020)	216.936*** (58.697)	8.621*** (2.434)	0.048** (0.021)
Income inequality (top 20%/bottom 20%)	27.963 (273.278)	51.741*** (11.089)	0.130 (0.094)	361.951 (259.809)	53.377*** (10.792)	0.117 (0.095)
Ranking of health behavior	−4009.759*** (1266.221)	−39.098 (51.493)	1.043** (0.433)	−3532.086*** (1319.690)	−114.793** (55.453)	0.718 (0.484)
Ranking of clinical care	14,938.194*** (1624.012)	417.203*** (65.931)	0.308 (0.556)	6149.187*** (1694.529)	229.807*** (70.498)	0.046 (0.620)
Ranking of physical environment	9652.010*** (3091.713)	424.520*** (125.305)	1.280 (1.061)	9726.676*** (3065.375)	148.456 (127.314)	−0.030 (1.123)
Rural (%)	−56.125*** (6.589)	−1.695*** (0.268)	−0.005** (0.002)	−15.012*** (7.319)	−1.325*** (0.305)	−0.007*** (0.003)
Log of median income	3934.547** (1566.786)	390.483*** (63.852)	2.683*** (0.535)	7039.542*** (1591.677)	269.310*** (66.533)	1.951*** (0.583)
Female (%)	−379.818*** (73.285)	11.079*** (2.973)	0.089*** (0.025)	−176.951** (71.711)	14.857*** (2.980)	0.089*** (0.026)
Sixty-five plus (%)	−346.762*** (46.764)	4.393** (1.904)	0.082*** (0.016)	−218.368*** (46.027)	5.743*** (1.933)	0.078*** (0.017)
Some college (%)				−170.376*** (20.837)	−4.852*** (0.871)	−0.008 (0.008)
Unemployment rate				−168.874 (125.605)	−6.591 (5.268)	0.004 (0.046)
Black (%)				214.009*** (15.104)	8.992*** (0.632)	0.020*** (0.006)
Asian (%)				112.031* (63.666)	9.055*** (2.694)	0.026 (0.023)
Hispanic (%)				218.611***	3.265***	−0.002

Table 4 (continued)

	Model 4.1 Confirmed cases per million	Model 4.2 Deaths per million	Model 4.3 Fatality rate	Model 4.4 Confirmed cases per million	Model 4.5 Deaths per million	Model 4.6 Fatality rate
Long commute—driving alone				(16.905) −36.269**	(0.710) 2.347***	(0.006) 0.015***
State level constant	22,400,000***	53,579***	2.247***	(15.055) 15,000,000***	(0.630) 46,821***	(0.006) 2.13***
Constant	(4,940,237) −11,609.888	(11,835) −5018.165***	(.532) −33.960***	(3,569,253) −50,650.394***	(10,629) −3680.811***	(0.516) −25.576***
LR test vs. linear model	(18,910.193) 623.55***	(770.496) 561.20***	(6.462) 361.07***	(18,973.029) 359.72***	(792.881) 453.00***	(6.951) 318.92***
Groups (States)	51	51	51	51	51	51
Observations (counties)	3075	3075	3062	3075	3075	3062

Standard errors are robust and shown in parenthesis * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Fatality rate ((case per million/death per million) * 100). The large and significant test results for LR test vs. linear model show that the multi-level estimation is more appropriate

Table 5 OLS, spatial lag, spatial error, and SAC models estimate on COVID-19 confirmed cases until the end of July (without ethnicity)

	OLS	Spatial Lag Model	Spatial Error Model	SAC Model
Poverty rate	403.748***	330.549***	355.088***	355.131***
Income inequality (top 20%/bottom 20%)	(95.961)	(56.909)	(58.462)	(58.546)
	3.228	− 117.936	− 270.548	− 263.127
Ranking of health behavior	(322.525)	(261.764)	(263.035)	(263.472)
	− 3494.850**	− 3727.932***	− 3750.301***	− 3767.761***
Ranking of clinical care	(1592.705)	(1220.860)	(1258.314)	(1259.937)
	15,083.133***	13,390.209***	13,728.461***	13,779.31***
Ranking of physical environment	(2250.001)	(1562.536)	(1654.707)	(1652.647)
	9795.149***	7210.5**	9893.472***	9731.312***
Rural (%)	(2842.045)	(2951.480)	(3237.799)	(3247.716)
	− 56.018***	− 56.848***	− 55.523***	− 55.818***
Log of median income	(7.620)	(6.338)	(6.488)	(6.511)
	4172.761*	1396.902	404.951	482.627
Female (%)	(2199.564)	(1529.002)	(1608.016)	(1608.819)
	− 403.444***	− 451.919***	− 451.116***	− 452.788***
Sixty-five plus (%)	(154.253)	(70.000)	(69.489)	(69.876)
	− 338.934***	− 320.733***	− 369.637***	− 368.077***
	(57.793)	(45.344)	(47.142)	(47.169)
State FE	**Yes**	**Yes**	**Yes**	**Yes**
Constant	− 2572.391	24,786.915	42,380.715**	41,165.781**
	(24,835.687)	(18,298.229)	(19,267.775)	(19,254.16)
Spatial lag parameter		0.37***		.028
		(0.0270)		(0.049)
Spatial error parameter			0.483***	0.457***
			(.028)	(.047)
R^2	0.41	0.40	0.40	0.40
AIC	63,737	63,541	63,506	63,506
BIC	64,093	63,909	63,874	63,874
N	3075	3075	3075	3075

Standard errors are robust and shown in parenthesis $*p<0.1$; $**p<0.05$; $***p<0.01$

spatial dependence effect of adjacent counties, we apply the spatial lag model, the spatial error model, and the spatial autoregressive combined model (SAC) to examine whether the effect of the poverty rate on COVID-19 infections remains robust across all models (Guliyev 2020; Sun et al. 2020; Narayanan et al. 2020). The spatial lag model explores how the COVID-19 infections in a county are influenced by the COVID-19 infections in neighboring counties. The spatial lag parameter reflects the association between the average COVID-19 inflection of neighboring counties and the COVID-19 inflections of a given county. The spatial error model examines how the model residuals of a given county are associated with that in its neighboring counties' average model residuals. The SAC model is a simultaneous estimate of the spatial lag and spatial error models. Table 5 presents OLS and the three spatial models without ethnicity, and Table 6 presents those models with ethnicity. The

Table 6 OLS, spatial lag, spatial error, and SAC models estimate on COVID− 19 confirmed cases until the end of July (with ethnicity)

	OLS	Spatial Lag Model	Spatial Error Model	SAC Model
Poverty rate	207.651**	180.799***	211.198***	215.642***
Income inequality (top 20%/ bottom 20%)	(96.723) 334.367	(57.062) 231.487	(57.169) 134.61	(56.808) 125.462
Ranking of health behavior	(323.194) − 2504.923	(253.561) − 2660.651**	(251.819) − 2591.773**	(249.817) − 2533.821*
Ranking of clinical care	(1624.129) 6030.034**	(1325.776) 5296.637***	(1309.865) 4715.073***	(1298.506) 4551.406***
Ranking of physical environment	(2563.789) 10,654.596***	(1658.906) 8827.736***	(1711.947) 9364.565***	(1708.966) 9663.716***
Rural (%)	(2704.545) − 11.511	(2989.858) − 16.923**	(3178.11) − 13.37*	(3199.239) − 12.165*
Log of median income	(7.342) 7577.098***	(7.202) 5463.144***	(7.269) 5838.467***	(7.264) 5969.190***
Female (%)	(2123.528) − 201.327	(1586.528) − 251.05***	(1602.243) − 261.765***	(1602.618) − 255.891***
Sixty-five plus (%)	(166.634) − 197.411***	(69.976) − 199.265***	(68.883) − 201.604***	(68.489) − 197.854***
Some college (%)	(59.315) − 166.916***	(46.131) − 155.793***	(47.613) − 151.168***	(47.576) − 148.596***
Unemployment rate	(33.7) − 108.081	(20.607) − 34.973	(20.853) − 30.185	(20.752) − 31.715
Black (%)	(184.805) 201.049***	(125.345) 165.104***	(132.928) 188.345***	(133.596) 192.303***
Asian (%)	(19.636) 195.336*	(15.399) 181.541***	(17.657) 177.942***	(18.248) 178.167***
Hispanic (%)	(100.163) 227.376***	(66.057) 204.642***	(68.839) 278.257***	(68.962) 292.339***
Long commute—driving alone	(28.097) − 45.236***	(17.365) − 41.945***	(19.694) − 35.032**	(20.294) − 33.241**
	(14.606)	(14.913)	(15.339)	(15.340)
State FE	**Yes**	**Yes**	**Yes**	**Yes**
Constant	− 48,341.59** (24,244.458)	− 26,623.672 (18,729.563)	− 27,400.575 (19,004.028)	− 28,121.418 (18,999.238)
Spatial lag parameter		.276*** (.026)		− .103** (.043)
Spatial error parameter			.459*** (.029)	.542*** (.04)
R^2	0.48	0.48	0.48	0.48
AIC	63,319	63,210	63,125	63,120
BIC	63,711	63,615	63,529	63,524
N	3075	3075	3075	3075

Standard errors are robust and shown in parenthesis *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

results in Table 5 reveal that three spatial models have better model fits (i.e., AIC and BIC) than the OLS model. The OLS also overestimates the coefficient of the poverty rate in comparison with the three spatial models. When we controlled for ethnicity in Table 6, the results illustrate that the three spatial models still have better model performance than the OLS model. However, the OLS model shows more conservative estimate of the coefficient of the poverty rate than the spatial error model and the SAC model. In short, spatial models may perform better model fits than the OLS model. However, we see that the significant positive effect of the poverty rate on COVID-19 is confirmed in all cases and the results are robust across all models.

6 Conclusion

This paper studies the impact of the poverty rate on the spread of COVID-19 viruses and related deaths caused by the virus across space and time using the best available data as of July 2020. Poverty leading to more physical interaction, incapacity to provide access to health care treatment, and job-related public behavior highly correlated with more infections and deaths in the U.S. Poverty related to limited economic activities is not the dominant factor leading to the containment of this infectious disease. Although we did not measure these poverty-related factors—hypermobility versus limited economic activity directly, we consider that these are the two main possibilities that can be driven by poverty in the context of infectious disease. This is an important contribution in the literature that provides systematic evidence in the developed world which is analogous to the evidence found in the developing world. Therefore, the local government and the Federal Government must be prepared to minimize the disproportional costs to poor people and minorities. Solutions are not only in creating infrastructure to provide easy access to health care services to all but also better coordination among housing, social welfare programs, food access, and other social protections so that public health responses are immediately effective and there are enough incentives for low-income people to maintain physical distancing and adjust job-related public behavior.

We must acknowledge some limitations and we should keep in mind while we explore these results. To detect a causal relationship between poverty and COVID-19, it would be better if we could have panel data where we would control for the time fixed effect to avoid any role of unusual time trend issues (i.e., spike in unemployment rate for business closures) and county fixed effect to control for both observable and unobservable time-invariant factors that can cause differences in impacts of COVID-19. Although our multi-level data allowed us to control for state fixed effects and, to minimize the suspicion of "wave effect" related to the timing of the pandemic, we also conducted the robustness analysis. We did this by disaggregating COVID-19 data by month, deploying multi-level and spatial autoregressive model estimation but it is not equivalent to a panel data analysis. We also did not have county-level COVID-19 hospitalization data, which is an alternative measure of infectious disease. It would also be better if we had infection and death data by age, gender, and ethnic groups. Future research should integrate hospitalization and more closely disaggregated COVID-19 data and hence causal control, and this would

provide potential reinforcement of our interpretations. Future analysis should also focus on more micro-level information that discerns more about incapacity to provide adequate health care measure and jobs-related exposure behavior that together are correlated with poverty. This could be done by linking employment information, economic sectors by county such as food processing, meat slaughter, and packaging plants, location of homes for the aged, or separate age effects from poverty since they are likely highly correlated. It would appear from our analysis that the public mobility data between counties applied to poverty regions might also be of value.

Declarations

Conflict of interest The authors declare that they donot have any conflict of interest to declare.

References

Adda J (2016) Economic activity and the spread of viral diseases: evidence from high frequency data. Q J Econ 131(2):891–941

Adda J, Banks J, Von Gaudecker H-M (2009) The impact of income shocks on health: evidence from cohort data. J Eur Econ Assoc 7(6):1361–1399

Barmby T, Larguem M (2009) Coughs and sneezes spread diseases: an empirical study of absenteeism and infectious illness. J Health Econ 28(5):1012–1017

Brodeur A, Gray DM, Anik I, Suraiya B 2020. "A literature review of the economics of COVID-19." Global labor organization. https://www.econstor.eu/bitstream/10419/222316/1/GLO-DP-0601.pdf.

Campos MC, Dombrowski JG, Phelan J, Marinho CRF, Hibberd M, Clark TG, Campino S (2018) Zika might not be acting alone: using an ecological study approach to investigate potential co-acting risk factors for an unusual pattern of microcephaly in brazil. PLoS One 13(8):e0201452

Charnes A, Haynes KE, Hazleton JE, Ryan MJ (1975) A hierarchical goal-programming approach to environmental land use management. Geogr Anal 7(2):121–130

Chin T, Kahn R, Li R, Chen JT, Krieger N, Buckee CO, Balsari S, Kiang MV (2020) US county-level characteristics to inform equitable COVID-19 response. MedRxiv. https://doi.org/10.1101/2020.04.08.20058248

Cordes J, Marcia MC (2020) Spatial analysis of COVID-19 clusters and contextual factors in New York city. Spat Spatio-Tempor Epidemiol 34:100355

Collins KR (2020) "Inequality and surveillance during the COVID-19 pandemic." The Vienna Circle, September 23, 2020. https://medium.com/the-vienna-circle/inequality-and-surveillance-during-the-covid-19-pandemic-c786dac594a5.

Ettner SL (1996) New evidence on the relationship between income and health. J Health Econ 15(1):67–85

Gold J, Marcus J (2020) "Colleges are getting ready to blame their students." The Atlantic, July 21, 2020, sec. Ideas. https://www.theatlantic.com/ideas/archive/2020/07/colleges-are-getting-ready-blame-their-students/614410/.

Goolsbee A, Luo NB, Nesbitt R, Syverson C (2020) "COVID-19 Lockdown policies at the state and local level. SSRN Electron J. https://doi.org/10.2139/ssrn.3682144

Guliyev H (2020) Determining the spatial effects of COVID-19 using the spatial panel data model. Spat Stat 38:100443

Haynes KE, Kulkarni R (2021) Modeling region based regimes for COVID-19 mitigation: an inverse gompertz approach to coronavirus infections in the USA, New York, and New Jersey. Reg Sci Policy Pract. https://doi.org/10.2139/ssrn.3629898

Haynes KE, Kulkarni R, Li M-H, Siddique AB (2020) The impact of differing COVID-19 mitigation policies: three natural experiments. SSRN Electron J. https://doi.org/10.2139/ssrn.3702606

Kim J, Kwan M-P (2021) The impact of the COVID-19 pandemic on people's mobility: a longitudinal study of the us from march to september of 2020. J Transp Geogr 93:103039

Li AY, Hannah TC, Durbin J, Dreher N, McAuley FM, Marayati NF, Spiera Z, Ali M, Gometz A, Kostman JT, Choudhri TF (2020) Multivariate analysis of factors affecting COVID-19 case and death rate in U.S. counties: the significant effects of black race and temperature. MedRxiv. https://doi.org/10.1101/2020.04.17.20069708

Lima MS, Rosati A, Iglesias SP (2020) "Brazil hands out so much covid cash that poverty nears a low." Bloomberg.Com, September 2, 2020. https://www.bloomberg.com/news/articles/2020-09-02/brazil-hands-out-so-much-covid-cash-that-poverty-nears-a-new-low.

Liu L, Miller HJ, Scheff J (2020) The impacts of COVID-19 pandemic on public transit demand in the United States. Plos One 15(11):e0242476

Markina SS, Maksimova NM, Vitek CR, Bogatyreva EY, Monisov AA (2000) Diphtheria in the Russian federation in the 1990s. J Infect Dis 181:27–34

Massey DS, Gross AB, Shibuya K (1994) Migration, segregation, and the geographic concentration of poverty. Am Socio Rev 425–445. https://doi.org/10.2307/2095942

Moore M, Gelfel B, Adeyemi Okunogbe CP (2017) "Identifying future disease hot Spots." Rand Health Quarterly 6 (3). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5568150/.

Mukherji N (2020) The social and economic factors underlying the impact of COVID-19 cases and deaths in US counties. MedRxiv 5:1–195

Narayanan RP, James Nordlund R, Pace K, Ratnadiwakara D (2020) Demographic, jurisdictional, and spatial effects on social distancing in the United States during the COVID-19 pandemic. Plos One 15(9):e0239572

Oster E (2005) Sexually transmitted infections, sexual behavior, and the HIV/AIDS epidemic. Q J Econ 120(2):467–515

Otani K, Deng Y, Herrmann PA, Kurz RS (2020) Patient satisfaction, quality attributes, and organizational characteristics: a hierarchical linear model approach. J Patient Exp 7(5):801–806

Qiu Y, Chen X, Shi W (2020) Impacts of social and economic factors on the transmission of coronavirus disease (COVID-19) in China. MedRxiv 28:1012

Randolph SE (2008) Tick-borne encephalitis incidence in Central and Eastern Europe: consequences of political transition. Microbes Infect 10(3):209–216

Redding DW, Atkinson PM, Cunningham AA, Iacono GL, Moses LM, Wood JLN, Jones KE (2019) Impacts of environmental and socio-economic factors on emergence and epidemic potential of Ebola in Africa. Nat Commun 10(1):1–11

Rollston R, Galea S (2020) COVID-19 and the social determinants of health. SAGE Publications Sage CA, Los Angeles, CA

Ruhm CJ (2000) Are recessions good for your health? Q J Econ 115(2):617–650

Ruhm CJ (2003) Good times make you sick. J Health Econ 22(4):637–658

Ruhm CJ (2005) Healthy living in hard times. J Health Econ 24(2):341–363

Saxe R (2020) "Contact tracing and COVID-19: lessons from HIV." American Civil Liberties Union, May 2020. https://www.aclu.org/news/hiv/contact-tracing-and-covid-19-lessons-from-hiv/

Schafft KA (2006) Poverty, residential mobility, and student transiency within a rural New York school district. Rural Sociol 71(2):212–231

Semenza JC, Giesecke J (2008) Intervening to reduce inequalities in infections in Europe. Am J Public Health 98(5):787–792

Shilova MV, Dye C (2001) The resurgence of tuberculosis in Russia. Philos Trans Royal Soc London. Series B: Biol Sci 356(1411):1069–1075

Siddique AB (2016) Comparative advantage defying development strategy and cross country poverty incidence. J Econ Dev 41(4):45–78. https://doi.org/10.35866/caujed.2016.41.4.003

Siddique AB (2021) Does ethnic diversity hurt fiscal sustainability? Econ Bull 41(2):532–543

Stoilova Y, Popivanova N (1999) Epidemiologic studies of leptospiroses in the Plovdiv region of Bulgaria. Folia Med 41(4):73–79

Suhrcke M, Stuckler D, Suk JE, Desai M, Senek M, McKee M, Tsolova S et al (2011) The impact of eco-
     nomic crises on communicable disease transmission and control: a systematic review of the evidence.
     PLoS ONE 6(6):e20724. https://doi.org/10.1371/journal.pone.0020724

Suk JE, Manissero D, Büscher G, Semenza JC (2009) wealth inequality and tuberculosis elimination in
     Europe. Emerg Infect Dis 15(11):1812

Sun F, Matthews SA, Yang T-C, Ming-Hsiao Hu (2020) A spatial analysis of COVID-19 period prevalence
     in US counties through June 28, 2020: Where geography matters? Ann Epidemiol. https://doi.org/10.
     1016/j.annepidem.2020.07.014

Wissel BD, Van Camp PJ, Kouril M, Weis C, Glauser TA, White PS, Kohane IS, Dexheimer JW (2020)
     An interactive online dashboard for tracking COVID-19 in U.S. counties, cities, and states in real
     time. J Am Med Inform Assoc 27(7):1121–1125. https://doi.org/10.1093/jamia/ocaa071

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published
maps and institutional affiliations.