Income Inequality and Adverse COVID-19 Outcomes in US Counties

By Henri Vilandre
(8856963)

Major Paper presented to the Department of Economics of the University of Ottawa
in partial fulfillment of the requirements of the M.A. Degree
Supervisor: Professor Rose Anne Devlin
ECO 6999

University of Ottawa
April 2021

Abstract: This paper exploits US county-level variations in income inequality and a variety of demographic factors to test the empirical association between inequalities and COVID-19 health outcomes between January and December 2020. The results suggest that the distribution of income is a robust and statistically significant predictor of adverse health outcomes, which is consistent with the most recent literature. In addition, this paper argues that inequality ten years prior to the pandemic is a better predictor of confirmed cases and mortality, implying that causal pathways present with a lag. Furthermore, this effect is mitigated for counties that experienced a decrease in inequality and amplified for counties that have become more unequal, suggesting that public policy aimed at addressing inequality will directly and/or indirectly improve health outcomes, including for COVID-19.

Keywords: Income Inequality, COVID-19, Health, Negative Binomial Model, United States.

Acknowledgements

I would like to thank my M.R.P supervisor Rose Anne Devlin, whose expertise, patience and insightful advice have proved invaluable to the elaboration of this research paper. I also wish to thank the University of Ottawa for years of fascinating courses and inspiring professors. Finally, I wish to extend my gratitude to Sminem, Ph.D., whose affable banter gave me the original idea for this paper.

I. Introduction

Income inequality has received much attention over the years. Widening economic inequalities worldwide and their complex ramifications on society have piqued the interest of multidisciplinary researchers and policy makers (IMF, 2015; Mayhew & Wills, 2019). They have been linked - causally and otherwise – to a sweeping array of phenomena ranging from crime and civil conflict to democracy, economic growth, and health outcomes (Elgar & Aitken, 2011; IMF, 2015). The United States accounted for a quarter of global confirmed cases of the SARS-CoV-2 virus and a fifth of associated deaths as of March 2021. It also is the most unequal country of the G7. Recent literature suggests that inequality plays an important role in the country's resilience to the virus (Oronce *et al.*, 2020; Karmakar *et al.*, 2021). This paper contributes to this literature by analyzing the role of historical inequalities on present health outcomes. It also provides initial evidence that the trend in income inequality (rising or decreasing) might help explain the relationship between health and income gaps.

The initial motivation behind this analysis is to examine whether the distribution of income at the county-level influenced COVID-19 outcomes. For decades, considerable research has linked income, be it relative or absolute, to adverse health outcomes (Marmot et al., 1991; Wilkinson, 1992; Kawachi et al, 2002). Exploiting US state-level variations in the distribution of income, Kaplan et al. (1996) found a significant association between inequalities and a large number of health outcomes, thus echoing Wilkinson's (1992) notion that relative income may adversely affect health outcomes and mortality. Several authors have upheld the importance of considering relative income alongside absolute income to study population health outcomes (Wilkinson & Pickett, 2015). To quote De Maio (2014, p.1), the income inequality hypothesis (IIH), or relative income hypothesis, "asserts that an individual's health is influenced not only by their own level of income, but by the level of inequality in the area in which they live". As such, the IIH suggests that in developed countries, economic disparities play a determining role in health outcomes, more so than absolute income. The answer as to how the distribution of income affects health lies both in the underlying mechanisms that allow inequality to rise in the first place, as well as in the pervasive social effects of rising inequalities in a community. For example, Kahn *et al.* (2000) suggest that policy frameworks conducive to increasing inequality are associated with under investment in safety nets and public services, thus placing sub-groups of the population at risk of not being able

to care for their health optimally. Elgar *et al.* (2020) suggest that high levels of inequality deteriorate interpersonal trust and negatively affect social cohesion, both of which are important in the elaboration and implementation of public health policy. Bor *et al.* (2017) implicate the eroded social fabric resulting from inequality as it can decrease confidence in public institutions and increase chronic stress among sub-groups of the population.

This paper intends to apply the IIH to US counties' COVID-19 infection and mortality rates, by empirically testing the extent to which past and current distribution of income predict the resilience of a community when faced with a public health crisis. To the best of my knowledge, this research represents the first attempt to analyze the role of past inequalities on the risk of infection and mortality in the context of COVID-19. It also explores the role of 'trends' in income inequality, measured by inequality over two time periods, to determine the role of persistent inequality and changing inequality, on health.

One reason why I am able to undertake this analysis is the availability of a rich amount of data on county-specific variables in the United States. Often, the IIH is applied to *between* country analyses but several factors such as cultural, political, and geographical differences restrict the research to a few countries that are 'similar' in these regards. But heterogeneity issues remain. I am also able to improve upon the small number of papers that examine the role of inequality in health outcomes in the United States by incorporating data on past and current measures of inequality, age, race, medical insurance, and political factors to empirically assess how 3 089 geographical units coped during first 328 days of the pandemic.

Exploiting variations in the large number of US counties, data from a variety of public sources and COVID-19 data up to the start of vaccinations, I find that income inequality is positively correlated with the virus's spread and mortality. One novel finding is that income inequality ten years prior to the start of the pandemic is a better predictor of health outcomes than current inequalities. By comparing US counties with an upward and downward trend in income inequality over the last decade, it appears that reducing inequalities is beneficial to health outcomes regardless of the initial or current level. Moreover, rising inequality accentuates the deleterious effects of the income inequality hypothesis, thus suggesting that inequality trends play a role in the IIH.

II. Income Inequality and Health

Much is known of the link between health outcomes and poverty. In 1931, Sydenstricker reported higher incidence and mortality of the Spanish flu among working classes in the United States. Poverty is linked to reduced life expectancy in a widely cited paper by Wilkinson (1992) and mental health deterioration and unhealthy behaviours such as drug use (Hollingsworth *et al.,* 2017*).* Studies have found that lower job status and environmental factors associated with poorer neighbourhoods are linked to higher stress-related morbidity and chronic illnesses such as asthma, hypertension, obesity, and cardiovascular diseases (Oates *et al*., 2018; Bambra, 2016; Bambra, 2011).[1] In the United States, poverty rates are closely linked with negative health outcomes both at the state and county-level using self-perceived measures of health status (Mackenbach *et al.*, 2005), life expectancy (Cutler *et al.,* 2006), and various medical-health biomarkers (Carrieri & Jones, 2017). Although the consensus is well established of the deleterious impact of poverty on health, the relationship with income inequality and health remains more ambiguous with research differing on the magnitude and causal mechanisms.

To address the problems arising from income inequality, it is important first to distinguish it from related concepts such as poverty. The most widely used measure of income inequality is the Gini index, which measures the statistical dispersion of income in a population/group. If everyone receives the same share of income, the Gini index takes a value of 0, whereas if a single individual or household receives the totality of the income, the Gini index is 1. Income inequality and poverty do not necessarily go hand in hand. For instance, Ireland and Mongolia have a significantly lower Gini index than the United States despite having, respectively, a lower and higher poverty rate than the United States (World Bank Database, 2021). From the mid 1900s to the 1970s, the United States witnessed unprecedented growth and shared prosperity, however, since the 1980s a trend in inequality has emerged: the country's economic growth has not been evenly shared across income tiers. Analysis by the Pew Research Center (2020) suggests that the median US household income has increased by 49% since 1970, but this increase is unevenly distributed across income categories and the number of adults in the middle-class has shrunk by a

[1] The increase in chronic illnesses resulting from prolonged exposure to stress is often referred to as the *Allostatic Load Hypothesis,* first described by McEwen and Stellar in a 1993 article entitled "Stress and the individual. Mechanisms leading to disease"

fifth. According to the latest US Census Bureau[2] data, 48% of aggregate income goes to the four lower quintiles while 23% of it goes to the top 5%.[3] Between 2006 and 2018, the upward trend of income inequality is ubiquitous and statistically significant across all US states, although the magnitude varies between states (see Figure A).

Income is a way of amassing wealth and acquiring assets both of which increase resilience in times of crisis. A recent study by Saez & Zucman (2020) suggests that wealth inequality in the US has adopted an even sharper trend than income inequality as the capital share of income plays a steadily growing role, they also point out that traditional methods of measuring inequality fail to account for 100% of national income. In an increasingly globalized world, where offshoring and hiding assets has become easier, income and wealth inequalities are likely higher than figures reported by the Census Bureau.

Several papers published in the wake of the Covid-19 pandemic have called attention to the effect of the pandemic on inequality, with low-wage workers and marginalized communities bearing the brunt of the initial shock and facing decreased job prospects currently and going forward (Furceri et al., 2020; Bonacini et al., 2020; Stabile *et al.*, 2020). The Covid-19 recession is by all indicators the most unequal in recent US history, high-paying jobs and jobs requiring degrees rebounded swiftly close to their baseline levels, while low-paying jobs continue to lag behind their pre-pandemic levels (Long *et al.*, 2020). Layered onto this asymmetric economic impact of Covid across income levels, is the possibility that income inequality actually contributes to the health impact of this virus.

As the Covid-19 pandemic spread rapidly in the United States, socio-economic inequalities were amplified through the unequal impact of social-distancing measures and unemployment (Bonacini et al., 2021; Blundell et al., 2020). With the rising death toll, a growing body of research is pointing to the role of socio-economic disparities to explain inequalities in Covid-19 incidence and mortality, dampening claims that the virus does not discriminate between rich and poor.[4] Although the virus does not discriminate on a biological level, its spread and health impact in the

[2] "Table A-4. Selected Measures of Household Income Dispersion" available at https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-inequality.html

[3] In 1970, around 57% of aggregate income went to bottom 80% earners, and the top 5% earned a share equivalent to 16%.

[4] In April 2020, World Health Organization Director General claimed that SARS-CoV-2 "does not discriminate between rich nations and poor". In March 2020, UK cabinet minister Michael Cove said the "virus does not discriminate" in response to British Prime Minister Boris Johnson contracting the virus. The phrase has also been used by several local institutions such as the Alberta government, the Minnesota Department of Health, and others.

United States seem to be closely linked with socio-economic, ethnic, and geographical inequalities. Minorities are disproportionally affected by the virus, especially people of Hispanic, Afro-American and Native descent (CDC 2020). As of March 30, 2021, there were 30 million confirmed cases and over 540 000 deaths in the USA.

The leading theory on income inequality and health is the *income inequality hypothesis*[5] (IIH). Defined by De Maio (2014), the hypothesis "asserts that an individual's health is influenced not only by their own level of income, but by the level of inequality in the area in which they live." The IIH goes beyond the distinction between poor and non-poor by studying the notion that the distribution of income inequality within a population is a determinant of overall health.

The literature is still debating the pathways through which income inequality might affect a population's overall health; it is likely a combination of several mechanisms and idiosyncratic factors proper to a nation or subnational unit. One such mechanism described by Kahn *et al.* (2000), and gaining traction in recent studies, is the neo-material pathway, which asserts that the rise in income inequality is associated with the systematic underinvestment in social infrastructures such as education, health services, food security, and public transportation. Another widely studied mechanism points to the erosion of the social fabric and trust, which decreases confidence in public institutions and increases chronic stress among sub-groups of the population (Bor *et al.,* 2017). Echoed in a recent study by Elgar *et al.* (2020), this erosion of trust is particularly relevant to the resilience of a population during epidemiologic crises – such as with H1N1, SARS and Covid-19 – as trust is an important determinant of compliance with social distancing measures and public health guidelines. In all cases, it is generally agreed that the pathways linking the health and inequality are forged over time and may take years to materialize (Leigh & Jencks, 2007; Lynch *et al.*, 2005). This lagged relationship can impede policy measures to improve health in the short-term by delaying the actual magnitude of the relationship.

It is the more economically equal countries, rather than the richest ones, that tend to be healthier (Wilkinson & Pickett, 2006; Wilkinson & Pickett, 2015; De Maio, 2014; Kim, 2019). Some studies have found little or inconclusive support for the IIH in industrialized nations when analyzing *between* country differences (Lynch *et al.*, 2004; De Vogli *et al.* 2005; Avendano, 2012). Pop *et al.* (2013) argues that the reason for the sometimes-inconclusive findings in support of the IIH in *between* country analyses of industrialized nations stems from the small-N problem, usually

[5] Sometimes referred to as the *relative-income hypothesis* or the *weak income-inequality hypothesis.*

restricted to 20-40 countries depending on the studies. The small sample size leads to failure to account for regional variance as well as heterogeneity in physical geography and policy environments, even among relatively homogeneous OECD countries.[6] Kim (2019) circumvents the issues caused by the chronic small numbers problem by using data for 292 OECD regional units and finds statistically significant evidence that income inequality negatively affects average life expectancy, child mortality and old-age mortality. According to Kragten & Rözer (2017), another explanation for the lack of consensus in the literature is the wide variety of data and statistical methods used, including the use of self-reported measures of health, different health indicators as proxies for overall health, as well as linear and non-linear models. In this paper, I investigate the IIH using county-level observations in the United States (3,089 observations) using the incidence and death rates of Covid-19, addressing the small-N issue, and mitigating possible regional/cultural divergences in self-reported health.

In a paper on the presence of IIH in Canada and the United States, Dunn *et al.* (2006) point out that there is a "peculiar" absence of geographic consistency in the IIH literature. They present evidence that at lower-scale geographic units (i.e., county, and metropolitan areas) some confounding outliers blur the income inequality-health relationship by embodying specific cultural and historical trajectories counter to the IIH, such is the case with the influence of "university towns" where health is unexpectedly good despite high inequality. The authors argue that "rather than dismissing the importance of income inequality and health because it does not appear to exist at all times and in all places" (p. S16) research should focus on the conditions under which the relationship hold, and the similarities between places where there is evidence for IIH.

There is much less debate about the IIH when moving away from *between* countries analysis and using a more granular geographical level. When studying country-specific effects of income inequality on health outcomes, research points towards a "threshold" effect, in the sense that countries with relatively low inequality present with minimal or null coefficient estimators, while highly unequal countries lend credence to the IIH. As such, significant effects have been found in China (Pei & Rodrigez, 2006; Du *et al.*, 2019), Brazil (Cavalini & De Leon, 2006) and Italy (Vogli *et al.*, 2005) yet smaller or null effects have been found in the Scandinavia area (Bockerman *et al.*, 2009) and Canada (Latif, 2015).

[6] For example, the United States territory in 1000 times larger than that of Iceland; Switzerland is four times richer than Turkey.

Recent literature on the existence of the IIH in the United States is surprisingly sparse. In a state-level study, Matthew & Brodersen (2018) found evidence that income inequality is negatively related to behavioural, physical, and mental health outcomes such as heavy drinking, obesity, exercise, diabetes, heart attack, heart disease, physical and mental health problems, and depression. They find that the absolute poverty effect is larger than the income inequality effect. In stark opposition to most of the literature, a cross-sectional study of obesity at the county-level by Fan *et al.* (2016) suggests that mixed-income neighbourhoods may be beneficial for reducing obesity as poorer individuals benefit from higher service quality and informal social control. A possible caveat to this study is the self-selection of poorer individuals choosing to live in mixed-income neighbourhoods, which could be correlated with factors associated with lower obesity risks.

Association Between Inequalities and COVID-19

Following an outbreak in China's Hubei province, in March 2020, the world seemingly came to a grinding halt when countries implemented social-distancing measures and lockdowns *en masse* to slow the spread of the virus. The viral spread marched on despite targeted public health policies; the dust settled, reliable data became available, and trends became discernable. It became evident that socio-economic factors played a large role in both the spread and mortality of the virus. As early as April 2020, reports of disproportionate infection and mortality among certain demographic/socio-economic groups began to surface. Based on data for the first four months of 2020, Oronce *et al.* (2020) found that a one Gini point increase at the state-level was associated with a 27% increase in Covid-19 mortality, and marginally associated with the incidence of the virus (+13%; 95% CI: +0% and +30%).

Initial studies were often restricted to a few metropolitan areas given that the early surges in cases happened mostly in large, crowded cities with international airports such as New York, Chicago, Los Angeles, and Miami. Brown & Ravaillion (2020) challenge the traditional epidemiologic models of homogeneous populations in a given context. Using county-level cross-sectional data, the authors argue that higher per capita income increases the effectiveness of social-distancing when controlling for inequality and other socio-economic variables. They also describe an ambiguous relationship between health outcomes and the share of elderly in a population,

suggesting that the number of infections is reduced through greater compliance with social-distancing even though the death rate conditional on infections is greater in counties with a large elderly population. Another study, using a similar econometric framework, highlighted the increased risk of both infection and mortality posed by social and demographic factors such as housing, belonging to a racial/ethnic minority and type of employment (Karmakar *et al.*, 2021). Although the association between race/ethnicity and Covid health outcomes is almost ubiquitous in the literature, Liao & De Maio (2021) interact racial make-up with inequality at the county-level and provide evidence that income inequality is a larger determinant of health outcomes even though both are so closely linked that the incremental effect of inequality is hard to measure precisely, especially in counties with a significant African American population. And lastly, partisan differences in the United States have been linked to social-distancing compliance, one study found that a 10% increase in regional viewership of conservative media led to a 1.3 percentage point decrease in the propensity to stay at home during the first lockdown (Simonov *et al.,* 2020).

Looking at causal pathways, Bambra et al. (2020) draws on two previous pandemics, the Spanish flu and H1N1 outbreak, along with Covid-19 to argue that the relationship between socio-economic and Covid inequalities are related to existing inequalities in chronic disease and social determinants of health. The authors demonstrate that income inequality prior to Covid-19 is closely linked with a heightened risk of noncommunicable diseases and chronic conditions that increase the risk of mortality from the novel coronavirus.

To the best of my knowledge, no study explores the impact of prior income inequality as a risk factor for Covid infection and death rates, even though several earlier studies have indicated the likelihood of a lagged effect of inequality on health (Leigh & Jencks, 2007; Lynch *et al.*, 2005). This paper starts to fill this gap by examining empirically the extent to which fluctuations in income inequality contributes to adverse Covid outcomes.

III. Data

This study uses data amalgamated from a variety of public sources and represents the first time that a historical measure of inequality is incorporated into these data sources. It expands upon Liao and De Maio (2021) by extending the COVID-related data to account for the second wave of

cases in the US, and by incorporating a measure of previous inequality. The result is a cross-sectional dataset of 3,089 [7] county-level observations in 50 states as well as for the District of Columbia. The nine major sources are the Centers for Disease Control and Prevention (CDC) and USAfacts.org for cumulative Covid-related deaths and cases; the American Community Survey for demographic and inequality data; the US Census Bureau for population density; the Bureau of Economic Analysis for per capita income figures; GitHub for county-level voting patterns; and the HealthData.gov reports for robustness checks using poverty rates. The data cover the first 328 days of the pandemic, from the first confirmed case in the US on January 22, to the date of the first vaccination roll-out on December 14, 2020. The decision to limit the analysis to when the first vaccine was administered stems from the potential political dimensions of vaccine compliance and the divergent state strategies for distribution.

Table 1 summarizes the key statistics for the analysis. The two outcome variables are cumulative deaths and confirmed cases. The key independent variables are income per capita, and Gini index measures for two different years: 2010 and 2019. The Gini index (or Gini coefficient) is a measure of statistical dispersion of income which ranges from 0 (where everyone receives an equal share of total income) and 1 (where one recipient/household receives all the income), to render it interpretable in a log-linked regression the Gini was transformed to a 100-point scale. The inclusion of a measure of the counties' previous level of inequality, namely its 2010 Gini coefficient, is an important aspect of this paper.

Three elements could pose an issue to the model's robustness: i) if county-level inequality did not vary between 2010 and 2019, ii) if the patterns of inequality were not random, for example if only the most unequal counties in 2010 witnessed a reduction of inequality and, iii) if the changes in inequality followed a similar trend nationwide, meaning that counties stayed in the same inequality quantile. The first issue can be eliminated as evident in Figure 3, which compares the 2010 values of the Gini index against their 2019 values; the majority of counties have experienced an increase in Gini, but significant variation remains, the correlation coefficient across counties being 0.644. To check for the second and third issues, I test for the inter-quintile mobility of counties. In total 54% of counties (1,678 out of the 3,089) changed their inequality quintile

[7] Omitted from this are 53 counties from Alaska, Hawaii, and Virginia. This represents a fraction of the three states' counties and are omitted for reasons of different geographic and population delimitations across the sources used in this paper.

between 2010 and 2019.[8] Of those in the most unequal quintile in 2010, 39% were no longer in that quintile in 2019. Of those in the least unequal quintile in 2010, 43% were no longer there in 2019.

Also included in table 1, a battery of controls including population density per square kilometre; the share of population that is male, younger than 20 years of age, older than 65 years; the racial/ethnic composition; the share of inhabitants without medical insurance; the number of days since a county's first confirmed case; a county's share of votes that went to the republican party in the 2016 presidential elections. Population density had a mean of 98.2 inhabitants per sq. km (range, 0.01 – 27 600); the mean of younger than 20 was 24.4% (range, 7.8 – 44.9) while it stood at 19.8% for older than 65 (range, 4.9 – 58.2); The mean of Black population was 9.2% (range, 0-86.56%); the mean Hispanic population was 9.8% (range, 0.65%-96.4%); the mean Native population was 2.4% (range, 0%-92.4%); The medical insurance rates ranged from 1.7% to 45.6%; and on average 255.8 days had passed since the first confirmed case of the virus, with only 2 counties never having a case. Table 6 provides all the variables' definitions and sources.

IV. Identification Strategy

Data on cumulative Covid cases and deaths are count data as they are composed of non-negative integers arising from counting (as opposed to ranking) the number of incidences. For count variables when a Poisson-type distribution is observed, a Poisson regression model is generally used. Poisson regressions work under the assumption that the mean and variance of the distribution will be of equal or similar value. When this assumption is too restrictive and the variance far exceeds the mean of the distribution parameter, a negative binomial regression is a more appropriate model (Lawless, 1987). Both outcome variables of this paper show signs of over-dispersion: the mean count of cases is 5,178 and the variance is 313,204,691 while the mean count of deaths is 95 and with a variance of 146,242. Additionally, following the econometric framework detailed by Payne *et al.* (2018), I ran a Pearson chi-squared test to measure the goodness-of-fit of a Poisson model. The test yields, for both outcome variables, a Pearson chi-squared substantially

[8] 36% have changed by 1 quintile and 18% have changed by 2 or more quintile

above one.[9] Given the over-dispersed properties of both my outcome variables, this paper estimates the coefficients of the predictor variables with a mixed-effect negative binomial model using maximum likelihood. The mixed-effect model includes a state-level random intercept to account for potential correlations between counties of the same state as a result of, for instance, cultural similarities and lockdown measures. And a fixed effect is included to control for unobserved variables that may contribute to counties' inequality and income levels.

Figures 1 and 2 illustrate the frequency distribution of the COVID-19 incidence and mortality across all counties, expressed per 100 000 population. Both outcomes are skewed to the right and appear to follow an over-dispersed Poisson distribution, best characterized as a negative binomial distribution, as is often the case for *count* variables in epidemiology. There is significant heterogeneity in county population size, therefore an offset variable representing the natural log of population size for each county is added, thus allowing to model the count variables as rates. An offset variable in a Poisson-type regression is a predictor variable whose coefficient is fixed at 1. A simplified version of a negative binomial model, where the dispersion parameter is nil (i.e. a Poisson model), can be written as:

$$\ln \lambda = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$$

Where the predictor variables x_j are given and the coefficients α_j are to be estimated. If Y is a dependent variable that counts the cumulative number of cases for a given population ω, this means that Y / ω is the observed case rate. Since $E(Y/\omega) = \lambda / \omega$, we can rewrite the basic model as such:

$$\ln \lambda = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n + ln\, \omega$$

The term $\ln \omega$ is the offset variable, which allows us to interpret the dependent variable (here, cases and deaths by counties) as a rate.

In equation (1) and (2), *Case* and *Death* is the county-level cumulative cases and deaths. Both dependent variables are fitted with the offset variable $\ln(Pop)$, where $\ln(Pop)$ is the natural

[9] Specifically, the Pearson chi-squared for incidence rate is 2,331,791 and 137,973for the death rate. A Pearson's chi-squared of 1 means the data is not over-dispersed.

log of a county's population. The $\alpha's$ and $\beta's$ are the parameters to be estimated. The independent variable *inequality* is the Gini index, both the current Gini index and its 2010 value are included; *income* is the county's per capita income in thousands of dollars; X is a vector of county-specific control variable such as the proportion of under 20 years of age, an age group that has a negligible mortality rate but has been linked to higher viral spread in a community; over 65 years old which are most at risk of dying once infected representing around 80% of Covid related deaths in the US (Johns Hopkins Medicine, 2021); without medical insurance; racial/ethnic makeup of the population; population density; number of days since the first confirmed COVID-19 case and; the share of votes for the Republican Party in the 2016 elections. In equation (2) an additional parameter is estimated, with $\ln(Cases)$ representing the counties' infection rate following the assumption that more cases per capita will lead to more deaths.

$$Case_{cs} = \exp\,(1 \times \ln(Pop) + \alpha_0 + \alpha_1 inequality_{cs} + \alpha_2 income_{cs} + X'_{cs}\alpha_3 + \mu_s + \mathcal{E}_{cs}) \quad \textbf{(1)}$$

$$Death_{cs} = \exp\,(1 \times \ln(Pop) + \beta_0 + \beta_1 inequality_{cs} + \beta_2 income_{cs} + \beta_3 \ln\,(Case_c) + X'_{cs}\beta_4 + \sigma_s + \epsilon_{cs}) \quad \textbf{(2)}$$

Death is associated with infection rates, but as Brown and Ravallion (2020) point out, controlling for ln(*Cases*) poses an endogeneity problem as the latent characteristics of counties that tend to increase infection rates are likely to be correlated with the severity of cases or the healthcare system, for example. To circumvent this issue, I insert the predicted value of the case rate estimated in (1) into equation (2). The error terms $\mathcal{E}_{cs}$ and ϵ_{cs} includes the over-dispersion parameters of the negative binomial distribution to account for the Poisson over-dispersion of both outcome variables. Since counties are nested within states, the terms μ_s and σ_s are the state-level random intercepts. All predictor variables are log-linked so the coefficients are reported as risk ratios (RR). Standard errors are robust and clustered at the state-level. This approach is in line with the most recent literature on the relationship between socio-economic factors and COVID-19 outcomes (Karmakar *et al.*, 2021; Liao & De Maio, 2021; Brown & Ravallion, 2020).

To explore further the impact of inequality on outcomes, I run two additional specifications. The first includes only those counties that experienced a decrease in their Gini index (inequality) over the last decade and the second, includes only those that experienced an increase in their Gini

index. This allows for a comparison in Covid outcomes between geographical units that have witnessed a favourable fluctuation in income inequality against those that have witnessed a deterioration in equality. It has been hypothesized that rising inequalities negatively affect health through an erosion of social cohesion, trust and an increased risk of chronic stress in subgroups of the population (Bor *et al.,* 2017; Elgar *et al.* 2020). Should these mechanisms be present in US counties, I expect to find adverse health outcomes of higher magnitude for counties that have witnessed a rise in inequality.

I conduct a variety of other robustness checks and sensitivity analyses. Firstly, the death rate conditional on infection is falling as the medical community's understanding of the virus grows and as the initial shock to medical supplies such as ventilators faded (CDC, 2020). Because surges in the spread of the virus happened at different times in different states, table 8 excludes the 150 counties with the highest and lowest incidence and mortality rates to test for the model's robustness to medical understanding of the virus rather than the population's resilience. Secondly, while in my main specification, *income per capita* is included to capture the direct income effect on health outcomes, an additional robustness check substitutes it with a measure of poverty. Poverty rates were not included in the main model specifications because they are, by construct, substantially explained by a combination of inequality and income per capita, presenting a collinearity problem. For example, if a county has a high Gini index (inequality) and a low per capita income, poverty will be high. Table 7 shows that poverty, in line with the literature, is a statistically significant predictor of incidence and mortality, yet the coefficient estimate remains smaller than that of inequality. And lastly, Table 9 excludes the counties in the top and bottom Gini index quintile. A total of 1 232 counties are excluded to test that the results are not overly influenced by counties concentrated in the extremes of the inequality distribution. This three-part sensitivity analysis suggests that the findings of this paper are robust to the exclusion of extreme cases both for the outcome variables and the main dependent variables, as well as with alternative measures of a community's average richness.

V. Results

I begin by ensuring that the results of my regressions are consistent with previously published studies that use a shorter timeframe of COVID-19 observations. This also enables me to verify that the COVID-19 and inequality relationship previously established still hold. In a paper

also using a mixed-effect negative binomial regression with Covid data spanning January to July 2020, Karmakar *et al.* (2021) found that when controlling for density and testing rates, a 1% increase in Gini index was associated with a 3% and 4% increase in the incidence and mortality, respectively. Using data for the first 200 days of the pandemic, Liao and De Maio (2021) found that a 1% rise in a county's income inequality corresponds with a 1.9% and 3% increase in the incidence and mortality. Liao and De Maio controlled for density, number of days since the first confirmed case, age, race and several state-level variables such as the adoption of Medicaid, the governors' political affiliation and gender. Of the state-level variables, only Medicaid was statistically significant, reducing the risk of incidence by 32%. This paper does not include state-level variables and uses county-level medical insurance rates and political vote instead, the assumption being that state characteristics will be captured through the state-level intercept and the nesting of counties within their state.

The estimated coefficients reported in table 2 are similar in sign but smaller in magnitude to those of Liao and De Maio (2021): A 1% increase in Gini index is associated with a 1.31% and 1.41% increase in the risk of incidence and mortality, respectively, at a statistically significant level. The income per capita does not affect the risk of incidence but appears to reduce the risk of mortality though not at a significant level. The coefficient estimate for the share of people aged 65 and above corroborates the ambiguous relationship between the share of elderly and COVID-19 outcomes described by Brown and Ravallion (2020): a 1% increase in the share of elderly reduces the risk of incidence by around 0.94% but increases the risk of mortality by 4.26%, this is likely due to the increased effectiveness/compliance of social distancing measures and the presence of a self-protection effect among the elderly. The coefficient estimate for the share of uninsured suggests that a larger share of people without medical insurance reduces the risk of incidence, I interpret this as an indication that uninsured people are less likely to be tested or seek medical attention when presented with mild symptoms; the coefficient is no longer significant when predicting the risk of mortality. There could also be unobserved variables correlated with being uninsured such as the prevalence of health institutions in a county or the share of young adults. It is also possible that the irregular adoption of Medicaid at the state-level obstructs the estimation of the true coefficient.

The results of table 3 suggest that a county's past level of inequality increases the risk of incidence and mortality with a higher magnitude than current inequality, even rendering the

coefficient estimates for current inequality on mortality barely statistically significant (see table 2 for comparison). Controlling for all variables in the vector X of equation (1) and for current Gini value, a 1% increase in past inequality corresponds to a 1.5% rise in both incidence and mortality (p-value < 0.01). When running the regression without controlling for current Gini value, a 1% rise in 2010 Gini corresponds with a 1.8% increase in the risk of incidence (p-value < 0.01) and a 1.9% rise in risk of mortality (p-value < 0.01). Income per capita seems to have a negligible impact on Covid incidence but might slightly reduce the risk of mortality by 0.31% though not statistically significant (p-value=0.18). These findings are not completely foreign to the literature on the IIH. There has been evidence of a lagged effect of inequality on health, with some potential causal pathways such as chronic stress, noncommunicable diseases and food insecurity that can take several years to materialize into health risks (Leigh & Jencks, 2007; Bambra *et al.*, 2020).

The results in tables 2 and 3 suggest that past inequality, defined as inequality ten years prior, is a better predictor of a community's resilience to the COVID-19 pandemic than current levels of inequality. In fact, the coefficient estimates for 2010 inequality on the risk of incidence and mortality are, respectively, 32% and 34% larger than the coefficient estimates for 2019 inequality. In both regressions, per capita income did not play a statistically significant role in assessing the risk of incidence and mortality, though the p-values are smaller for the mortality rate. This could imply that an individual's income does little to protect individuals from infection but slightly reduces the risk of death, the effect could also be captured by other variables such as population density and demographic make-up of the population. In fact, all variables for race, gender, Republican vote and age make-up of a county have statistically significant coefficient estimates. The share of Afro-Americans and elderly increasing the risk of mortality by 2.2% and 4.2% respectively, despite having a negative or negligible effect of the risk of incidence.

I divide my data in two categories based on whether a county has witnessed an overall rise or fall in inequality between 2010 and 2019.[10] Roughly one quarter of counties (778 out of 3089) witnessed a decrease in inequality during that period. Table 4 replicates the regression in table 2 excluding all counties with an increase in inequality between 2010 and 2019. The coefficient estimates for current inequality suggest that inequality plays a role in the risk of incidence (1.06%

[10] Six counties had the same Gini value in 2010 and 2019. I used the 2014 Gini measure to place them in a category: if 2014 Gini is higher than the 2010 and 2019 value, the county is placed in the "increase in inequality" category. And *vice versa*

increase, p-value<0.01) yet is negligible and statistically insignificant for the risk of mortality. Interestingly, a 1% increase in income per capita is estimated to reduce the risk of mortality by 1% (p-value<0.00). This contrasts with the earlier tables where income per capita did not have a statistically significant effect on incidence or mortality when all counties were included. In other words, for counties that witnessed a decrease in inequality over the last decade, income per capita seems to play an important role in reducing the risk of mortality, while inequality solely increases the risk of infections.

Table 5 includes only counties where inequality increased over the last decade. The results are similar in sign to when all counties were considered. A 1% increase in the Gini value is associated with a 1.7% and 2.3% rise in the risk of incidence and mortality, respectively, and at a statistically significant level. Both coefficients are larger than when all counties were included, 62% greater in the case of mortality. Income per capita does not affect either outcome variables in a statistically significant way.

VI. Discussion

The income inequality hypothesis (IIH) whereby a community's level of inequality, not just level of income, plays a determining role in its health is still the subject of debate in the economic literature. It is generally accepted that evidence for the IIH is not always present because the underlying mechanisms through which inequality affects health are not well understood and likely vary depending on the cultural, geographic, and demographic context, particularly with *between* country analyses. This paper is restricted to a single country with a certain degree of shared culture, demographics and importantly, quasi-centralized policy making. In the case of the United States' COVID-19 outcomes, the IIH appears to be present both for the viral spread and the risk of mortality. Not only are current levels of inequality a robust predictor of a county's health resilience, but the results also suggest that past levels of inequality may be the better predictor of both outcome variables, with a larger estimated coefficient than per capita income or current inequality.

Distinguishing counties by whether their inequality levels have improved or deteriorated underscores the notion that the IIH is not applicable everywhere regardless of context. It appears that the process of reducing inequality mitigates the effect of current inequality on the risk of mortality, with income per capita becoming a better predictor. However, if counties experience

deteriorating inequality, the strength of the IIH is amplified, regardless of the average income of a county's inhabitants. If predictions about a sharp rise in inequality resulting from the pandemic and the Covid recession prove to be right, this does not bode well for future pandemic preparedness.

This paper highlights the importance of considering both inequality and patterns in inequality, as factors affecting a population's health outcomes. As to why counties with a decreasing inequality exhibit a different expression of the IIH than those on a different trend, the neo-material hypothesis could offer clues. The neo-material pathway, as suggested by Kahn *et al.* (2000), hypothesizes that the policy framework and societal structure that is conducive to rising income inequality are themselves associated with systematic under investments in social infrastructure such as education, health services, food security, and public transportation. In turn impacting the social gradient of overall health. Related to this hypothesis and increasingly present in the literature is the deterioration of trust and social cohesion resulting from high levels of inequality in a community (e.g., Elgar *et al.*, 2020).

Another explanation for why inequality might play a large role in a community's ability to withstand the deleterious health ramifications of COVID-19 might be directly linked with the association inequality has with co-morbidities. COVID-19 is a highly transmissible virus with a relatively low lethality. Lethality is, however, strongly associated with co-morbidities. It is estimated that 94% of Covid related deaths in the US had at least one co-morbidity despite *only* 6 out 10 American adults living with one or more co-morbidities (Auwaerter, 2021). A recent meta-analysis of hospitalized patients found striking evidence of the prevalence of existing co-morbidities in COVID-19 deaths. Cardiovascular diseases increase the likelihood of death by a factor of 2.25, hypertension by a factor of 1.82, and diabetes by a factor of 1.48 (Ssentongo *et al.*, 2020). Other studies have suggested that diabetes is the best predictor of mortality rates in different countries, after adjusting for confounders (Corona *et al.*, 2021). The strong relationship between co-morbidities and COVID-19 does not diminish the transferability of the empirical results in this paper. Co-morbidities are an overall strong indicator of a population's general health and likely to be associated in a similar way to future epidemics.

VII. Conclusion

This paper draws on the recent literature identifying income inequality as a risk factor for COVID-19 outcomes and demonstrates that eleven months into the pandemic, the relationship holds. This research also brings forth the importance of including a community's past level of inequality when assessing its health vulnerabilities. The public health policy implications are straightforward, more unequal counties have experienced higher infection and death rates even when controlling for income and poverty. Moreover, counties that experienced a decrease in inequality since 2010 have been able to mitigate the risks associated with current levels of inequality. The results of this analysis are not causal, and although the data for inequality predates the pandemic, there are likely confounders correlated with both inequality and health outcomes. Besides co-morbidities like diabetes, such confounding factors could potentially be social cohesion, trust and suboptimal investments in infrastructure or safety nets to protect the most vulnerable. More research would be needed to determine the pathways through which the IIH presents itself in the United States.

This study has several limitations. First, it does not include measures of co-morbidities or negative health habits that could be linked with inequality. No causal pathway is proposed to explain how inequality relates to a community's health resilience, rather, it implicitly assumes that a deterioration in the fair distribution of income generates a plethora of phenomena that can lead to unfavourable health outcomes, specifically in the case of COVID-19. Secondly, the United States has idiosyncrasies that may not be transferable to other nations, such as the lack of universal healthcare. This reduces the policy makers' incentives to seriously promote/implement general health policies such as junk food taxes, nicotine addiction campaigns, subsidized physical activities for marginalized communities, environmental decontamination and fair access to healthy food or medical practitioners, to name a few. Far-reaching and integrated health public policies could mitigate the risk of health conditions arising from income inequality and poverty.

VIII. References

Auwaerter, Paul G. (2021), "Coronavirus COVID-19 (SARS-CoV-2)", *Johns Hopkins ABX Guide*, The Johns Hopkins University, accessed March 15, 2021 at www.hopkinsguides.com/hopkins/view/Johns_Hopkins_ABX_Guide/540747/all/Coronavirus_COVID_19__SARS_CoV_2_

Avendano, Mauricio (2012), "Correlation or causation? Income inequality and infant mortality in fixed effects models in the period 1960–2008 in 34 OECD countries", *Soc Sci Med.*, Vol. 75(4), pp. 754–760. Doi: https://doi.org/10.1016/j.socscimed.2012.04.017

Bambra, Clare (2011), *"*Work, worklessness and the political economy of health inequalities." *J Epidemiol Community Health.* Vol. 65(9):746-50. http://dx.doi.org/10.1136/jech.2009.102103

Bambra, Clare (2016). *"Health divides: where you live can kill you", Policy Press*, 320 pages.

Bambra, Clare; Riordan, Ryan; Ford, John; Matthews, Fiona (2020), "The COVID-19 pandemic and health inequalities", *J. Epidemiol Community Health*. Vol. 74(11), pp. 964-968. doi: 10.1136/jech-2020-214401

Blundell, Richard; Costa Dias, Monica; Joyce, Robert; Xu, Xiaowei (2020), "COVID-19 and Inequalities". *Fiscal Studies,* Vol. 41, pp. 291-319. https://doi.org/10.1111/1475-5890.12232

Böckerman, P.; Johansson, E.; Helakorpi, S.; Uutela, A. (2009), "Economic inequality and population health: looking beyond aggregate indicators", *Sociology of Health and Illness*, Vol. 31(3), pp. 422–440. doi: https://doi.org/10.1111/j.1467-9566.2008.01144.x

Bonacini, Luca; Gallo, Giovanni; Scicchitano, Sergio (2021), "Working from home and income inequality. Risks of a 'new normal' with COVID-19", *Journal of Population Economics*, Vol 34(1). https://doi.org/10.1007/s00148-020-00800-7

Bor, Jacob; Cohen, Gregory H.; Galea, Sandro (2017), "Population health in an era of rising income inequality: USA, 1980-2015", *The Lancet,* Vol. 389, pp. 1475-1490. doi 10.1016/S0140-6736(17)30571-8

Brown, Caitlin; Ravallion, Martin (2020), "Inequality and the Coronavirus: Socio-economic Covariates of Behavioral Responses and Viral Outcomes Across Us Counties", *NBER Working Paper,* No. w27549. Available at SSRN: https://ssrn.com/abstract=3658845

Carrieri, Vincenzo; Jones, M. Andrew (2017), "The Income–Health Relationship 'Beyond the Mean': New Evidence from Biomarkers", *Health Economics,* Vol.26, pp.937– 956. doi: 10.1002/hec.3372.

Cavalini, L.T.; De Leon, A.C. (2008), "Morbidity and mortality in Brazilian municipalities: a multilevel study of the association between socio-economic and healthcare

indicators", *International Journal of Epidemiology*, Vol. 37(4), pp. 775–783. Doi: https://doi.org/10.1093/ije/dyn088

Centers for Disease Control and Prevention (2020), "COVID-19 Hospitalization and Death by Race/Ethnicity". Accessed 03/02/2021. https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html

Corona, Giovanni; Pizzocaro, Alessandro; Vena ,Walter; *et al. (*2021) "Diabetes is most important cause for mortality in COVID-19 hospitalized patients: Systematic review and meta-analysis", *Reviews in endocrine & metabolic disorders*, Vol. 22 (1), doi:10.1007/s11154-021-09630-8

Cutler, David; Deaton, Angus; Lleras-Muney, Adriana (2006), "The Determinants of Mortality", *Journal of Economic Perspectives,* Vol. 20(3), pp. 97-120. doi: 10.1257/jep.20.3.97

De Maio, Fernando (2014), "Income Inequality Hypothesis.", *The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society* https://doi.org/10.1002/9781118410868.wbehibs312

De Vogli Roberto, Mistry Ritesh, Gnesotto Roberto, Cornia Giovanni (2005) "Has the relation between income inequality and life expectancy disappeared? Evidence from Italy and top industrialised countries", *Journal of Epidemiology and Community Health,* Vol. 59(2), pp.158-62. doi: 10.1136/jech.2004.020651.

Du, Hongfei; King, Ronnel B.; Chi, Peilian (2019), "Income inequality is detrimental to long-term well-being: A large-scale longitudinal investigation in China", *Social Science & Medicine*, Vol. 232, pp.120-128

Dunn, James; Schaub, Peter; Ross, Nancy (2007), "Unpacking Income Inequality and Population Health: The Peculiar Absence of Geography", *Canadian Journal of Public Health*, vol. 98, pp. S10–S17. https://doi.org/10.1007/BF03403722

Elgar, Frank J; Aitken, N (2011), "Income inequality, trust and homicide in 33 countries", *European Journal of Public Health,* Vol 21, pp. 241–246. https://doi.org/10.1093/eurpub/ckq068

Elgar, F.J.; Stefaniak, A; Wohl, M.J.A (2020), "The trouble with trust: Time-series analysis of social capital, income inequality, and COVID-19 deaths in 84 countries", *Social Science & Medicine,* Vol. 263, pp. 113365. doi: https://doi.org/10.1016/j.socscimed.2020.113365

Fan, Jessie; Wen, Ming; Kowaleski-Jones, Lori (2016), "Tract- and county-level income inequality and individual risk of obesity in the United States", *Social Science Research,* Vol. 55, pp. 75-82. Doi: https://doi.org/10.1016/j.ssresearch.2015.09.008

Hollingsworth, Alex; Ruhm, Christopher; Simon, Kosali (2017), "Macroeconomic conditions and opioid abuse", *Journal of Health Economics,* Vol. 56, pp. 222-23. https://doi.org/10.1016/j.jhealeco.2017.07.009

International Monetary Fund (2015), "Causes and Consequences of Income Inequality: A Global Perspective", *Strategy, Policy, and Review Department,* June 2015

Kahn, R.S.; Wise, P.H.; Kennedy, B.P.; Kawachi, I. (2000), "State income inequality, household income, and maternal mental and physical health: cross sectional national survey", *British Medical Journal*, Vol. 321(7272), pp. 1311–1315.

Kaplan, GA.; Pamuk E.R.; Lynch J.; Cohen RD; Balfour JL, (1996) "Inequality in income and mortality in the United States: analysis of mortality and potential pathways.", *BMJ,* Vol. 312(7037), pp. 999-1003. doi: 10.1136/bmj.312.7037.999.

Karmakar, Monita; Lantz Paula, Tipirneni Renuka (2021), "Association of Social and Demographic Factors With COVID-19 Incidence and Death Rates in the US", *JAMA Netw Open.* Vol. 4(1), doi: 10.1001/jamanetworkopen.2020.36462

Kawachi, Ichiro; Subramanian S.V.; Almeida-Filho Naomar, (2002), "A glossary for health inequalities", *Journal of Epidemiology & Community Health,* Vol. 56, pp. 647-652. Doi http://dx.doi.org/10.1136/jech.56.9.647

Kim, Ki-tae (2019), "Revisiting the Income Inequality Hypothesis With 292 OECD Regional Units", *International Journal of Health Services*, Vol. 49(2), pp. 360-370. doi: https://doi.org/10.1177/0020731418814105

Kragten, Nigel; Rözer, Jesper (2017) "The Income Inequality Hypothesis Revisited: Assessing the Hypothesis Using Four Methodological Approaches", *Social Indicators Research*, Vol.131, pp. 1015–1033. https://doi.org/10.1007/s11205-016-1283-8

Latif, Ehsan (2015), "Income Inequality and Health: Panel Data Evidence from Canada", *The B.E. Journal of Economic Analysis & Policy*, Vol. 15(2), pp. 927-959. doi: https://doi.org/10.1515/bejeap-2013-0184

Lawless, Jerald F. (1987), "Negative Binomial and Mixed Poisson Regression", *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, Vol. 15(3),pp. 209-225. Doi: https://doi.org/10.2307/3314912

Leigh, Andrew; Jencks, Christopher (2007), "Inequality and mortality: long-run evidence from a panel of countries" *Journal of Health Economics*, Vol. 26(1), pp.1–24.

Liao, Tim; De Maio, Fernando (2021) **"**Association of Social and Economic Inequality With Coronavirus Disease 2019 Incidence and Mortality Across US Counties"**,** *JAMA Netw Open,* Vol. 4(1):e2034578. doi: 10.1001/jamanetworkopen.2020.34578

Long, Heather; Van Dam, Andrew; Fowers, Alyssa; Shapiro, Leslie (2020), "The COVID-19 recession is the most unequal in modern U.S. history", *The Washington Post,* The Covid Economy. Retrieved from https://www.washingtonpost.com/graphics/2020/business/coronavirus-recession-equality/

Lynch, John; Davey Smith, George; Harper, Sam, *et al.* (2004) "Is income inequality a determinant of population health? Part 1. A systematic review", *Milbank Quarterly*, Vol. 82(1), pp. 5–99. doi: 10.1111/j.0887-378X.2004.00302.x

Lynch, John; Harper, Sam; Kaplan, George A.; Davey Smith, George (2005), "Associations between income inequality and mortality among US states: the importance of time period and source of income data", *American Journal of Public Health*, Vol. 95(8), pp.1424–1430

Mackenbach, Johan; Martikainen ,Pekka; Looman, Caspar *et al.* (2005), "The shape of the relationship between income and self-assessed health: an international study", *International Journal of Epidemiology*, Vol. 34(2), pp. 286–293. doi: https://doi.org/10.1093/ije/dyh338

Marmot, Michael G.; Smith, George D.; Stansfeld, Stephen A.; et al. (1991) "Health inequalities among British civil servants: the Whitehall II study", *Lancet*, Vol. 337(8754), pp. 1387-93. doi: 10.1016/0140-6736(91)93068-k

Matthew, Pravin; Brodersen, Donka Mirtcheva (2018), "Income inequality and health outcomes in the United States: An empirical analysis", *The Social Science Journal*, Vol. 55(4), pp. 432-442. Doi: https://doi.org/10.1016/j.soscij.2018.05.001

Mayhew, Ken; Wills, Samuel (2019) "Inequality: an assessment", *Oxford Review of Economic Policy*, Vol. 35(3), pp. 351–367, https://doi.org/10.1093/oxrep/grz009

Oates, Gabriella; Jackson, Bradford; Partridge, Edward (2018), "Sociodemographic Patterns of Chronic Disease: How the Mid-South Region Compares to the Rest of the Country", *American Journal of Preventive Medicine*. Vol. 52 (1.1), pp. 31-39. https://doi.org/10.1016/j.amepre.2016.09.004

Oronce, Carlos Irwin A.; Scannell, Christopher A.; Kawachi, Ichiro; Tsugawa, Yusuke (2020) "Association Between State-Level Income Inequality and COVID-19 Cases and Mortality in the USA", *Journal of General Internal Medecine*, Vol 35, pp.2791-2793. doi: 10.1007/s11606-020-05971-3

OECD (2020) "OECD Income Distribution Database". Accessed on February 12, 2021. http://www.oecd.org/social/income-distribution-database.htm

Payne, Elizabeth H.; Gebregziabher, Mulugeta; Hardin, James; *et al.* (2018) "An empirical approach to determine a threshold for assessing overdispersion in Poisson and negative binomial models for count data", *Communications in statistics: Simulation and computation*, Vol. 47(6), pp.1722-1738. doi:10.1080/03610918.2017.1323223

Pei, X.; Rodriguez, E. (2006)," Provincial income inequality and self-reported health status in China during 1991-7", *Journal of Epidemiology and Community Health*, Vol.60(12), pp. 1065–1069. doi: 10.1136/jech.2005.043539

Pew Research Center (January 2020), "Most Americans Say There Is Too Much Economic Inequality in the US, but Fewer Than Half Call It a Top Priority" retrieved online from https://www.pewsocialtrends.org/2020/01/09/views-of-economic-inequality/

Pop, Ioana A.; van Ingen, Erik; van Oorschot, Wim (2012), "Inequality, Wealth and Health: Is Decreasing Income Inequality the Key to Create Healthier Societies?", *Social Indicators Research*, Vol.113, pp. 1025–1043. https://doi.org/10.1007/s11205-012-0125-6

Qureshi, Zia (2020), "Report: Tackling the inequality pandemic: Is there a cure?", *Brookings*, Reimagining the global economy: Building back better in a post-COVID-19 world, retrieved from https://www.brookings.edu/research/tackling-the-inequality-pandemic-is-there-a-cure/

Saez, Emmanuel; Zucman, Gabriel (2020), "The Rise of Income and Wealth Inequality in America: Evidence from Distributional Macroeconomic Accounts", *Journal of Economic Perspectives*, Vol. 34(4), pp. 3-26. doi: https://doi.org/10.1257/jep.34.4.3

Simonov, Andrey; Sacher, Szymon; Dubé, Jean-Pierre; Shirsho, Birwas (2020), "The Persuasive Effect of Fox News: Non-Compliance with Social Distancing During the COVID-19 Pandemic", *NBER Working Papers*, No. 27237

Ssentongo, Paddy; Ssentongo, Anna; Heilbrunn, Emily; Ba, Djibril; Chinchilli, Vernon (2020), "Association of cardiovascular disease and 10 other pre-existing comorbidities with COVID-19 mortality: A systematic review and meta-analysis." *PLoS One*, Vol. 15(8), e0238215. doi: 10.1371/journal.pone.0238215

Stabile, Mark; Bénédicte Apouey and Isabelle Solal (2020), "COVID-19, inequality, and gig workers", *VoxEU*, CEPR: 1 April 2020.

Sydenstricker, Edgar (1931), "The incidence of influenza among persons of different economic status during the epidemic of 1918", *Public Health Reports (1896-1970)*, Vol. 46, No. 4 (Jan. 23, 1931), pp. 154-170. https://doi.org/10.2307/4579923

Thebault R, Tran AB; Williams V. (2020), "African Americans are at higher risk of death from coronavirus", The Washington Post, 7 April 2020. https://www.washingtonpost.com/nation/2020/04/07/coronavirus-is-infecting-killing-black-americans-an-alarmingly-high-rate-post-analysis-shows/?arc404=true

Whitehead, Margaret; Pennington, Andy; Orton, Lois; Nayak, Shilpa; Petticrew, Mark; Sowden, Amanda; White, Martin (2016), "How could differences in 'control over destiny' lead to socio-economic inequalities in health? A synthesis of theories and pathways in the living environment", *Health & Place*, Vol 39, pp. 51-61. https://doi.org/10.1016/j.healthplace.2016.02.002

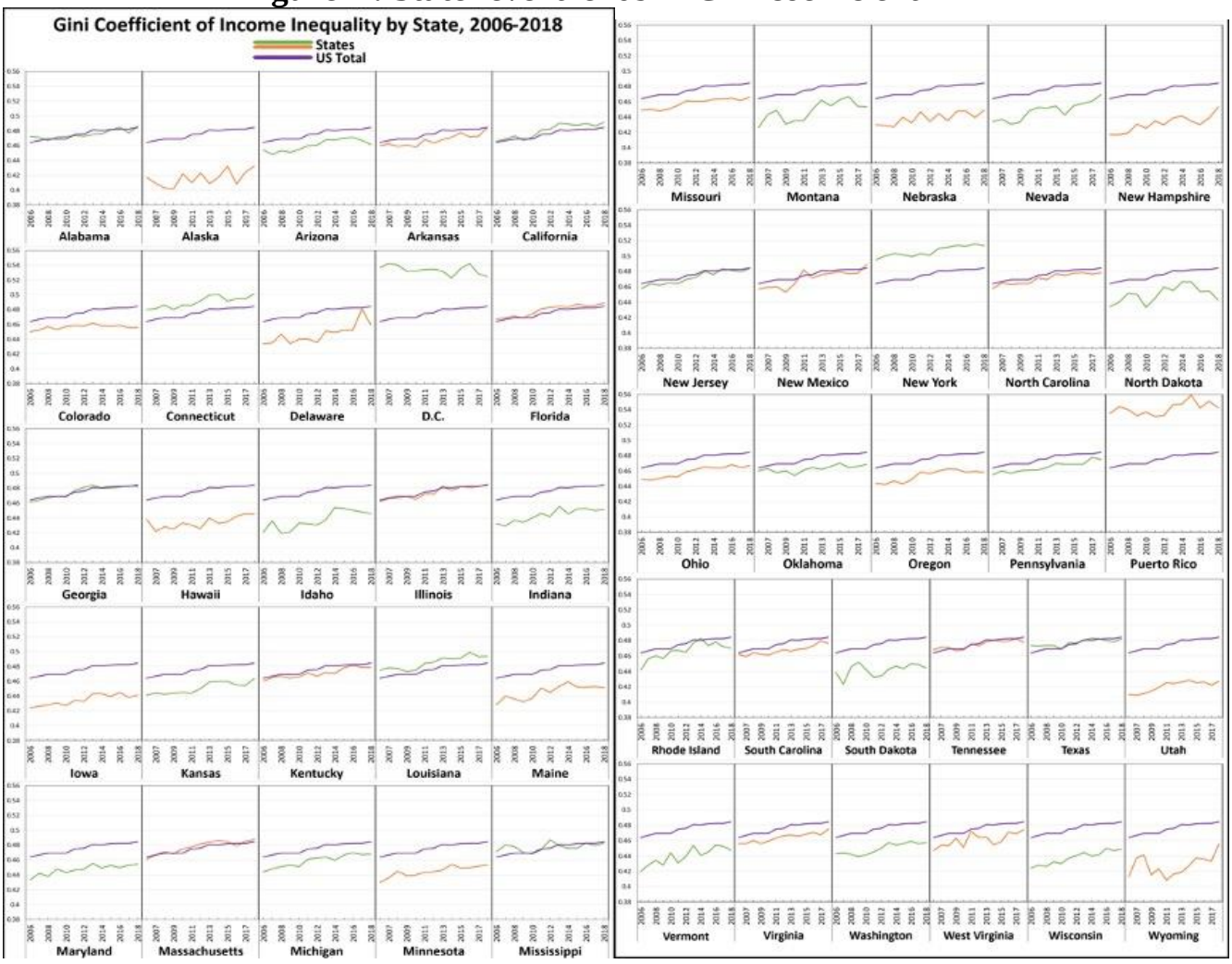
Wilkinson, R.G. (1992), "Income distribution and life expectancy", *British Medical Journal*, Vol. 304(6820), pp. 165–168. https://doi.org/10.1136/bmj.304.6820.165

Wilkinson, R.G; Pickett, Kate (2006), "Income inequality and population health: a review and explanation of the evidence", *Social Science & Medicine*, Vol. 62(7), pp.1768-84. Doi: https://doi.org/10.1016/j.socscimed.2005.08.036

World Bank (2016), "Poverty and Shared Prosperity 2016: Taking on Inequality", *Washington, DC: World Bank*. doi:10.1596/978-1-4648-0958-3

IX. Figures and Tables

Figure A: State-level trends in Gini coefficient



Source: SSTI analysis of Census Bureau, American Community Survey 1-year estimates

Figure 1

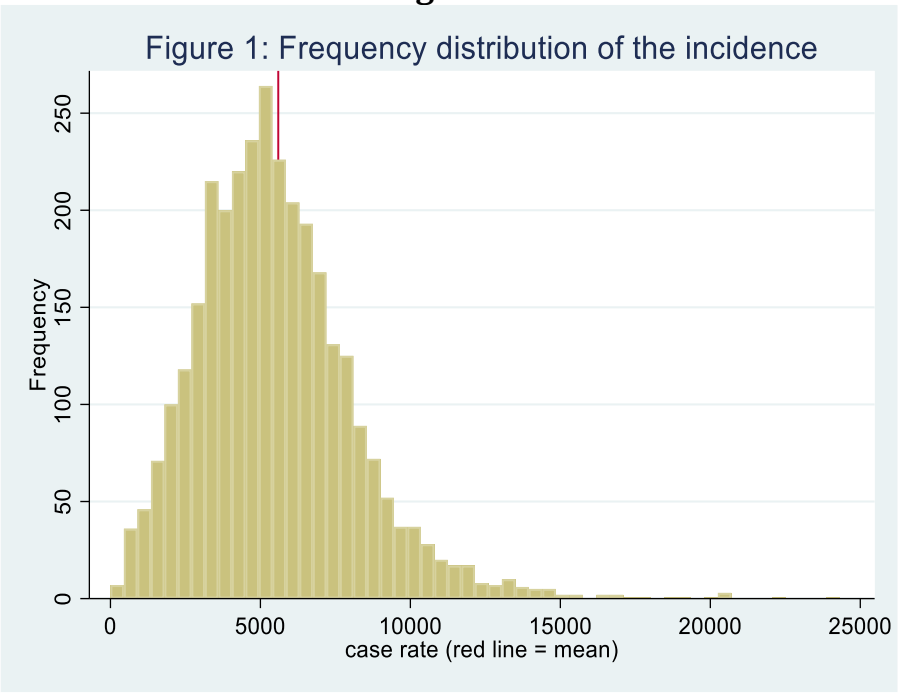


Figure 2

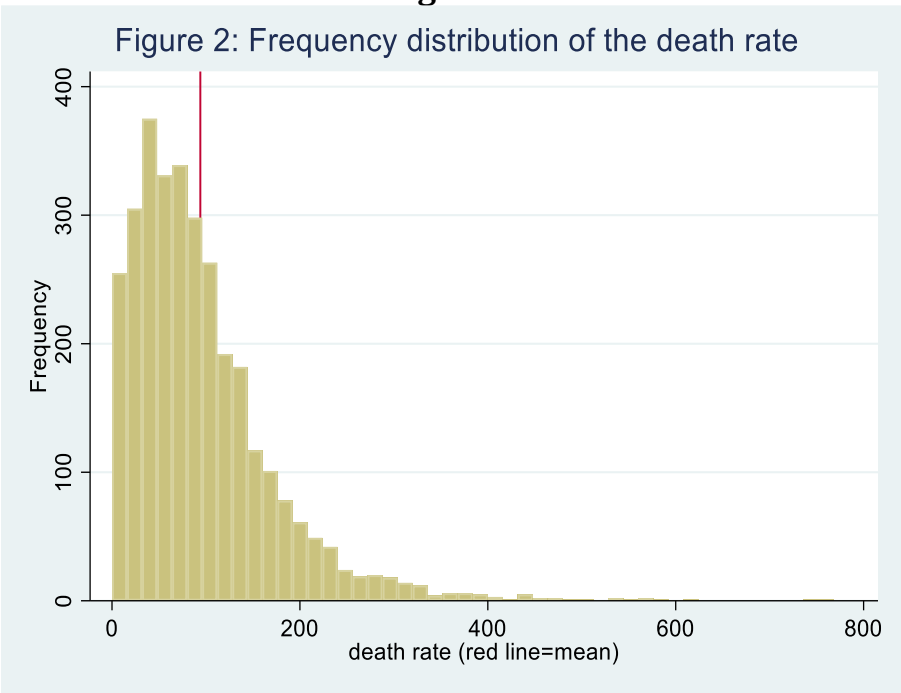
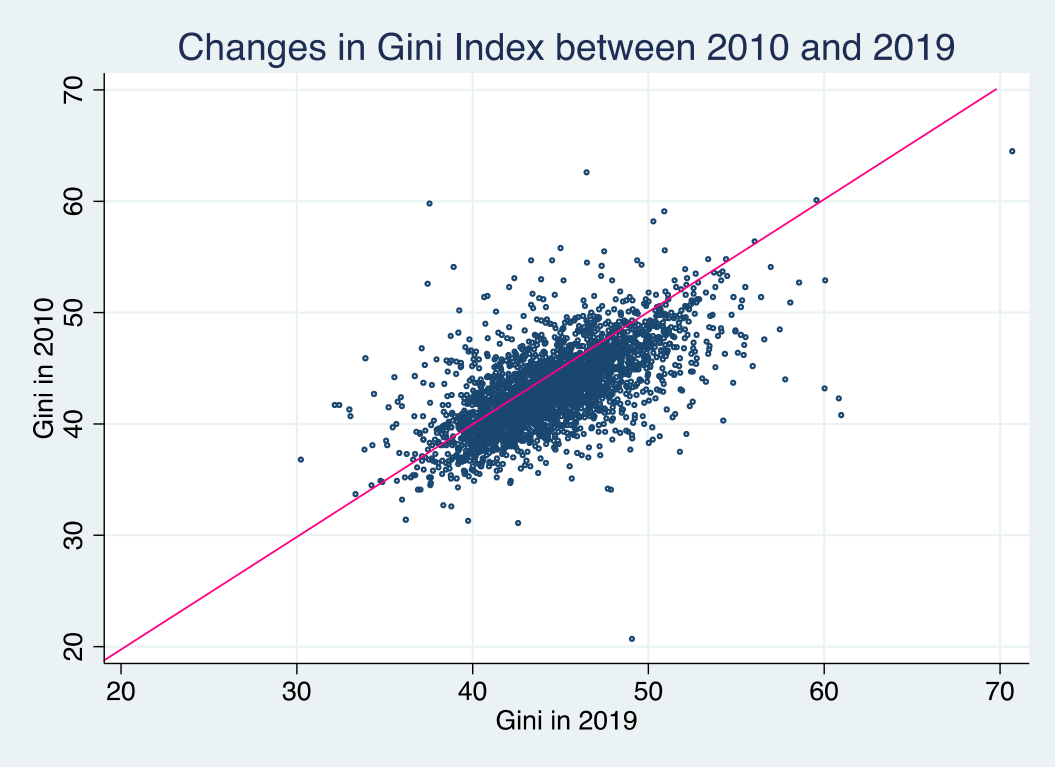


Figure 3



*The red line is the 45 ° line

Table 1: Descriptive Statistics for 3 089 US counties

Variable	Mean	Std. Dev.	Min	Max
Cumulative deaths	94.9	382.4	0	8345
Cumulative cases	5178.9	17697.6	0	532730
Death per 100k	95.5	78.1	0	758.7
Incidence per 100K	5 534.3	2 650.3	0	23 956.4
Days since first case	255.8	30.1	0	328.0
Pop. Density (sq. km)	98.2	688.8	0.014	27 602.3
% male	50.1	2.2	43.0	73.5
% under 20 years	24.4	3.6	7.8	44.9
% over 65 years	19.8	4.8	4.9	58.2
% Black	9.2	14.3	0	86.6
% Hispanic	9.8	13.9	0.65	96.4
% Native	2.4	7.7	0	92.4
Uninsured rate	10.1	5.1	1.7	45.6
Per capita income (2019)	45 933	13 250	19 472	229 825
Gini index (2019)	44.5	3.7	30.2	70.7
Gini index (2010)	43.1	3.6	20.7	64.5
Gini change 2010-2019	1.40	3.1	-22.25	28.4
County % vote Rep	63.7	15.5	4.12	95.3

Table 2: Mixed-effect negative binomial regression to estimate the relationship between current inequality and Covid incidence and mortality

	(1)		(2)	
Variables	**Case Rate**	P-value	**Death Rate**	P-Value
Gini 2019	1.0131	0.00	1.0141	0.00
Income per capita	1.0001	0.92	0.9982	0.25
Ln (Cases)	NA	-	1.0048	0.86
Male	1.0292	0.00	1.0128	0.13
% under 20	1.0127	0.07	1.0358	0.00
% over 65	0.9899	0.02	1.0426	0.00
% Black	1.0096	0.00	1.0224	0.00
% Native	1.0084	0.00	1.0187	0.00
% Hispanic	1.0149	0.00	1.0184	0.00
# Days since first	1.0011	0.05	1.0014	0.11
Density	1.0000	0.01	1.0001	0.00
% Uninsured	0.9916	0.00	0.9957	0.22
% voted GOP	1.0092	0.00	1.0151	0.00
Constant	0.0019	0.00	0.0000	0.00
Number of Obs.	3089		3089	
Wald χ^2	691.8		1076.2	

Exponentiated coefficients mean coefficients are reported as Risk Ratios

Table 3: Mixed-effect negative binomial regression to estimate the relationship between current and past inequality and Covid incidence and mortality

	(1)		(2)	
Variables	**Case Rate**	P-value	**Death Rate**	P-Value
Gini 2019	1.0061	0.08	1.0065	0.13
Gini 2010	1.0150	0.00	1.0156	0.00
Income per capita	0.9998	0.83	0.9979	0.18
Ln (Cases)	NA	-	1.0025	0.93
Male	1.0309	0.00	1.0141	0.10
% under 20	1.0143	0.04	1.0371	0.00
% over 65	0.9897	0.01	1.0419	0.00
% Black	1.0092	0.00	1.0220	0.00
% Native	1.0081	0.00	1.0186	0.00
% Hispanic	1.0145	0.00	1.0182	0.00
# Days First	1.0011	0.05	1.0015	0.10
Density	1.0000	0.01	1.0001	0.00
% Uninsured	0.9906	0.00	0.9943	0.10
% voted GOP	1.0095	0.00	1.0155	0.00
Constant	0.0012	0.00	0.0000	0.00
Number of Obs.	3089		3089	
Wald χ^2	723.1		1427.9	

Exponentiated coefficients mean coefficients are reported as Risk Ratios

Table 4: Mixed-effect negative binomial regression: model only includes counties with income inequality decrease (2010-2019)

	(1)		(2)	
Variables	**Case Rate**	P-value	**Death Rate**	P-Value
Gini 2019	1.0106	0.01	1.0050	0.65
Income per capita	0.9983	0.22	0.9910	0.00
Ln (Cases)	NA	-	0.9727	0.37
Male	1.0276	0.03	1.0050	0.68
% under 20	1.0134	0.20	1.0514	0.00
% over 65	0.9895	0.04	1.0434	0.00
% Black	1.0074	0.00	1.0213	0.00
% Native	1.0080	0.01	1.0120	0.02
% Hispanic	1.0142	0.00	1.0180	0.00
# Days First	1.0013	0.06	1.0034	0.06
Density	1.0000	0.03	1.0001	0.00
% Uninsured	0.9940	0.27	0.9918	0.37
% voted GOP	1.0078	0.00	1.0121	0.00
Constant	0.0028	0.00	0.0000	0.00
Number of Obs.	778		778	
Wald χ^2	391.8		170.7	

Exponentiated coefficients mean coefficients are reported as Risk Ratios

Table 5: Mixed-effect negative binomial regression: model only includes counties with income inequality increase (2010-2019)

	(1)		(2)	
Variables	**Case Rate**	P-value	**Death Rate**	P-Value
Gini 2019	1.0166	0.00	1.0229	0.00
Income per capita	1.0009	0.42	1.0008	0.64
Ln (Cases)	NA	-	1.0215	0.42
Male	1.0297	0.00	1.0150	0.14
% under 20	1.0129	0.06	1.0279	0.00
% over 65	0.9889	0.02	1.0386	0.00
% Black	1.0107	0.00	1.0243	0.00
% Native	1.0084	0.00	1.0216	0.00
% Hispanic	1.0142	0.00	1.0174	0.00
# Days First	1.0011	0.07	1.0005	0.56
Density	1.0000	0.00	1.0001	0.00
% Uninsured	0.9898	0.00	0.9934	0.06
% voted GOP	1.0105	0.00	1.0188	0.00
Constant	0.0015	0.00	0.0000	0.00
Number of Obs.	2 311		2 311	
Wald χ^2	661.9		715.4	

Exponentiated coefficients mean coefficients are reported as Risk Ratios

Table 6: Variable Definitions and Sources

Variable Name	Definition	Source
Cases (or Incidence)	Number of confirmed cases of COVID-19 per county between January 22 and December 14, 2020	CDC/USAFacts.org U.S Census Bureau
Deaths (or Mortality)	Number of confirmed deaths associated to COVID-19 per county between January 22 and December 14, 2020	CDC/USAFacts.org U.S Census Bureau
Gini	Gini Index of income inequality by county for 2010 and 2019	U.S Census Bureau (American Community Survey 2019 and 2010, tables B19083)
Income per capita	Personal income divided by population	Bureau of Economic Analysis
Male	Percent male population, 2019	U.S. Census Bureau
% under 20	Percent population under age 20, 2019	U.S. Census Bureau
% over 65	Percent population age 65 & over, 2019	U.S. Census Bureau
% Black	Percent Black population, 2019	U.S. Census Bureau
% Native	Percent Native population, 2019	U.S. Census Bureau
% Hispanic	Percent non-white Hispanic population, 2019	U.S. Census Bureau
# Days since first	Number of days since 1st case confirmed in a county	CDC/USAFacts.org
Density	Population density per km^2, 2019	U.S Census Bureau
% Uninsured	Percent of population under 65% without medical insurance	HealthData.gov
% voted GOP	Percentage of votes that went for the GOP in the 2016 elections, by county	GitHub

X. Robustness checks

Table 7: Replicating Table 3 with Poverty Rate Instead of Income per Capita

	(1)		(2)	
Variables	**Case Rate**	P-value	**Death Rate**	P-Value
Gini 2019	1.0040	0.28	1.0015	0.72
Gini 2010	1.0139	0.00	1.0119	0.01
Poverty Rate	1.0044	0.05	1.0103	0.05
Ln (Cases)	NA	-	1.0132	0.60
Male	1.0300	0.00	1.0150	0.09
% under 20	1.0137	0.04	1.0368	0.00
% over 65	0.9894	0.01	1.0434	0.00
% Black	1.0086	0.00	1.0211	0.00
% Native	1.0074	0.00	1.0174	0.00
% Hispanic	1.0143	0.00	1.0177	0.00
# Days First	1.0012	0.04	1.0015	0.11
Density	1.0000	0.00	1.0001	0.00
% Uninsured	0.9890	0.00	0.9921	0.01
% voted GOP	1.0094	0.00	1.0158	0.00
Constant	0.0014	0.00	0.0000	0.00
Number of Obs.	3 088		3 088	
Wald χ^2	745		1172	

Table 8: Replicating Table 3 Excluding the Top and Bottom 150 Counties in Terms of Death Rate

	(1)		(2)	
Variables	**Case Rate**	P-value	**Death Rate**	P-Value
Gini 2019	1.0048	0.22	1.0070	0.16
Gini 2010	1.0167	0.00	1.0123	0.01
Income per capita	1.0003	0.75	0.9988	0.42
Ln (Cases)	NA	-	1.0149	0.47
Male	1.0325	0.00	1.0026	0.73
% under 20	1.0102	0.03	1.0255	0.00
% over 65	0.9876	0.00	1.0343	0.00
% Black	1.0087	0.00	1.0207	0.00
% Native	1.0059	0.00	1.0155	0.00
% Hispanic	1.0144	0.00	1.0162	0.00
# Days First	1.0001	0.78	1.0006	0.42
Density	1.0000	0.33	1.0001	0.01
% Uninsured	0.9910	0.00	0.9964	0.23
% voted GOP	1.0093	0.00	1.0158	0.00
Constant	0.0017	0.00	0.0000	0.00
Number of Obs.	2 788		2 788	
Wald χ^2	641.9		763.4	

Table 9: Replicating Table 3 Excluding Counties in the Top and Bottom Quintiles of Gini Coefficient

	(1)		(2)	
Variables	**Case Rate**	P-value	**Death Rate**	P-Value
Gini 2019	1.0062	0.29	1.0151	0.14
Gini 2010	1.0152	0.00	1.0176	0.02
Income per capita	0.9990	0.38	0.9993	0.76
Ln (Cases)	NA	-	0.9936	0.82
Male	1.0344	0.00	1.0110	0.31
% under 20	1.0136	0.02	1.0212	0.07
% over 65	0.9873	0.00	1.0287	0.00
% Black	1.0091	0.00	1.0231	0.00
% Native	1.0083	0.00	1.0181	0.00
% Hispanic	1.0160	0.00	1.0204	0.00
# Days First	1.0005	0.40	1.0002	0.87
Density	1.0001	0.00	1.0003	0.00
% Uninsured	0.9855	0.00	0.9969	0.58
% voted GOP	1.0103	0.00	1.0170	0.00
Constant	0.0014	0.00	0.0000	0.00
Number of Obs.	1 856		1 856	
Wald χ^2	914		510.3	