Martin Luther University Halle-Wittenberg
School of Economics and Business
Chair of Econometrics

Bachelor's Thesis

Spatial heterogeneity in Covid-19 prevalence in the US. Are Democratic regions different from Republican regions?

Marges Pinderi

Submitted on: 06.10.2020

Abstract

Covid-19 became the biggest pandemic of the last decades and has had a large impact on the public health and economy of almost every country. In this paper, I focus on the way politics has driven the outbreak in the United States of America. I analyze the link between the political majority of counties in the USA and the prevalence of the Covid-19 confirmed cases and deaths in the USA. Using data on a county-level, I find that counties with a Democratic majority have a higher death rate than the Republican ones taking into account relevant factors but I see no significant difference in the spread of the confirmed cases when we control the other variables in the model. However, in the counties whose first outbreaks took place in the late summer, a connection between a Republican majority and the high number of confirmed cases can be observed. I find that counties with Republican governors have higher numbers in the confirmed cases and the deaths, compared to counties governed by Democratic governors.

Keywords: Covid-19, Party, politics, affiliation, spread, USA

Contents

List of Figures

List of Tables

1 Introduction

Covid-19 has spread at a very fast rate during 2020 and became a tremendous global health problem. It has caused more than 1 000 000 confirmed deaths up until September 2020, only months after the disease was discovered and started to emerge (JHU, 2020). The United States has been hit harder than any other country taking into account the absolute numbers with more than 200 000 victims. In this paper, I analyze the scale of impact the virus has had on geographical regions with different political stances in the country. I take into account data on county-level and analyze the relationship on both the prevalence of the disease in these communities and the number of people who have lost their lives because of it.

This paper investigates this possible link with the help of econometric models. I will account for other factors that might affect the spread of disease like the time period since the outbreak first began, whether we have to do with a metropolitan area or not and the per capita annual income of the counties. The Covid-19 related data has been obtained from the John Hopkins University's public dataset. Other data concerning income, population, and residual density were taken from US governmental websites and have been gathered via censuses and official estimations. For the information related to the the counties' political affiliation, I am using data from the 2016 presidential elections made public from the MIT election lab.

Despite the fact that Covid-19 has emerged less than one year before this paper was written, a substantial amount of research is already available on the disease and SARS-CoV-2, the virus which causes the illness. Li et al. (2020), Zhang et al. (2020) confirm the high virulence of the disease and conclude that mandating face coverings can be a very decisive factor in the way the pandemic takes shape. So far, the Republican governed states have been more hesitant to put these kinds of measures in place or communicate guidelines to the public (Lyu & Wehby, 2020). Nevertheless, to my knowledge there is no paper that analyzes the differences between regions on the basis of their vote casting. A greater amount of literature on the general health differences between Republican and Democratic voters is available and will be discussed in the next section.

The closest conclusion of voters in this paper is based on the way they voted on a county level in the 2016 election, nevertheless the state voting patterns might also give the regions a party identity. For this reason this paper tries to capture the differences concerning both of these political factors. For the first style of political identification, the party identity of the county is determined on the basis of the last presidential election in 2016 majority vote winner. The second style will focus on the party that the state governor of each county has. In the United States, it is governors, who are responsible for the implementation of certain measures which might have an effect on the containment of the disease. Therefore, this style of identification will also capture the county level differences based on the state wide party identification.

Concerning the political majorities of each county based on the data from the 2016 elections, I find that the Republican majority counties have a significantly lower death rate per confirmed cases than Democratic ones but this significance does not hold once we control for the initial outbreak in the country. On the other hand, there is no significant difference between the counties with a Democratic majority and those with a Republican majority concerning the prevalence of the confirmed cases. However when we investigate for the recent outbreaks, the Republican majority counties account for a greater number of confirmed cases. Registered deaths were statistically significantly greater for Democratic counties early in the pandemic but this relationship does not hold for the counties, which had their outbreaks later in the summer. Concerning the state governor's association with the outbreak situation, a Republican governor is significantly linked to both greater number of confirmed cases and greater number of deaths.

The structure of my paper is as follows: Firstly, in the second section, I give a brief background on the disease since the beginning of the outbreak and introduce the literature on the topic. In this chapter, I include the academic knowledge regarding the relationship between the political affiliation of the individuals with their behavior and health, which might be reflected in the spread of Covid-19 and its mortality. Later in the third chapter, I will present the data that I use in my analysis along with some descriptive statistics. The fourth chapter introduces the study framework and contains my econometric models. In the fifth chapter I present the results and interpret them for the reader. In the last part I give my conclusions of the analysis.

2 Background and Literature review

2.1 Background of the virus and its spread in the United States

Covid-19, the novel coronavirus disease caused from the SARS-CoV-2 virus, was officially declared a pandemic from the World Health Organization on March 11, 2020. The first cases were recorded late in 2019 in Hubei province, China. The cases were initially connected to a meat market, therefore it was firstly reported from the local authorities for a non human to human transmission (WHO, 2020a). Later on January, as the numbers of cases started to increase in China, more information was reported from the authorities. It was confirmed that the new coronavirus disease could spread between humans and on 30th of January it was declared as a Public Health Emergency of International Concern. This declaration is used in rare occasions and raised the alarm for countries across the world. This, along with an increase in the number of cases in mainland China and abroad would make governments worldwide start taking precautions. These precautions included banning flights from China and ramping up their testing capacities. Despite the alert, many countries started to experience uncontrolled outbreaks since February, notably Italy in Europe and Iran in the Middle East. On the same time, a large outbreak in South Korea was gradually controlled (Kim et al., 2020) and so did the one in China eventually through the drastic measures which started at the end of January (AlTakarli, 2020) despite already claiming thousands of lives.

Later in March, Covid-19 continued its spread globally and was officially declared to be a pandemic by the World Health Organization. It was during this month that the disease increased its prevalence aggressively in almost every major country. Despite confirming its first case since mid January (Holshue et al., 2020), The United States had managed to avoid a large outbreak until late March, when the numbers of new cases and confirmed deaths started to increase at a very fast rate (JHU, 2020). Figure 1 shows the accumulated number of Covid-19 cases since the first cases in the country. After a stable first month, I notice a sudden jump which continues during the months after. This increase, however, did not occur homogeneously geographically.

The New York state was the country's epicenter of the outbreak at the beginning of the epidemic. New York City, being an international hub and a congested urban area (Stier et al., 2020) contained the perfect environment for the virus to quickly spread throughout its residents. The state was struck very hard, making its authorities impose a strict lock-down and increase other capacities like testing and later contact tracing which resulted in an impressive slowdown of the infections during the next months. As of September 2020, the state has one of the lowest infection and death rates in the country despite still having the largest accumulative death numbers as shown in Figure 2. Unfortunately, the virus was able to spread to other regions and states in the rest of the country. As of the moment this paper is written, September 2020 the daily confirmed cases are soaring and most of the states have already been hit very hard as Figure 3 shows. During this month, states like Texas, California, and Florida contain the most severely hit regions as can be seen in Figure 4. There are many conversations being conducted on the data around political and demographic aspects along with the governing effects during this pandemic, this study is mostly concerned on the difference between the regions with a Republican majority compared to the ones with a Democratic one.

2.2 Literature review

There are currently different studies on the geographical and demographic parts of the country's outbreak. Sun et al. (2020) add spatial econometric models along the aspatial ones, with the former ones being better at predicting, Karaye and Horney (2020) find that regions in worse social conditions fare worse with the spread of the disease, while Moore et al. (2020) show the racial disparities in the spread of Covid-19. Grossman et al. (2020) analyze the governors' early actions to the pandemic and their constituents' response to their calls, finding both that the Democratic governors acted earlier and Democratic leaning counties responded stronger to their guidelines.

Duckitt and Sibley (2007) define social conservatives as people who might perceive the world as unstable or dangerous. During a global pandemic, when a greater amount of uncertainty and risk can be present, these individuals might tend to be skeptical towards national and international institutions which are responsible and able to inform the public. Instead, they often rely on their own judgement, which might contradict the one coming from the institutions. The current Republican administration, whose supporter base is made of conservative voters, decided to pull the country out from the World Health Organization (Rauhala et al., 2020), an institution that represents the highest international authority in the common institutional image during this pandemic. This might have added more to the skepticism these voters have toward health authorities. Furthermore, both previous studies (Baumgaertner et al., 2018) and recent polls (Sanders, 2020; Sparks & Langer, 2020) support this behaviour from the members of this party. They show that voters of the Republican party in United States are more skeptical about official information and more prone in believing debunked theories. Calvillo et al. (2020) find these differences in perceptions to exist for the current coronavirus situation and suggest that they might come from the media (McCloskey et al., 2020) and the political leaders. This partisanship in beliefs might affect the behavior of the residents and the decisions of their authorities. This paper is at my knowledge unique at discussing the spatial heterogeneity from a politics perspective in the United States. It adds to the studies which report a difference in the perceptions of the voters from the two main parties have about the disease.

Chinazzi et al. (2020) conclude that travel bans are significant only at the beginning of outbreaks in certain areas and later they only have a modest effect on the reduction of the number of infections if not followed by other measures. Therefore, my analysis takes into account only domestic factors after the epidemic started to take shape in the country. Hence, community spread is assumed to be the main driver of the disease prevalence, with measures such as social distancing and face masks (Zhang et al., 2020) being the most effective tools in controlling the spread of the disease. Lyu and Wehby (2020) support the findings that face covering mandates have slowed community transmission and reduced thousands possible cases. A dense research is still ongoing in hopes of understanding the virus and its spreading patterns better. One of the biggest challenges is the variance of symptoms of Covid-19, lately it is found that the virulence in asymptomatic cases is considerably high (Tan et al., 2020). This finding makes the containment of the disease even more challenging and puts weight on implementing the aforementioned tools from the authorities and on the changes to the daily behavior from the broad public. People need to adjust their behavior individually, despite not showing any symptoms. This paper takes a look on whether the politics' differences between regions are linked to a difference in the prevalence of the virus. My initial assumption from the literature is that Republicans will tend to spread the virus more into their communities, hence that regions with a majority of Republican voters will have a higher prevalence than Democrat majority regions. I have similar assumptions for counties with Republican state leadership to have larger prevalence of the virus than counties which have Democrat state leadership.

Covid-19 is more dangerous to people who are already in a poor health condition compared to otherwise healthier individuals (Fang et al., 2020; Kassir, 2020; Xia et al., 2020) . From a health difference perspective, both aggregate and micro studies have found that Republicans report better health than democrats (Subramanian & Perkins, 2010). The differences in health reported by individuals would get larger between stronger supporters of the parties. The same study finds that not only do democrats report worse health but they also tend to smoke more. Individuals with bad conditions and smokers are associated with higher severity and a greater number of deaths from the disease (WHO, 2020b). This suggests that republicans might tend to get less sick and die in fewer numbers from Covid-19 compared to Democrats. On the other hand, during the past election cycles, there has a been a shift of voters who report poorer health towards the Republican party (Wasfy et al., 2020). Assumptions on mortality from Covid-19 based on voting patterns will be difficult to make before the analysis due to these conflicting reporting in the literature.

3 Data

3.1 Sources and variables

This study is based on a county-level aggregated data analysis. The data is retrieved from multiple sources and is later compiled and merged with the help of statistical software. The John Hopkins University has made available an online database which updates daily and is the source for the Data related to Covid-19 in my study. It is a large database which produces each day new reports with worldwide data on the disease based on information coming from public institutions, hospitals and the web. From there I directly extract two variables and calculate two others. All of this four variables will be listed and explained in the next paragraph. MIT Election Data and Science Lab is another source of this paper. It is a lab which collects, analyzes and shares data and research in the political science sphere of the United States of America. The Census Bureau of the United States is the source of my population and statistical area data. Finally, the per capita income data for each county comes from the Bureau of Economic Analysis.

As of 2020, there are 3143 counties in the USA. For the second part of the description analysis in subsection 3.2 I will use almost all of them to give a first look into the question of this paper. However, after compiling all the variables together for the rest of the empirical analysis aided by the econometric models, I use 1571 of these counties. The detailed reasoning of this will be presented at the end subsection 4.1. Figure 5 shows my final data frame. This is how it looks like after it is compiled together and sorted by descending number of confirmed cases as of September 23, 2020. This sample consists of these variables:

- **County** : The name of each county in the data-frame. Due to the fact that there can be many counties with the same name, they can be used as an ID only combined with the name of the state they belong to.

- **State** : The name of the state the observation is located. There are 50 states in the United States and each is made of between 3 (Delaware) and 254 (Texas) counties. Together with the counties they serve for identification purposes in the database.

- **Confirmed** : The accumulated number of confirmed cases for each county. Each person tested, who positive in the county since the start of the pandemic is included. The actual number of infected people is believed to be higher, especially in the beginning of the pandemic due to widespread infections, asymptomatic cases and limited testing (Havers et al., 2020). This variable is extracted from the JHU online databases.

- **Deaths** : The people registered as deceased due to Covid-19. Excessive deaths suggest this might be as well underestimated (Lu, 2020). As of September 2020, Kings County in New York, also known as Brooklyn, has more than 7300 registered deaths, having the largest number in the country. This variable is extracted from the JHU online databases.

- **Repwin** : A dummy variable which takes the value 1 if the Republican presidential candidate won more votes than their Democrat opponent in the 2016 elections or value 0 otherwise. Republicans have won 2597 counties out of a total of 3083 in the 2016 elections and 1241 out of the 1571 observations of my final database. Own calculation from the data provided by MIT Election Data and Science Lab's 2016 presidential election result database.

- **Pop** : The population for each county. I have used it in my analysis in order to calculate the **Cases_10ths** and **Deaths_10ths** variables for each county. Estimated from The Census Bureau of the United States based on their 2010 census. Their next census is being conducted in 2020 but has not finished yet.

- **daysfrom100** : The number of days since the 100^{th} confirmed case in each county. I have calculated this value with the help of a time series data set also provided by the John Hopkins University and which gives the accumulated number of confirmed cases for each day since the start of the pandemic. Kings County in the Washington state is the first one to have reached the 100^{th} case and therefore takes the largest value.

- **Metropol_Area**: A dummy variable which takes the value 1 if the county is part of one Metropolitan Statistical Area. Each Metropolitan Statistical Area must have at least one urban living area with more than 50 000 population (U.S Census Bureau, 2015). This variable is extracted from the US Census Bureau.

- **Income**: Per Capita Personal Income of each county. It is made of Employment earnings, transfer payments (government payments made to individuals), and investment income (dividends, interest, and rent (Kinghorn, 2007). Extracted from the Bureau of the Economic Analysis.

- **cases_10ths**: Confirmed Covid-19 cases per 10 000 residents.

- **deaths_10ths**: Registered Covid-19 deaths per 10 000 residents.

- **rep_gouv**: A dummy variable which takes value 1 if the **State** has a Republican governor and 0 if it is governed by a Democrat.

- **death_rate**: represents the ratio of people who died after testing positive $\frac{\textbf{Deaths}}{\textbf{Confirmed}}$.

3.2 Descriptive statistics of the data

Table 1 shows the descriptive statistics of the counties used for the final analysis lastly updated on September 23, 2020. It is a data set comprised of 1571 counties. It is made of only counties which already have experienced at least 80 confirmed cases. The minimum number of confirmed cases is 81 and is recorded in the Lewis County, Missouri which has a population of 9776 and the maximum number is 263 333 in Los Angeles, California. As for the number of lives lost, New York City's counties have the highest tolls, with the Kings County which is also known as the Brooklyn borough having recorded 7319 deaths and with over 20 thousand deaths in the entire city. King County in Washington has reached the 100^{th} case since the 2^{nd} of March which corresponds to 205 days before the data was updated. 59.3% of the counties presented here belong in states where Republican governors lead and 79% of those have cast a majority vote for the Republican presidential candidate in the last elections. Orleans County in New York has the largest death rate in the data set with a 54 deceased out of 325 confirmed cases, or a 16.6% rate. Chattahoochee County in Georgia has the highest rate of infections at 14.92%, with 1 628 of its 10 907 residents having tested positive. Another county in Georgia, Hancock has 497 deaths per 100 000 residents, the highest death prevalence in a county of this data frame. In this county 42 people have died so far from Covid out of a population of 8457 people.

When looking at the data from the entire country(voting data could be merged with Covid-19 data for 2989 counties)here are some interesting results:

- Total number of deaths: 192 610 people
- Total number of confirmed cases: 6 619 414
- Total number of deaths in Republican counties: 60 217
- Total number of confirmed cases in Republican counties: 3 889 746

We notice Republican majority counties to have more than half of the total confirmed cases but less than a third of the total deaths. Table 2 gives a better insight of the variables that we are interested. Here, I have tabulated three important averages which are not the same between the regions with different party majorities. We can see that all of the 3 variables; cases per 10 thousands residents, deaths per 10 thousands residents and the death rate are higher in the counties with a majority of Democrats. This differences however varies in the scale. Cases per 10 thousands residents are somehow close, with 217 confirmed cases per 10 thousand residents in the blue counties and 185.68 in the Republican counties. This might be attributed to the late surge in the cases in the Republican states. However, Democrats have double the rate of deaths in their communities and a 50% higher death rate. The contrast between the differences between the mean confirmed number of cases and the mean number of deaths might come for two reasons. First, The Democratic regions in the Washington state and along the east coast in states like New York and New Jersey were hit the first in March and April when the testing capacity was not sufficient. The percentage of positive tests during these months averages near 20% and dropped later around 5% when most of the Republican regions got hit (JHU, 2020). Due to this, undetected cases might have affected disproportionately the Democratic counties. The other reason might be because of differences in the actual mortality between the regions. Moore et al. (2020) find that there exists a mortality disparity in the national level and the literature suggested that Republicans had a better health than Democrats. However, the last argument might not hold with the shifting voting patterns (Wasfy et al., 2020).

4 Estimation strategy and the models

4.1 The strategy and assumptions

This study uses econometric models to investigate the differences in the prevalence of Covid-19 between counties with a majority of Democratic candidate voters and the counties with a Republican voters majority. This analysis is separated in two parts. The first part will try to capture the link between majority of the votes in counties and the prevalence of the disease. This will be made focusing on the **repwin** variable of the data frame. The second part will try to capture the link between the political party that the governor of each county affiliates to and the prevalence of the disease in the county. The prevalence of the disease will be measured for both, the number of confirmed cases and the number of deaths. These two variables will be adjusted according to the population of their counties.

The dependent variables for the upcoming models will be:

- **the number of cases**
- **the number of deaths**
- **the number of cases per 10 thousands residents**
- **the number of deaths per 10 thousands residents**

The last two variables are used to take into account the population each county has and therefore account for the relative burden the disease has on the communities.

From the independent variables, I control for:

- **daysfrom100**: I control for this variable because counties which have had their first cases earlier in the outbreak have a higher accumulated number of cases.
- **Income per capita**: (Raifman & Raifman, 2020) find that lower income people are more at risk from Covid-19. One of the reasons might be because many of these people tend to work on activities classified as essential where you need to be present despite the scale of the outbreak in the community, often rely on public transport. Another reason might be because people who live on lower income have worse medical care access.

- **Metropol_area**: Dense urban areas tend to contribute to a faster spread of contagious disease and tend to have higher mobility rates. They also tend to have higher ethnic minority rates which also have been infected at a higher rate (Moore et al., 2020; Raifman & Raifman, 2020).

For the purpose of this analysis, I use only counties that already have at least 80 confirmed cases in my data. I do this for two main reasons; firstly, to better observe regions which already had an outbreak, where the behaviour of their residents would make an actual difference. The counties, which were yet to be introduced to the virus from outside, would have no infections driven from the local behaviour of its residents. Secondly, these counties represent nearly 18 percent of all the observations in the sample but only 1.1 percent of the entire population of the United States, therefore they might put more weight in the model compared to the population that they actually represent.

I adjust for three categories depending on when they had their outbreaks. The first and default is since the beginning of the pandemic. The second is for counties which have had their 100^{th} case after March 17. In this way I exclude the countries which had the most infections at the start of the Pandemic but later did better. The third is for counties which had their 100^{th} case after July 5. This category is for counties which had no outbreaks until later in the summer. For this categories of counties, there was much more information available on how to protect from the virus and use the tools that I have previously mentioned. The extra set of information is available for both the private citiziens and the authorities.

4.2 County majority vote winner

I will be using multiple regression linear models for this analysis. Model 1 is a starting equation and gives a general view on the relationships these variables will have with the prevalence of the cases of the new coronavirus disease in the counties around the country. I construct Model 2 with a similar method as Model 1, but with raw number of deaths as the dependent variable.

$$\begin{aligned} ln(Cases) = &\beta_0 + \beta_1 DaysFrom100thCase + \delta_0 RepWin + \beta_2 ln(Population) \\ &+ \delta_1 MetropolArea + U \end{aligned} \tag{1}$$

$$
\begin{aligned}
ln(Deaths) = &\beta_0 + \beta_1 DaysFrom100thCase + \delta_0 RepWin + \beta_2 ln(Population) \\
&+ \delta_1 MetropolArea + U
\end{aligned} \tag{2}
$$

The population effect however might be better captured if we use cases and deaths per 10 thousands residents as dependent variables. Therefore I construct models 3 and 4. Now I will also also use the per capita income as an explanatory variable.

$$
\begin{aligned}
\ln(casesper10thousands) = &\beta_0 + \beta_1 DaysFrom100thCase + \delta_0 RepWin + \\
&\beta_2 \ln(Income) + \delta_1 MetropolArea + U
\end{aligned} \tag{3}
$$

$$
\begin{aligned}
deathsper10thousands = &\beta_0 + \beta_1 DaysFrom100thCase + \delta_0 RepWin \\
&+ \beta_2 \ln(Income) + \delta_1 MetropolArea + U
\end{aligned} \tag{4}
$$

To check the rate of people who have died after contracting the virus I construct model 5. This can be helpful to draw a relation from a health perspective and compare it to the literature.

$$
\begin{aligned}
DeathRate = &\beta_0 + \beta_1 DaysFrom100thCase + \delta_0 RepWin \\
&+ \beta_2 \ln(Income) + \delta_1 MetropolArea + U
\end{aligned} \tag{5}
$$

4.3 State governor vote winner

To capture the differences between the counties with a Democratic governor and a Republican one, I construct similar models. Models 6 and 7 are constructed to check whether a difference in these two types of counties exist. These two models will help us see whether there exists a relationship between the state leadership of a county and its Covid-19 prevalence. The literature suggests we might observe one (Baccini & Brodeur, 2020).

$$
\begin{aligned}
\ln(casesper10thousands) = &\beta_0 + \beta_1 DaysFrom200thCase + \delta_0 RepGov + \\
&\beta_2 \ln(Income) + \delta_1 MetropolArea + U
\end{aligned} \tag{6}
$$

$$(deathsper10thousands) = \beta_0 + \beta_1 DaysFrom200thCase + \delta_0 RepGov + \beta_2 \ln(Income) + \delta_1 MetropolArea + U \tag{7}$$

Lastly, I create Model 8 to check whether there is a link between the death rate a county has and the party of its governor.

$$DeathRate = \beta_0 + \beta_1 DaysFrom100thCase + \delta_0 RepGov + \beta_2 \ln(Income) + \delta_1 MetropolArea + U \tag{8}$$

5 Results

5.1 Results on the county level vote winner

Table 3 shows the coefficients when regressing the Model 1 and Model 2. There we observe a negative relation between having a Republican majority and the absolute number of confirmed cases and deaths for the counties in the data set. A Republican county, while controlling for the other variables in the model has 6.7% less confirmed cases and 29.4% less registered deaths than a Democratic county. However, I observe a statistical significance only for the number of deaths and not for the number of confirmed cases. The absence of a clear statistical significance might be attributed to the conflicted effect that the early wide-spreads in the democratic counties had compared to what the literature suggests. Literature hinted that Republicans would use the tools which slow the spread of the disease less than the Democrats.

Models 3 and 4, which use the cases per 10 thousand residents and deaths per 10 thousand residents as dependent variables and add Income per capita as a control variable, might give a better representation of these relationships. Table 4 gives the coefficients for Model 3 in three subsets. Column (1) takes into account every county in the data frame, column (2) only counties which had their 100^{th} confirmed case less than 190 days before and column (3) only those counties with their 100^{th} confirmed case less than 80 days before. I add those 2 other regressions for two reasons. First, to take into account for learning new information from science and authorities. As the pandemic unfolded more information became available and people learned that tools like masks were effective, this was not a given or encouraged at the start of the pandemic (Jingnan et al., 2020). Second, some states were badly affected in the beginning in urban areas when testing and personal protective equipment were scarce. Table 4 (1) shows that while taking into account every county, Republican majority counties are linked to a slightly and insignificantly lower number of confirmed cases in their communities compared to the Democratic majority ones. However, for counties which had their 100^{th} less than 190 days before (March 17,2020 or later), this relationship is reversed. Republican majority counties are linked to a 4.3% higher number of confirmed cases per 10 thousand residents. This is not statistically significant. Finally, when considering the data for counties which had their

100^{th} case less than 80 days before (July 5, 2020 or later) we observe a positive statistically significant relationship. The literature suggested that Republicans might tend to not adhere to behavioral adjustments and this is supported from counties which had their outbreaks later but not when taking every county on consideration. Income has a negative relationship with the number of confirmed cases for each group of the counties analyzed. An increase in 1 percent of the income in a given county while controlling for the other variables consists to 0.8 percent statistical significant less cases per 10 thousand residents. Similar relationships also exist for counties which had their 100^{th} case later than March, 17 or July, 5. Surprisingly to my initial assumption, metropolitan areas have a negative relationship with the number of cases per 10 thousand residents. However, Miller et al. (2020) find that rural counties have experienced harsh outbreaks in the United States and this might be a reason for this relationship.

I use the same type of analysis for Model 4. Table 5 gives the results when this model is regressed. I observe a negative relationship between Republican counties and deaths per 10 thousand residents. When considering every county in the analysis, a Republican majority county has 2.283 deaths less per 10 thousand residents than a Democratic one. For counties which have had their 100^{th} case after March, 17, Republican majority counties have 1.533 less cases per 10 thousand residents than Democratic majority counties. Both of these coefficients are statistically significant. However, for counties which had their 100^{th} case after July, 5 this kind of relationship does not hold anymore. This, once more supports the fact that Democratic regions were hit harder in the beginning of the pandemic and later started to recover. Here, an one percent increase in the income per capita of a county consists in a 4.474 decrease in the number of deaths per 10 thousand residents. This is statistically significant and observed similarly when regressed for counties which started to have their outbreaks later. The literature in income disparity and covid-19 prevalence supports this relationship (Raifman & Raifman, 2020).

The Model 5's results in Table 5 help us to display the differences we saw between the death rates of the Republican majority counties and the Democrat majority ones. The Republican majority counties have a 44.7% statistically significant lower death rate than the Democrat majority counties for all 1571 observations. However, when regressing only on the 1489 counties which had their 100^{th} case after March, 17 this relationship losses its statistical significance and halves in quantity. This suggests that this sample is however significant at 5% for the income variable. This might be explained due to the fact that the first hard hit counties were mostly located in Washington, New York, New Jersey and Massachusetts, states which have a relatively high income and vote for the Democratic county. In overall, the death rate seems to be connected with the political majority a county has, but this seems to not be explained by health reasons discussed in the literature since it does not hold for the counties which were impacted later by the coronavirus but rather from the harsh outbreak in the beginning with many undetected cases and limited capacities in treatment.

5.2 Results for the governor vote winner

This part of the study examines the relationship between the partisanship in the state leadership and the prevalence of Covid-19 in the county level. To check this association, I analyze the coefficients of Models 6 and 7 given at table 7. Here, I notice a stark relationship between the dependent variables and the governor's party of a county. Controlling for the other variables, a county with a Republican governor has 46.4% more cases per 10 thousand residents than a county with a Democratic governor. Conversely, a county with a Republican governor has almost 1.2 more deaths per 10 thousand residents than a county with a Democrat governor. These coefficients are both statistically significant and supported from the literature that partisanship influences the political leaders. Once again we notice a negative relationship of income and the prevalence of Covid-19. In this model an increase of 1 percent of income shows a decrease of 0.67 % in cases per 10 thousand residents and 0.3287 less deaths per 10 thousand residents. Both of these coefficients are significant.

Table 8 shows no statistically significant relationship between the party a governor has and the proportion of people who die after confirmed positive.

6 Conclusion

A lack of in epidemiological background or topic specific literature has made this study challenging and interesting at the same time, nevertheless the models in this paper could capture a disparity in the politics of this pandemic. Democratic regions were hit disproportionately at the beginning from imported cases and this added some complexity to the model. Nevertheless, I believe that the approach with examining sub-samples based to when the counties had their first cases helped account for it to some extent. This study added to the literature which focuses on the social disparities during this pandemic and to the differences between the Democrats and Republicans in their well-being and behaviour adjustment. It also added to the leadership differences during this pandemic.

The statistical differences between the death rates are explained from the initial outbreak in the Democratic regions at a time when testing and treatment were limited. This means that infected people are not dying in higher rates in Democratic regions when controlling for other factors such as income. This, however needs to be proven from a study with a better epidemiological expertise. To find whether Democrats are dying at a different rate than Republicans, a micro-level data analysis could be more helpful. This, because determining relationships for individual partisanship based on the aggregated data can be misleading, especially when some regions are more partisan than others.

Analogously, the statistical differences in the cases per capita in communities between regions with different political majorities need to be further analyzed to better understand more details. Both, the county voting patterns and the state voting patterns seem to be related to differences in the cases prevalence in the counties. Republican leadership in the state level, however is related to a statistically significant higher number of deaths in communities. Spatial factors which I did not include in this analysis might play a role for this and might be a further study interest.

The analysis on the connection between the governor of a county and its Covid-19 prevalence might shed a light to a pattern of state governing which is working better than the other. Acccording to my findings, the Democrat governors are doing a better job at controlling the spread. This might help both the rest of the Republican states and the federal authorities for a better response in the United States. A larger-scale study for the international spread of the virus can be made and we might observe how political orientations are affecting the spread and the burden of the disease.

Lastly, I want to conclude that I hope these findings helped the readers think about the relationship of political partisanship to sensitive issues like the course of the spread of a pandemic. Hopefully, the conversation will extend beyond in order for us to understand this situation better and maybe manage similar situations better in the future.

References

AlTakarli, N. S. (2020). China's response to the covid-19 outbreak: A model for epidemic preparedness and management. *Dubai Medical Journal*, *3*(2), 44–49.

Baccini, L., & Brodeur, A. (2020). Explaining governors' response to the covid-19 pandemic in the united states. *IZA Discussion Paper*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3579229

Baumgaertner, B., Carlisle, J. E., & Justwan, F. (2018). The influence of political ideology and trust on willingness to vaccinate. *PloS one*, *13*(1), e0191728.

Bureau of Economic Analysis. (2018). Personal Income by County, Metro, and Other Areas. BEU. Retrieved August 20, 2020, from https://www.bea.gov/data/income-saving/personal-income-county-metro-and-other-areas

Calvillo, D. P., Ross, B. J., Garcia, R. J., Smelter, T. J., & Rutchick, A. M. (2020). Political ideology predicts perceptions of the threat of covid-19 (and susceptibility to fake news about it). *Social Psychological and Personality Science*, 1948550620940539.

Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., Pastore y Piontti, A., Mu, K., Rossi, L., Sun, K., Viboud, C., Xiong, X., Yu, H., Halloran, M. E., Longini, I. M., & Vespignani, A. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (covid-19) outbreak. *Science*, *368*(6489), 395–400. https://doi.org/10.1126/science.aba9757

Duckitt, J., & Sibley, C. G. (2007). Right wing authoritarianism, social dominance orientation and the dimensions of generalized prejudice. *European Journal of Personality: Published for the European Association of Personality Psychology*, *21*(2), 113–130.

Fang, L., Karakiulakis, G., & Roth, M. (2020). Are patients with hypertension and diabetes mellitus at increased risk for covid-19 infection? *The Lancet. Respiratory Medicine*, *8*(4), e21.

Grossman, G., Kim, S., Rexer, J., & Thirumurthy, H. (2020). Political partisanship influences behavioral responses to governors' recommendations for covid-19 prevention in the united states. *Available at SSRN 3578695*.

Harvard. (2018). County Presidential Election Returns 2000-2016. Harvard Dataverse. https://doi.org/10.7910/DVN/VOQCHQ

Havers, F. P., Reed, C., Lim, T., Montgomery, J. M., Klena, J. D., Hall, A. J., Fry, A. M., Cannon, D. L., Chiang, C.-F., Gibbons, A., Et al. (2020). Seroprevalence of antibodies to sars-cov-2 in 10 sites in the united states, march 23-may 12, 2020. *JAMA Internal Medicine*.

Holshue, M. L., DeBolt, C., Lindquist, S., Lofy, K. H., Wiesman, J., Bruce, H., Spitters, C., Ericson, K., Wilkerson, S., Tural, A., Diaz, G., Cohn, A., Fox, L., Patel, A., Gerber, S. I., Kim, L., Tong, S., Lu, X., Lindstrom, S., … Pillai, S. K. (2020). First case of 2019 novel coronavirus in the united states [PMID: 32004427]. *New England Journal of Medicine*, *382*(10), 929–936. https://doi.org/10.1056/NEJMoa2001191

JHU. (2020). Covid-19 dashboard by the center for systems science and engineering (csse) at johns hopkins university. John Hopkins University. Retrieved September 23, 2020, from https://github.com/CSSEGISandData/COVID-19

Jingnan, H., Aubrey, A., & Wroth, C. (2020). Should we all be wearing masks in public? health experts revisit the question. National Public Radio. Retrieved September 25, 2020, from https://www.nytimes.com/interactive/2020/08/12/us/covid-deaths-us.html

Karaye, I. M., & Horney, J. A. (2020). The impact of social vulnerability on covid-19 in the us: An analysis of spatially varying relationships. *American journal of preventive medicine*, *59*(3), 317–325.

Kassir, R. (2020). Risk of covid-19 for patients with obesity. *Obesity Reviews*, *21*(6).

Kim, J.-H., Ah-Reum An, J., & Jackie Ho, S. (2020). Emerging COVID-19 success story: South Korea learned the lessons of MERS. Ariadne Labs. Retrieved August 15, 2020, from https://ourworldindata.org/covid-exemplar-south-korea#licence

Kinghorn, M. (2007). Key aspects of per capita personal income. *Indiana Business Review*, *82*(1), 6.

Li, Y., Zhang, R., Zhao, J., & Molina, M. J. (2020). Understanding transmission and intervention for the covid-19 pandemic in the united states. *Science of the Total Environment*, *748*, 141560.

Lu, D. (2020). The true coronavirus toll in the u.s. has already surpassed 200,000. The New York Times. Retrieved September 15, 2020, from https://www.nytimes.com/interactive/2020/08/12/us/covid-deaths-us.html

Lyu, W., & Wehby, G. L. (2020). Community use of face masks and covid-19: Evidence from a natural experiment of state mandates in the us: Study examines impact on covid-19 growth rates associated with state government mandates requiring face mask use in public. *Health Affairs*, *39*(8), 1419–1425.

McCloskey, B., Zumla, A., Ippolito, G., Blumberg, L., Arbon, P., Cicero, A., Endericks, T., Lim, P. L., & Borodina, M. (2020). Mass gathering events and reducing further global spread of covid-19: A political and public health dilemma. *The Lancet*, *395*(10230), 1096–1099.

Miller, I. F., Becker, A. D., Grenfell, B. T., & Metcalf, C. J. E. (2020). Disease and healthcare burden of covid-19 in the united states. *Nature Medicine*, *26*(8), 1212–1217.

Moore, J. T., Ricaldi, J. N., Rose, C. E., Fuld, J., Parise, M., Kang, G. J., Driscoll, A. K., Norris, T., Wilson, N., Rainisch, G., Et al. (2020). Disparities in incidence of covid-19 among underrepresented racial/ethnic groups in counties identified as hotspots during june 5–18, 2020—22 states, february–june 2020. *Morbidity and Mortality Weekly Report*, *69*(33), 1122.

Raifman, M. A., & Raifman, J. R. (2020). Disparities in the population at risk of severe illness from covid-19 by race/ethnicity and income. *American journal of preventive medicine*.

Rauhala, E., Demirjian, K., & Olorunnipa, T. (2020). Trump administration sends letter withdrawing u.s. from world health organization over coronavirus response. Retrieved September 5, 2020, from https://www.washingtonpost.com/world/trump-united-states-withdrawal-world-health-organization-coronavirus/2020/07/07/ae0a25e4-b550-11ea-9a1d-d3db1cbe07ce_story.html

Sanders, L. (2020). The difference between what republicans and democrats believe to be true about covid-19. Retrieved August 23, 2020, from https://today.yougov.com/topics/politics/articles-reports/2020/05/26/republicans-democrats-misinformation

Sparks, S., & Langer, G. (2020). Unlikely to be vaccinated against the coronavirus; republicans, conservatives especially: Poll. Retrieved August 23, 2020, from https://abcnews.go.com/Politics/27-vaccinated-coronavirus-republicans-conservatives-poll/story?id=70962377

Stier, A. J., Berman, M. G., & Bettencourt, L. M. A. (2020). Covid-19 attack rate increases with city size. *medRxiv*. https://doi.org/10.1101/2020.03.22.20041004

Subramanian, S., & Perkins, J. (2010). Are republicans healthier than democrats? *International journal of epidemiology*, *39 3*, 930–1.

Sun, F., Matthews, S. A., Yang, T.-C., & Hu, M.-H. (2020). A spatial analysis of the covid-19 period prevalence in us counties through june 28, 2020: Where geography matters? *Annals of epidemiology*.

Tan, J., Liu, S., Zhuang, L., Chen, L., Dong, M., Zhang, J., & Xin, Y. (2020). Transmission and clinical characteristics of asymptomatic patients with sars-cov-2 infection. *Future Virology*, (0).

U.S Census Bureau. (2015). Core based statistical areas (CBSAs), metropolitan divisions, and combined statistical areas (CSAs). Retrieved August 20, 2020, from https://www.census.gov/programs-surveys/metro-micro/about.html

U.S Census Bureau. (2018). Annual Resident Population Estimates. Retrieved August 20, 2020, from https://www2.census.gov/programs-surveys/popest/datasets/2010-2018/counties/totals/?C=N;O=D

Wasfy, J. H., Healy, E. W., Cui, J., & Stewart III, C. (2020). Relationship of public health with continued shifting of party voting in the united states. *Social Science & Medicine*, 112921.

WHO. (2020a). "coronavirus disease 2019 (covid-19) situation report – 94". World Health Organization. Retrieved August 28, 2020, from https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200423-sitrep-94-covid-19.pdf

WHO. (2020b). Smoking and covid-19: Scientific brief, 30 june 2020. World Health Organization. Retrieved August 20, 2020, from https://apps.who.int/iris/handle/10665/332895

Xia, Y., Jin, R., Zhao, J., Li, W., & Shen, H. (2020). Risk of covid-19 for cancer patients. *Lancet Oncol*, *21*(4).

Zhang, R., Li, Y., Zhang, A. L., Wang, Y., & Molina, M. J. (2020). Identifying airborne transmission as the dominant route for the spread of covid-19, *117*(26), 14857–14863. https://doi.org/10.1073/pnas.2009637117

Tables and Figures

Table 1: Descriptive statistics for the final group of counties data

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Confirmed	1,571	3,898.538	12,290.800	81	414	2,707.5	263,333
Deaths	1,571	115.050	448.313	0	6	60	7,319
repwin	1,571	0.790	0.407	0	1	1	1
pop	1,571	189,563.000	469,881.200	3,829	31,005	158,015	10,039,107
daysfrom100	1,571	123.396	52.022	0	82	171	205
Metropol_area	1,571	0.638	0.481	0	0	1	1
Income	1,571	45,662.410	14,238.870	26,195	37,931	49,623.5	251,728
cases_10ths	1,571	192.375	133.416	4.746	102.426	250.783	1,492.619
deaths_10ths	1,571	4.455	4.846	0.000	1.385	5.581	49.663
rep_gouv	1,571	0.593	0.492	0	0	1	1
death_rate	1,571	0.024	0.021	0.000	0.010	0.030	0.166

Source: Data from JHU, BEA, US Census Bureau, MIT Election Data Lab

Table 2: Differences between the counties

	repwin	cases_10ths	deaths_10ths	death_rate
1	0.00	217.53	6.87	0.03
2	1.00	185.68	3.81	0.02

Source: Data from JHU, MIT Election Data Lab

Table 3: Results. Models 1,2

	Dependent variable:	
	log(Confirmed)	log(Deaths)
	(1)	(2)
repwin	-0.067	-0.294^{***}
	(0.041)	(0.061)
daysfrom100	0.009^{***}	0.015^{***}
	(0.0004)	(0.001)
Metropol_area	-0.004	0.062
	(0.034)	(0.051)
log(pop)	0.712^{***}	0.591^{***}
	(0.019)	(0.029)
Constant	-2.010^{***}	-5.221^{***}
	(0.194)	(0.287)
Observations	1,571	1,510
R^2	0.824	0.735
Adjusted R^2	0.824	0.734
Residual Std. Error	0.597 (df = 1566)	0.871 (df = 1505)
F Statistic	$1{,}833.481^{***}$ (df = 4; 1566)	$1{,}043.069^{***}$ (df = 4; 1505)
Note:		*p<0.1; **p<0.05; ***p<0.01

Source: Data from JHU, BEA, US Census Bureau, MIT Election Data Lab

Table 4: Results. Model 3.

	Dependent variable:		
	log(cases_10ths)		
	(1)	(2)	(3)
days since 100^{th} case		<190	<80
repwin	-0.006	0.043	0.394^{***}
	(0.041)	(0.044)	(0.141)
daysfrom100	0.006^{***}	0.006^{***}	0.005^{***}
	(0.0003)	(0.0003)	(0.001)
Metropol_area	-0.070^{**}	-0.069^{**}	0.138^{**}
	(0.034)	(0.034)	(0.065)
log(Income)	-0.881^{***}	-1.024^{***}	-0.782^{***}
	(0.069)	(0.075)	(0.178)
Constant	13.812^{***}	15.298^{***}	12.262^{***}
	(0.733)	(0.801)	(1.913)
Observations	1,571	1,489	376
R^2	0.198	0.207	0.120
Adjusted R^2	0.196	0.205	0.111
Residual Std. Error	0.607 (df = 1566)	0.610 (df = 1484)	0.618 (df = 371)
F Statistic	96.780^{***} (df = 4; 1566)	97.089^{***} (df = 4; 1484)	12.685^{***} (df = 4; 371)

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Source: Data from JHU, BEA, US Census Bureau, MIT Election Data Lab

Table 5: Results. Model 4

	Dependent variable:		
	(deaths_10ths)	deaths_10ths	
	(1)	(2)	(3)
days since 100^{th} case		<190	<80
repwin	-2.283^{***}	-1.533^{***}	-0.022
	(0.294)	(0.280)	(0.647)
daysfrom100	0.035^{***}	0.032^{***}	0.009
	(0.002)	(0.002)	(0.006)
Metropol_area	-0.803^{***}	-0.891^{***}	0.709^{**}
	(0.245)	(0.220)	(0.297)
log(Income)	-4.474^{***}	-6.292^{***}	-4.144^{***}
	(0.496)	(0.483)	(0.816)
Constant	50.326^{***}	69.310^{***}	45.397^{***}
	(5.274)	(5.125)	(8.758)
Observations	1,571	1,489	376
R^2	0.191	0.206	0.086
Adjusted R^2	0.189	0.204	0.076
Residual Std. Error	4.365 (df = 1566)	3.902 (df = 1484)	2.830 (df = 371)
F Statistic	92.298^{***} (df = 4; 1566)	96.090^{***} (df = 4; 1484)	8.715^{***} (df = 4; 371)
Note:			*p<0.1; **p<0.05; ***p<0.01

Source: Data from JHU, BEA, US Census Bureau, MIT Election Data Lab

Table 6: Results, Model 5.

	Dependent variable:	
	(death_rate) $*100$	
	(1)	(2)
repwin	-0.447^{***}	-0.198
	(0.132)	(0.131)
daysfrom100	0.011^{***}	0.010^{***}
	(0.001)	(0.001)
Metropol_area	-0.139	-0.174^{*}
	(0.110)	(0.103)
log(Income)	0.085	-0.516^{**}
	(0.222)	(0.227)
Constant	0.496	6.775^{***}
	(2.364)	(2.405)
Observations	1,571	1,489
R^2	0.105	0.073
Adjusted R^2	0.103	0.071
Residual Std. Error	1.956 (df = 1566)	1.831 (df = 1484)
F Statistic	46.155^{***} (df = 4; 1566)	29.389^{***} (df = 4; 1484)
Note:		*p<0.1; **p<0.05; ***p<0.01

Source: Data from JHU, BEA, US Census Bureau, MIT Election Data Lab

Table 7: Results. Models 6,7

	Dependent variable:	
	log(cases_10ths)	deaths_10ths
	(1)	(2)
rep_gouv	0.464***	1.199***
	(0.030)	(0.232)
daysfrom100	0.006***	0.040***
	(0.0003)	(0.002)
Metropol_area	−0.073**	−0.823***
	(0.032)	(0.247)
log(Income)	−0.670***	−3.287***
	(0.065)	(0.505)
Constant	11.286***	34.492***
	(0.683)	(5.335)
Observations	1,571	1,571
R^2	0.307	0.174
Adjusted R^2	0.305	0.172
Residual Std. Error (df = 1566)	0.564	4.411
F Statistic (df = 4; 1566)	173.164***	82.373***
Note:	*p<0.1; **p<0.05; ***p<0.01	

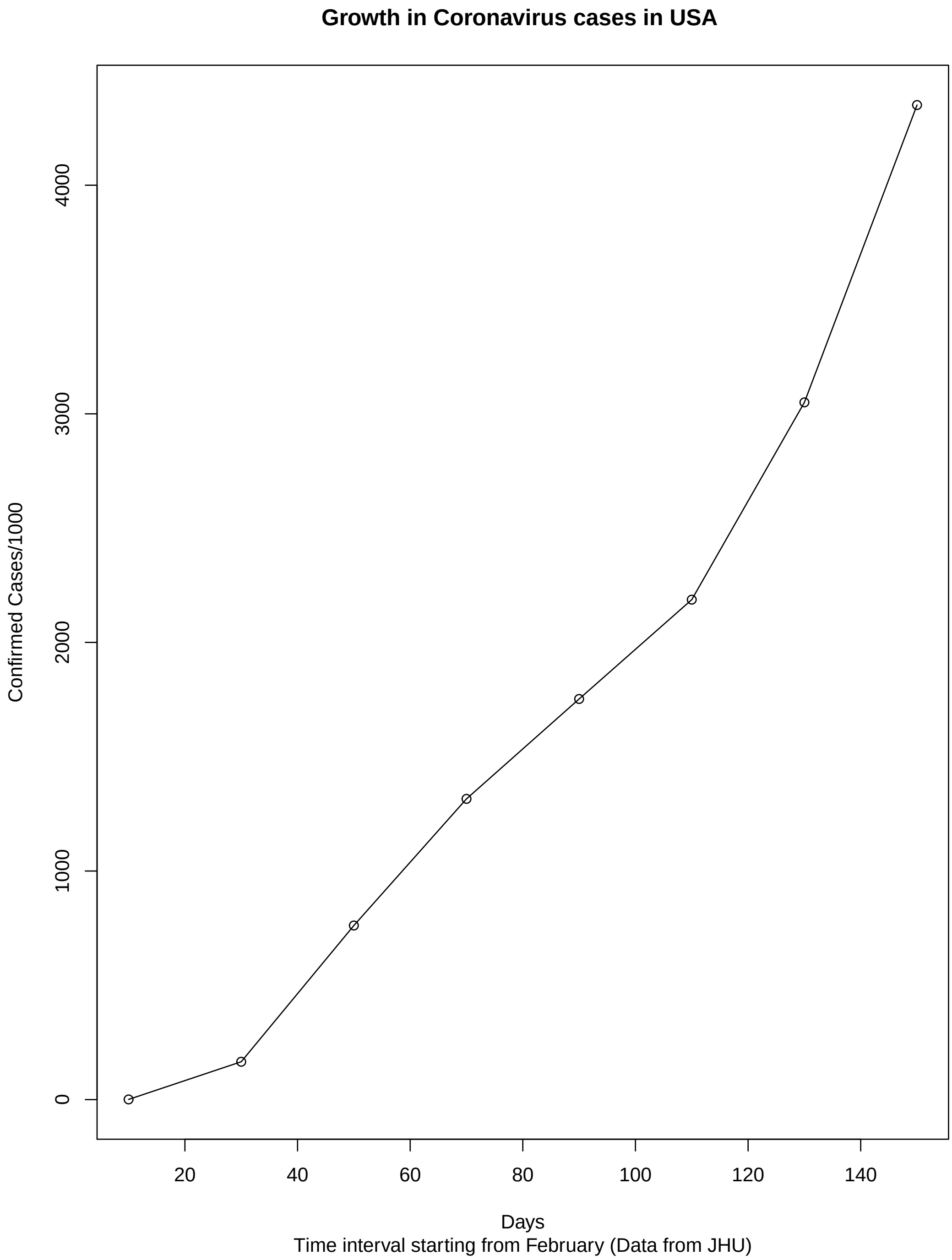
Source: Data from JHU, BEA, US Census Bureau, MIT Election Data Lab

Table 8: Results. Model 8

	Dependent variable:
	(death_rate) $*100$
rep_gouv	-0.141
	(0.103)
daysfrom100	0.013^{***}
	(0.001)
Metropol_area	-0.140
	(0.110)
log(Income)	0.147
	(0.225)
Constant	-0.578
	(2.374)
Observations	1,571
R^2	0.100
Adjusted R^2	0.098
Residual Std. Error	1.962 (df = 1566)
F Statistic	43.492^{***} (df = 4; 1566)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Source: Data from JHU, BEA, US Census Bureau, MIT Election Data Lab

Figure 1: Increase in covid-19 cases



Source: JHU, own visualization

Figure 2: Cumulative deaths for each state

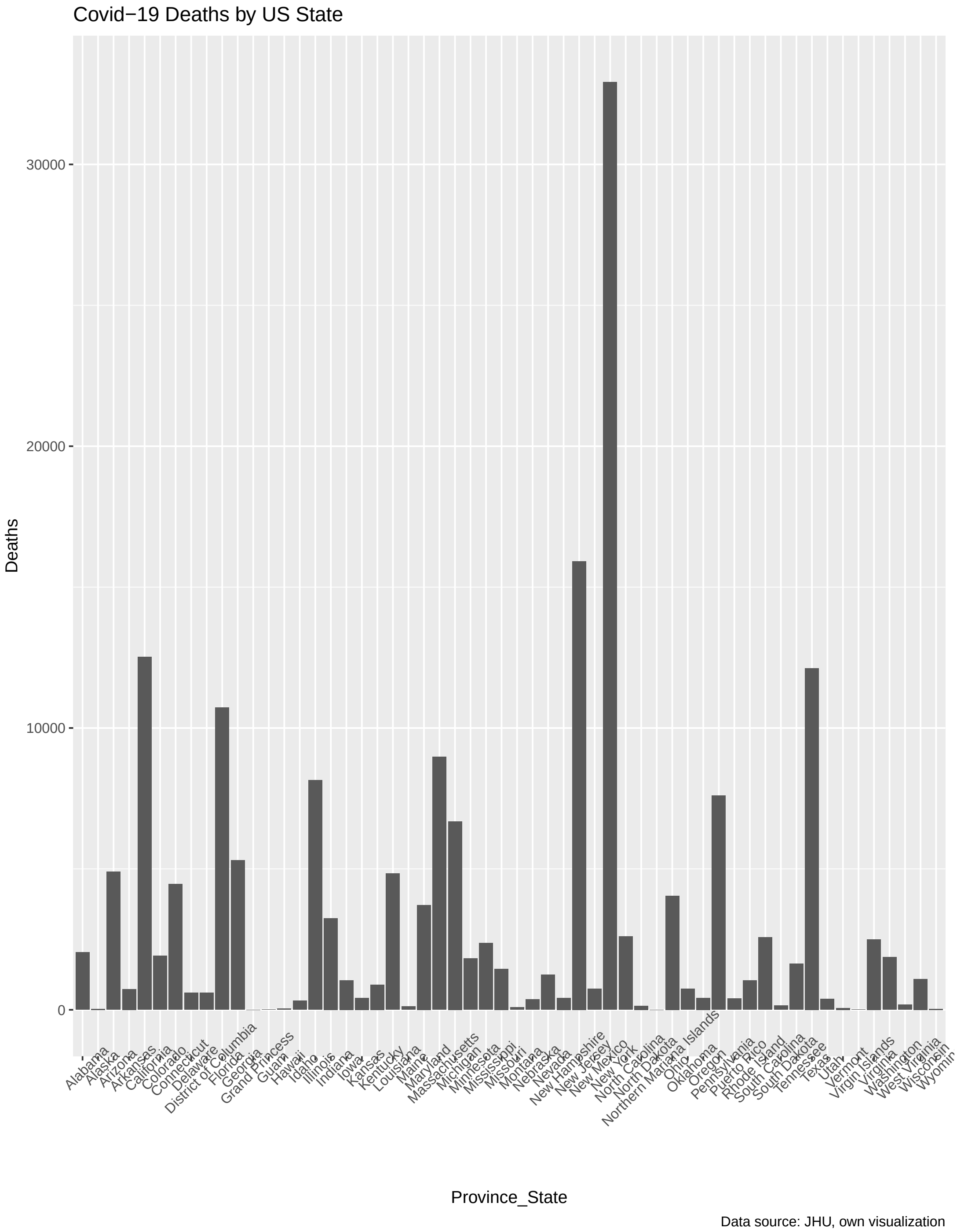


Figure 3: Cumulative confirmed cases for each state

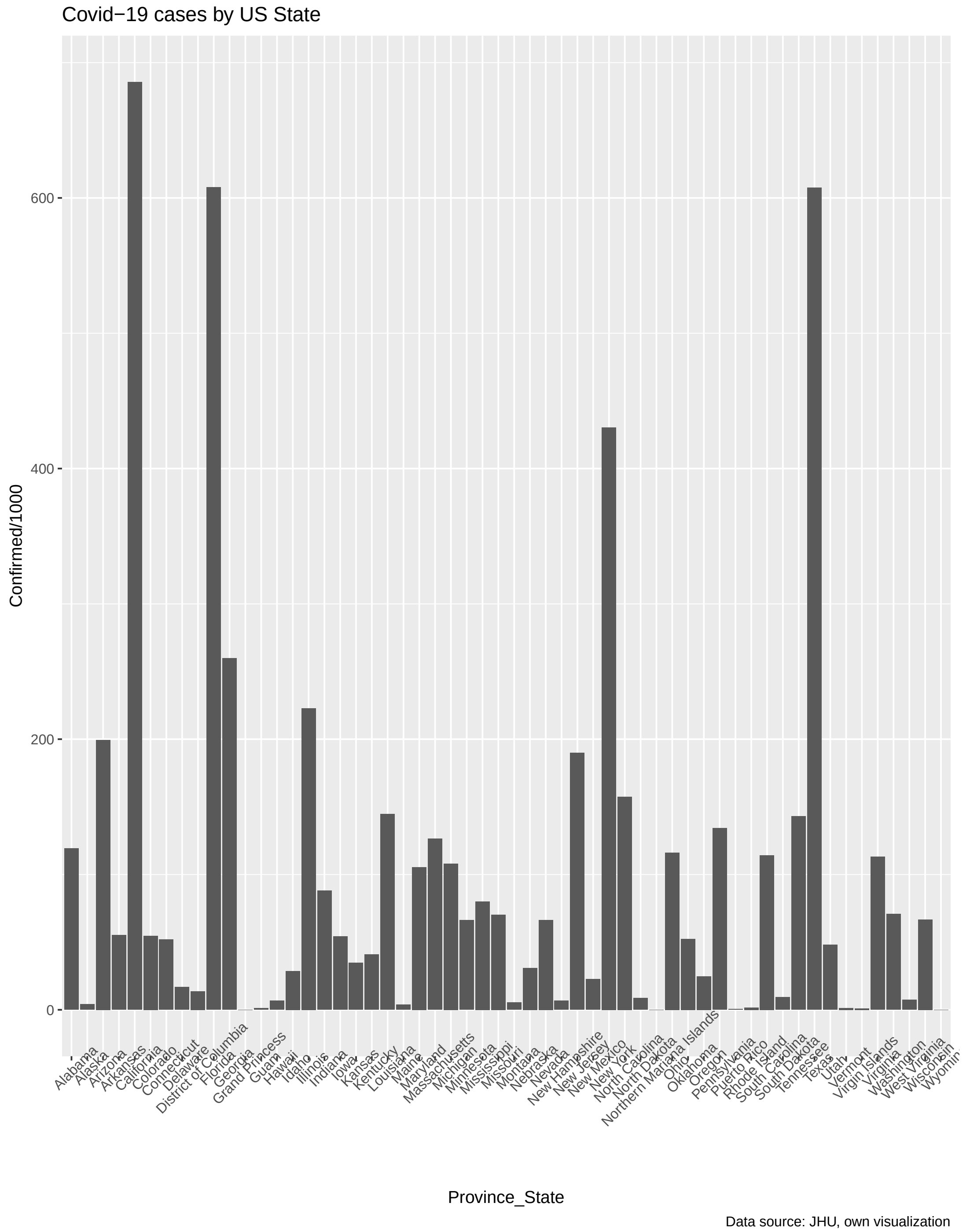


Figure 4: Daily new cases for each state, 8 day average

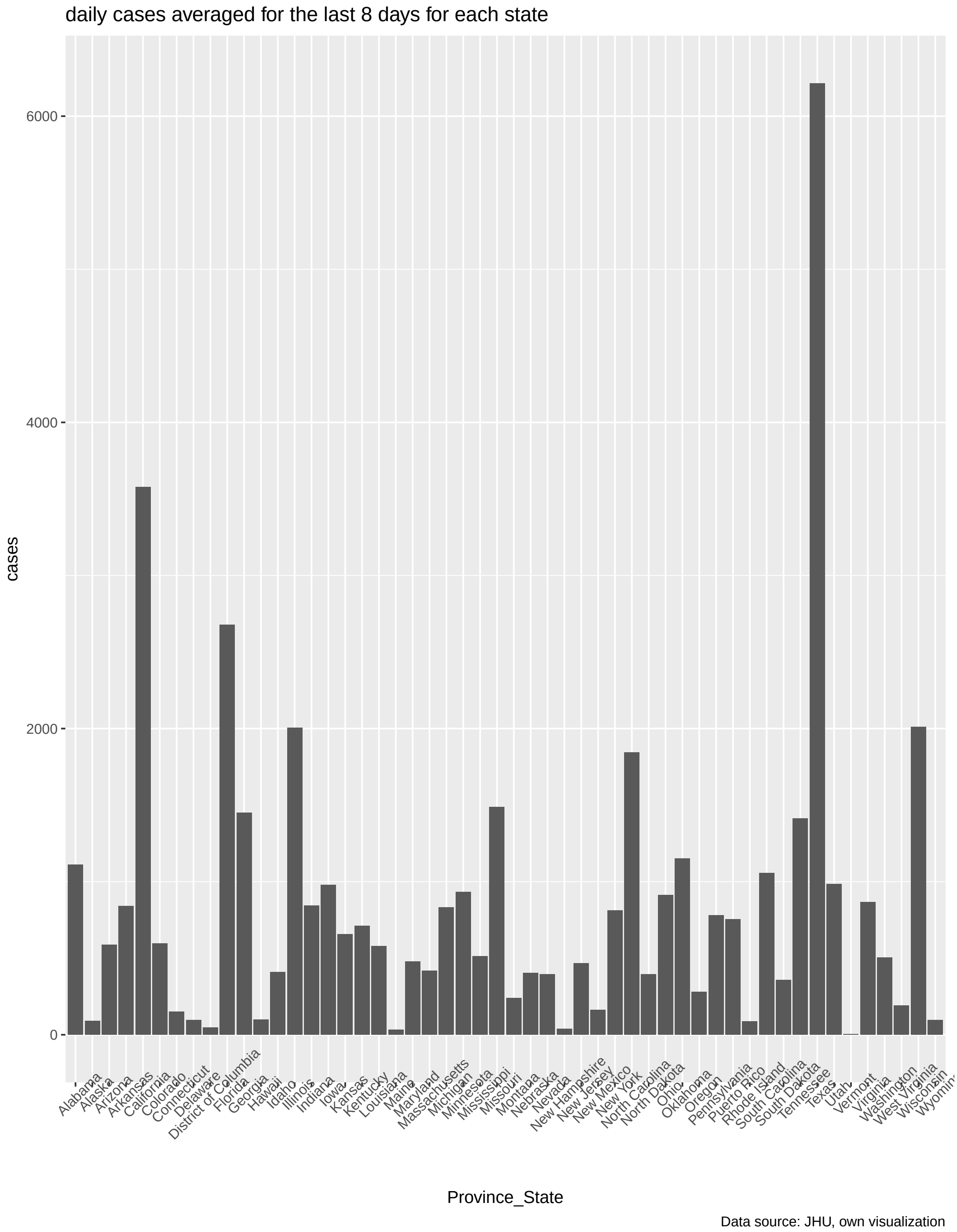


Figure 5: Preview of the database sorted by number of deaths

county	state	Confirmed	Deaths	repwin	pop	daysfrom100	Metropol_area	Income	cases_10ths	deaths_10ths	rep_gouv	death_rate
Los Angeles	California	263333	6423	0	10039107	198	1	62224	262.30720	6.3979794	0	0.024391170
Miami-Dade	Florida	167880	3127	0	2716940	196	1	50022	617.90102	11.5092715	1	0.018626400
Cook	Illinois	140623	5180	0	5150233	198	1	62205	273.04202	10.0577974	0	0.036836079
Maricopa	Arizona	140409	3304	1	4485414	192	1	47694	313.03465	7.3660982	1	0.023531255
Harris	Texas	139017	2535	0	4713325	191	1	56474	294.94465	5.3783688	1	0.018235180
Dallas	Texas	78723	1106	0	2635516	193	1	58993	298.70052	4.1965217	1	0.014049261
Broward	Florida	76146	1343	0	1952778	195	1	50269	389.93680	6.8773819	1	0.017637171
Queens	New York	71881	7244	0	2253858	199	1	49777	318.92426	32.1404454	0	0.100777674
Kings	New York	67510	7319	0	2559903	198	1	52192	263.72093	28.5909271	0	0.108413568
Clark	Nevada	64895	1353	0	2266715	195	1	47090	286.29537	5.9689904	0	0.020849064
Riverside	California	57695	1173	0	2470546	189	1	40637	233.53137	4.7479383	0	0.020331051
Bexar	Texas	53794	1264	0	2003554	188	1	46058	268.49289	6.3087893	1	0.023497044
San Bernardino	California	53121	908	0	2180085	185	1	40316	243.66481	4.1649752	0	0.017093052
Bronx	New York	52710	4943	0	1418207	197	1	37376	371.66648	34.8538683	0	0.093777272
Orange	California	52538	1176	0	3175692	192	1	69268	165.43796	3.7031299	0	0.022383798
Tarrant	Texas	47917	698	1	2102515	189	1	51239	227.90325	3.3198336	1	0.014566855
Nassau	New York	46388	2201	0	1356924	199	1	89839	341.86145	16.2205105	0	0.047447616
Suffolk	New York	46208	2011	1	1476601	197	1	68617	312.93491	13.6191158	0	0.043520602
Palm Beach	Florida	45743	1308	0	1496770	191	1	79760	305.61142	8.7388176	1	0.028594539

Data source: Data from JHU,BEA,US Census Bureau,MIT Election Data Lab