**Fatal Political Beliefs: Can Support for the Republican Party Explain COVID-19 Deaths in the US? A Difference-in-Difference Analysis**

**A Project Presented to**

**The Faculty of the Department of Industrial and Systems Engineering**

**San José State University**

**In Partial Fulfillment of the Requirements**

**For the Degree Master of Science in Engineering Management**

**By**

**Vinithra Nagaraj**

**April 2022**

**SAN JOSÉ STATE UNIVERSITY**

**The Undersigned Project Committee Approves the Project Titled**

**Fatal Political Beliefs: Can Support for the Republican Party Explain COVID-19 Deaths in the US? A Difference-in-Difference Analysis**

**By**

**Vinithra Nagaraj**

**APPROVED FOR THE DEPARTMENT OF INDUSTRIAL & SYSTEMS ENGINEERING**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Professor Niranjani Patel, Department of Industrial and Systems Engineering**

**Table of Contents**

1. Introduction

1.1 Literature Review

1. COVID-19 Vaccine Hesitancy and Political Beliefs in the US
2. Data

3.1 Vaccination Rates

3.2 COVID-19 Deaths

3.3 Measures of Republican Party Support

3.4 Demographic Data

3.5 Normalization of Variables

1. Empirical Analysis

4.1 The Effect of Full Vaccination on COVID-19 Deaths Per 100K

4.2 The Effect of Republican Vote Shares on Vaccination Rates

4.3 The Effect of Republican Vote Shares on COVID-19 Deaths Per 100K

4.4 Difference-in-Difference Estimates of Effects of Republican Support on Deaths

1. Conclusions and Recommendations

**List of Figures**

Figure 1: Vaccination Doses Administered Per 100 People, High and Upper Middle-Income Countries, (June 2021 - April 2022) 12

Figure 2: Heatmap of Fully Vaccinated as a % of Population at the County Level (June 2021 - April 2022) 16

Figure 3: COVID-19 Deaths Per 100K as of December 2021 at the US County Level 18

Figure 4: County-Level 2016 and 2020 Presidential Election Vote Share for the Republican

Party 20

Figure 5: County-Level Map grid of the Demographic Variables 23

Figure 6: Correlation Plot of COVID Deaths Per 100K, Republican Share, Density, Income, Percent Population Above 65, and Percent White Population 23

Figure 7: Directed Acyclic Graph - Causal Inference Diagram 26

Figure 8: Correlation between COVID Deaths Per 100K and Fully Vaccinated % of

Population 27

Figure 9: Correlation between Vote Share for the Republican Party and Fully Vaccinated % of Population 32

Figure 10: Correlation between Vote Share for the Republican Party and COVID

Deaths Per 100K 34

Figure 11: Correlation between Change in Vote Share for the Republican Party between 2016-2020 and COVID Deaths Per 100K 37

**List of Tables**

Table 1: Summary Statistics of Vaccination Rates 15

Table 2: Summary Statistics of the COVID Deaths per 100K across US Counties 18

Table 3: Summary Statistics of the Demographic Variables 21

Table 4: The Effect of Full Vaccination on COVID-19 Deaths Per 100K at the

County Level 30

Table 5: The Effect of 2020 Republican Vote Shares on COVID-19 Deaths Per 100K at the County Level 32

Table 6: The Effect of Republican Vote Shares on COVID-19 Deaths Per 100K at the County Level 35

Table 7: Difference-in-Difference Estimates of the Effect of Republican Party Support on Normalized COVID-19 Deaths Per 100K 38

**Fatal Political Beliefs: Can Support for the Republican Party Explain COVID-19 Deaths in the US? A Difference-in-Difference Analysis**

Vinithra Nagaraj, San Jose State University[[1]](#footnote-1)

**Abstract**

Despite the vaccine being available, free of cost, only 65% of Americans have decided to get fully vaccinated against COVID (Our World in Data, World Bank, 2022). Vaccine hesitancy has large-scale implications for the limits of public response to pandemics. Recently, public discourse and news media within the right-wing of the US political spectrum have pushed the narrative against vaccine mandates and/or cast shadows on vaccines’ legitimacy. This study investigates the relationship between support for the Republican party and reduced vaccination rates and increased deaths from COVID-19, after controlling for demographic variation in counties, including income, race, age, and population density. Based on the latest available public data, this study finds that for a one standard deviation increase in Republican vote share at the county level in the 2020 Presidential election, vaccination rates in 2021 and 2022 decreased by 0.7-0.8 standard deviations, or between 8 and 12 percentage points, and COVID-19 Deaths Per 100K increased by 0.4 standard deviations, or 54 additional deaths per 100,000 people. Using a difference-in-difference method, this study identifies the significant marginal causal impact of increasing support for the Republican party from 2016-to 2020 on COVID-19 deaths per 100K, even for counties with high vaccination rates.

**1 Introduction**

This study aims to find the causal relationship between political beliefs and COVID-19 deaths. In the US, the Republican Party and its leadership have communicated counterintuitive messages about the pandemic, including downplaying the severity of illness, downplaying the safety and importance of vaccination, and disregarding safety measures. While the literature on this topic has addressed the opinions of supporters of the Republican party suggesting lower risk perception of COVID-19 due to political beliefs, no study has yet established the relationship between these beliefs and eventual fatality from COVID-19. Using the latest COVID-19 deaths data, vaccination rate data, election results from 2016 and 2020, demographic county-level data on income, race, population density, and age, and a difference-in-difference model, this study finds that a one standard deviation increase in the support for the Republican Party in a county between 2016 and 2020 caused 5.5 additional COVID-19 deaths in that county, even if the county had a high vaccination rate. In this section, the problem statement, the literature, data, methods, and results are briefly summarized. In Section 1.1, the literature review discusses in detail the prior studies on this topic and their findings; in Section 2, the problem statement and the background of the topic are detailed narratively; in Section 3, the data sources, summary statistics, and data visualization are detailed; in Section 4, the empirical methodology and results are discussed, and Section 5 concludes and provides some potential improvements for a similar study and recommendations.

Based on 2016 county-level presidential vote shares, research has found that as Trump voter share rises, individuals search less for information on the COVID-19 virus, and engage in less social distancing behavior, as measured by smartphone location patterns. (Barrios and Hochberg, 2021). The literature on the link between COVID-19 spread and political beliefs has largely been silent on the marginal effect of political beliefs on COVID-19 deaths after controlling for these potentially confounding factors, largely due to lack of data. The methodological contribution of this study is to control for observable confounding variables in addition to the factors considered in the literature thus far, using the latest data on COVID-19 and demographics and a difference-in-difference model.

In Section 3, the data used in the study are discussed in detail. From CDC (Center for Disease Control), cumulative COVID-19 deaths were pulled as of December 2021, and the cumulative full vaccination rates were pulled from four dates, 1 June 2021, 1 Sep 2021, 1 Jan 2022, and 13 April 2022. This process helps capture changes in percentage acceptance of vaccinations against the virus over the span of the pandemic, which is used to evaluate potential influential parameters and partially control for the effects of time and confounding events during the 2021-2022 period, like changing COVID-19 policies, travel rates and as more information was released regarding the vaccine and its effectiveness.

From the data, vaccination rates showed slow progress in Republican states with minimal improvements even in April 2022, where these states still lagged Democrat majority counties in New England, the West Coast, New Mexico, and other large cities. Looking at the demographic differences, one would expect that sparsely populated areas would be the safest during a pandemic. However, COVID-19 Deaths Per 100K are the highest in rural, sparsely populated counties in North Dakota, South Dakota, Utah, Texas, Oklahoma, Georgia, Alabama, Mississippi, and Montana. The plausible explanation could be these counties also tend to lean Republican, and the marginal effect of Republican support is so negative, that it undoes the advantages of income, race, population density, and location in reducing the impact of COVID-19.

In Section 4, the causal mechanism and the methodology is discussed. The hypothesis detailed here is not that political leaning themselves will cause someone to die, but it is that political beliefs that come associated with a political party and a worldview cause certain risky behavior that leads to fatality. While vaccine hesitancy in itself is deadly for COVID-19, this only explains a piece of the puzzle, that political beliefs and support for the Republican Party are also additionally causally linked to more COVID-19 deaths. In Section 4, the analysis shows that support for the Republican party reduces vaccination rates in the counties and increases deaths per 100K.

In Section 4.4, the difference-in-difference estimator shows that for a county with both one standard deviation increase in Republican vote share from 2016 and comparatively one standard deviation above the mean vaccination rates as of January 2022, there are 0.04 standard deviations or 5.5 additional COVID-19 deaths per 100K. This shows that even accounting for high vaccination rates, time trends in vote shares for the Republican Party are significantly related to more COVID-19 deaths, which must be related to lower COVID-19 risk perception stemming from political beliefs. This finding confirms the hypothesis that political beliefs led to more deaths during the COVID-19 pandemic.

**1.1 Literature Review**

The population proportion of U.S. residents willing to be vaccinated has fluctuated significantly in 2020 and 2021. In a survey conducted in 2020, an astounding 39% of respondents had indicated that they would not get the vaccine, and only 46% said they might be open to vaccination once others start getting it and more information becomes available (Wood and Schulman, 2021). Since then, as of March 2022, [65% of the US has been fully vaccinated](https://ourworldindata.org/covid-vaccinations?country=USA), but these proportions and time periods for vaccine adoption vary significantly across locations and other demographic factors.

One of these factors that impact variation in COVID-19 risk perception is political communication and encouragement (Barrios and Hochberg, 2021). Based on 2016 county-level presidential vote shares, current research has found that as Trump voter share rises, individuals search less for information on the virus, and engage in less social distancing behavior, as measured by smartphone location patterns. (Barrios and Hochberg, 2021). Therefore, it is safe to say that at least in the US, there may be an association between political beliefs and COVID-19 transmissions. This is confirmed by differences in Google Trends search volumes for COVID-19 risks across states that voted for Clinton versus Trump in 2016 (Timoneda and Vallejo Vera, 2021).

In addition to risk perception, there is evidence that even the understanding of the severity and forecast of the pandemic is impacted by partisan political ideology (Freira et. al, 2021). These studies were performed in 2020 and 2021 during the pandemic and therefore were not able to observe fatality rates and transmission rates across different waves of different COVID-19 variants (Allcot et. al, 2020).

There are several key factors that, in addition to political ideology, are likely to drive risk perception and behavior to follow COVID-19 safety protocols and vaccination (Allcot et. al, 2020). These include demographic factors, primarily income, race, population density, the transmissibility of different variants of COVID-19, as well as access to resources and information. The contributions of this project to this literature from a data perspective are discussed in Section 3.

From a methodological standpoint, the literature currently consists of reduced-form observational studies using small surveys (Freira et. al, 2021; Wood and Schulman 2021); internet search history or information content (Wood and Schulman, 2021; Barrios and Hochberg, 2021); some political indicators from 2016 presidential elections (Adolph et. al, 2021; Allcot et. al, 2020) and the latest COVID-19 transmission data available at the time the research was conducted.

The literature on the link between COVID-19 spread and political beliefs has largely been silent on the marginal effect of political beliefs on COVID-19 spread after controlling for these potentially confounding factors, largely due to lack of data. Given that factors like location, age, race, and income covary significantly with political ideology, it is likely that the marginal impact of political ideology on COVID-19 spread is overstated in the current research. The methodological contribution of this study aims to control for observable confounding variables in addition to the factors considered in the literature thus far, using a difference-in-difference causal inference model. The details of the research design are explained further in Section 4.

**2 COVID-19 and Political Beliefs about COVID-19 Vaccination**

Coronavirus, officially known as the SARS-CoV-2, was declared a pandemic by the World Health Organization in March 2020. Despite equitable access made available to vaccines, and with consistent warnings about the dependability of the public to get fully vaccinated to avoid new variants from emerging, the US lays low in the total number of vaccine doses administered.

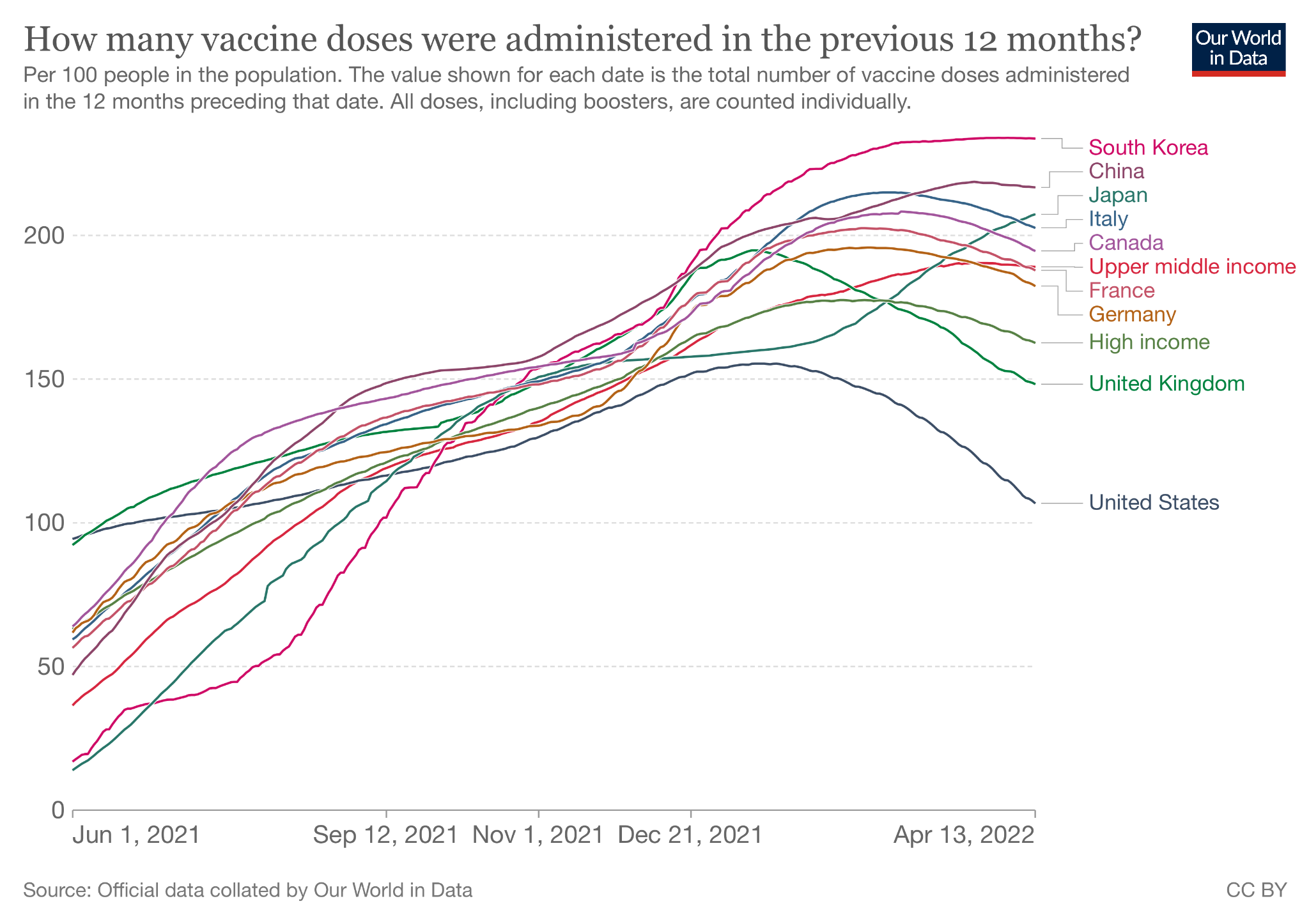


Figure 1: Vaccination Doses Administered Per 100 People, High and Upper Middle-Income Countries, (June 2021 - April 2022)

Global comparison of the number of vaccine doses administered per 100 people in concentration was sourced from World in Data (Ritchie et al., 2020), as shown in Figure 1. The representation of this data was filtered to visualize the trend in High and Upper Middle-Income countries in the last year, June 2021- to April 2022, where the United States data lies lower than all other states since September 2021.

Politicians in the right-wing have shown resistance to COVID-19 vaccine mandates since its availability, leading the Republicans to staunchly oppose getting immunizations. Activists have been protesting in Washington, California, New York, Colorado, and Texas, believed to be preserving their constitutional rights, comforted by the Party’s leader failing to urge the requirement. Conservative media has been spreading misinformation fostering skepticism about vaccines, further causing many GOP (Grand Old Party) supporters to cast doubt against getting inoculated against the virus.

A notable annual event, the Sturgis Motorcycle rally in 2020 in Meade, South Dakota, had 61% of the attendees fly in from counties outside of South Dakota. Despite several public health department warnings, confirmed COVID cases among the residents of these counties were reported to have increased by 199% within 14-days from the day of the event (​​Carter et al., 2021). Epidemiologic investigations conducted in the neighboring state Minnesota tracked down rally-related cases in one-third of its counties and caused the secondary and tertiary transmission of the virus in the neighboring states. It is important to note that South Dakota is known to have had the longest Republican streak in the country which made it difficult to enforce any COVID mitigation mandates even during the midst of the pandemic.

**3 Data**

This project addresses the gap in the literature by exploiting newly available detailed public data, vaccination trends including full vaccination rates through the span of the pandemic, data across different waves of COVID-19, and data on political trends from the 2020 presidential elections. The holistic COVID data is noteworthy as its transmissibility and fatality rates, including the Omicron and Delta variants, present interesting statistical opportunities to test the hypothesis detailed in the next section. COVID-19 deaths and vaccination data by date have been taken, from the Center of Disease Control. County-level voting history from the years 2016 and 2020 national elections has been collected from public databases via the MIT Election Lab. Public demographic data for income, age, ethnicity, population density, and other similar data at the county level have been collected fromthe US Census Bureau. This project requires the integration of pandemic, political and demographic data to study the effect of independent variables on COVID deaths while controlling for all the constants.

**3.1 COVID-19 Vaccination Rates in the US**

One of the key pieces of evidence to direct this study are the vaccination trends by demographics. CDC has made available the most accurate vaccination datasets which are fetched through their public API. This study utilizes the advantage of the US Census Bureau being the common data source for the CDC to curate the vaccination completeness data at the county level. This data particularly improves the quality of the results as they include data across all the variants of COVID-19 and the recommended vaccines and booster shot completeness numbers when made eligible. This categorical data has vaccination dates in the rows and the completeness numbers and percentages in the columns. For the purpose of this study, cumulative full vaccination (Vaccine 1 and 2) completeness rates were used, not including the booster shots recently rolled out, in order to avoid data completeness issues associated with the newer, and more varied distribution and availability of booster shots.

In Section 3.5, a detailed discussion can be found on the transformation of these variables for the purposes of this study, including z-score normalization. To visualize the trends of vaccination rates, the cumulative full vaccination rates were pulled from four dates, 1 June 2021, 1 Sep 2021, 1 Jan 2022, and 13 April 2022. This process helps capture changes in percentage acceptance of vaccinations against the virus over the span of the pandemic, which is used to evaluate potential influential parameters and partially control for the effects of time and confounding events during the 2021-2022 period, like changing COVID-19 policies, travel rates and as more information was released regarding the vaccine and its effectiveness.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **N** | **Mean** | **Std. Dev.** | **Min** | **Pctl. 25** | **Pctl. 75** | **Max** |
| Vaccination %  Jun 2021 | 3046 | 28.327 | 13.783 | 0 | 22 | 37.575 | 99.9 |
| Vaccination %  Sep 2021 | 3046 | 35.415 | 16.24 | 0 | 28.6 | 45.8 | 99.9 |
| Vaccination %  Jan 2022 | 3046 | 47.476 | 12.482 | 0 | 39.6 | 55 | 95[[2]](#footnote-2) |
| Vaccination %  April 2022 | 3046 | 51.074 | 12.006 | 11.1 | 42.6 | 58.1 | 95 |

Table 1: Summary statistics of Vaccine penetration data

Looking at the summary statistics of the distribution of vaccination rates at the county level across these four dates, it was observed that the mean value as shown in the above Table 1, shows sizable improvement in vaccination rates from 28% in June 2021 to 51% in April 2022 for people fully vaccinated in the US counties during the selected period. Another important observation from this summary table is that the minimum percentage of people vaccinated has been 0 for three consecutive time periods which were rural counties in the states of Texas and Georgia, with large Republican vote shares. This minimum improved to 11.1% by 13 April 2022.

In order to observe the regional progress noticeably, this data was transcribed on 2D maps to visualize the changes over time and help with a visual comparison with other demographic variables.

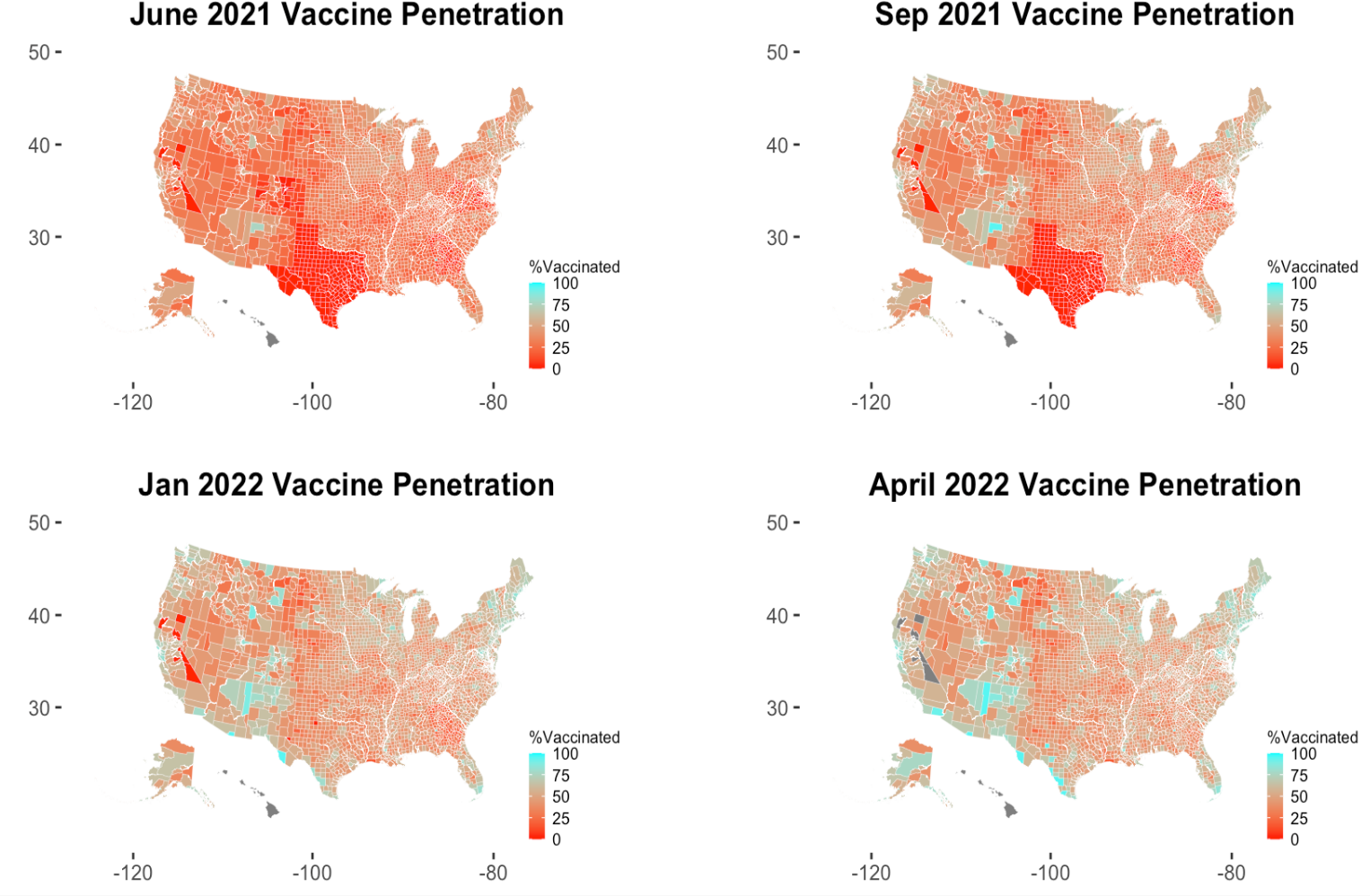


Figure 2: Heatmap of Fully Vaccinated as a % of Population at the County Level (June 2021 - April 2022)

The Albers Projection parameter (ArcGIS, 2021) is used in the map functions for visualizing three-dimensional maps as flat figures. Figure 1 below shows four North America maps at the selected time frames over the span of the pandemic to study the progress of fully vaccinated percentages. The color palette chosen denotes these percentages across counties, where Red denotes the least percentages of people who are fully vaccinated against COVID, and turquoise denotes the highest percentages of people at the county level. It is to be noted that in the data from 13 April 2022, some of the counties (less than 0.003 percent of the data) were not available yet. The following packages in R were used in the preparation of these plots: usmap, mapdata, urbnmapr, ggplot.

Figure 2 demonstrates improvements in vaccine penetration across US counties at different stages of the pandemic by showing the relational effects. The red zones evidently seen in large sections show slow progress in Republican states like Texas, Georgia, North and South Dakota, Idaho, Wyoming, West Virginia, Missouri, Alabama, Mississippi, and Florida which started as bright reds in the initial phases of the pandemic denoting lowest vaccination percentages with minimal improvements in the last stages, where these states still lag behind Democrat majority counties in New England, the West Coast, New Mexico and in large cities. The next sections will discuss how these variations in the vaccination rates relate to COVID-19 Deaths Per 100K, income, race, age, and population density.

**3.2 COVID-19 Deaths Per 100K**

This is the key variable of interest for this study as it is the dependent variable because the goal of this study is to investigate the relationship between political beliefs and fatality. CDC which gives access to the most accurate numbers reported for COVID-19 deaths was used. Cumulative deaths that occurred until 21 Dec 2021 were extracted to account for the holistic effects of the virus. Deaths per 100K people were calculated by dividing the County level death numbers by Population and multiplying that by 100k. This measure helps to report the number of deaths resulting from COVID for every 100k people in each county. This rate is calculated to aid the comparative study between the varied population groups across counties.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **N** | **Mean** | **Std. Dev.** | **Min** | **Pctl. 25** | **Pctl. 75** | **Max** |
| COVID Deaths Per 100K | 3046 | 291.40 | 139.03 | 0 | 192.47 | 374.25 | 1109.57 |

Table 2: Summary statistics of the COVID Deaths per 100K across US counties

The above Table 3 shows that the standard deviation of 139 deaths per 100K is relatively small when compared to the range from 0 to 1110 deaths per 100K observed. This denotes that the data is not widespread and is concentrated around the mean.

**3.2.1 Geospatial representation of COVID deaths**

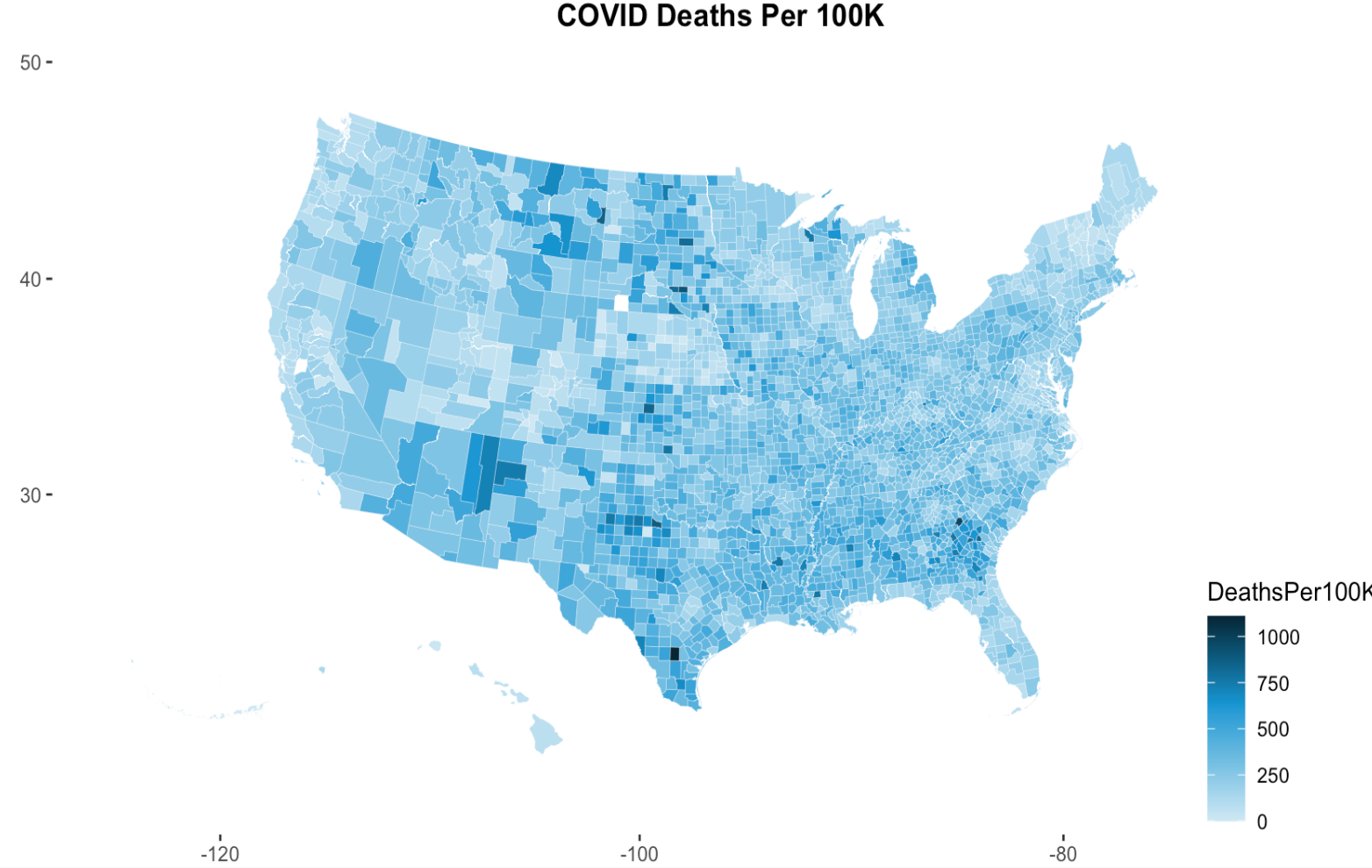


Figure 3: County-Level impact of COVID

The first striking detail about this map is that COVID-19 Deaths Per 100K are the highest in rural, white, sparsely populated counties in North Dakota, South Dakota, Utah, Texas, Oklahoma, Georgia, Alabama, Mississippi, and Montana. These are counties that are heavily Republican, as shown in Figure 4.

**3.3 Measure of Support for the Republican Party**

To measure political partisanship at the county level, this study has leveraged the 2020 presidential election results published in MIT election Data Science and Lab. The number of people who voted for Trump was extracted as the indicator of the Republican share in each county. This is vital for this study as there were divergent messages sent out from the two parties about the severity of the virus whose effect is what this study aims to measure as the pandemic response by the public. Similar to the earlier field capturing the COVID deaths, the RVotes (Republican votes) measure was calculated per 100K for every county to support the comparisons across these social entities. An additional field Republican share was calculated for each county as a percentage of total votes to highlight the partisan difference measure.



Figure 4: County-Level 2020 Presidential Election Vote Share for the Republican Party

The map above in Figure 4 shows the Republican share in each county highlighting the republican majority states in shades of red and liberals in shades of blue. Just eyeballing the vaccination rate maps and this Republican share map, the Republican states seem to coincide denoting their vaccine hesitancy. This will be further reviewed in the analysis section that follows.

**3.4 Demographic Data**

The required demographic variables for this study such as population density, personal per capita income, age, and race at the county level are sourced from the credible government website, United States Census Bureau. The data is a combination of results from the Census and Surveys conducted by the Bureau, in addition to the administrative data provided by the federal, state, and local governments.

The Bureau reuses the administrative data from agencies across the nation which is required by law to reduce the burden on the public respondents to their surveys. This study is interested in these variables which add the elements of heterogeneity while sophisticating the statistical comparisons. These fields were synthesized from the year 2019 Census results and have been geo-assigned to COVID-19 data which is discussed in the next section. The reason to avoid using population and other demographic measures from the estimates for the years 2020 and 2021 was that the data would include the effect of the pandemic.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **N** | **Mean** | **Std. Dev.** | **Min** | **Pctl. 25** | **Pctl. 75** | **Max** |
| Population | 3046 | 104258.8 | 334712.0 | 98 | 11033.5 | 67995 | 10,081,570.0 |
| Population Density | 3046 | 97.3 | 695.75 | 0.1 | 6.5 | 43.1 | 27806.8 |
| Population Aged 65 Years and Older (%) | 3046 | 18.9 | 4.6 | 3.2 | 15.9 | 21.3 | 56.7 |
| Percentage of White Population alone, Not Hispanic or Latino | 3046 | 76.7 | 19.9 | 0.7 | 64.9 | 92.6 | 100.0 |
| Per Capita Income  (Personal) | 3046 | 48821.6 | 12882.3 | 21087 | 40885.3 | 53732.3 | 220,645.0 |

Table 3: Summary Statistics of Demographic Variables

The descriptive statistics of the variables sourced from the Census Bureau are shown in Table 3 above. The observed number of counties is 3,046 and the spread seen in the table reveals some quick facts about the data collected. The wide range of population from 98 to 10,081,570 ensures that even the most rural and the densest counties of the United States are included. The population density was calculated as a quotient of population size over the land area (in square miles) which again is observed to be a wide range between 0.057 and 27806.767. These wide ranges of data improve the accuracy of the mean values while making it easy to spot outliers which can skew the data. The age dataset was utilized to include the percentage of senior residents in each county of age 65 and more. This is a critical variable for this study as it helps to account for the increased impact of the deadly virus on older adults. Though the maximum value is 56.7%, the mean value is about 19.0%. The 25th and the 75th percentile show a tight distribution around the mean and are between 16.0% and 21.0%.

To support the political partisanship measure, which is discussed in the following section in detail, race was included from demographic data sources to extract the percentage of the white population. It is not surprising to see the maximum number as 100.0%, but what's more appealing to this study is the value of the mean of about 77.0%. The standard deviation being about 20.0% shows we have a large spread of data away from the mean. Lastly, as an affordability measure which is telling in aspects to measure people’s ability to secure themselves during the pandemic, personal per capita income was included at the county level.

**3.4.1 Visualization of the demographic data**

To position the above data statistics in context, geographical representation is leveraged. This allows for viewing the demographic data in US maps layered as counties.

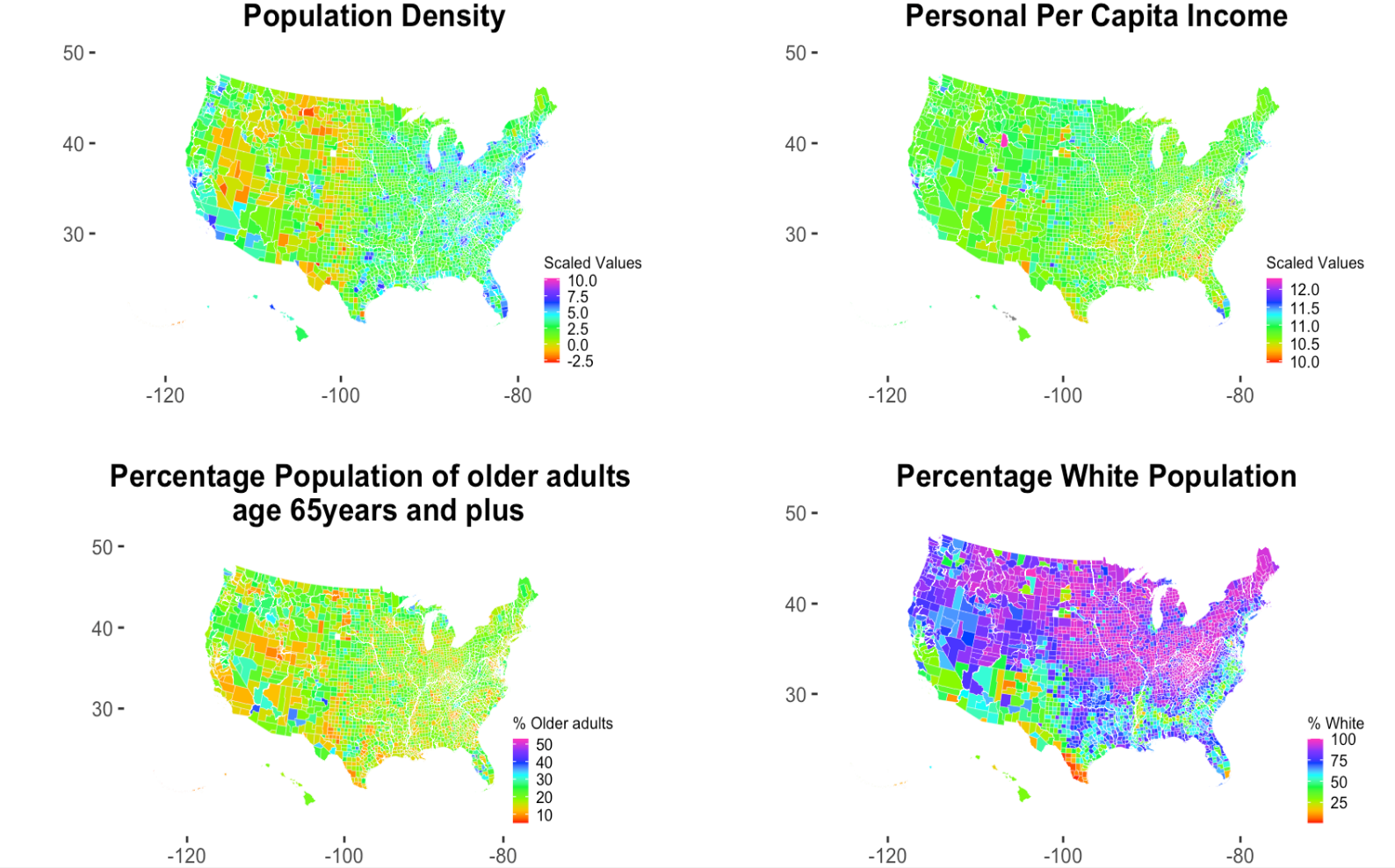


Figure 5: County-Level Maps of Demographic Data



Figure 6: Correlation Plot of COVID Deaths Per 100K, Republican Share, Density, Income, Percent Population Above 65, and Percent White Population

In the plots above, we can note interesting correlations between income, race, density, and location. We can note the east coast and large cities have high density, high incomes, and in the case of cities, high levels of diversity. From Figure 2 and Figure 3, we can draw interesting comparisons between COVID fatality, Republican support, and demographic differences. Rural, white, and sparsely populated counties tend to be Republican. These are characteristics that would make it easier for someone to avoid COVID. Being in a sparsely populated rural area means that the spreading of COVID through public transportation, closely built apartment buildings, and crowds in public areas would be limited. Additionally, the effect of race on structural racism and reduced access to health care means that counties with a high percentage of the white population would likely have easier access to healthcare and reliable COVID-19-related information. Similarly, in wealthier counties, people will find it easier to protect themselves against COVID and potentially also work from home to a greater extent. They are also financially protected from COVID, as there are income-related impacts from the pandemic as well. However, what we see in Figure 3, the COVID-19 Deaths Per 100K distribution across the US, is that rural, white, sparsely populated counties have some of the highest fatality rates. The plausible explanation could be these counties also tend to lean Republican, and the marginal effect of Republican support is so negative, that it undoes the advantages of income, race, population density, and location in reducing the impact of COVID-19. Now, as seen in Figure 9, this could be the effect of the vaccination rates on COVID-19 deaths.

**3.5 Normalization of Variables**

In order to standardize the relational comparisons of the multivariate data used in this study to a common scale, they are normalized as shown in the equation below.

(1)

In the equation above, a variable X, let’s say Republican vote share at the county level, is normalized by subtracting the mean of the variable X from the observation from County i and dividing this difference with the standard deviation of X, or . This will convert our data into z-scores, i.e., the relative position of as a measure of , i.e, in relation to the remaining X in other counties. This is essential for the relative positioning of the available data when doing like-for-like comparisons between two counties.

It also allows for easier interpretation of regression models, i.e., one can compare the impact of a one standard deviation change in X on some multiple or fraction of a standard deviation of dependent variable Y, which in this study would be COVID-19 Deaths Per 100K or vaccination rates. The variables normalized are vaccination percentages at the four different dates of the pandemic, Republican share in 2020, Population density, Personal per capita income, Percentage population at and above age 65, Percentage population that is white, change in Republican vote share from 2016 to 2020 and finally the COVID deaths recorded per 100K, all at the county level. The R packages dplyr and tidyverse are utilized for the transformation of these variables into scaled values.

**4 Empirical Analysis**

The primary aim of this study is to identify the additional impact of political beliefs on COVID-19 deaths in the United States. For this, the causal mechanism was identified as vaccine hesitancy. Consequently, it was first required to study if vaccines reduced COVID deaths and proceed to discover the pathways (causes) through which the hesitant behavior was fostered. Before joining the variables of interest with available evidence and examining correlations and regressions to confirm the causal hypothesis, DAGs (Directed Acyclic Graphs) are used to study the relationships in the causal structures. This causal model is a coded form of the prior assumptions made for this study, whose output shows paths and unidirectional arrows for the relationship among the variables included.

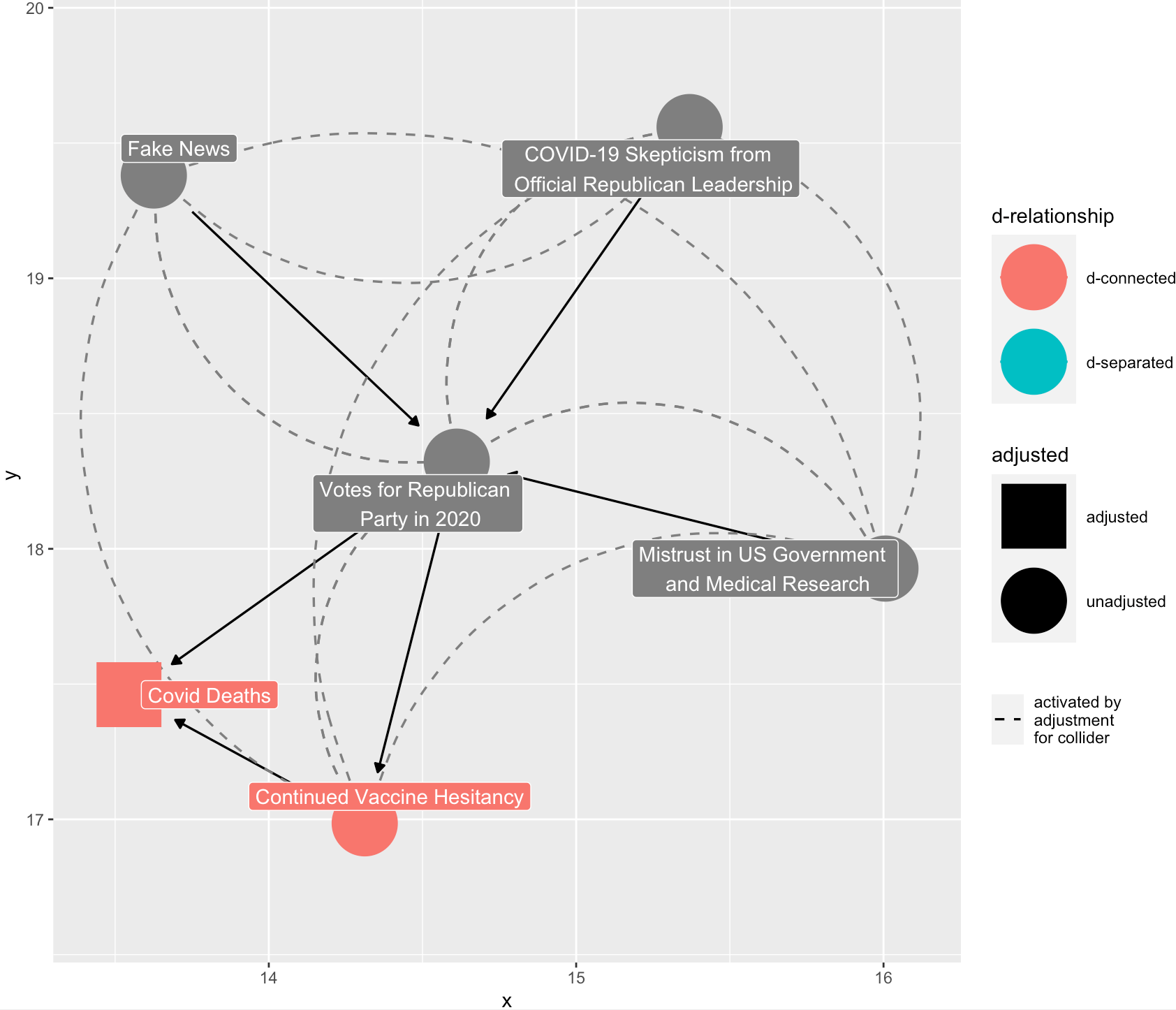


Figure 7: Directed Acyclic Graph - Causal Inference Diagram

In the following sections we study the relationship of the variables stage by stage, for which the output of the DAG shown in Figure 3 above serves as a guide. The D relationship shown in the legend is the unidirectional arrows denoting the causal effects of the outcome, COVID Deaths. The first stage of this analysis is to study the relationship between Vaccine acceptance rates in the US and COVID deaths and proceed to find the relationship between political beliefs and vaccine penetration. This will answer the unseen causal effects of political belief on COVID deaths. The figure above also includes additional variables that are causal reasoning for political beliefs, such as fake news and misinformation spread by the political parties. The dotted lines show the interconnected causations for political beliefs and its tendency on vaccine hesitancy while adjusting for the collider bias.

The conclusion is not that political leanings themselves will cause someone to die, whereas it is observed that political beliefs that come associated with a political party and a worldview would cause certain risky behavior that will lead to fatality. This project aims to find that while vaccine hesitancy in itself is deadly for COVID-19, this only explains a piece of the puzzle and that political beliefs and support for the Republican Party are also additionally causally linked to more COVID-19 deaths.

**4.1 The Effect of Vaccination on COVID-19 Deaths Per 100K**

Deaths per 100k, the dependent variable is calculated by dividing the number of COVID deaths in each county by its total population (as on 2019 Census results), times 100,000. The initial regression models include vaccination rates as of Sep 2021, Jan 2022, and April 2022 as the independent variables, while controlling for confounding variables. For example, in a county like Santa Clara in California, there are 1,932 deaths against 1,927,852 people which is relatively less severe in ratio when compared to a county like Dickey in North Dakota where there are 48 deaths against 4,872 people.

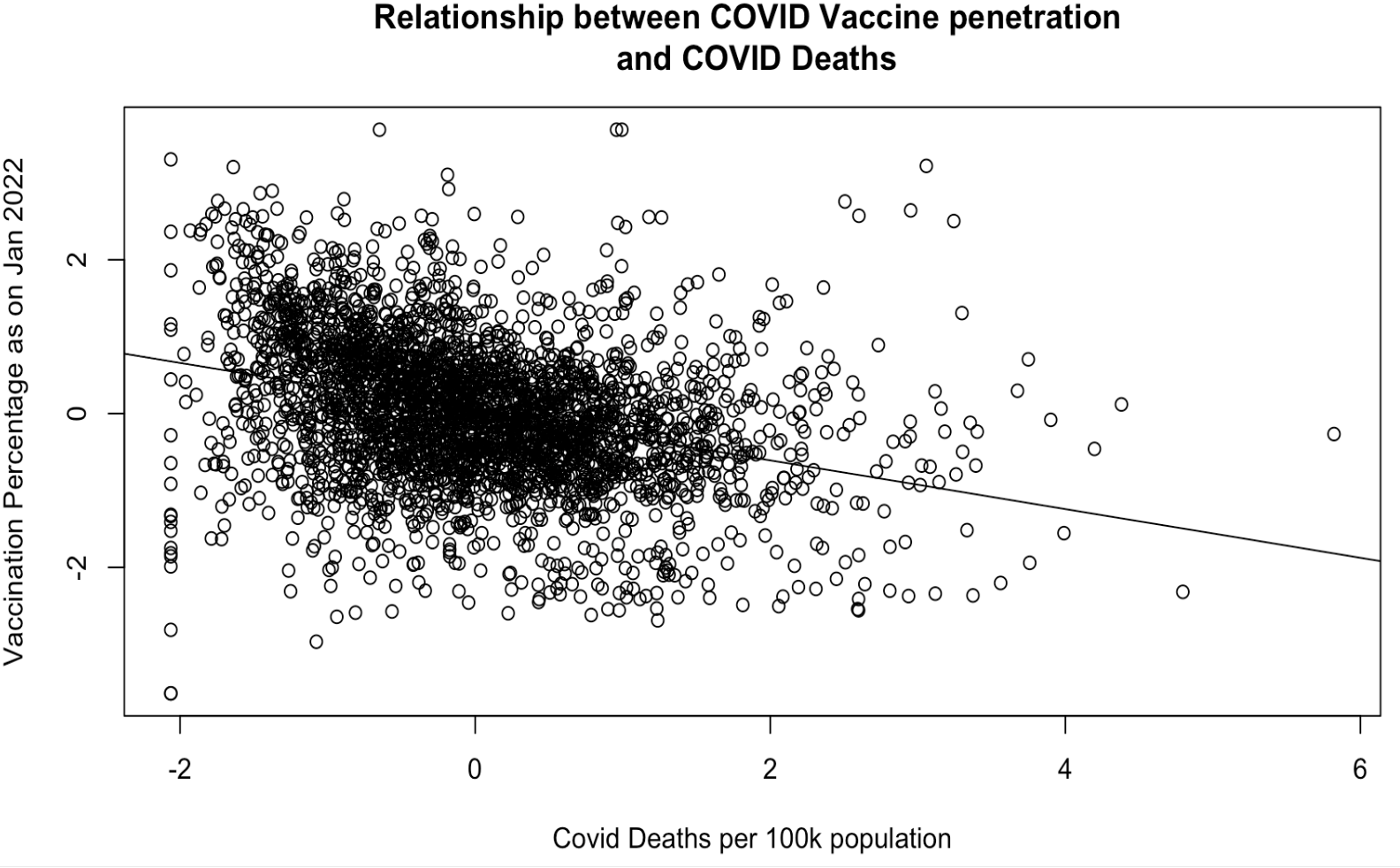


Figure 8: Correlation between COVID Deaths Per 100K and Fully Vaccinated % of Population

The plot shown in Figure 8 above clearly visualizes the inverse correlation as verified from the regression table in Table 4. The plot shows the normalized values of both these variables precisely show the effect of one standard deviation increase or decrease in vaccination percentage over deaths. The underlying assumption of this study is that, including demographic data as control variables, we control for every other parameter that affects COVID deaths between two counties with the only significant difference being the vaccine penetration percentage.

In order to effectively capture the severity of COVID across such varied characteristics of the US counties, the models include population density, per capita income, age, and race. Population density, as explained in the Data section was sourced from the US Census Bureau, which controls the element of social connectedness. This is a key variable that impacts how transmissible COVID in a region is, as denser regions have an increased likelihood of COVID cases. Per capita income is an indicator of people’s ability to have the resources to protect themselves against COVID. These may include having houses that are farther from higher population centers, getting groceries delivered home, afford COVID Protective gears, working in industries that provide luxury to avoid in-person attendance, etc. The percentage of older adults was included to account for the severity of COVID’s impact on them. A county with a higher percentage of older adults with more COVID deaths reported should be identified among other counties with a lower percentage of older adults and increased deaths. Lastly, the percentage white population is included. These control variables ensure that comparison between a test county A and control county B, with the only difference being the percentage of vaccination rates, captures the marginal effects of vaccines in reducing deaths.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Sep 2021 | Jan 2022 | April 2022 |
| % Vaccinated Sep 21 | -0.26 \*\*\* |  |  |
|  | (0.02) |  |  |
| % Vaccinated Jan 22 |  | -0.27 \*\*\* |  |
|  |  | (0.02) |  |
| % Vaccinated April 22 |  |  | -0.26 \*\*\* |
|  |  |  | (0.02) |
| Population Density | 0.06 \*\*\* | 0.05 \*\* | 0.05 \*\* |
|  | (0.02) | (0.02) | (0.02) |
| Per capita income | -0.22 \*\*\* | -0.18 \*\*\* | -0.18 \*\*\* |
|  | (0.02) | (0.02) | (0.02) |
| % Population 65 and + | 0.20 \*\*\* | 0.21 \*\*\* | 0.21 \*\*\* |
|  | (0.02) | (0.02) | (0.02) |
| % White Population | -0.19 \*\*\* | -0.28 \*\*\* | -0.30 \*\*\* |
|  | (0.02) | (0.02) | (0.02) |
| N | 3046 | 3046 | 3046 |
| R2 | 0.22 | 0.22 | 0.21 |

Table 4: The Effect of Full Vaccination on COVID-19 Deaths Per 100K at the County Level

The regression results, as expected, show that vaccines reduce COVID deaths at an economically and statistically significant level. To precisely capture the effect of one standard deviation increase or decrease in vaccine penetration between two counties, we use the standard deviation of vaccination rates. The standard deviation of Vaccine penetration as of Sep 2021 as referred in Table 1 is 16.24, and the standard deviation of COVID deaths per 100K as referred in Table 2 is 139.027. Using the coefficients observed in the regression results, for one standard deviation increase in vaccine penetration, there is 0.26 decrease in COVID deaths.

As an illustrative example, as per our model,

1 standard deviation increase in vaccination rates => 0.29 fewer COVID Deaths per 100K (2)

=> 0.26 \* 1 standard deviation of Deaths per 100K

= 0.26 \* 139.02

= 36.14 fewer COVID Deaths per 100K

This says that comparing a county A with 16% higher vaccination rate as of Sep 2021 as compared to a county B, there were 36.14 fewer COVID deaths per 100K in county A. County A and County B here will hypothetically have similar income, race, population density and age distribution, thereby making them suitably comparable in every other way except for vaccination rates. Similar calculations applied to other two time periods of vaccine penetrations percentages show, with a 12.4% difference in vaccination rates as of Jan 2022, there were 37.53 fewer COVID deaths per 100K and a 12% difference in vaccination rates as of April 2022, there were 36.14 fewer COVID deaths per 100K. The coefficients are fairly similar at all three time periods, so the relationship is maintained all through the pandemic.

**4.2 The Effect of Republican Vote Shares on Vaccination Rates**

The next relationship to analyze as per the DAG created in Figure 7 is the relationship between political beliefs and Vaccine acceptance rates. The Republican share of voters in each county as explained in the data section is used for this analysis against the dependent variable vaccination percentages in the three selected time periods.

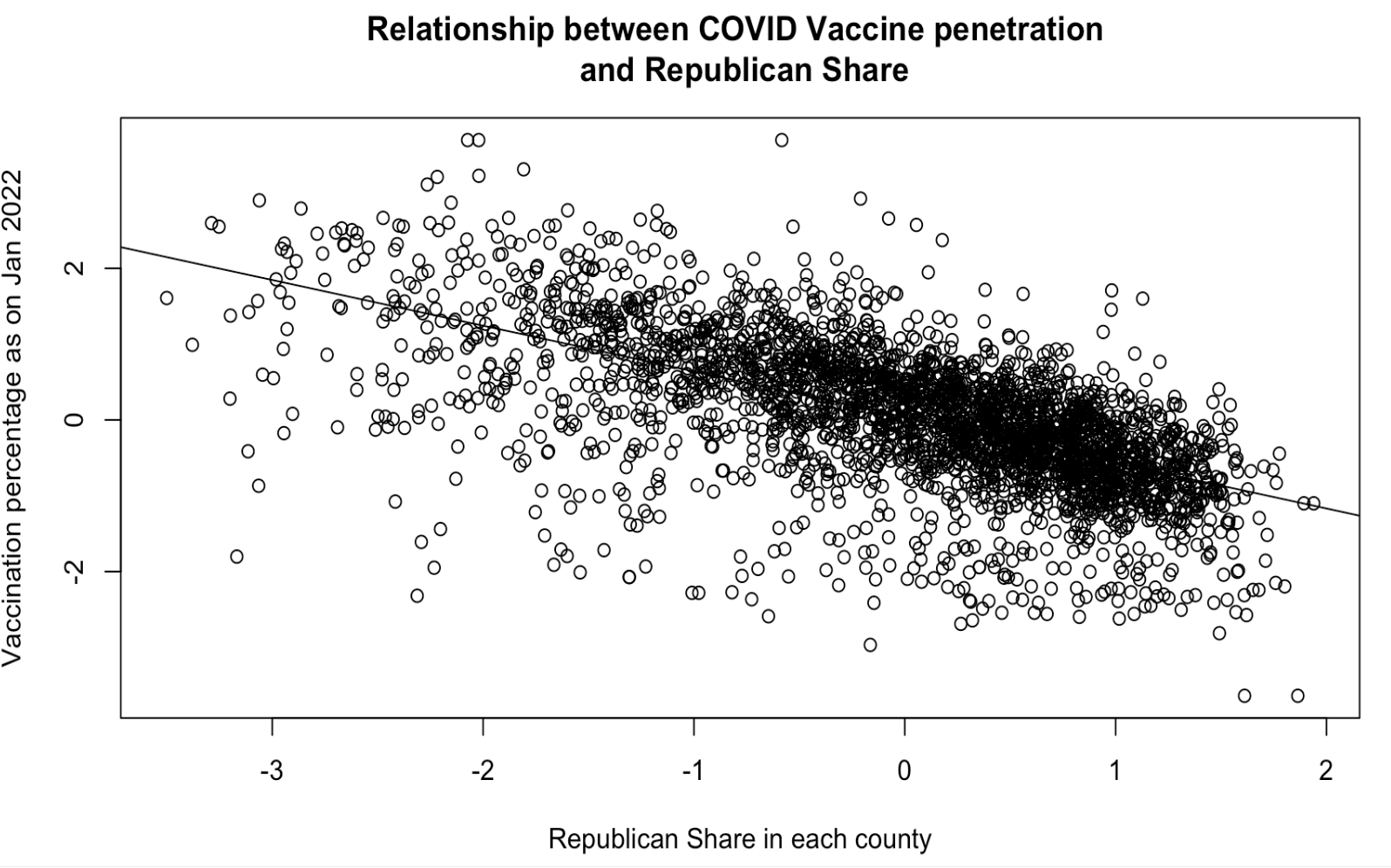


Figure 9: Correlation between Vote Share for the Republican Party and Fully Vaccinated % of Population as of January 2022

The graph above evidently shows the inverse relationship between Republican share in a county and the percentage of fully vaccinated. Higher the vaccination percentages captured in areas lower are Republican concentration in them. This graph answers the initial assumption made those political beliefs are an influential factor in the vaccination acceptance rates.

The regression model is again controlling for the demographic variables to have unbiased results when comparing counties with varied characteristics.

Table 5: The Effect of 2020 Republican Vote Shares on Vaccination Rates in September, January, and April[[3]](#footnote-3)

|  |  |  |  |
| --- | --- | --- | --- |
|  | Sep 2021 | Jan 2022 | April 2022 |
| Republican Vote Share | -0.74 \*\*\* | -0.64 \*\*\* | -0.69 \*\*\* |
|  | (0.02) | (0.02) | (0.02) |
| Population Density | -0.01 | -0.03 \* | -0.04 \*\* |
|  | (0.01) | (0.01) | (0.01) |
| Per Capita Income | 0.09 \*\*\* | 0.24 \*\*\* | 0.25 \*\*\* |
|  | (0.01) | (0.01) | (0.01) |
| % Population 65+ | 0.02 | 0.06 \*\*\* | 0.04 \*\*\* |
|  | (0.01) | (0.01) | (0.01) |
| % White Population | 0.54 \*\*\* | 0.17 \*\*\* | 0.09 \*\*\* |
|  | (0.02) | (0.02) | (0.02) |
| N | 3046 | 3046 | 3046 |
| R2 | 0.50 | 0.49 | 0.55 |

The variables included in the regression have been standardized, hence the results show the difference in 1 standard deviation change on either side of the regression equation. The negative coefficients against the Republican share are a clear indication of the inverse relationship between Republican shares and full vaccination rates. To evaluate the actual numbers of changes in vaccination percentages as on selected time periods, we use the standard deviation of Republican share across counties, 0.158. The standard deviation of the vaccination percentages as of Sep 2021 is 16%, Jan 2022 is 12.4% and April 2022 is 12%, as referred to in Table 1, which is used in the below equations. As illustrative examples,

1 σ increase in Republican share => 0.74 σ decrease in % vaccinated as on Sep 2021 (3)

= 0.74 \* 1 standard deviation of % vaccinated

= 0.74 \* 16.0

= 11.84 percentage point decrease in % vaccinated as on Sep 2021

1 σ increase in Republican share => 0.64 σ decrease in % vaccinated as on Jan 2022 (4)

= 0.64 \* 1 standard deviation of % vaccinated

= 0.64 \* 12.4

= 7.9 percentage point decrease in % vaccinated as on Jan 2022

1 σ increase in Republican share => 0.69 σ decrease in % vaccinated as on April 2022 (5)

= 0.69 \* 1 standard deviation of % vaccinated

= 0.69 \* 12.0

= 8.3 percentage point decrease in % vaccinated as of April 2022

To summarize, this study finds a potentially causal relationship between Republican vote shares and vaccination against COVID-19, at both statistically and economically significant levels. For a county that had a Republican vote share that was one standard deviation above the national average (15 percentage point lean toward the Republican party), its full vaccination rate goes down by 0.64 to 0.74 standard deviations (7.89 to 11.8 percentage point decrease) even after controlling for income, race, age, and population density.

**4.3 The Effect of Republican Party Support on COVID-19 Deaths Per 100K**

To further study the indirect relationship between Republican shares in counties and COVID deaths reported in those areas, the regression model included COVID deaths per 100K as the dependent variable against Republican share. The correlation plot to view this relationship is shown in Figure 10.

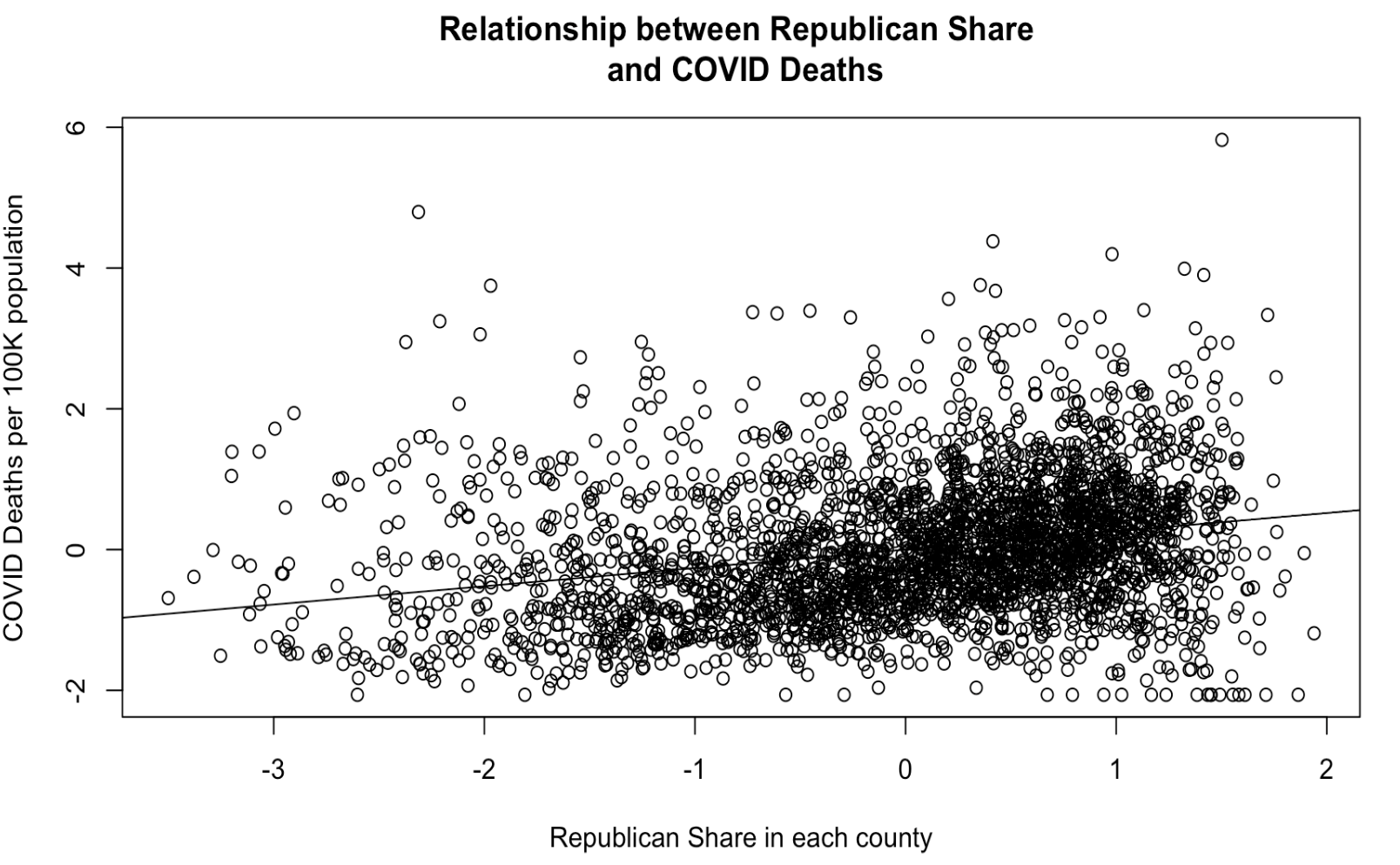


Figure 10: Correlation between Vote Share for the Republican Party and COVID Deaths Per 100K

Both variables in the correlation plot have been normalized, so we have to think about them as relative changes. There is a slightly positive relationship between deaths and Republican vote share based on this correlation plot. Similar to previous models run, the demographic control variables are added to the model presented below.

Table 6: The Effect of Republican Vote Share on COVID-19 Deaths Per 100K[[4]](#footnote-4)

|  |  |
| --- | --- |
|  | **Normalized COVID-19 Deaths Per 100K** |
| Republican Share | 0.39 \*\*\* |
|  | (0.02) |
| Population Density | 0.08 \*\*\* |
|  | (0.02) |
| Per Capita Income | -0.19 \*\*\* |
|  | (0.02) |
| % Population 65+ | 0.18 \*\*\* |
|  | (0.02) |
| % White Population | -0.42 \*\*\* |
|  | (0.02) |
| N | 3046 |
| R2 | 0.26 |

The positive coefficient against Republican share in the results of the model as shown in Table 6, shows that there is an increase in the number of COVID deaths per 100K. The effect of change in one standard deviation is evaluated using the below equation,

1 σ increase in Republican Share = > 0.39 increase in 1 σ of COVID deaths per 100K (6)

= 0.39 \* σ of COVID deaths per 100K

= 0.39 \* 139.02

= 54.21 increase in COVID deaths per 100K

This says that the 15-percentage point increase in Republican share results in an increase of 54.21 COVID deaths per 100K observed. The next step in the analysis is to study the effect of time trends, i.e., the change in Republican support between 2016 and 2020 on COVID-19 Deaths Per 100K while also differencing out the vaccination rates to tease out the causal impact of support for the Republican Party on COVID related fatality.

**4.4 Difference-in-Difference Estimates of Effects of Republican Support on Deaths**

The main goal of this section is to pressure-test the results discussed in Section 4.3 to check if they hold if controlled for the effect of time trends and vaccination rates. First, is the correlation plot between the change in support for the Republican party between the 2016 and 2020 elections and COVID-19 Deaths in December 2021.

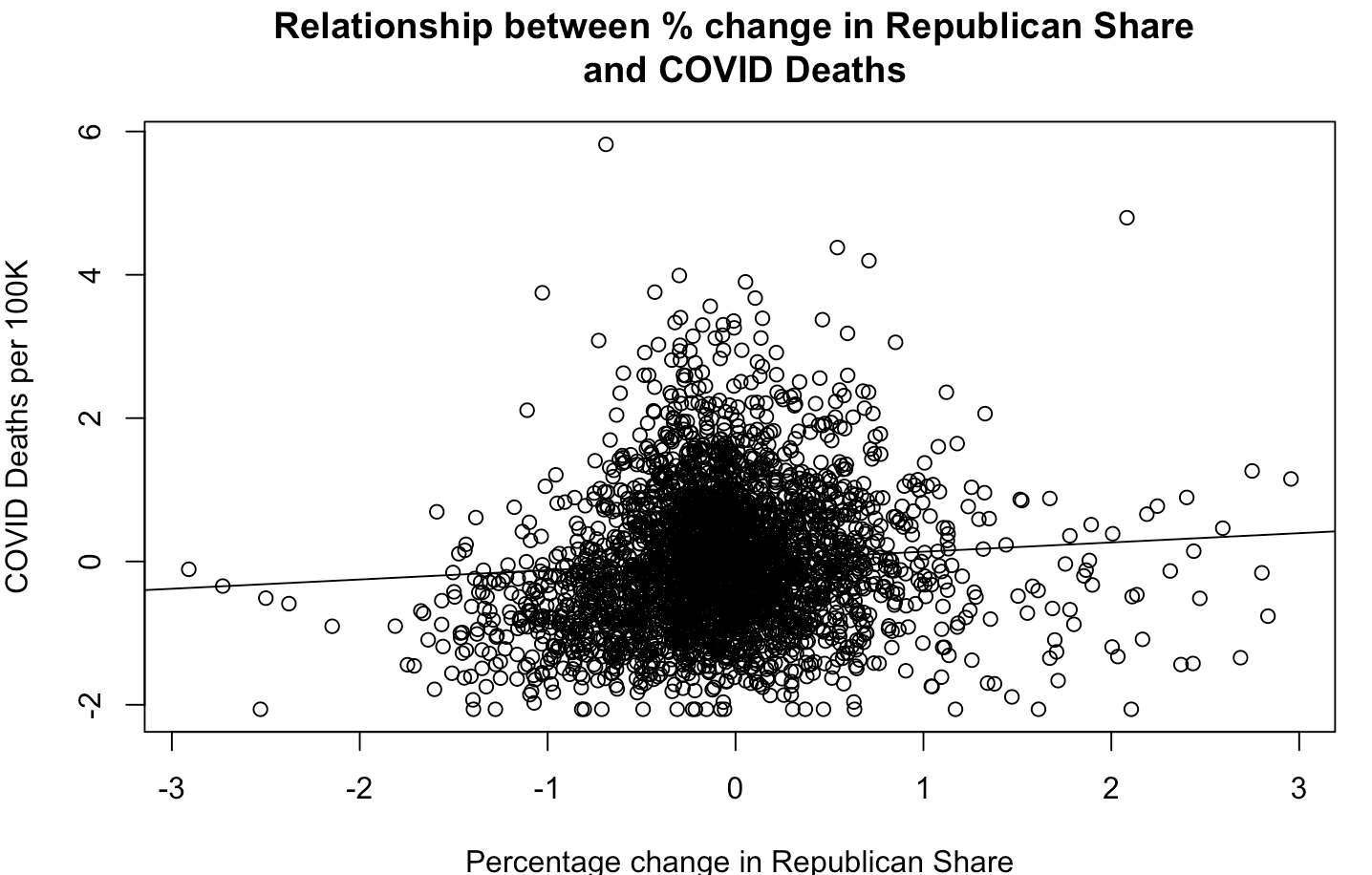


Figure 11: Correlation between Change in Republican Share from 2016 -2020 and COVID-19 Deaths as of December 2021

The plot above shows the only mild correlation between changing vote shares and COVID-19 fatality. As the distribution of variation in changing vote shares is concentrated around 0, the correlation is not visible to the naked eye. To investigate the causal relationship, the primary model for the analysis are normalized difference-in-difference regression (7) where the dependent variable is deaths per 100K per county c, which is modeled as a function of intercept , the percentage difference in Republican vote share between 2016 and 2020 presidential elections , vaccination rates as of January 2022, the interaction term between change in the Republican vote share and the vaccination rate, which gives us the difference-in-difference estimate , demographic control vector X, which includes county-level characteristics like income distribution, age distribution, race distribution, population density, etc., and the idiosyncratic error term .

(7)

In addition to the model above, a second difference-in-difference model (8) is run using only the 2020 Republican vote share to illustrate the effect of the time trend captured through the change in Republican share in (7). The variables in (8) mean the same as in (7) except that they only account for the 2020 vote share for the Republican party.

(8)

The table below shows the results of the models (7) and (8).

Table 7: Difference-in-Difference Estimates of the Effect of Republican Party Support on Normalized COVID-19 Deaths Per 100K[[5]](#footnote-5)

|  | Normalized COVID Deaths Per 100K | Normalized COVID Deaths Per 100K |
| --- | --- | --- |
| 2020 Republican Share | 0.32 \*\*\* |  |
|  | (0.02) |  |
| % Change in Republican Share 2016-2020 |  | 0.02 |
|  |  | (0.02) |
| Population Density | 0.08 \*\*\* | 0.03 \* |
|  | (0.02) | (0.02) |
| Per Capita Income | -0.15 \*\*\* | -0.17 \*\*\* |
|  | (0.02) | (0.02) |
| % Population 65 and + | 0.18 \*\*\* | 0.21 \*\*\* |
|  | (0.02) | (0.02) |
| % White Population | -0.41 \*\*\* | -0.26 \*\*\* |
|  | (0.02) | (0.02) |
| % Vaccinated as of Jan 2022 | -0.09 \*\*\* | -0.27 \*\*\* |
|  | (0.02) | (0.02) |
| 2020 Republican Share \* | 0.04\* |  |
| % Vaccinated as of Jan 2022 | (0.02) |  |
| % Change in Republican Share 2016-2020 \* |  | 0.04\* |
| % Vaccinated as of Jan 2022 |  | (0.02) |
| N | 3046 | 3046 |
| R2 | 0.27 | 0.23 |

The difference-in-difference models shown above control for two differences: one is the vaccination rates across counties in December 2021, and the second, is the key control difference in Republican vote share from 2016 to 2020. Now the hypothetical county A’s treatment is a one standard deviation increase in Republican vote share from 2016 to 2020 measured against county B which had its Republican vote share change at the same rate as the mean change. If the vaccination rates between county A and county B are constant, then the marginal effect of increasing Republican share on COVID-19 deaths per 100K is not significantly different from zero. The marginal effect of a one standard deviation increase in vaccination rates is associated with a 0.27 standard deviations fewer COVID-19 deaths per 100K when comparing between counties with the same change in Republican support.

The difference-in-difference estimator shows for a county with both one standard deviation increase in Republican vote share from 2016 and comparatively one standard deviation above the mean vaccination rates as of January 2022, there are 0.04 standard deviations more COVID-19 deaths per 100K. This shows that even accounting for high vaccination rates, time trends in vote shares for the Republican Party is significantly related to more COVID-19 deaths, which have to be related to lower COVID-19 risk perception stemming from political beliefs. This finding is the first of its kind and shows the causal link between political beliefs and fatality during the COVID-19 pandemic. This result is a key contribution to the literature which has thus far focused on survey data on beliefs about COVID-19 and effects on perceptions on vaccination but has not linked political beliefs to deaths.

**5 Conclusion and Recommendations**

This study presented evidence that showed support for the Republican party is causally linked to COVID-19 deaths. There might be other causal mechanisms correlated with support for the Republican party that was not controlled for like generally more risky behavior, general mistrust in science, general mistrust of elite mass media with denial to abide by mask mandates, but the framing of the phrase “political beliefs” in this paper encompasses these trends. In addition, since in Section 4.4, only the increase in support between 2016 and 2020 is used, these other variables within the county are accounted for, at least in theory. The other omitted variables may include specific dates when vaccine mandates were announced, whether the local leaders are democrats or republicans, viewership of the Fox News, viewership of content related to conspiracy theories against risks of COVID. As these variables are heavily correlated with political beliefs which we are controlling as a factor of Republican share in each county, adding these variables to a future study will only add stronger evidence to the results obtained. There are other ways to corroborate the analysis in this study, including performing a difference-in-difference analysis using variations in announcements of COVID restrictions in neighboring counties and using trends in vaccination rates to identify before and after periods.

Using public data from the CDC, MIT Election Lab, the Bureau of Economic Analysis, and the US Census Bureau, this study finds that support for the Republican party is associated with lower vaccination rates and higher COVID-19 fatality rates. Even after accounting for time trends and vaccination rates, support for the Republican party is shown to increase COVID-19 deaths at a statistically significant level, which confirms the hypothesis that political beliefs led to more deaths during the COVID-19 pandemic.

**References**

(Adolph et al., 2021) Adolph C, Amano K, Bang-Jensen B, Fullman N, Wilkerson J (2021) Pandemic Politics: Timing State-Level Social Distancing Responses To COVID-19. J Health Politics Policy Law 46(2):211–233

(Ajzenman et al., 2020) Ajzenman N, Cavalcanti T, Da Mata D (2020). More Than Words: Leaders’ Speech And Risky Behavior During a Pandemic. SSRN 3582908

Albers. (n.d.). Retrieved April 28, 2022, from https://pro.arcgis.com/en/pro-app/latest/help/mapping/properties/albers.htm

(Allcot et al., 2020) Allcott H, Boxell L, Conway J, Gentzkow M, Thaler M, Yang D (2020) Polarization and Public Health: Partisan Differences In Social Distancing During The Coronavirus Pandemic. J Public Econ 191:104254

(Barrios and Hochberg, 2021) John M. Barrios, Yael V. Hochberg, Risk Perceptions And Politics: Evidence From The COVID-19 Pandemic, Journal Of Financial Economics, Volume 142, Issue 2, 2021, Pages 862-879, ISSN 0304-405X

(​​Carter et al., 2021) Widespread severe acute respiratory syndrome coronavirus 2 transmission among attendees at a large motorcycle rally and their contacts, 30 US jurisdictions, August–September, 2020. Retrieved April 28, 2022, from <https://doi.org/10.1093/cid/ciab321>

CDC Covid Data tracker. (n.d.). Retrieved April 28, 2022, from <https://covid.cdc.gov/covid-data-tracker/#county-view?list_select_state=all_states&list_select_county=all_counties&data-type=Vaccinations&metric=Administered_Dose1_Pop_Pct&null=Vaccinations>

Covid-19 vaccinations in the United States,county. (n.d.). Retrieved April 28, 2022, from <https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh>

(Freira et al., 2021) Freira, L., Sartorio, M., Boruchowicz, C. et al. (2021) The Interplay Between Partisanship, Forecasted COVID-19 Deaths, And Support For Preventive Policies. Humanit Soc Sci Commun 8, 192 (2021).<https://doi.org/10.1057/s41599-021-00870-2>

(Ritchie et al., 2020). Coronavirus pandemic (COVID-19). Retrieved April 28, 2022, from <https://ourworldindata.org/coronavirus>

(Sabahelzain et al., 2021) Majdi M Sabahelzain, Kenneth Hartigan-Go, Heidi J Larson,The Politics Of Covid-19 Vaccine Confidence, Current Opinion In Immunology, Volume 71, 2021, Pages 92-96, ISSN 0952-7915

(Timoneda and Vallejo Vera, 2021) Timoneda JC, Vallejo Vera S (2021) Will I Die Of Coronavirus? Google Trends Data Reveal That Politics Determine Virus Fears. PLoS ONE 16(10): e0258189.<https://doi.org/10.1371/journal.pone.0258189>

(Wood and Schulman, 2021) Stacy Wood and Kevin Schulman. 2021. Beyond Politics — Promoting Covid-19 Vaccination In The United States, New England Journal of Medicine, February 2021.<https://doi.org/10.1056/NEJMms2033790>

1. I’d like to acknowledge the guidance of Dr. Niranjani Patel, my family, and friends for aiding me in completing this research article. I’d also like to acknowledge San Jose State University and the Department of Industrial & Systems Engineering for the opportunity to perform original research as part of the Engineering Management program. [↑](#footnote-ref-1)
2. In later stages of cumulative vaccination rate data collection, counties with vaccination rates above 95% were coded as 95%, whereas in the previous rounds in June and September 2021, counties above 95% were coded accurately, i.e., the actual rate of vaccination. As you can see in the first two rows of Table 1, the maximum rate of full vaccination was 99.9% - in counties with a high proportion of protected-class government employees (military bases). This classification change did not affect a large portion of our data, (less than 0.1 percent of the sample). [↑](#footnote-ref-2)
3. All control variables, including Republican Share, Population Density, Per Capita Income, % Population 65+, % White Population have all been normalized. [↑](#footnote-ref-3)
4. All control variables, including Republican Share, Population Density, Per Capita Income, % Population 65+, % White Population have all been normalized. [↑](#footnote-ref-4)
5. All control variables, including Republican Share, Population Density, Per Capita Income, % Population 65+, % White Population have all been normalized. [↑](#footnote-ref-5)