Distribution Agreement

Xinxuan (Phyllis) Guo December 7, 2021

The Analysis of COVID-19 Physical and Mental Health Impacts in the United States

By

Xinxuan (Phyllis) Guo

Seunghwa Rho

Adviser

The Department of Quantitative Theory and Methods

Seunghwa Rho

Adviser

Jessica Barber

Committee Member

Shivani A. Patel

Committee Member

2021

The Analysis of COVID-19 Physical and Mental Health Impacts in the United States

By

Xinxuan (Phyllis) Guo

Seunghwa Rho

Adviser

An abstract of

a thesis submitted to the Faculty of Emory College of Arts and Sciences

of Emory University in partial fulfillment

of the requirements of the degree

Bachelor of Science with Honors

The Department of Quantitative Theory and Methods

2021

Abstract

The Analysis of COVID-19 Physical and Mental Health Impacts in the United States

By Xinxuan (Phyllis) Guo

As the COVID-19 pandemic has been ongoing for about two years now, people around the globe have gradually adapted to the ”new normal” of wearing masks, social distancing, quarantine and lockdowns. This project aims to explore the physical and mental health impacts of COVID-19 in the United States, and how they collectively impact the well-being of American people. To identify factors associated with physical health and vaccination progress, linear regressions were applied to state-level dataset comprising national COVID-19 vaccine completion rate and coronavirus-related mortality rate. We retrieved data from the Centers for Disease Control and Prevention and 2018 American Community Survey. What we found out is that Americans with underlying diabetes disease and those who have seen an increasing number of deaths and positive cases around them seem more likely to receive vaccines over time, and the three types of vaccines administered within the U.S. are playing a role in reducing the death and case numbers.

In terms of psychological health, past studies have shown the significant mental and emotional impacts such as anxiety, depression, uncertainty and stigma brought by a global pandemic when the cause or progression of the disease and outcomes are unclear. In current study, the researchers analyzed Google Trends and national suicide rate data as two ways to measure the Americans' psychological stability during the pandemic. The result indicates that although the national suicide rate has decreased in year 2020 and the popularity of depression-related search terms haven't changed much, we can't easily draw conclusion that COVID-19 hasn't mentally impacted the Americans. There're many social movements happening at the same time with the COVID-19 outbreak, which could be important confounding variables. For future study, these variables need to be controlled in order to find out the real psychological fluctuation happened in the U.S.

The Analysis of COVID-19 Physical and Mental Health Impacts in the United States

By

Xinxuan (Phyllis) Guo

Seunghwa Rho

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences

of Emory University in partial fulfillment

of the requirements of the degree

Bachelor of Science with Honors

The Department of Quantitative Theory and Methods

2021

Acknowledgement

I would like to thank my committee chair and advisor, Dr. Seunghwa Rho, for guiding me and helping me refine this project to its current state. Her continual encouragement and insightful comments, especially the more critical ones, pushed me to critically think about these research questions and their implication, and ultimately brought this research work to a higher level.

I would also like to thank the other committee members, Dr. Jessica Barber and Dr. Shivani A. Patel, for their valuable advice and enlightening comments. Our continuous discussions throughout the project brought to light new avenues of inquiry.

Finally, I would like to thank my family and friends who have emotionally and mentally supported me throughout my project. I'd like to especially thank my parents. As it turns out, it was extremely hard to get back to work after losing loved ones during the pandemic and I would like to devote my work to my father who has been fighting against cancer like a warrior until the end of life, and my mother who has been going through the most difficult time with me together and always unconditionally supports me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

The present coronavirus disease 2019 (COVID-19) was first identified in December 2019 in Wuhan, China, and has resulted in an ongoing pandemic and a havoc on the human civilization. As of September 30, 2021, the U.S. death toll from COVID-19 has surpassed 721,000 and more than 44 million Americans have been infected. With the efforts of researchers and scientists, on December 11, 2020, the U.S. Food and Drug Administration (FDA) issued the first emergency use authorization (EUA) for use of the Pfizer-BioNTech COVID-19 vaccine in persons aged 16 years and older for the prevention of COVID-19 and the first deliveries of Pfizer vaccines began on December 14. The Moderna vaccine was later issued on December 18 for individuals aged 18 years and older, and began its distribution soon afterwards. On February 27, 2021, the FDA issued the EUA for the Janssen vaccine, which was the third vaccine for COVID-19 prevention and only required 1 dose, but the FDA later included information about a very rare and serious type of blood clot in people who received Janssen vaccine in the EUA. All three vaccines authorized in the U.S. have been proved to reduce the risk of COVID-19 infections and are highly effective at preventing severe illness and death from the disease (Lipsitch, Dean, 2020). However, although the vaccines brought a promising layer of protection to Americans (Forni, Mantovani, 2021), it soon became a problem when the COVID-19 vaccine allocation started in

January in the country (Islam et al., 2021). In mid-January, Centers for Disease Control and Prevention (CDC) began distributing COVID-19 vaccinations to the general population 65 years of age and over and individuals with underlying medical condition, which are two groups of population that're considered as high-risk individuals to coronavirus. Within days, vaccine sites reported that vaccines were not being administered to local communities, but rather to more advantaged people traveling to their sites from surrounding counties (Goldhill, 2021). This misallocation of vaccines created racial, ethnic, and socioeconomic disparities, and could potentially worsen the COVID-19 mortality rate as the population who are more vulnerable can't receive treatment more efficiently and quickly (Jean-Jacques, Bauchner, 2021). Therefore, this paper analyzes and shows a big picture of the willingness of individuals with existing medical conditions to receive COVID-19 vaccination from an aggregated, state level dataset, and explores how COVID-19 mortality rate and vaccination rate in the U.S. have affected each other. In other words, I would like to examine the effect of COVID-19 vaccines on coronavirus-related mortality rate within the U.S. to find out the vaccines' efficacy.

In addition, the highly contagious nature of COVID-19 has not only caused huge public health emergencies such as large number of infections, health care system overload, and lockdown in many states (Hartnett, Kite-Powell et al., 2020), but also significantly impacted people's psychological well-being (Bojdani, Rajagopalan et al. 2020). Decreased social interaction during lockdown and constant worry and uncertainty brought by the pandemic can easily cause symptoms of major depressive disorder, mood disorder or anxiety disorder (Usher, Durkin, and Bhullar, 2020). Prior studies have shown that the lockdown and quarantines are positively correlated with the feelings of insecurity, confusion, emotional isolation and stigma in individuals, which might further impact the population's mental health condition from the country level (Dubey, Biswas, et al., 2020). In this case, this paper also examines how the COVID-19 pandemic has impacted Americans' mental health, specifically in year 2020.

1.2 Motivation

Given facts that coronavirus is novel and COVID-19 protocols such as wearing masks and city lockdown are uncommon in Americans' daily lives, how they have collectively influenced people's physical and mental health from an aggregated, state level remains largely unknown. Current studies showed that COVID-19 mortality rate is negatively associated with test number and government effectiveness from analyzing the global COVID-19 data (Liang, Tseng, Ho, and Wu, 2020). There has also been studies using linear regressions to predict the number of COVID-related deaths from a global perspective (Ghosal, Sengupta, Majumder, and Sinha, 2020). Researches that are specifically focused on analyzing the pandemic in the U.S. are limited and it is unclear how Americans' physical and mental health have been impacted. Better understanding of this question provides insight on how to encourage more Americans to receive vaccines and how to improve our current psychological health services during extreme situations. Existing pandemic impact measures of current studies mostly assess the external factors reflected by the population's behavior such as the number of vaccine dose 1 distributed. For example, according to CDC, as of October 17, 2021, there are 57.2% of the U.S. population who are fully vaccinated, and 66.1% have received at least 1 dose (Centers for Disease Control and Prevention, 2021). Validated measures for assessing individual's internal intention and willingness to receive COVID-19 vaccines have not yet established in these researches. Given the nascent stage of research on this topic, exploratory studies that take a bottom-up approach — using available data to generate insights rather than answer a specific question — may be particularly useful.

The objective of this project is from two perspectives. From the vaccination progress and physical health aspect, as prior studies and CDC have proved that certain existing medical conditions will make individuals more vulnerable to coronavirus and more easily develop serious symptoms (Erener, 2020), the researcher would like to explore how existing medical conditions play a role in making decisions to receive COVID-19 vaccines for people who have past or underlying medical history. Understanding this questions will help the doctors and scientists better protect the vulnerable groups and offer them effective incentives to receive the vaccines. In addition, the researcher is interested in whether there're any associations between population's willingness to receive vaccine and how COVID-19 mortality has changed around them. The underlying hypothesis is if people see the increasing number of coronavirus-related cases and deaths around them, they're more likely to get vaccinated to protect themselves and their family and friends. This study might sound intuitive but the researcher would like to test the hypothesis with a concrete dataset to find out if the reality aligns with our expectations. Finally, I am interested in how vaccines have helped reduce COVID-19 mortality increment so far by running regression models and finding the association between COVID-19 mortality and the vaccine increment in certain time periods. The last study examines the effectiveness and efficacy of the COVID-19 vaccines at different national vaccination rates, or different time periods.

In terms of mental health, it's proved that the COVID-19 pandemic is associated with distress, anxiety, fear of contagion, depression and insomnia in the general population (Killgore, Taylor, Cloonan, and Dailey, 2020). Social isolation, anxiety, fear of contagion, uncertainty, chronic stress and economic difficulties may lead to the development or exacerbation of depressive, anxiety, substance use and other psychiatric disorders in vulnerable populations including individuals with pre-existing psychiatric disorders and people who reside in high COVID-19 prevalence areas. Stress-related psychiatric conditions including mood and substance use disorders are associated with suicidal behavior

(Sher, 2020). Therefore, the researcher delve into the question that during year 2020, how the U.S. Google key term search trend and national suicide rate have changed, which I assume are two representations of the population's mental health condition. For Google search trend, three key words are picked, including ”depression”, ”anxiety”, and ”mental health”. I'm interested in seeing from 2016 to 2020 the search popularity trend changes of these three terms. The hypothesis is if we see an increase in year 2020 compared to other year, then we can say Americans' mental condition were influenced during the pandemic. However, we cannot make causal conclusions that the mental health fluctuation is due to the pandemic, because there are some important confounding variables I couldn't control for in the study due to data availability. Therefore, like mentioned before, this research is by no means a causal-analysis study, but simply offers us some descriptive data and a big picture of what's going on in the country's mental health. Future studies and richer datasets are needed in order to find out the causal relationship between the COVID-19 pandemic and Americans' psychological health. By doing so, it provides a better understanding for factors that might influence people's psychological state during the pandemic with some more robust evidence.

Chapter 2

Data Description

2.1 Vaccination and Physical Health

With global health pandemics like COVID-19, the population's physical health such as positive case number and mortality rate over time is always of interests, so in this project, the researcher is interested in finding out how the coronavirus and pandemic have influenced the U.S. mortality rate and if the vaccine has played a role in saving people's life. Therefore, the first aspect we would like to examine is the physical health impact of COVID-19, specifically, the death and case number trending behaviors and how the vaccine has influenced them.

2.1.1 Data Source

The U.S. COVID vaccination data in this project was collected from the CDC COVID-19 Vaccinations in the United States. It is daily data based on county level. In this project, I aggregated the county-level data to state-level in order to match the level of demographic information that is at the state level. Data was collected in a CSV file and analysed in Python 3.8.2 software and R Studio. In this dataset, the available variables include geographical information (state, county), time series information (year, month, day), and vaccine series completion rate in different age groups (older than 12-year-olds, 18-year-olds, and 65-year-olds). The time span is from January 18, 2021 to October 17, 2021. In the vaccination dataset, there are two vaccine variables included, vaccine distributed and vaccine administered, which are worth clarifying their meanings here. Vaccine distribution is the process of shipping vaccines to provider locations, as directed by jurisdictions, Federal agencies, and pharmacy partners who are enrolled in the COVID-19 Vaccination Program. Vaccine delivery is the last part of the distribution process, which represents the vaccine doses that have arrived at their destination. Whereas vaccine administration refers to vaccines that are administered by public health jurisdictions, Federal entities, healthcare providers, long-term care facilities, employers, retail pharmacies, and other businesses to various populations throughout the U.S. This includes managing vaccine inventories, tracking vaccine doses given to recipients, creating vaccine records, scheduling vaccine appointments, sending appointment reminders, and other administrative functions (CDC COVID-19 Vaccinations in the United States, 2021). In this project, I chose to use vaccine administered variable as it more accurately reflects the number of population who actually received the vaccine.

The U.S. COVID-19 death data was collected from CDC COVID Data Tracker, which is daily data based on county level. In current project, it was aggregated and analyzed from state level as well in order to match demographic information. The time span is from April 1, 2020 to October 17, 2021.

The U.S. demographic information data was collected from 2018 American Community Survey by the U.S. Census Bureau. It is a state-level demographic dataset including gender ratio, poverty level, age group proportion and many other variables.

All available data points are further explained in Table 2.1: Physical Health: Existing Data Overview.

Variable Name	Definition	Time series	Cross section
COVID mortality	Total COVID-related deaths per 100K population in the state	Daily	county-level
Dose 1 cumulative	Percent of population who received first dose of vaccines cumulatively	Daily	county-level
Dose 2 cumulative	Percent of population who received second dose of vaccines cumulatively	Daily	county-level
Obesity	Percent of the population with obesity in the state	NA	state-level
Heart Disease	Percent of the population with heart disease in the state	NA	state-level
Chronic Kidney Disease	Percent of the population with chronic kidney disease in the state	NA	state-level
Age65over	Percent of population over 65 years old in the state	NA	state-level
Poverty	Percent of population in poverty in the state	NA	state-level

Table 2.1: Physical Health: Existing Data Overview

2.1.2 Preprocessing

In order to make the datasets consistent with the demographic data which is state level, the researcher chose to aggregate vaccine and death data on the state level. I then combined daily, state-level vaccine and death data together. On top of that, demographic information such as race, gender, age and poverty of each state was merged.

Missing data (NA) existed in the vaccination dataset for some days. Therefore, instead of using the daily data, the researcher cuts data by national vaccination percentages to represent time. I selected 10 data points, or 10 snapshots of the country. In other words, 10 daily snapshots were picked on the state level, which would be the snapshot up to a specific day when the national vaccine rate is 1%, 5%, 10%, 15%, 20%, 25%, 30%, 40%, 50%, 55%. I then subtracted between two vaccination periods to find out the increment of vaccines and deaths during a certain time frame. By doing this, it allows us to examine how vaccine distribution and COVID mortality rate have changed during a specific time period, such as a time when the national vaccine rate increased from 30% to 40%.

The underlying reason of doing this is because as prior studies suggest, when measuring vaccine effectiveness, there are two complementary forms of evaluation. One is measuring direct effect of the vaccine on the vaccinated individual; the other is measuring the overall effect of the vaccination program on an entire population (Rossman, Shilo, Meir, Gorfine, Shalit, Segal, 2021). Most past researches have been focusing on the vaccine effectiveness from individual perspective. To our knowledge, very few studies thus far have analyzed the effect of the vaccination campaign on the patterns of pandemic dynamics at the population level. As the U.S.implements a COVID-19 vaccine campaign on national level, I think that this quantification might be of interest to explore. In their study, Rossman, Shilo et al. (2021) examined the national immunization program in Israel, concluding that for individuals aged 60 years and older, the efficacy of vaccines started to show only after the population vaccine completion rate reaches 85% (both doses). In Israel, after two months into the vaccine program where the national vaccination rate reached 30%, there was an approximately 77% drop in cases, a 45% drop in positive test percentage, a 68% drop in hospitalizations and a 67% drop in severe hospitalizations compared to peak values. This implies that there are important turning points for national vaccine completion rate.

This new dataset was combined with the demographic data and can be found in Table 2.2: Physical Health: Created Data Overview.

Variable Name	Definition
National fully vaccinated percentage	percentages of 1, 5, 10,15, 20, 25, 30, 40, 50, 55
Date	Jan. 25, Feb. 18, March 10, March 26, April 9, April 16, April 29, May 26, Aug. 2, Sept. 20 (in year 2021)
Covid mortality increment	COVID-related death increments per 100K population between 2 vaccination periods
Dose 1 increment	Percent of additional population who received first dose of vaccines between 2 vaccination periods
Dose 2 increment	Percent of additional population who received second dose of vaccines between 2 vaccination periods

Table 2.2: Physical Health: Created Data Overview

2.2 Mental Health

Another research interest of this project is how the pandemic influenced population's mental health. The uncertainty caused by the public health emergencies might affect people's emotional stability, causing unhealthy behaviors (such as excessive substance use and depressive symptoms) in people who contract the disease and in the general population. Extensive research in disaster mental health has established that emotional distress is ubiquitous in affected populations — a finding certain to be echoed in populations affected by the COVID-19 pandemic (Pera, 2020). The search interest of mental health related terms on Google before and after the outbreak of COVID-19 pandemic reveals how public's concern is affected by the pandemic, and its impact to mental health of people in the U.S. The COVID-19 crisis may also increase suicide rates during and after the pandemic (Sher, 2020). Mental health consequences of the COVID-19 crisis including suicidal behavior are likely to be present for a long time and peak later than the actual pandemic.

2.2.1 Data Source

The Google mental health search term scores were collected from Google Trends. The researcher selected three mental health related search terms to represent how Americans' mental health fluctuated over time, which are ”mental health”, ”depression” and ”anxiety”. The rationale of selecting these three terms specifically is based on a prior study (Knipe, Gunnell, Evans, John, Fancourt, 2021). Knipe et al. used relative search volumes (RSV) for the topics depression, anxiety, self-harm, suicide, suicidal ideation, loneliness, and abuse, which were obtained from Google Trends to find out how these search terms associate with people's mental health. The reason why I only chose three terms instead of all seven terms in current project is because the Google Trends dataset is not rich and large enough to run an equation with seven regressors. In order to preserve the accuracy of results, I decided to selected only three terms for the study.

The search trends are from October 2016 to September 2021 on the country level. The Google Trend search score does not represent the actual search volume numbers, but rather an index ranging from 0-100. The numbers represent the search interest relative to the highest point in the list of the selected terms in a certain region and time. A value of 100 is the peak popularity of the term, whereas a value of 50 means that the term is half as popular.

Scores of 0 mean that a sufficient amount of data was not available for the selected term. Google Trends normalizes search data to make comparisons between terms easier.

The U.S. suicide rate quarterly data was downloaded from National Center for Health Statistics Mortality dashboard, which contains the quarterly national suicide rate from 2017 to 2020.

2.2.2 Preprocessing

It's true that population's mental health fluctuates during a year. Figure 2.1 shows how the popularity of three search keywords changed over past 5 years, indicating that there seems to have seasonality issues every year and we can see a pattern. For example, if we look at the green line that represents "mental health" search trend in the graph, it seems like there're drops at the end of each year. One possible explanation could be due to the holiday season – people are able to have a break from their work and gather together with family and friends, which helps reduce their mental and emotional stress. In this case, in order to take care of seasonality issues for Google trend search, 4 quarterly binary variables were added to control for the seasonal change.

Figure 2.1: Changes of search terms in last 5 years

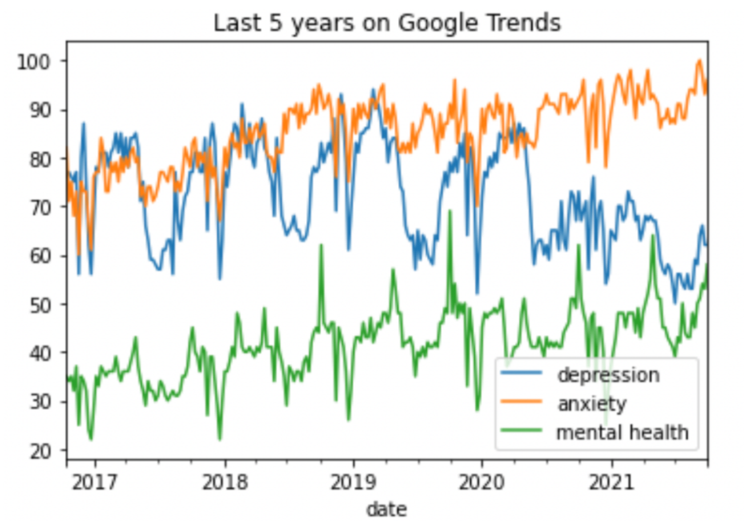


Table 2.3 presents the existing data in the mental health dataset.

Variable Name	Definition	Time series
Date	Weekly data from 2016/10 to 2021/9	Weekly
Depression	Google trend search score for keyword “depression” (out of 100)	Weekly
Anxiety	Google trend search score for keyword “anxiety” (out of 100)	Weekly
Mental Health	Google trend search score for keyword “mental health” (out of 100)	Weekly
Mortality Rate	Deaths per 100K population during the quarter	Quarterly
Q1-Q4 binary variables	1 when quarter is 1, 0 when quarter is 2,3,4	Quarterly
COVID binary variable	1 when year is 2020, 0 when it’s other time periods	Quarterly

Table 2.3: Mental Health: Existing Data Overview

Chapter 3

Vaccination and Physical Health

3.1 Study 1: Vaccination and Medical History

3.1.1 Methods and Models

When distributing the vaccines, it is important to make sure the vulnerable population receives it first. Therefore, during the vaccine rollout, was state-level vaccination coverage responsive to population's medical vulnerability? Did people with underlying medical conditions react to get more vaccines? Simple linear regressions were first applied to investigate the correlation between COVID-19 vaccine increment and individuals with pre-existing medical conditions, because the increment offers a better understanding of how people have reacted within certain time frames. The goal was to examine whether the state as a whole reacted to receive more vaccines, which to some extent reflected the intentions of individuals with underlying medical conditions to get vaccines. The equation below reflects whether people with underlying medical conditions are recognizing they should receive vaccines or not. The i subscript in the equation implies state and t subscript implies time. This equation was measured for each time period or at each vaccination completion rate for 10 times, resulting 10 correlation coefficients for each medical condition.

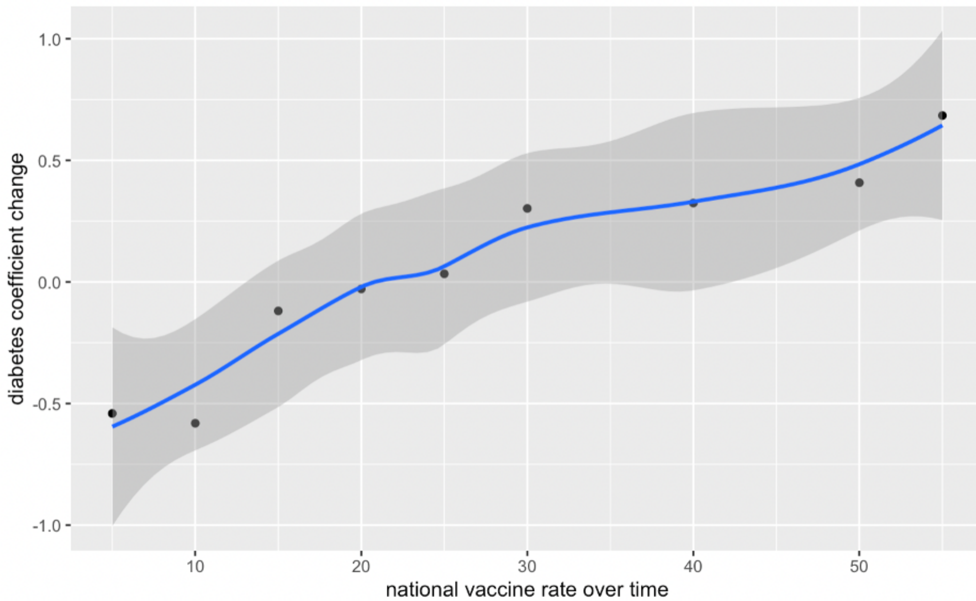
$$\text{Dose1Inc}_{it} = \beta_0 + \beta_1 obesity_i + \beta_2 heart_i + \beta_3 diabetes_i + \beta_4 kidney_i + u_{it}$$

Correlation coefficients and the p-value of coefficients for independent variables were calculated and reported.

3.1.2 Results

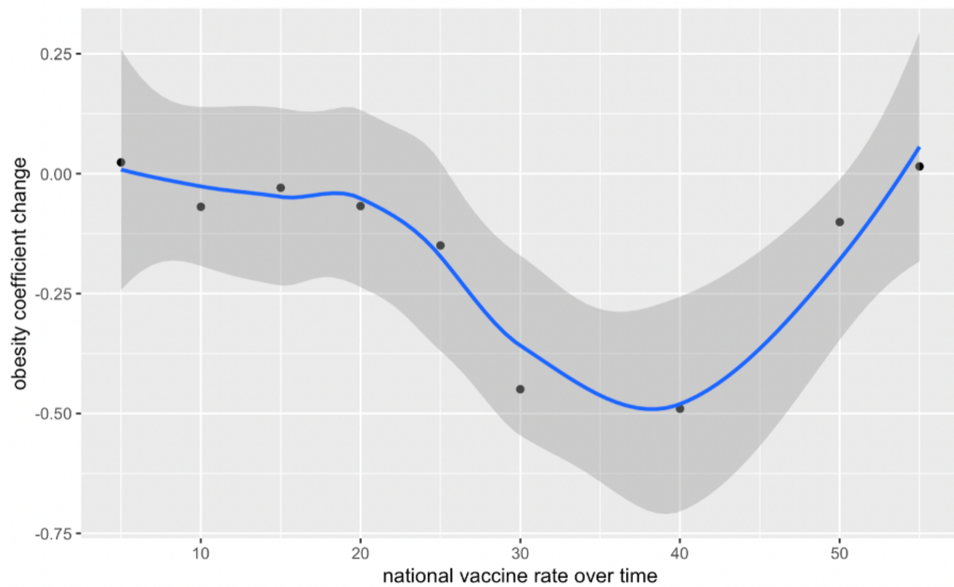
Based on the result, we can tell that for diabetes, the trend is increasing over time with β values starting from negative (-0.54) to positive value (0.68). This implies that as time goes by, the state with large proportion of diabetes patients are reacting and sending out more COVID-19 vaccines. Although the number here is practically small, but considering not everyone in the state has diabetes, it shows that individuals with diabetes are recognizing they're more vulnerable so they take actions to get vaccines, and that's why the trending behavior is towards positive. This trend can be seen from Figure 3.1. It has the cumulative national vaccine rate as x axis, which also reflects time as the rate increases, and the β value over time as y axis.

Figure 3.1: Diabetes coefficient changes over time



For obesity, the beta value is negative for most part, and as for most part it is not significant as the confidence interval contains 0. The negative beta value over time shows that states with large proportion of obesity seem to have decreasing increment of vaccine, implying they didn't show reactions to send out more vaccines. However, it's worth noticing that since the result is not significant, we can't find strong evidence to draw the conclusion.

Figure 3.2: Obesity coefficient changes over time



For heart disease and chronic kidney disease, the confidence intervals contain 0 for most part, implying that the changes here are never really significant. Therefore, there's no strong evidence to show that states with high proportion of heart disease and chronic kidney disease are reacting to send out more vaccines.

Figure 3.3: Heart disease coefficient changes over time

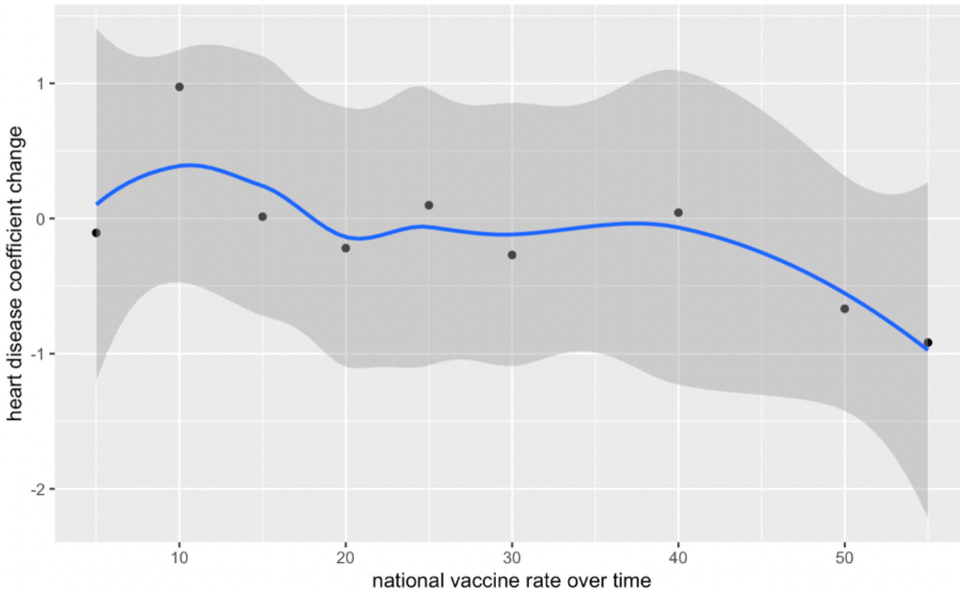
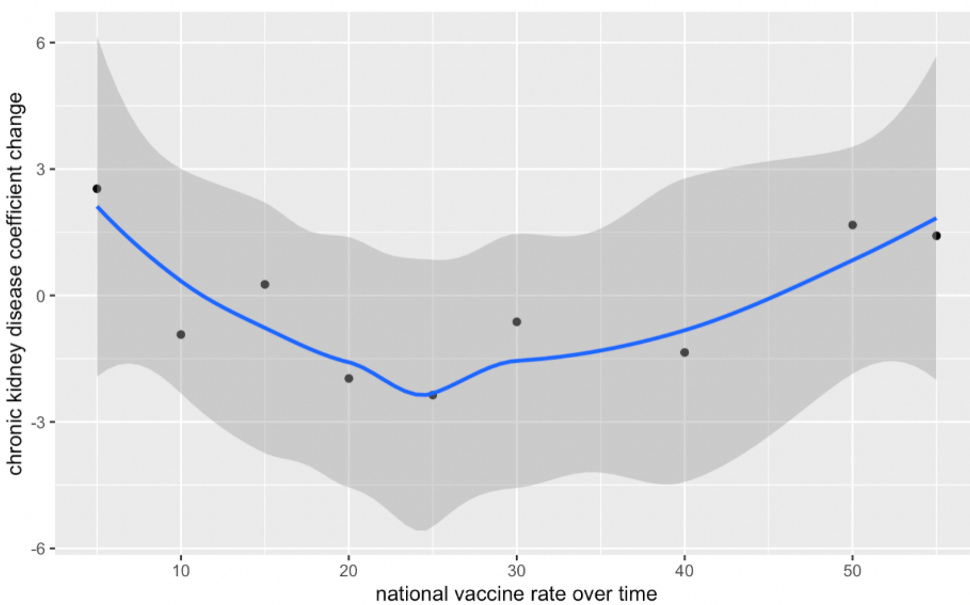


Figure 3.4: Chronic kidney disease coefficient changes over time



3.1.3 Discussion

Study 1 implies that among diabetes, obesity, heart disease and chronic kidney disease, only states with high proportion of diabetes seem to react to administer more vaccines, which is intuitive based on previous studies' results that diabetes patients are more vulnerable to coronavirus and more likely to develop severe symptoms once infected the virus (Erener, 2020). Therefore, this group of population reacted more significantly compared to other groups. However, for people with three other diseases, we don't find strong evidence that they're reacting to the pandemic.

It's necessary to notice that the diseases included in this study share co-morbidity, so we need to recognize the conclusion based on the result might be biased because of the co-morbidity of diseases.

In a past study investigating the predictors of death among COVID-19 patients from worldwide open access data, the researcher concluded that "males, advanced age, hypertension patients, diabetes mellitus patients, and patients located in America were the independent risk factors of death among COVID-19 patients" (Albitar, Ballouze, Ooi, Ghadzi, 2020). This aligns with our results that individuals with diabetes are more likely to react to receive vaccines. However, extra attention and future research is required for patients with the other three diseases that are obesity, heart disease and chronic kidney disease.

In addition, the analysis in this study was very general as the dataset is based on state level. There're many other factors that could influence the vaccine rollout efficiency, such as urbanicity, socioeconomic status, race, etc. We cannot know clearly whether vaccine increment is due to the fact that people with underlying medical conditions are receiving the vaccine, or the other confounding factors. It only gives out a big picture of the vaccination progress in the country.

3.2 Study 2: Population's Willingness to Be Vaccinated

3.2.1 Methods and Models

The primary research question in Study 2 is whether and how the population reacted to the recent change in COVID-related deaths or cases around them. Would it be considered as a motivation for people to get vaccinated? The researcher used four models to analyze the relationship between COVID mortality rate and vaccination progress. For all models, we controlled the age group of 65-year-olds or over because it influences both dependent and independent variables. This is because this group of population received vaccines earlier than other groups and they're more vulnerable to the virus.

In the first model, the dependent variable is the increment of dose 1 in percentage and the independent variable is the mortality rate increment (death out of 100K), both of which happen in the same time period. In the second model, the dependent and independent variables remain the same, except that the mortality rate increment is from the past time period. This time-lagged model shows how the mortality from past time period affects people's vaccination willingness at current moment. The dependent and independent variables in the following models are both used as increment values because they reflect how people react to the most recent COVID-19 death around them. In the third model, the dependent variable remains the same but the independent variable is the cumulative morality rate up to a certain time, which represents how the cumulative COVID-19 death number affects people's vaccine willingness.

$$\text{Dose1Inc}_{it} = \beta_0 + \beta_1 MortalityInc_{it} + \beta_2 age65over_i + u_{it}$$

$$\text{Dose1Inc}_{it} = \beta_0 + \beta_1 MortalityInc_{it-1} + \beta_2 age65over_i + u_{it}$$

$$\text{Dose1Inc}_{it} = \beta_0 + \beta_1 MortalityCml_{it} + \beta_2 age65over_i + u_{it}$$

3.2.2 Results

The coefficient of mortality rate increment is practically small but statistically significant (p-value=0.007), and the positive sign of the coefficient shows that as the mortality rate goes up, the dose 1 increment goes up, implying that people are reacting to the recent COVID deaths around them, which made them more likely to get vaccines.

Table 3.1: Study 2: Model 1 Results

	Dependent variable
	Coefficient
Mortality rate inc	0.033***
	(0.007)
age65over	0.742
	(0.302)
Note:	*p<0.1; **p<0.05; ***p<0.01

Comparing the results from the two models, it shows that people are reacting to the most recent change in the mortality. The change in mortality in the previous period has practically almost no effect with the coefficient estimate 0.002 while the change in the same time period which represent the most recent change in had a size of 0.033 which is significant at 5% significant level.

Table 3.2: Study 2: Model 2 Results

	Dependent variable
	Coefficient
Mortality rate inc	0.002
	(0.900)
age65over	0.086
	(0.349)
Note:	*p$<$0.1; **p$<$0.05; ***p$<$0.01

In the last model, the coefficient of the mortality rate tells a similar story by showing that as cumulative mortality rate goes up, the vaccine increment goes up, but it's practically very small (0.003), and using the cumulative mortality rate gives a less significant result with p-value=0.190.

Table 3.3: Study 2: Model 3 Results

	Dependent variable
	Coefficient
Mortality rate cml	0.003
	(0.190)
age65over	0.093
	(0.203)
Note:	*p<0.1; **p<0.05; ***p<0.01

3.2.3 Discussion

The results collectively supported our hypothesis, indicating that people are more likely to react to receive the vaccine when they see the recent death around them increasing, not the death from past time period nor the cumulative death up to a certain point. It is logical since individuals are trying to protect themselves and people around them from coronavirus by receiving the vaccine. However, as a previous study suggests, it's important to work with Americans' hesitancy to receive vaccine in order to increase the national vaccine completion rate (Coustasse, Kimble, Maxik, 2021). Admittedly, it would be more intuitive to use survey method if we would like to gauge individuals' vaccine willingness, but due to time limits of current study, the survey method could not be served as a way to collect data.

3.3 Study 3: Vaccination Progress and COVID Mortality

3.3.1 Methods and Models

In Study 3, the researcher aims to find out how vaccine has affected the change of COVID-19 mortality rate in the U.S. Can we say the vaccine is effective in reducing the death number? In this study, the researcher used all time periods stacked together, so I aim to find out in general how vaccines have influenced COVID mortality rate.

There're five models created to help answer the question. For the first two models, the dependent variable is mortality rate increment in a time period and the independent variable is cumulative vaccine dose 1 up to a certain time point. I controlled age group of 65-year-olds or over and poverty as confounding factors. Cumulative dose 1 is used here because in general vaccine has a long-lasting effect, and I'm interested in seeing how the collective, or total, vaccines administered can affect COVID mortality rate in a certain time frame. Using cumulative instead of increment value for independent variables makes more sense in this case. For the third and fourth models, variables are the same except that the cumulative dose 1 has time lag here, showing how the vaccine administered in the past affects current mortality rate. Last model is for dose 2. These are only applicable for vaccines that are required for 2 doses, which only included Pfizer-BioNTech and Moderna COVID-19 vaccines in the U.S. We naturally have decreasing number of observations here because individuals who received Janssen would not be considered in this model as it only requires 1 dose. In this case, the significance for dose 2 would naturally decrease in this model.

$$\text{CovidMortalityInc}_{it} = \beta_0 + \beta_1 Dose1Cml_{it} + \beta_2 age65over_i + u_{it}$$

$$\text{CovidMortalityInc}_{it} = \beta_0 + \beta_1 Dose1Cml_{it} + \beta_2 age65over_i + \beta_3 poverty_i + u_{it}$$

$$\text{CovidMortalityInc}_{it} = \beta_0 + \beta_1 Dose1Cml_{it-1} + \beta_2 age65over_i + u_{it}$$

$$\text{CovidMortalityInc}_{it} = \beta_0 + \beta_1 Dose1Cml_{it-1} + \beta_2 age65over_i + \beta_3 poverty_i + u_{it}$$

$$\text{CovidMortalityInc}_{it} = \beta_0 + \beta_1 Dose2Cml_{it} + \beta_2 age65over_i + u_{it}$$

3.3.2 Results

The coefficient of β_1 from Model 1 is negative and practically small (p-value=0.02), which tells a story that as the collective vaccine number, or the national vaccine rate, increases, the COVID-19 mortality decreases in response.

Table 3.4: Study 3: Model 1 Results

	Dependent variable
	Coefficient
Dose 1 Cml	-0.088^{**}
	(0.020)
age65over	0.422
	(0.257)
Note:	*p<0.1; **p<0.05; ***p<0.01

In Model 2, with controlling both age and poverty, the coefficient of β_1 still has negative sign, consistent with the interpretation in Model 1. However, as I treated poverty as a control variable, it becomes an important factor to influence the COVID mortality rate with p-value smaller than 0.01 and positive β value. This shows that based on the data I have, the state's poverty level is a better predictor of COVID mortality rate than vaccine distributed in that state.

Table 3.5: Study 3: Model 2 Results

	Dependent variable
	Coefficient
Dose 1 Cml	-0.062^{*}
	(0.093)
age65over	0.306
	(0.392)
poverty	1.093^{***}
	(0.000)
Note:	*p<0.1; **p<0.05; ***p<0.01

In Model 3 and 4, when using the time-lagged model, they indicate the same story that when the vaccination rate, specifically dose 1 rate, goes up, the COVID mortality goes down. It shows evidence that COVID-19 vaccines seem to be helpful in reducing the mortality rate.

Table 3.6: Study 3: Model 3 Results

	Dependent variable
	Coefficient
Dose 1 Cml	-0.051
	(0.161)
age65over	1.015^{*}
	(0.094)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 3.7: Study 3: Model 4 Results

	Dependent variable
	Coefficient
Dose 1 Cml	-0.075^{*}
	(0.038)
age65over	0.308
	(0.387)
poverty	1.090^{***}
	(0.000)
Note:	*p<0.1; **p<0.05; ***p<0.01

In Model 5, the coefficient of β_1 is still negative and practically small (p-value=0.556), which tells the same story that as the cumulative dose 2 are distributed, the COVID mortality rate decreases. Because I have a smaller number of observations for populations who received dose 2, the p-value in this model is larger compared to previous models, implying lower significance and less accurate result.

Table 3.8: Study 3: Model 5 Results

	Dependent variable
	Coefficient
Dose 2 Cml	−0.024
	(0.556)
age65over	0.383
	(0.307)
Note:	*p<0.1; **p<0.05; ***p<0.01

3.3.3 Discussion

Based on the result, I can conclude that vaccine plays a role in reducing COVID-19 mortality in the U.S., and poverty level in each state is strongly correlated with mortality rate. But we should keep in mind that the regression results only give us a basic idea and we need to improve our model for an accurate COVID mortality prediction. In addition, when measuring the real-life effectiveness of COVID-19 vaccine, it's important to notice that it might be different from the clinical trial. Particularly, the logistics of refrigeration, storage, transportation and on-site administration of the vaccines could have been imperfect, thus lowering effectiveness. Furthermore, it is possible that older individuals, who were prioritized earlier in the vaccination program, could have a reduced or belated response to the vaccination due to a deterioration in both innate and adaptive immune function, also termed immunosenescence, as was previously shown for other vaccines. In this case, all of these need to be taken into consideration when interpreting the results. Finally, a prior study has concluded that COVID-19 mortality is negatively associated with test number and government effectiveness by analyzing global COVID-19 data (Liang, Tseng, Ho, Wu, 2020), and therefore, future research direction can examine the relationship between the state government effectiveness and COVID-19 mortality.

Chapter 4

Mental Health

4.1 Study 4: Google Search Trend

It is not surprising that a global pandemic would spark human's anxious and unsafe feelings. With the COVID-19 outbreaks, when the cause or progression of the disease and outcomes are unclear, rumors grow and close-minded attitudes eventuate (Ren et al. 2020). Anxiety and fear related to infection can lead to acts of discrimination. People from Wuhan, China were targeted and blamed for the COVID-19 outbreak and Chinese people have since been stigmatized internationally. As a result, the psychological issues caused by the pandemic are worth examining.

4.1.1 Methods and Models

The primary research question in Study 4 is how the pandemic has affected Americans' mental health before, during and after 2020, which is measured by Google search trends. First, the researcher applied simple regression models to find out the association between the search terms and the COVID-19, or year 2020, controlling the quarterly binary variables to eliminate the seasonal noise. The three models below show the relationship between the pandemic and how the search interests of ”depression”, ”anxiety” and ”mental health” have changed before, during and after the pandemic.

$$\text{anxiety}_t = \beta_1 COVID_t + \beta_2 DQ1_t + \beta_3 DQ2_t + \beta_4 DQ3_t + \beta_5 DQ4_t + u_t$$

$$\text{depression}_t = \beta_1 COVID_t + \beta_2 DQ1_t + \beta_3 DQ2_t + \beta_4 DQ3_t + \beta_5 DQ4_t + u_t$$

$$\text{mental health}_t = \beta_1 COVID_t + \beta_2 DQ1_t + \beta_3 DQ2_t + \beta_4 DQ3_t + \beta_5 DQ4_t + u_t$$

4.1.2 Results

Based on the result, we can tell that the COVID binary variable is significant for ”anxiety” and ”mental health” with small p-value, meaning that there seems to be evidence that during the pandemic, people tend to search these 2 terms more than normal time. For search term ”depression”, the result is not significant, so we can’t find strong evidence about how the pandemic has influenced the frequency of people googling ”depression”. The four quarterly dummy variables are significant with small p-values. It shows that during Quarter 4, Americans will google ”anxiety” the least often, which might be explained by the holiday or Christmas season.

Table 4.1: Study 4 Model Results

	Dependent variable
	Coefficient
Anxiety	4.301***
	(0.001)
Depression	−0.825
	(0.716)
Mental Health	3.392**
	(0.034)
Note:	*p<0.1; **p<0.05; ***p<0.01

When comparing the coefficients to a neutral word (”project”), whose change of the search popularity or the coefficient is 1.1697. The absolute value is larger than the coefficient for ”depression” but smaller than those for ”anxiety” and ”mental health”. It implies that although the coefficients is practically small, when we compare to changes of other words, they’re relatively fair numbers.

In conclusion, the researcher didn’t find strong evidence that people felt more anxious or depressed during the COVID-19 period compared to other periods based on Google Trends result.

4.1.3 Discussion

In year 2020, many social and political events happened in the U.S., such as the Black Lives Matter movement and the U.S. presidential election, which could swing the public's emotions and cause psychological stress to people. As a result, in this study there are many confounding variables we cannot control, and the accuracy of the result might not accurately reflect solely due to the pandemic. Nevertheless, it is still important to conduct this study as most existing studies focused on global mental health conditions during the COVID-19 pandemic, which included the data from many other countries. Analyzing Google Trends search terms within the U.S. can tell us specifically about what's happening in this country, compared to prior studies on a global perspective.

In addition, it's worth noticing that based on prior studies, Google Trends was not proved to be a good predictor of variables of interest (Knipe, Gunnell, Evans, John, Fancourt, 2021), so for future study, it would be better to use survey method to find out population's mental health fluctuation.

4.2 Study 5: Suicide Rate

4.2.1 Methods and Models

Suicide rates have been steadily growing in the U.S. over the last two decades. From 1999 through 2017, the age-adjusted suicide rate in this country grew 33% from 10.5 to 14.0 per 100K population. For women, the rate grew 53% from 4.0 in 1999 to 6.1 in 2017. For men, the rate grew 26% from 17.8 in 1999 to 22.4 in 2017. from Figure 4.1, we can see clearly that the U.S. presents an increasing suicide rates, which may become a significant public health issue. This is why the researcher used the national suicide rate to represent the national mental health condition. The primary research question is how the pandemic has affected Americans' mental health, which is assumed to be reflected by the national suicide rate change. The researcher applied a simple regression model to find out the association between the national suicide rate change and the COVID-19, or year 2020. The following model analyzed quarterly national suicide rate from 2017 to 2020, which has suicide rate as dependent variable and COVID binary as independent variable to explore this question.

$$\text{suicide rate}_t = \beta_1 COVID_t + \beta_2 DQ1_t + \beta_3 DQ2_t + \beta_4 DQ3_t + \beta_5 DQ4_t + u_t$$

4.2.2 Results and Discussion

Based on the model result, the negative coefficient -2.183 of COVID binary variable indicates that the suicide rate actually decreases in year 2020, and it is significant due to the small p-value. This somehow implies that people's mental health became better during the pandemic because the suicide rate went down.

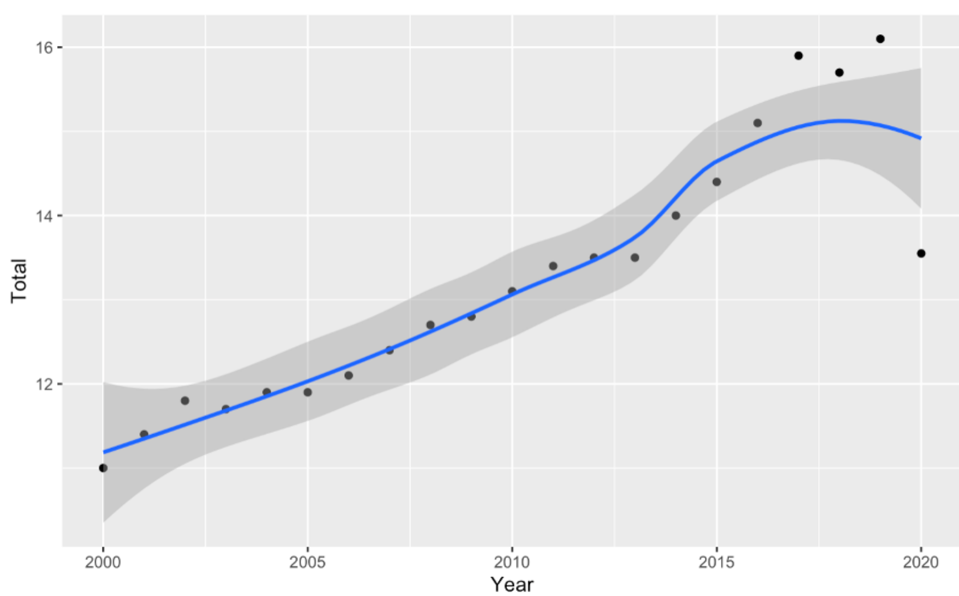
Table 4.2: Study 5 Model Results

	Dependent variable
	Coefficient
COVID	-2.183^{***}
	(0.001)
Note:	*p<0.1; **p<0.05; ***p<0.01

However, when looking at the yearly national suicide rate from 2000 to 2020, the trending behavior shows tthat in past year before the pandemic, the suicide rate is increasing faster than before. Although the result of the study shows that in year 2020 the national suicide rate decreases, the trending behavior is not convincing because there is a huge increase in national suicide rate starting from 2017 to 2019. We can't know clearly whether the suicide rate decrease in 2020 is due to the general improvement on population's mental health or the rate just returns back to normal. Future research are needed to find out what caused the sudden increase of the national suicide rate. Additionally, another possible explanation of the suicide rate drop in 2020 might be due to the dataset not fully update for year 2020. It would provide a better understanding if I can collect suicide rate data in 2021 or years after 2020 for future study.

In addition, for future study, it will be interesting to look at the COVID and quarter interaction terms to see how the independent variables influence each other.

Figure 4.1: National suicide rate changes over past 20 years



Chapter 5

Conclusion

5.1 Limitation

There are several limitations to the present study. First, the data used in this project is solely based on COVID-19 death and case numbers reported by CDC. Although the trends might not be influenced much, inaccurate reporting and the rapid increases in cases and deaths may have influenced the validity and significance of our model. It would add more validity to the study if the researcher could compare data from different sources and combine the most related variables to make a new dataset. Another concern related to only using U.S. state-level data is that due to a relatively small sample size for each group, there might not be enough variation in some independent variables of interest. A replication project of the current study should extend the scale of the sample and reexamine the relationship between vaccine distribution rate and different demographic variables for each state. This limitation could also be improved by using individual-level vaccine data. As a result, we will be able to analyze personal information of those who received vaccines and have a clear understanding of vaccine distribution efficiency. In addition, the limitation of this project falls mainly on the data sources. The data is an aggregated, state level data, which makes it hard to pinpoint down on one individual and tell whether this exact person received vaccine or not. Last but not least, there are many confounding variables that we can't control for in the mental health part of the study. For example, many other social events happened in the U.S. in year 2020, such as the Black Lives Movement and presidential election. They all can cause the population's mental health fluctuation. This problem can be solved by sending out mental health survey to individuals, collecting answers about their psychological state in last 1, 3, 6 and 12 months.

for google search trend, Few studies have assessed the validity of this approach

5.2 Future Directions

For future studies, it would produce more insightful analysis if we can collect individual vaccination information and COVID mortality. These analysis becomes much more powerful and valid as the sample size becomes larger. Also, there are many other aspects of the pandemic future research could focus on and explore, not just physical and mental health these two aspects. Future research can focus on broader and more diverse aspects of a pandemic period. Additionally, an interactive geographical map showing how the vaccine rate and mortality rate within a specific state changes over time would be helpful to show the vaccine distribution efficiency during a certain time frame.

References

Lipsitch, M., Dean, N. E. (2020). Understanding COVID-19 vaccine efficacy. Science, 370(6518), 763-765.

Forni, G., Mantovani, A. & on behalf of the COVID-19 Commission of Accademia Nazionale dei Lincei, Rome. (2021). COVID-19 vaccines: where we stand and challenges ahead. Cell Death Differ 28, 626–639. https://doi.org/10.1038/s41418-020-00720-9

Islam, M. R., Oraby, T., McCombs, A., Chowdhury, M. M., Al-Mamun, M., Tyshenko, M. G., et al. (2020). Evaluation of the United States COVID-19 vaccine allocation strategy. PLoS ONE, 16(11). https://doi.org/10.1371/journal.pone.0259700

Goldhill, O. (2021, March 4). In Palm Beach, Covid-19 vaccines intended for rural Black communities are instead going to wealthy white Floridians. STAT.

Jean-Jacques, M., Bauchner, H. (2021). Vaccine Distribution—Equity Left Behind? JAMA Network, 325(9), 829–830. doi:10.1001/jama.2021.1205.

Hartnett, K. P., Kite-Powell, A., DeVies, J., Coletta, M. A., Boehmer, T. K., Adjemian, J., Gundlapalli, A. V., & National Syndromic Surveillance Program Community of Practice (2020). Impact of the COVID-19 Pandemic on Emergency Department Visits – United States, January 1, 2019-May 30, 2020. MMWR. Morbidity and mortality weekly report, 69(23), 699–704. https://doi.org/10.15585/mmwr.mm6923e1

Bojdani, E., Rajagopalan, A., et al. (2020). COVID-19 Pandemic: Impact on psychiatric care in the United States. Psychiatry Research, 289(0165-1781), 113069. https://doi.org/10.1016/j.psychres.2020.113069

Dubey, S., Biswas, P., Ghosh, R., Chatterjee, S., Dubey, M. J., Chatterjee, S., Lahiri, D., Lavie, C. J. (2020). Psychosocial impact of COVID-19. Diabetes & metabolic syndrome, 14(5), 779–788. https://doi.org/10.1016/j.dsx.2020.05.035

Usher, K., Durkin, J., Bhullar, N. (2020). The COVID-19 pandemic and mental health impacts. International journal of mental health nursing, 29(3), 315–318. https://doi.org/10.1111/inm.12726

Ghosal, S., Sengupta, S., Majumder, M., Sinha, B. (2020). Linear Regression Analysis to predict the number of deaths in India due to SARS-CoV-2 at 6 weeks from day 0 (100 cases - March 14th 2020). Diabetes metabolic syndrome, 14(4), 311–315. https://doi.org/10.1016/j.dsx.2020.03.017

Liang, L. L., Tseng, C., Ho, H. J., Wu, C. (2020). COVID-19 mortality is negatively associated with test number and government effectiveness. Sci Rep 10, 12567. https://doi.org/10.1038/s41598-020-68862-x

Ehde, D. M., Roberts, M. K., Humbert, A. T., Herring T. E., Alschuler, K. N. (2021). COVID-19 vaccine hesitancy in adults with multiple sclerosis in the United States: A follow up survey during the initial vaccine rollout in 2021. Multiple Sclerosis and Related Disorders, 54 (2211-0348), 103163. https://doi.org/10.1016/j.msard.2021.103163

Killgore, W., Taylor, E. C., Cloonan, S. A., & Dailey, N. S. (2020). Psychological resilience during the COVID-19 lockdown. Psychiatry research, 291, 113216. https://doi.org/10.1016/j.psychres.2020.113216

Erener, S. (2020). Diabetes, infection risk and COVID-19. Molecular Metabolism, 39 (101044). https://doi.org/10.1016/j.molmet.2020.101044

Sher L. (2020). The impact of the COVID-19 pandemic on suicide rates. QJM : monthly journal of the Association of Physicians, 113(10), 707–712. https://doi.org/10.1093/qjmed/hcaa202

Rossman, H., Shilo, S., Meir, T. et al. (2021) COVID-19 dynamics after a national immunization program in Israel. Nat Med 27, 1055–1061.

Pera, A. (2020). Cognitive, Behavioral, and Emotional Disorders in Populations Affected by the COVID-19 Outbreak. Frontiers in psychology, 11, 2263. https://doi.org/10.3389/fpsyg.2020.02263

Albitar, O., Ballouze, R., Ooi, J. P., & Sheikh Ghadzi, S. M. (2020). Risk factors for mortality among COVID-19 patients. Diabetes research and clinical practice, 166, 108293. https://doi.org/10.1016/j.diabres.2020.108293

Cheng, C., Cheung, M. W. (2005). Psychological responses to outbreak of severe acute respiratory syndrome: a prospective, multiple time-point study. Journal of personality, 73(1), 261–285. https://doi.org/10.1111/j.1467-6494.2004.00310.x

Ren, S. Y. , Gao, R. D., Chen, Y. L. (2020). Fear can be more harmful than the severe acute respiratory syndrome coronavirus 2 in controlling the corona virus disease 2019 epidemic. World Journal of Clinical Cases, 8 (4), 652–657.

Coustasse, A., Kimble, C., Maxik, K. (2021). COVID-19 and vaccine hesitancy: A challenge the United States must overcome. Journal of Ambulatory Care Management, 44(1), 71-75.

Knipe, D., Gunnell, D., Evans, H., John, A., Fancourt, D. (2021). Is Google Trends a useful tool for tracking mental and social distress during a public health emergency? A time-series analysis. Journal of Affective Disorders, 294, 737-744.