GÖTEBORGS UNIVERSITET

DEPARTMENT OF POLITICAL SCIENCE

THE ROLE OF BENEFITS AND COSTS

The COVID-19 pandemic effect on presidential election-2020 in the United States

Author: Aniya Samandi

Master's Thesis:	30 credits
Programme:	Master's Programme in Political Science
Date:	2021-05-25
Supervisor:	Ann-Kristin Kölln
Words:	12 385

Abstract

The year of 2020 was the most challenging year in modern history, where it will be remembered as the year that the COVID-19 pandemic took place. The COVID-19 pandemic led to worldwide lockdowns, where people all over the world had to adjust to a new lifestyle. The United States of America held a presidential election during the pandemic, which questioned the increase of the COVID-19 cases across the country. This master's thesis will investigate if the COVID-19 pandemic influenced the presidential election-2020 in the United States. In order to investigate this topic, this quantitative study will use a cross-sectional analysis to display if the COVID-19 pandemic influenced the voter turnout in presidential election-2020. The theoretical framework for this master's thesis is the rational choice theory, as this model investigates the rational behavior of how some citizens vote and if they vote. The theoretical framework will also investigate the benefits and costs of voter turnout. The rational choice theory suggests higher voter turnout in presidential election-2020, where it became a necessity for the states in the U.S. to have postal ballot available, to lower the increasement of COVID-19 cases. This quantitative study displays that the COVID-19 pandemic had a negative effect on voter turnout. However, the effect of COVID-19 pandemic on voter turnout was not statistically significant, where the effect was not large enough for statistical significance.

Keywords: COVID-19, Rational choice model, Voter turnout, U.S. presidential elections.

Table of contents

Table of figures and tables

FIGURES

TABLES

Abbreviations

CDC – Centers for Disease Control and Prevention

GDP – Gross Domestic Product

NBC News – National Broadcasting Company News

NCSL – National Conference of State Legislatures

OLS – Ordinary Least Squares

WHO – World Health Organization

1. Introduction

The year of 2020 was a turbulent year for many individuals around the world for different reasons. American citizens faced challenging issues such as police brutality against Afro-American citizens, increased unemployment and emerging conflicts with China (Reinfeldt, 2020, p 7). However, one major indicator that affected individuals around the world and American citizens is the COVID-19 pandemic, that is still ongoing yet to this day. This opened the discussion of how the COVID-19 pandemic would impact the presidential election of 2020 in the United States. The presidential election of 2020 was held amid the COVID-19 pandemic. Therefore, it was important for this master's thesis to analyze the if the COVID-19 pandemic did affect the voter turnout during the presidential election-2020, which resulted in the highest voting turnout in 120 years (Ibid, 2020, p 351). This is a scientific problem that has limited research, where the topic is still under development by different scholars. One reason why the voter turnout was high is maybe counterintuitive, which is another reason to look more closely at the effects of the pandemic. This indicates that the relationship between the COVID-19 pandemic and the voter turnout from presidential election-2020 needs to be investigated further. Therefore, the research question for this master's thesis is: *Did the COVID-19 pandemic have an impact on the voter turnout during the presidential election of 2020 in the United States?*

This master's thesis uses the rational choice model as a theoretical framework to discuss the potential benefits and costs of voting turnout during the pandemic in the U.S. I argue that the pandemic poses a cost to the individual because it is a potential threat, where this will be illustrated as the benefits and costs. The benefits and costs are the reason for the circumstance for explaining if the American voters was willing to vote during the COVID-19 pandemic. Previous scholars have not investigated this topic from a rational choice perspective, where this master's thesis will be an addition to existing literature. Based on the theoretical framework, I want to test if voting turnout was lower in states with higher COVID-19 cases. I also want to test if the postal ballot availability increases the voting turnout during the COVID-19 pandemic. The results display that the COVID-19 pandemic had a negative effect on voter turnout. However, the results were not statistically significant, where I argue that there are not enough cases or states for the relationship to be statistically significant.

1.1 Disposition

This master's thesis will be divided into five parts in order to investigate if the COVID-19 pandemic influences voting turnout during the presidential election-2020 in the United States. The topic has been introduced along with the research question above. Chapter two will present the background of the topic, where the background and information about the COVID-19 pandemic and the presidential election of 2020 and 2016 will be discussed. Chapter three will present the theoretical framework and previous research of this master's thesis. The previous research discusses the COVID-19 pandemics impact during an election in 2020 and how this master's thesis will be an addition to existing literature. Chapter four will present the statistical method and data that is used to investigate if the COVID-19 pandemic had an impact on the voter turnout during the presidential election-2020 in the United States. Chapter five will present the results and analysis for this master's thesis. Lastly, this master's thesis will end with a conclusion.

2. Background

In order to investigate if the COVID-19 pandemic had an impact on the voter turnout during the presidential election 2020 in the United States, the definition and the development of the COVID-19 virus need to be defined. Furthermore, this chapter will explain the development of the COVID-19 virus in the United States along with the presidential election of 2016 and 2020.

2.1 The COVID-19 virus

The 31st of December 2019 was the day that different cases of pneumonia in Wuhan, China was detected, where this incident quickly alarmed the World Health Organization (WHO) and started the process of identifying the cases of pneumonia (World Health Organization, 2021). The cases were highly contagious which spread quickly in China. In middle of January 2020, China shared with the world a scientific course of the COVID-19 virus and the first COVID-19 case was detected outside of China, which was in Thailand (Ibid, 2021). It wasn't until late January 2020 that human to human interaction was the main aim of how the cases in Wuhan escalated (Ibid, 2021).

However, nobody could predict that these cases of pneumonia would quickly escalate into a worldwide pandemic. The spread of the COVID-19 virus had a major outbreak in China and East Asia, where it escalated quickly into other regions of Europe and North America. Moreover, in 11th of March 2020 the WHO declared the COVID-19 virus as a worldwide pandemic (Ibid, 2021) where it is still ongoing yet to this day. However, former president Donald Trump declared the escalated COVID-19 virus as a national emergency in 13th of March 2020, where the first declaration was to ban individuals from traveling from Europe to the United States (AJMC, 2021). The only exception for travel to the United States was if the individuals were citizens of the United States (Ibid, 2021). In order to understand the spread of the COVID-19 virus in the United States, the amounts of cases during October 2020 will be displayed along with the spread of the COVID-19 virus among different states.

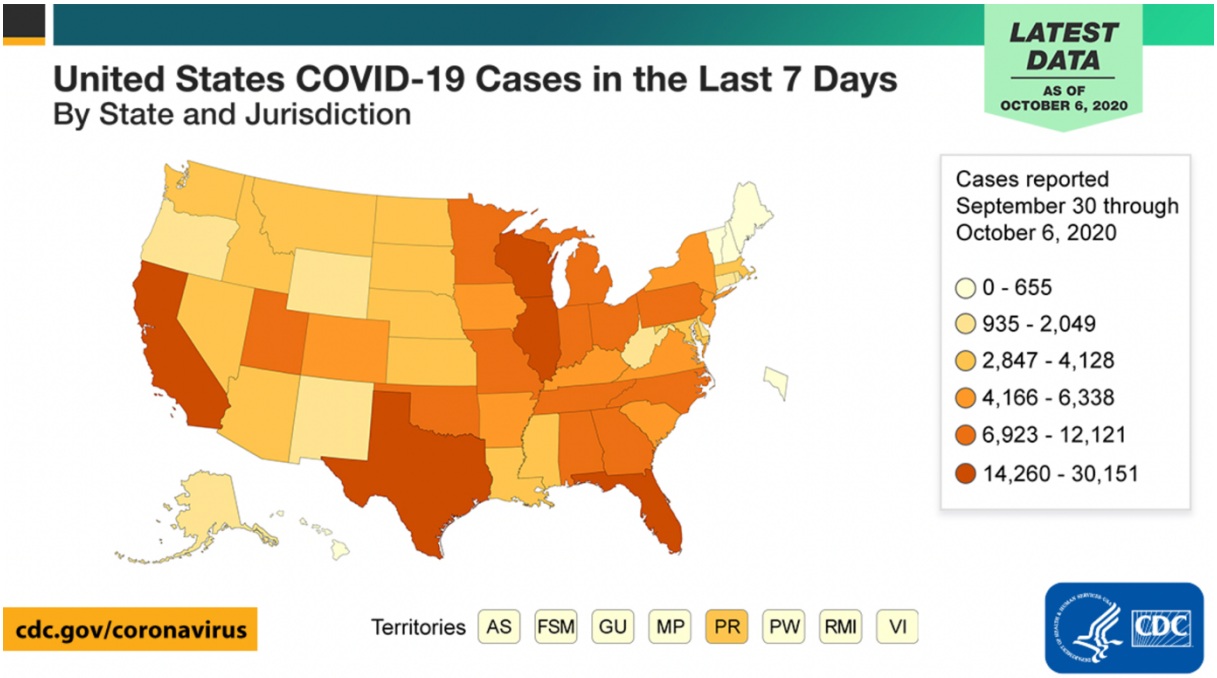

Figure 1. *Chart of COVID-19 cases the United States early October 2020.*

Source: Centers for Disease Control and Prevention.

The figure above displays a map of the amount of COVID-19 in October 2020 in the United States, where it is applied from Centers for Disease Control and Prevention (CDC). The United States had approximately 84 000 reported cases by the end of October 2020 (National Geographic, 2021) and had a total 1 908 493 million cases throughout the country by the end of October 2020 (Ibid, 2021). The map above displays that 5 states had between 14 260 – 30 151 new cases reported (red), 12 states had between 6 923 – 12 121 new cases reported (orange), 9 states had between 4 166 – 6 338 new cases reported (light orange), 13 stated had 935 – 2 049 new cases reported (yellow) and 9 states had 0 - 655 new cases reported (white) (CDC, 2020). Every state in the U.S. had increasing COVID-19 cases during October 2020, which is displayed in figure 1. Based on the map above, the amounts of COVID-19 cases in October 2020 increased more in 5 states: California, Florida, Illinois, Texas and Wisconsin (Ibid, 2020). The spread of the COVID-19 cases also started to expand in Alabama, Georgia, Indiana, Michigan, Minnesota, Missouri, North Carolina, Ohio, Oklahoma Pennsylvania, Tennessee and Utah (Ibid, 2020).

2.2 The presidential election-2020 in the United States

The presidential election-2020 in the United States took place 3rd of November 2020, where Joe Biden won with 306 electoral college votes against Donald Trump's 232 electoral votes (CNN, 2021). Each of the presidential candidates need 270 electoral votes to win the presidential election (Ibid, 2021). The presidential election-2020 had 156 million American citizens participating by voting, where this is the first time in American history that an election had more than 140 million American citizens casting their ballot (Council on foreign relationship, 2021). The presidential election-2020 had the highest voting turnout in 120 years, with the total of 66.7 percent of the voting turnout (Ibid, 2021).

2.2.1 Voter turnout on a state-wide level: election 2016 and election 2020

STATES	ELECTION 2020	ELECTION 2016
ALABAMA	63.1	58.65
ALASKA	68.8	60.64
ARIZONA	65.9	56.35
ARKANSAS	56.1	52.46
CALIFORNIA	68.5	53.8
COLORADO	76.4	67.86
CONNECTICUT	71.5	62.59
DELAWARE	70.7	63.86
FLORIDA	71.7	65.61
GEORGIA	67.7	60.0
HAWAII	57.5	42.52
IDAHO	67.7	59.68
ILLINOIS	67.0	60.11
INDIANA	61.4	56.18
IOWA	73.2	68.56
KANSAS	65.9	55.67
KENTUCKY	64.9	59.51
LOUISIANA	64.6	59.86
MAINE	76.3	69.92
MARYLAND	70.7	61.4
MASSACHUSETTS	72.1	65.35
MICHIGAN	73.9	64.59
MINNESOTA	80.0	74.16
MISSISSIPPI	60.2	53.51
MISSOURI	66.3	61.59
MONTANA	73.1	63.45
NEBRASKA	69.9	61.69
NEVADA	69.9	57.09
NEW HAMPSHIRE	75.5	70.31
NEW JERSEY	75.3	59.25

NEW MEXICO	61.3	54.8
NEW YORK	63.4	52.4
NORTH CAROLINA	71.5	64.57
NORTH DAKOTA	64.5	59.22
OHIO	67.4	64.55
OKLAHOMA	55.0	52.11
OREGON	75.5	66.85
PENNSYLVANIA	71.0	61.26
RHODE ISLAND	65.7	58.77
SOUTH CAROLINA	64.5	56.85
SOUTH DAKOTA	66.0	58.69
TENNESSEE	59.8	51.02
TEXAS	60.4	51.1
UTAH	69.2	46.4
VERMONT	74.2	64.66
VIRGINIA	73.0	65.5
WASHINGTON	75.7	62.46
WEST VIRGINIA	57.6	50.97
WISCONSIN	75.8	68.33
WYOMING	65.6	60.09

Table 1. *Voter turnout per state in percentage for election 2016 and 2020.*

Source: NBC News and Statista.

Table 1 displays the eligible voter turnout per state in presidential election-2020 and 2016 in percent, where the data is selected from NBC News (NBC News, 2021) and Statista (Statista, 2021). It is important to acknowledge that some states have more registered votes than other states, where some states are more populated. However, the state of Minnesota had the highest voting turnout during presidential election-2020, with 80 percent of eligible voter turnout (NBC News, 2021). The second state with the highest voter turnout was the state of Colorado, with 76.4 percent (Ibid, 2021). The third state with the highest voter turnout was the state of Maine with 76.3 percent (Ibid, 2021). These three states are not the largest states in the United States by area and are not the most populated (Ibid, 2021). The state with the lowest voter turnout in presidential election-2020 was the state of Oklahoma, with 55 percent of voter turnout (Ibid, 2021) and the second state with the lowest voter turnout was the state of Arkansas with 56.1 percent of voter turnout (Ibid, 2021). However, the lowest eligible voter turnout per state in presidential election-2020 was the average voter turnout in presidential election-2016.

The presidential election-2016 had a different outcome where Donald Trump from the Republican Party won against Hillary Clinton from the Democratic Party with 306 electoral votes against 232 (NBC News, 2020). Based on the voter turnout for presidential election-2016, we can acknowledge a difference in voter turnout comparing to presidential election-2020 in the United States. The voter turnout in percentage within the states was lower in presidential election-2016 comparing to election-2020, which is noticeable based on the eligible voter turnout that is displayed in table 1. The state with the highest voter turnout was the state of Minnesota, with the voter turnout of 74.16 percent (Statista, 2021). The state of Minnesota had also the highest eligible voter turnout in presidential election-2020 (NBC News, 2021). The state with the second-highest voter turnout was the state of New Hampshire, with the eligible voter turnout of 70.31 percent (Statista, 2021). New Hampshire had a voter turnout of 75.5 percent during presidential election-2020, which was average for the voter turnout in percent among the states. The third state with highest voting turnout was the state of Maine, where the voter turnout was at 69.92 percent (Ibid, 2021). However, the state with the lowest voting turnout was the state of Hawaii, with the eligible voter turnout of 42.52 percent (Ibid, 2021). The state of Hawaii increased the voter turnout in presidential election-2020 with 15 percent and was not the state with the lowest voting turnout, which was the state of Oklahoma, with 55 percent of voter turnout (NBC News, 2021). The presidential election of 2016 and 2020 had different outcomes in terms of voting turnout. It is important to acknowledge the increasement of voter turnout during presidential election-2020, which is noticeable based on voter turnout that was presented in table 1. Every state increased their voter turnout in presidential election-2020 comparing to the voting turnout from election-2016.

3. Theoretical framework and previous research

3.1 The rational choice model and the calculus of voting model

The citizens in the United States that is registered to vote, does vote for different reasons. Kaat Smets and Carolien van Ham (2013) investigated that different models of voter turnout measures factors that are associated with voter turnout. Smets and van Ham stated that there are six theoretical models that highlight voter turnout on individual level: the resource model, the mobilization model, the socialization model, the rational choice model, the psychological model and the political institutional model (Smets & van Ham, 2013, p 347). However, the theoretical framework for this master's thesis will be based on the rational choice model, which will explain how the COVID-19 pandemic had an impact on the voter turnout during presidential election-2020 in the United States. In order to understand the rational choice model, the calculus of voting model needs to be defined, where the terms benefit and costs is important for the theoretical framework. The calculus of voting model by John H. Aldrich (1993) was originated by Downs in 1957 and further developed by Riker and Ordershook in 1968 (Aldrich, 1993, p 251). The model justifies the possibility to estimate the voter turnout (Ibid, 1993, p 251). Therefore, the costs weigh more than the citizens reason to vote (Ibid, 1993, p 251), where the definition of the calculus of voting model is about the rational choice model. Based on these two models, it is understandable that some citizens vote, and some citizens does not vote (Riker & Ordershook, 1968, p 25).

The theoretical framework for this master's thesis will describe that the individual will vote in terms of the conditions of benefits and costs. It is important to understand why the citizen votes, which will describe the outcome of voting turnout in general. Therefore, the calculus of voting model was defined in order to understand the rational choice model and what purpose it has for the theoretical framework of this master's thesis. However, the rational choice model will be presented and explained more through a formula developed by André Blais, Robert Young and Miriam Lapp (Blaise et al., 2000, p 181). Furthermore, it is important for this master's thesis to present the equation of voting turnout Ibid, 2000, p 181):

$$R = BP - C$$

What is the definition of the equation above, and how is it relevant for this master's thesis? Blais, Young and Lapp stated that the equation above is based on four factors: net rewards for voting (R), instrumental benefits based on comparative outcomes (B), probability of votes and benefits (P) and costs of informed individuals regarding the election and voting (C) (Ibid, p 2000, p 181). The equation does also define that the rational individual will only vote if the instrumental benefits based on comparative outcomes and probability of votes and benefits weighs more than the costs of informed individuals regarding the election and voting (Ibid, 2000, p 182). The literature describes the rational choice model as a different model for every individual, which defined the benefits and the costs. The benefits weight more than the costs, which indicates that the citizens will step outside and vote. If an election in the United States were held during normal circumstances, which does not include a worldwide pandemic, this defines the benefits. The voter needs to prioritize and identify how much they care about the election, which is the costs. Moreover, the voter's circumstance of life would be different. Furthermore, it is important that the costs of voting and the time investments is investigated within the rational choice model, which is also known as general costs. It is stated by Smets and van Ham that citizens are more likely to vote during an election, if the election attract the citizens to cast their ballot for individual reasons (Smets & van Ham, 2013, p 352). However, it depends on the interests from the citizens to participate in the election (Ibid, 2013, p 352). For example, the citizens need to care for the election in order to cast their ballot, which has resulted in lower voting turnout through the past presidential elections in the United States (Ibid, 2013, p 352).

Furthermore, I argue that when an election is held during a pandemic the circumstance of voting becomes different. The specific costs were a large factor during the presidential election of 2020 in the United States. The specific costs define the option to vote by going to the ballot box, queue up to schools to cast your ballot, meeting people during election day along with the voters exposing themselves for catching and spreading the COVID-19 virus. The specific costs are crucial to acknowledge during a pandemic, where the pandemic is an overall cost in general. Therefore, I argue that the postal ballot is an option for the American citizens to not get affected by the COVID-19 virus during the presidential election-2020. The option of the postal ballot will reduce the amounts of COVID-19 cases, which is a simplistic understanding of how citizens interact during an election. The option for postal ballot does also reduce the costs,

especially the specific costs, where reducing the amounts of queues to the polling stations and meeting different people while casting your ballot.

3.1.1 Predictions of the rational choice model within presidential election 2020

In order to implement the rational choice model for this master's thesis, it is important to understand how the rational choice model operates in the presidential election-2020 in the United States. The formula for voting turnout developed by Blais, Young and Lapp (Blaise et al., 2000, p 181) was previously presented, where a new formula has been developed by me, which is designed for presidential election-2020 and the rational choice model. The formula that I have developed will propose two outcomes for the presidential election-2020 regarding voting turnout in the United States. Below you will find the formula:

$$R = P*B - P*C$$

It is important to understand how the benefits and costs operates during a presidential election, especially and election that were held during the COVID-19 pandemic. 'R' are the rewards for voting and if they are above 0 according to the model, the citizens will vote. However, the formula above also displays the importance of probability. Individuals in the United States are different affected by the COVID-19 virus. Therefore, there are two types of effects on probability where I have added one other probability to the formula above. The first probability was included in the formula from the start. One scenario is that all the citizen in the United States is nervous by catching the COVID-19 virus, which actively demonstrates the costs. However, some individuals in the United States are more nervous than others to catching the COVID-19 virus.

This is based on two factors: state and health conditions. The 'state' factor determines if the citizens are at high risk by catching the COVID-19 virus in the state and the 'health condition' factor determines the nervosity of the citizens by catching the COVID-19 virus. The United States is a large country by area, where they have different sizes of states, which identifies the higher or lower population in respective states. For example, the state of Alaska is the largest state in the United States with the population of 724 357 American citizens (World population review, 2021). However, the largest amount of population within a state is the state of California, with the population of 39 512 223 American citizens (Moving, 2021). Therefore, in

order to avoid catching the COVID-19 virus, it depends on where the citizens live and if they live in a state that is at a higher risk for catching the COVID-19 virus. The second factor, health condition, indicates how nervous people are to catching the COVID-19 virus in the United States. The citizens are nervous to expose themselves to the COVID-19 virus, especially during presidential election-2020. The health condition of the citizens is one factor that determines if they are nervous to expose themselves to the COVID-19 virus, where they don't want to get affected by the virus due to the health condition that they have. To sum up these factors that have been stated above, this is the probability affecting the costs which is presented in figure 2 below. The costs might be higher or lower for different individuals.

However, the probabilities and the costs have been identified, where it is important to display the costs from two different angles. Below you will find an illustration:

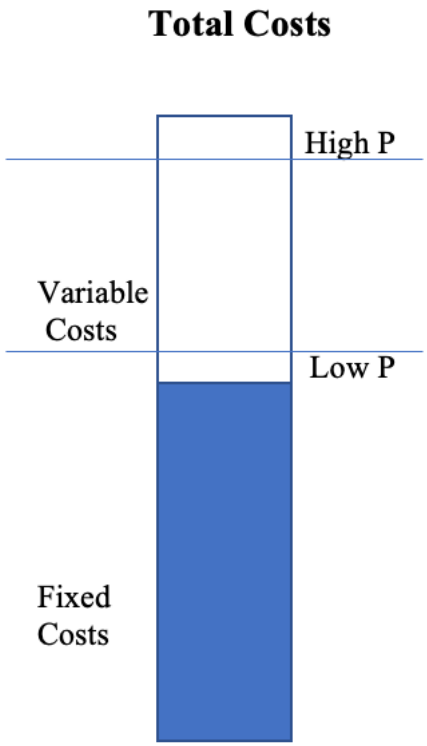

Figure 2. *Illustration of assumptions about the total costs of voting.*

Source: Self-made illustration.

Figure 2 displays the total costs of voting. My argument for this model is that we can split the total costs for voting up into two parts, which is the fixed and variable costs. The fixed costs

are displayed at the bottom of the diagram, which indicates the costs that are the same for every individual. For example, the fixed costs are the factor where every individual needs to go out physically and vote, unless they have a postal ballot. The variable costs have a probability attached to them due to where individuals live and their health conditions. Therefore, the variable costs explain the state and health conditions for the citizens. The arguments that I have made regarding the probabilities effect on the costs are crucial for the option of voting by postal ballot. My argument is that the postal ballot affects the costs. If we have the postal ballot as an option, then we can reduce the probability, where all citizens are nervous to catching the COVID-19 virus, to 0 percent. Furthermore, a new formula has been developed based on the second formula that was introduced above, which displays two options if the American citizens have the option to vote by postal ballot. Moreover, we can assume that if we multiply probabilities with benefits and costs, which will determine the citizens decisions to go out and vote. Below you will find the formula where the probabilities multiply with the benefits and costs:

$$P*B > P*C$$

The formula above displays the effect that the probability has on the benefits and costs during the presidential election-2020 in the United States. It depends on which state in the United States the American citizens live in, were catching the COVID-19 virus is at higher risk. However, the citizen's health condition does also play part in order to cast their ballot. The option to vote is now restricted where the option to vote on election day are a health risk. Furthermore, the presidential election-2020 in the United States was based on the population of 331 million individuals (Worldometer, 2021), where the environment for voting was on a larger scale and more individuals would queue up to polling stations or schools. Moreover, the social distancing would be an issue during the presidential election-2020 in the United States, where more individuals had access to vote. In summary, I argue based on the rational choice theory that the benefits are larger than the costs during the presidential election-2020 in the United States, because of the COVID-19 pandemic. The COVID-19 pandemic makes it harder for the citizens to vote when they are currently advised from the government to social distancing, which makes the citizens nervous of exposing themselves to catching the COVID-19 virus.

3.2 Hypothesis

The hypotheses for this master's thesis are derived from the theoretical framework and is a theoretical answer to the research question '*Did the COVID-19 pandemic have an impact on the voter turnout during the election of 2020 in the United States?'.* Therefore, the focus on the first hypothesis is if the voting turnout was lower in states with higher COVID-19 cases. The focus for the second hypothesis is also on the availability of postal ballots in some states, where if they increase voter turnout or not. Therefore, the hypotheses for this master's thesis are:

H1: '*Voting turnout was lower in states with higher COVID-19 cases'*
H2: '*COVID-19 has a conditional effect on voter turnout, where the postal ballot availability expects to increase the voter turnout in states'*

3.3 Previous Research

The existing literature have mainly investigated the relationship between the presidential election-2020 in the United States and the COVID-19 pandemic. Therefore, this master thesis will focus on four scientific research that displays the existing literature on the relationship between the COVID-19 pandemic and the elections in the United States during the year of 2020. It is important to acknowledge that this topic is still relevant and different aims of the topic is still investigated. This part of the master's thesis will investigate how these four scientific papers contributes to the research of voter turnout within the election of 2020 in the United States and the impact that the COVID-19 pandemic had. Furthermore, the existing literature written by Flanders, Flanders & Goodman (2020), Johnson, Pollock & Rauhaus (2020), Baccini, Brodeur & Weymouth (2020) and Landman & Splendore (2020) is a good foundation for investigate the relationship between voter turnout in the United States and the COVID-19 pandemic further. It is important to acknowledge that the previous research is not discussing the input of the rational choice model to the existing literature. In comparison to this study, previous scholars have not investigated this topic from a rational choice perspective. Therefore, this master's thesis will be an addition to existing literature.

The studies by Flanders, Flanders and Goodman and Andrew F. Johnson, Wendi Pollock and Beth Rauhaus have a similar theoretical approach but come to different conclusions. Flanders, Flanders and Goodman investigated the relationship between voter turnout on county-levels

along with the COVID-19 pandemic (Flanders et al., 2020, p 42) and Johnson, Pollock and Rauhaus discusses different political changes in terms of the voter turnout and the amounts of COVID-19 cases along with the increased COVID-19 cases that caused death (Johnson et al., 2020, p 249).

Flanders, Flanders and Goodman discussed the effect of the COVID-19 virus and how it could spread faster based on where the citizens live, how they interact and how they transport themselves (Flanders et al., 2020, p 42). However, Johnson, Pollock and Rauhaus discussed the evolution of mass casualty events in the United States, where they compared the ongoing pandemic with different pandemics in the past (Johnson et al., 2020, p 251). They compared the COVID-19 pandemic to the AIDS epidemic that began in 1981 until this present day, Plauge of Justinian which began in 541 AD and the Black Death which began in the year of 1348, where all these pandemics cause millions of people's lives (Ibid, 2020, p 251). Flanders, Flanders and Goodman investigated on a county-level how the COVID-19 virus occurrence influenced the primary election on March 10th and how the voter turnout differed (Flanders et al., 2020, p 42-43). However, Johnson, Pollock and Rauhaus investigated the statistics of COVID-19 cases and how the spread of the deadly virus quickly developed into a worldwide pandemic.

Baccini, Brodeur and Weymouth discussed the effect that the COVID-19 cases and deaths had by voting for Donald Trump in presidential election-2020 in comparison to 2016 (Baccini et al., 2020, p 1). Baccini, Brodeur and Weymoth investigates if the support for voting for Donald Trump was lower in the states that had higher COVID-19 cases (Ibid, 2020, p 1). The results indicate that the COVID-19 pandemic had a negative effect on the voters in the election of 2020 in comparison to the election of 2016 (Ibid, 2020, p 2). Flanders, Flanders and Goodman investigated voter turnout within counties that have higher risks for the spreading of the COVID-19 virus after the primary election took place (Flanders et al., 2020, p 42). However, Landman and Splendore investigated if the COVID-19 pandemic influenced open elections or democracy in general around the world (Landman & Splendore, 2020, p 1). The research starts with the comparison of how an election is upheld during a COVID-19 pandemic within the states in the United States. It was discussed in the paper that some states in the United States did postpone their primary election, where other countries in the world had a different approach. Furthermore, Landman and Splendore raised the question of how the COVID-19 pandemic is risking and affecting the countries for held open elections and how the postponing of the

election would affect the public health protection and electoral integrity (Ibid, 2020, p 2). In summary, this master's thesis is an innovation on existing literature by investigating voting turnout during the presidential election-2020 in the United States from a rational choice perspective and if the COVID-19 pandemic had an impact on the outcome of the election.

3.4 How will this master thesis be a contribution for the existing literature?

It is important to acknowledge that this topic is still relatively new and very limited research has been added so far. Previous research explains voter turnout from a different angle, and they are valid, where this master's thesis will focus on the rational choice theory and the individual in terms of benefits and costs. The pandemic is a large cost that was a factor that played part in voter turnout in during presidential election-2020. Therefore, this master's thesis will be an addition to the currently existing research, which includes investigating the voter turnout within the presidential election-2020 in the United States and if the COVID-19 pandemic had an impact on the outcome of the election. Based on the rational choice model, the benefits weights more than the costs which will determine if the individuals cast their ballot. The COVID-19 pandemic could have affected the voter turnout during presidential election-2020 by restricting citizens to cast their ballot by going to polling stations. This is also explained by the rational choice model, which indicates if the citizens will cast their vote or not during the COVID-19 pandemic. Therefore, the COVID-19 pandemic did have an impact on the voting turnout in presidential election-2020 in the United States. Since this master's thesis has developed a new research question and a new theoretical framework, I will test if either H1 or H2 holds.

4. Method and Data

This section will start by introducing the research design along with the dataset that was used in this master's thesis. It will continue to describe the variables that was used and how the variables were operationalized.

4.1 Research Design

The ideal research design for testing the research question and hypotheses for this master's thesis is a quantitative method (Jakobsen & Mehmetoglu, 2016, p 3), more specifically the cross-sectional analysis that implements the ordinary least squares (OLS) as a statistical method (Ibid, 2016, p 50). I find this research design ideal where this statistical method will investigate if the COVID-19 pandemic influenced voting turnout during presidential election-2020 on a state-level for one year and not states over time. However, this research design would not have been possible if states was measured over time, where the appropriate research design would be a panel data analysis (Ibid, 2016, p 227). The panel data analysis would be the best solution because this research design is adapted for measuring states, countries and samples over time (Ibid, 2016, p 228).

However, the ordinary least squares method has only been applied by Baccini, Brodeur & Weymouth from previous research that was discussed in this master's thesis. Even though they used this statistical method, they did a statistical analysis over time, which is different from this master's thesis. Moreover, this master's thesis investigates the voting turnout during presidential election-2020 in the states in the U.S., and not states over time, which is different from existing literature. Furthermore, it is important to acknowledge that the constructed data is on state level, even if the theoretical framework is based on an individual level. The data that is selected does not exist as an individual level data, where the selected data is on state level. Moreover, the theoretical framework displays the rationality of the individual before examining the voter turnout on state level.

4.2 Dataset

The dataset that is chosen for the quantitative method is based on the data that is created for this master's thesis. The compiled data for this dataset is from CDC, IDEA, NBC News, New York Times, NCSL and Statista. There was not a constructed dataset for the aim of this master's thesis, which was the reason why a new dataset was constructed, based on this topic. Therefore, this is an addition to existing literature where a new dataset has been constructed. The data from the dependent variable is selected from NBC News, the data from the independent variable is selected from CDC, the data from the moderating variable is selected from IDEA, NCSL and New York Times, and the data for the controlling variables is selected from Statista and NBC News. The topic of this master's thesis has limited research, which has been presented earlier in the thesis. Therefore, the most suitable option was to construct a dataset based on the data that was accessible for the voter turnout and relatable for the other contributing variables.

The data that is collected from CDC, IDEA, NBC News, New York Times, NCSL and Statista was inserted into a excel-file. In order to use the dataset in Stata, the excel file needs to be converted. The dataset was later applied into Stata, where the programme would read the numeric variables that were coded.

4.3 Dependent Variable: Voter Turnout

The dependent variable for this master's thesis is 'Voter turnout'. The data that is compiled for this variable is based on voter turnout per state in percent during presidential election-2020 in the United States. The voting turnout is displayed in table 1, that was presented earlier in this master's thesis. The measurement of the dependent variable is the share of the eligible voters that cast a ballot. The dependent variable is based on the data from NBC News (NBC News, 2021), where the data displays the voter turnout in the 50 states during the presidential election-2020. However, it is important to discuss the validity and the reliability of the variable, since it is collected and incorporated into a new dataset. The definition of validity is to identify if the variable intends to measure the data correctly or to the extent of how it should measure the data (Maslakci & Sürücü. 2020, p 2695). The measurement for the 'Voter Turnout' variable is valid based on the measurement of the data, where the data identifies the precise data of voter turnout of eligible voters, where it is calculated from voting ballots that is presented in percentage. The data is accurate based on where ballots that has been counted, which makes this data valid for

this master's thesis. While measuring the reliability of the variable, it indicates that the measurement of the variable has the same measurement of the data over time (Ibid, 2020, p 2707). The data of the 'Voter Turnout' is based on the eligible voter turnout of presidential election-2020, which only measures the voter turnout for one year. However, the measurement of the variable will remain the same if the variable is used again, where it will not change the outcome of the master's thesis, which makes the data reliable (Ibid, 2020, p 2712).

In order to estimate of the variable's central tendency, the mean, standard deviation, minimum and maximum value needs to be discussed (Jakobsen & Mehmetoglu, 2016, p 33). The mean value for the variable is 67.968 and the value for the standard deviation is 5.949359, which indicates how distributed the value is from the mean (Ibid, 2016, p 33). The minimum value for the variable is 55 percent, which was the state of Oklahoma, where the state had the lowest voter turnout in percentage. The maximum value for the variable is 80 percent, which was the state of Minnesota, where the state had the highest voting turnout in percentage from the presidential election-2020. The measure of the variable is also displayed in a histogram, that is displayed in figure 2 (see figure 2 in appendix).

4.4 Independent Variable: Covid-19

The independent variable for this master 's thesis is 'COVID-19'. The data that is compiled for this variable is selected from CDC's (Centers for Disease Control and Prevention) website, which displays the total new COVID-19 cases per state over time (CDC, 2021). The data displays total COVID-19 cases for the 50 states, where the new cases for October 2020 is collected. I argue that these statistics will highlight the importance of the new COVID-19 cases in October 2020 and if the pandemic influenced the presidential election-2020. This is based on the voting turnout numbers and new COVID-19 cases per state. However, the numbers of total new cases of COVID-19 per 100 000 inhabitants per state in the United States did not include from CDC's website, where adjustments had to be made for this variable. The numbers that occurred on CDC's website did only include total new cases for each state in October 2020. Therefore, the numbers had to be calculated so the total new cases of COVID-19 per 100 000 inhabitants was collected in the dataset. It was necessary to make these adjustments for the data to be valid, where the data for total new COVID-19 cases in October 2020 per 100 000 inhabitants was not available.

The population size for every state needed to be collected, where the data for the population size was established for every state from Statista's website (Statista, 2021). The data for the population size is presented in millions, where a graph displayed the resident population of the U.S. in 2020 per state (Ibid, 2021). Every state had different amount of total new COVID-19 cases in October 2020 and the population size. Since the total new cases of COVID-19 per 100 000 inhabitants was calculated, the validity of the variable is stronger where it measures the correct data, where the relationship with the dependent variable is specified (Maslakci & Sürücü, 2020, p 2695). The calculation for total new cases of COVID-19 per 100 000 inhabitants per state was:

Total new cases / population size = Z * 100 000 = total new cases of COVID-19 per 100 000 inhabitants

The formula above displays how the calculation for every state was applied to the dataset. The total new cases for COVID-19 were collected from CDC's website, which was presented earlier. For example, the state of Louisiana had total 19 191 new cases of COVID-19 in October (CDC, 2021) and had the population size of 4 650 000 million in the year of 2020. When you divide the numbers 19 191 and 4 650 000, the value of the outcome is 0.004127, which is displayed as 'Z' in the formula above. The outcome, which is 'Z', would later be multiplied with 100 000 (0.004127 * 100 000), which represents the 100 000 inhabitants. The final numbers are approximately 413 total new cases of COVID-19 per 100 000 inhabitants in the state of Louisiana. This formula was induced to calculate the total new COVID-19 cases in October 2020.

In order to estimate of the variable's central tendency, the mean, standard deviation, minimum and maximum value needs to be discussed (Jakobsen & Mehmetoglu, 2016, p 33). The mean value for the variable is 7774.84 and the value for the standard deviation is 584.1145. The minimum value for the variable is 68 cases per 100 000 inhabitants, which was the state of Vermont, which had the lowest new COVID-19 cases per 100 000 inhabitants during October 2020. The maximum value for the variable is 2964 cases per 100 000 inhabitants, which was the state of North Dakota, which had the highest amount of new COVID-19 cases during October 2020. The measure of the variable is also displayed in figure 3 (see figure 3 in appendix).

4.5 Moderating variable: Postal Ballot Availability

The moderating variable for this master's thesis is 'Postal Ballot Availability', which is an important variable for investigating the conditional effect of the focal relationship, which is the relationship between the dependent and the independent variable (Jakobsen & Mehmetoglu, 2016, p 111). Furthermore, to test H2, 'Postal Ballot Availability' is the most suitable moderating variable, where it will investigate if postal ballot availability has a conditional effect on COVID-19 and voter turnout. The discussion about postal ballot availability has been essential for this master's thesis and its access was different between various of states in the United States. The postal ballots are the counted ballots that the registered voters have cast. However, not every eligible voter had access to postal ballot, because it was depending on which state they were registered in. The data for the moderating variable is collected from National Conference of State Legislatures (NCSL), International Institute for Democracy and Electoral Assistance (IDEA) and New York Times.

The data displays how some states required additional excuses than COVID-19 for accessing postal ballot (Sullivan, 2020, p 20-22). This was inducted into the state legislatures, relating to absentee and mail voting (NCSL, 2020). This data is necessary to eventually elaborate the conditional effect of the dependent variable and the independent variable, which measures if access to postal ballot within the states effects the relationship between COVID-19 pandemic and voter turnout in presidential election-2020 in the United States. Therefore, the data for the moderating variable measures which states had access to postal ballot and which states allowed some voters to have access to postal ballots within the states. The data displays that there are 43 states where every eligible voter in the state have access to postal ballot, where seven states did not allow every eligible voter to have access to postal ballot (Gamio et al., 2020). According to Gamio, Love & Stevens (2020), the seven states that did not allow every eligible voter to have access to postal ballots was the states of Indiana, Louisiana, Mississippi, New York, South Carolina, Tennessee and Texas (Ibid, 2020). Some eligible voters in these states could send in a postal ballot. However, the seven states needed another reason than the COVID-19 pandemic to vote by postal ballot (Ibid, 2020). These states did not adjust their state legislative for voting during the COVID-19 pandemic (NCSL, 2020). This is an important discussion for this master's thesis, where I argue for the option of postal ballots was an advantage for preventing the increasement of COVID-19 cases during the presidential election-2020 and an advantage for every voter. Furthermore, the postal ballot availability became an advantage for American

citizens with health conditions, where the postal ballot availability would lower the increase of COVID-19 cases and the ability for the citizens to catching the COVID-19 virus. This also specifies the reliability for this variable, where the measurement of would still be the same if it was used in a different context, where this variable measure which states that had postal ballot available for every citizen in the United States (Maslakci & Sürücü. 2020, p 2707).

In order to correctly collect the data for the dataset, the seven states where every citizen does not have access to postal ballots are coded differently comparing to the states where every citizen had access to postal ballot. The seven states Indiana, Louisiana, Mississippi, New York, South Carolina, Tennessee and Texas, where only some voters had access to postal ballots were coded as "0" in the dataset. The other 43 states were coded as "1" in the dataset, which represented that every voter had access to postal ballots. However, the validity of the variable is based on the measurement of the variable, which is state legislature from NCSL. Therefore, the variable measures the data that is intended for this master's thesis and is a theoretical indicator to operationalize the variable into a regression model (Maslakci & Sürücü. 2020, p 2695).

4.6 Control Variables

The two control variables for this master's thesis are 'GDP per capita' and 'Tightness of the Race', which is measured on state-level. The controlling variables aim is to investigate if they have a spurious relationship with the dependent variable and the independent variable, which indicates that they have a cause on the focal relationship (Jakobsen & Mehmetoglu, 2016, p 236). The data for the controlling variables is collected from Statista (2021) and NBC News (2021). Previous research has indicated that there has been political agitation over economic situations in countries, especially during situations where health experts have preferred minimizing physical contact, for example during lockdowns (Johnson et alt., 2020, p 429). Therefore, the variable 'GDP per capita' is theoretical important for this master's thesis and a suitable control variable, where the variable will establish the economic value of each state during the COVID-19 pandemic in terms of affecting the voter turnout. The validity of the control variables measures the data that is expected from the relationship of other contributing variables. Therefore, the variables measure real gross domestic product of the United States on state-level, which is the theoretical definition of the variable while measuring for 'GDP per capita' (Maslakci & Sürücü. 2020, p 2695). The 'GDP per capita' variable would be an addition to existing literature where the variable measures the GDP from a different standpoint. A vast

amount of literature has investigated previous U.S. presidential election turnout on county-level and compared the estimated county support for presidential election-2020 (Baccini et alt., 2020, p 5). Therefore, the variable 'Tightness of the race' is a suitable control variable and theoretical important based on how close the presidential race was in vote share between the Democratic and Republican Party within each state. Previous scholars have also investigated the estimated county support for election 2020, where this variable estimate the vote share on state-level. The measurement of the variable measures how tight the race was between the Democratic and Republican Party, where it is the theoretical definition of how the variable is measured and used in the regressions (Maslakci & Sürücü. 2020, p 2695).

The control variable 'GDP per capita' displays the real gross domestic product of the United States by the third quarter of 2020 (Statista, 2021). The data is used from Statista, which displayed a figure of the real gross domestic product per state in 2020 (Ibid, 2021). The data that is used for the control variable measures the real gross domestic product from the third quarter of 2020, where it displays the timeline of when the presidential election-2020 took place. The measurement of the variable will remain the same when the variable is used for future research, which makes the data reliable based on the operationalization of the variable. It is important to display 'GDP per capita' as a control variable, while investigating the postal ballot availability's effect on voter turnout and total amount of COVID-19 cases in the United States. Therefore, 'GDP per capita' is a suitable control variable, where we can investigate if it has an additional effect on the focal relationship. It is important to acknowledge the central tendency of the variable to get more information of what is measured (Jakobsen & Mehmetoglu, 2016, p 33). The central tendency of the variable usually measures the mean, standard deviation, minimum and maximum value. The mean for the variable is the value of 366.319 and the value for standard deviation is 478.6862. The minimum value for the real gross domestic product per state is 28.52 billion U.S. dollars, which was the state of Vermont, and the maximum value is 2742.92 billion U.S. dollars, which was the state of California. The measure of the variable is also displayed in figure 4 (see figure 4 in appendix).

The control variable is 'Tightness of the race', where the data is collected from NBC News (NBC News, 2021). The data is based on the vote share of the presidential election-2016 from the Democratic and the Republican Party in percentage on state-level. As previous research investigated voting turnout on county-level for the Republican candidate Donald Trump (Baccini et alt., 2020, p 5), it was essential for this master's thesis to apply a variable based on

vote share on state-level from the presidential election-2016. The data is categorized in two columns per state, where the percentage of votes that voted for the Democratic candidate Hillary Clinton in a state and the percentage of votes that voted for the Republican candidate Donald Trump. The vote share of the percent is the data for the variable 'Tightness of the race'. The data is collected by calculating the vote share between Democratic and Republican Party within the states during presidential election-2016. For example, the state of Arkansas had 60.6 percent of the vote share voting for the Republican Party and 33.7 percent of the vote share voting for the Democratic Party (NBC News, 2021). The numbers 60,6 minus 33,7 becomes 26,9, which is the number that is used in the dataset for coding the state of Arkansas for 'Tightness of the race'. However, the estimation of the variable's central tendency, the mean, standard deviation, minimum and maximum value needs to be established (Jakobsen & Mehmetoglu, 2016, p 33). The mean value for the variable is 16.984 and the value for the standard deviation is 11.97213. The minimum value for the variable is 0.3 percent, which was the states of Michigan and New Hampshire, where the race regarding the vote share was essentially close. The maximum value for the variable is 46.3 percent, which was the state of Wyoming, where the vote share had the largest increase from presidential election-2016. The measure of the variable is also displayed in figure 5 (see figure 5 in appendix).

In order to understand the theoretical relationship, I will illustrate the relationship of the variables with a model. The model displays the relationship between the dependent, independent, moderating and controlling variables. Below you will find the illustrated model:

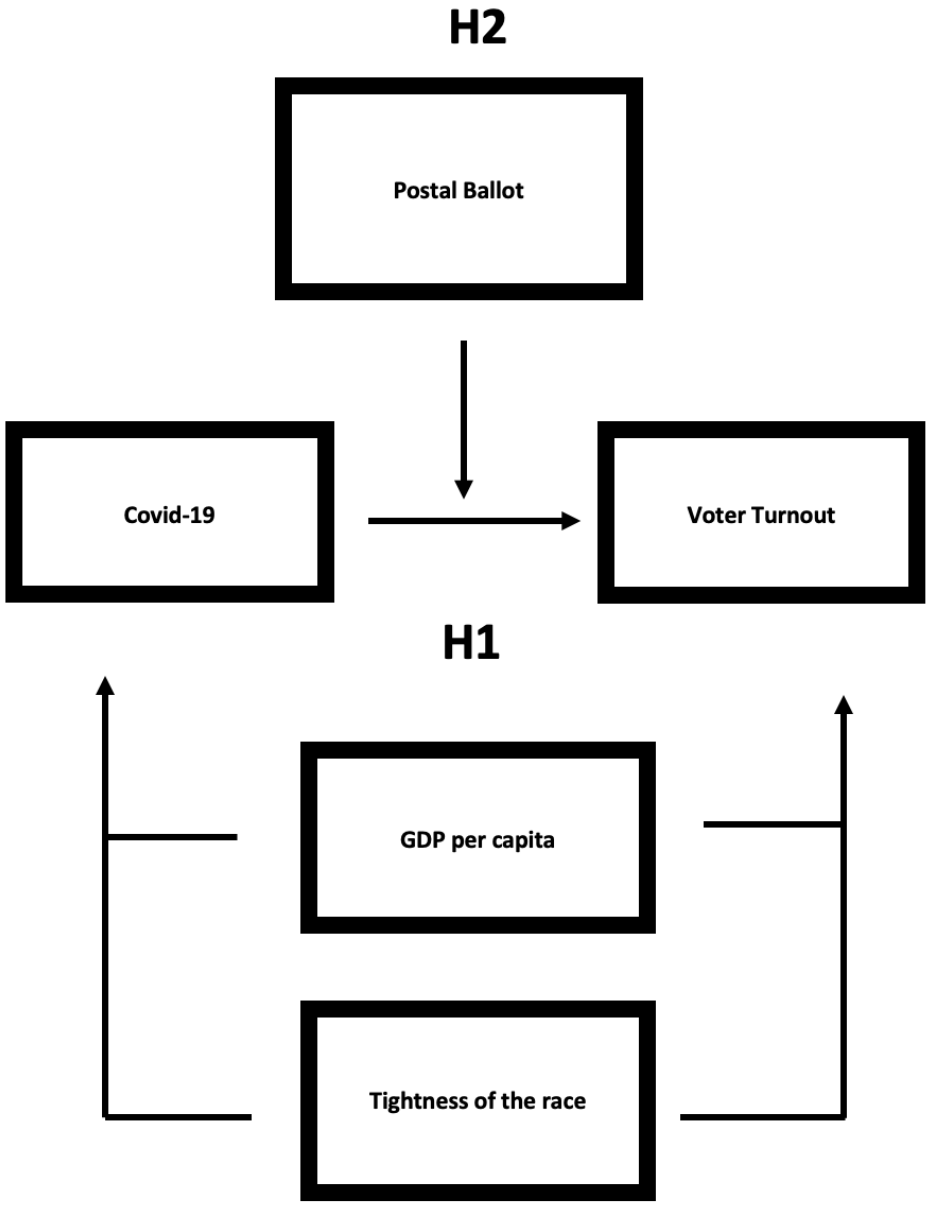

Figure 3. *A model of the interaction of the variables.*

The figure above displays the relationship of all the variables and how they will integrate into the regressions. The model for this master's thesis is important, where it highlights how the variables will be presented in the results. The model displays the research question, which indicates that the COVID-19 pandemic influenced voter turnout during presidential election-2020 in the United States. It also displays H1, '*Voting turnout was lower in states with higher COVID-19 cases'* and H2, '*COVID-19 has a conditional effect on voter turnout, where the postal ballot availability expects to increase the voter turnout in states.'* The spurious relationship that the controlling variables has with the focal relationship, are included in H2 (Ibid, 2016, p 236).

5. Results and Analysis

This section will present the results for this master's thesis, where the results will be displayed through a scatterplot, a regression table and a marginsplot. Below you will find the scatterplot that displays the relationship between 'COVID-19' and 'Voter Turnout'.

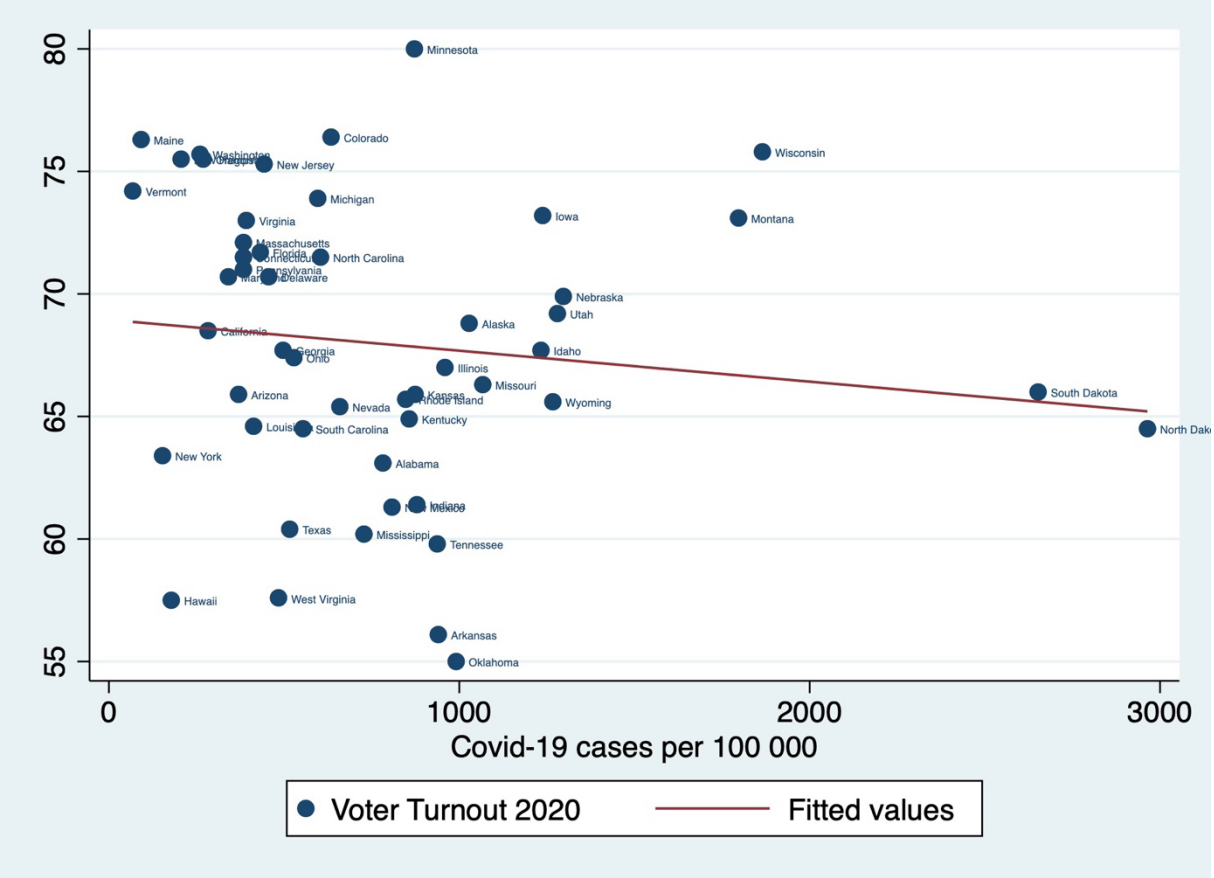

Figure 4. *The effect of COVID-19 on voting turnout.*

The figure above displays the focal relationship, which is the relationship between the dependent and independent variable. The x-axis displays the new COVID-19 cases per 100 000 inhabitants among the 50 states in October 2020 and the y-axis displays the voter turnout per state in percent. The figure displays a correlation between 'COVID-19' and 'Voter Turnout', which indicates that there is a negative effect between 'COVID-19' and 'Voter Turnout' in the presidential election-2020 in the United States. Moreover, we can discuss based on the figure above that some states have lower voter turnout in percentage while the COVID-19 cases are

higher. For example, the state of Oklahoma, have the lowest voter turnout among the states with 55 percent, but the state has the highest increasement of new COVID-19 cases with 991 cases per 100 000 inhabitants. However, the state of Tennessee had also increased new COVID-19 cases, with 937 cases per 100 000 inhabitants, and low voting turnout with 59.8 percent. The state of Tennessee is one of the seven states where all the citizens did not have access to postal ballot during presidential election-2020, where the state had low voter turnout but an increase in COVID-19 cases. Moreover, figure 4 indicates that there is a negative relationship between the COVID-19 pandemic and the voter turnout, but this is only measured between the dependent and the independent variable. Therefore, all the variables involved in the regressions need to be investigated and displayed in a table. Below you will find a regression table of all the variables involved.

	(1) Voter Turnout	**(2) Voter Turnout**	**(3) Voter Turnout**	**(4) Voter Turnout**	**(5) Voter Turnout**	**(6) Voter Turnout**
Covid-19	-0.00126 (-1.11)	-0.00132 (-1.03)	0.000408 (0.30)	-0.00162 (-1.14)	-0.00494 (-0.60)	-0.00343 (-0.48)
GDP per capita		-0.000245 (-0.16)	-0.000311 (-0.19)	0.000802 (0.46)		0.000450 (0.30)
Tightness of the race			-0.281*** (-4.39)			-0.264*** (-4.53)
Postal Ballot Availability				7.455** (3.25)	5.316 (0.97)	4.486 (0.93)
Covid-19 * Postal Ballot Availability					0.00323 (0.39)	0.00353 (0.49)
Intercept	68.94*** (53.12)	69.08*** (39.33)	72.54*** (42.72)	62.51*** (24.14)	64.99*** (12.24)	68.65*** (13.55)

N	50	50	50	50	50	50
R-squared	0.015	0.016	0.307	0.199	0.198	0.455
Adjusted R-squared	-0.005	-0.026	0.261	0.147	0.146	0.393

Standard errors are displayed in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2. *OLS-regression with all variables.*

The table above displays a regression table of all the variables that was used in the regressions. The regression table displays the independent variable, 'COVID-19', the dependent variable, 'Voter Turnout', the controlling variables, 'GDP per capita' and 'Tightness of the race', the moderating variable, 'Postal Ballot Availability' and the interaction term, which is the 'COVID-19' variable and the 'Postal Ballot Availability' variable. The OLS-regressions contain 50 observations, which corresponds to one observation of each state in the U.S. It is important to remember the aim of this master's thesis, which was to identify if the COVID-19 pandemic had an impact on the voting turnout during presidential election-2020 in the United States. This can be answered based on the regression table above. The regression table are divided into six models that adds variables for every model. However, the most important model to investigate through the regression table above are model 6, which contains of all the variables and the interaction term between 'COVID-19' and 'Postal Ballot Availability'.

Model 1 displays the relationship between 'COVID-19' and 'Voter Turnout' without any control variable, where it has a negative effect with the value of -0.00126. The standard error that is displayed in parentheses along with the non-existent star (*), indicates that it is non-significant. Model 2 displays the relationship between 'COVID-19' and 'Voter Turnout', while controlling for 'GDP per capita'. The value of model 2 is still negative with the coefficients -0.00132, which is not statistically significant. Model 3 includes the controlling variables, 'GDP per capita' and 'Tightness of the race' and is the only non-negative model of the OLS-regression, where the coefficients have a value of 0.000408. This indicates that COVID-19 does not have a negative effect on 'Voter Turnout' when the control variables are added to the regressions. The control variable 'GDP per capita' had negative coefficients in model 3 with

the value of -0.000311, which indicates that the domestic gross value in states declines when COVID-19 affects the voter turnout. However, the control variable 'Tightness of the race' is statistically significant in model 3 with the coefficient -0.281***, which indicates that the voter share of the Democratic and Republican Party has a correlation with the relationship of 'COVID-19 and 'Voter Turnout'.

Model 4 includes the controlling variable 'GDP per capita' and the moderating variable 'Postal Ballot Availability', where the model displays the coefficients of -0.00162, which indicates that there is a negative correlation between 'COVID-19' and 'Voter Turnout' when 'GDP per capita' and 'Postal Ballot Availability' is added. The control variable, 'GDP per capita', has a non-significant value of 0.00802, which indicates that domestic gross value in states does not have an effect when COVID-19 affects voter turnout. It is important to acknowledge that the moderating variable is positive and statistically significant, which is the only model that the moderating variable has a statistical significance. However, the moderating variable is displayed as a control variable in the regression table, where the interaction term is the most important relationship to investigate, which is presented in model 5 and 6.

Model 5 includes the dependent variable, the moderating variable and an interaction term, which consists of the 'COVID-19' variable and 'Postal Ballot Availability' variable, where it displayed the coefficient -0.00494, which indicates that 'COVID-19' had a negative non-significant effect on 'Voter Turnout' when the interaction term was added. The interaction term has a non-significant effect in model 5, with the coefficient of 0.00323. Model 6 consists of the independent variable, dependent variable, controlling variable, moderating variable and the interaction term, which is all the variables that was used for this master's thesis, where the coefficients -0.00343 was displayed. The control variable, 'Tightness of the race', is the only variable that is statistically significant in model 6 with the coefficients -0.264***, which indicates that the vote share has a negative impact when the interaction term is applied. The control variable 'GDP per capita' displayed the coefficient of 0.000450, which indicates that the real domestic gross products in the states does not have an impact on the relationship of 'COVID-19' and 'Voter Turnout'. The results for model 6 displays that there is no interaction between 'Postal Ballot Availability' and COVID-19 cases. Therefore, the effect of the amounts of COVID-19 cases on 'Voter Turnout' is not dependent of if postal ballot is available for the states or not.

Based on the results from the regression table above, we can now answer the research question: '*Did the COVID-19 pandemic have an impact on the voter turnout during the election of 2020 in the United States?'*. I argue based on the results that the COVID-19 pandemic did have a negative impact on voter turnout during presidential election-2020 in the United States, which is displayed in table 2. However, 'COVID-19' relationship with 'Voter Turnout' is not statistically significant, which makes H1, 'Voting turnout was lower in states with higher COVID-19 cases' not supported. The coefficient for the variables is not statistically significant, but the figure 4 displays that some states that have higher COVID-19 cases have lower voter turnout. I also argue that there are not enough cases or states for the relationship to be statistically significant, because the correlation is too weak for the relationship to be statistically significant, where the coefficient of 'COVID-19' and 'Voter Turnout' is -0.0001. Moreover, I argue for that H2, 'COVID-19 has a conditional effect on voter turnout, where the postal ballot availability expects to increase the voter turnout' is also not supported by the results. The moderating variable, 'Postal Ballot Availability' has a non-significant effect in table 2, which indicates that the relationship between 'Postal Ballot Availability' and 'COVID-19' has a small effect where the coefficients are at 0.00353. This displays that the effect is too small to detect. In order to investigate the small effect between 'COVID-19' and 'Postal Ballot Availability', a marginsplot is displayed below to present the conditional margins effect.

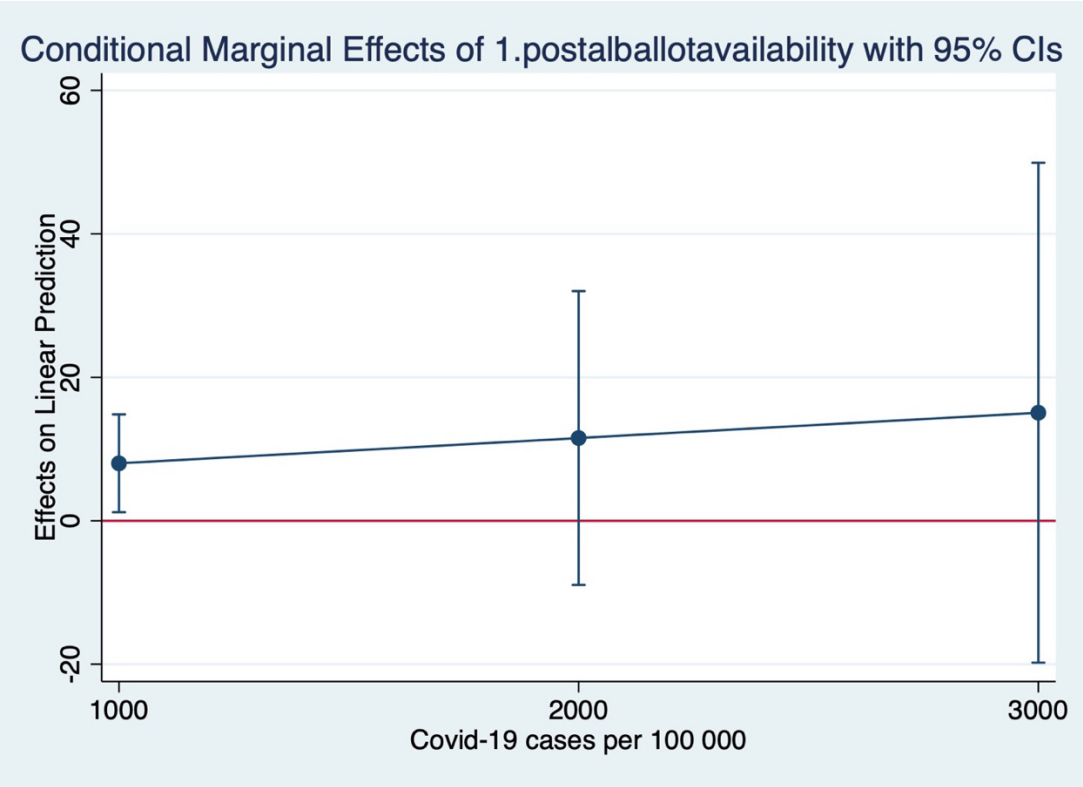

Figure 5. *Marginsplot of the conditional margins effect.*

The figure above displays a marginsplot, which is a graph that illustrate the interaction effect, which is 'COVID-19' and 'Postal Ballot Availability' (Ibid, 2016, p 117). The adjusted predictions for the moderating variable are displayed in figure 1 (see figure 1 in appendix), where it displays how the variable is coded into two groups. One group was coded where some voters in Indiana, Louisiana, Mississippi, New York, South Carolina, Tennessee and Texas did not have postal ballot available, where it is coded as '0'. The other group were the voters in the remaining 43 states in the United States had access to postal ballot availability, which is displayed as '1' in figure 1. The confidence interval identifies if the sampling method is assured or not (Hayes & Scott, 2021). Furthermore, the confidence interval displays that it is not significant. The marginsplot displays that COVID-19 cases decreases when postal ballot is available for every state in the United States, which is why the line increases. However, the states that had 1000 new COVID-19 cases per 100 000 inhabitants is significant with the confidence interval, which indicates that it has an effect if voters have postal ballot available. This means the states with two COVID-19 cases had higher voter turnout if the possibility of postal ballot was available. Furthermore, it will be non-significant for the states that have increased COVID-19 cases. I argue that there are few states in the dataset where postal ballots were not universally available, but this can't be detected since the observations are only at 50 states. Therefore, H2 is not supported which is stated above.

5.1 Diagnostics

In order to investigate if some regression models in table 2 differs from assumptions, different diagnostics tests have been identified to justify the variables that is used. The first test that was applied was the linktest, which identifies if the variables was used correctly in the regressions or if an additional independent variable should be added to the regression (Ibid, 2016, p 135). The linktest displayed that is not significant, with the value of 0.949 and the result for linearity is also non-significant. (Ibid, 2016, p 135). Therefore, a test for misspecification problem was applied, which indicated that the model is useable according to the non-significant test (Ibid, 2016, p 136). A DFBETA-test was later applied, which is a test for the states effect on value of the regression (Ibid, 2016, p 154). The test specifies if there is a difference in the regression while a specific state was enclosed or excluded from the model (Ibid, 2016, p 154). The DFBETA-test is displays in figure 6 (see figure 6 in appendix), where it displays that Hawaii, Montana, West Virginia and Vermont had a higher entity on the independent variable 'COVID-

19’ and Louisiana and Texas had a higher entity on the moderating variable, ‘Postal Ballot Availability’. The test displays that the test has higher unit above 2 or below -2. Therefore, a distance plot (Ibid, 2016, p 156) needed to be regressed, which is a graph that displays the states of the DFBETA-test more closely. While investigate for the state that had the largest impact, they were above the value 2, which indicates that the model does not need to change. A test for heteroscedasticity was applied in table 1 (see table 1 in appendix), which is a test for identifying the variables standard errors changes while observing the states over time (Ibid, 2016, p 150). The test displayed the value of 0.8054, which is defined and indicates that there is no problem with heteroscedasticity because the value is larger than 0.05 (Ibid, 2016, p 150). Lastly, a test for multicollinearity was applied in table 2 (see table 2 in appendix), which is a test for identifying a problem when two independent variables are correlated into the same regression (Ibid, 2016, p 146). There is no multicollinearity problem for any variable, where the VIF value is not over 5 (Ibid, 2016, p 147).

Conclusion

This master's thesis investigated if the COVID-19 pandemic influenced voter turnout from presidential election-2020 in the United States. The COVID-19 virus is a deadly infectious disease, which targeted the economy in the United States and had the largest declension since the Great Depression (Council on foreign relationship, 2021). This led to multiple lockdowns in the United States, during a time when citizens were facing a presidential election. Previous scholars have investigated that there has been political agitation over economic situations in countries, especially during situations where health experts have preferred minimizing physical contact, for example during lockdowns (Johnson et al., 2020, p 429). Therefore, it was essential for this master's thesis to apply the rational choice theory, where the COVID-19 pandemic makes it harder for the citizens to vote when they are currently advised from the government to social distancing, which makes the citizens nervous of exposing themselves to catching the COVID-19 virus. The pandemic is a large cost that was a factor that played part in voter turnout in presidential election-2020.

Moreover, it was essential for this master's thesis to use a cross-section analysis, where the voter turnout from the 50 states in the United States was investigated with new COVID-19 cases through OLS-regressions. The results displayed that COVID-19 pandemic had a negative effect on voter turnout during presidential election-2020. This master's thesis also included a moderating variable, 'Postal Ballot Availability', which identified that postal ballot availability influences voter turnout. Based on the coefficients in table 2, H1 *'Voting turnout was lower in states with higher COVID-19 cases'* is not supported, where I argue that there are not enough cases or states for the relationship to be statistically significant, because the correlation is too weak for the relationship to be statistically significant. The results also display that H2, '*Covid-19 has a conditional effect on voter turnout, where the postal ballot availability expects to increase the voter turnout in states'* is also not supported, where the relationship between 'Postal Ballot Availability' and 'COVID-19' has a small effect and there are not enough cases. The control variable, 'GDP per capita', were not statistically significant which indicates that the real domestic gross products in the states does not have an impact on the relationship of 'COVID-19 and 'Voter Turnout'. The control variable, 'Tightness of the race', was statistically significant but had a negative effect throughout the OLS-regressions.

The research question for this master's thesis was: '*Did the COVID-19 pandemic have an impact on the voter turnout during the election of 2020 in the United States?'.* The answer to the research question is that the COVID-19 pandemic did have a negative effect on voter turnout during presidential election-2020 in the United States, even if the correlation is not statistically significant. I argue for that the option for mail-in ballots set to all voters should be available for every state in the United States in order to reduce the number of COVID-19 cases during the election, which would also reduce the costs. Furthermore, future research should investigate this topic using a multilevel analysis, where the research should be investigated on individual level between the states. The multilevel analysis is different from the cross-sectional OLS-regressions that this master's thesis used. The multilevel analysis investigates two or three regressions models by using hierarchical data, which investigates on an individual-level (Jakobsen & Mehmetoglu, 2016, p 194). Therefore, the multilevel analysis is different from the OLS-regressions, where the OLS-regressions does not measure independent units (Ibid, 2016, p 223). By using this statistical method, an addition to future research will be made.

References

Articles & Books

Aldrich, John H. 1993. *Rational Choice and Turnout*. American Journal of Political Science, Feb., Vol. 37, No. 1, pp 246-278. Available from: https://www.jstor.org/stable/2111531?seq=1 [Accessed 16 February 2021].

Baccini, Leonardo, Brodeur, Abel, Weymouth, Stephen. 2020. *The COVID-19 Pandemic and the 2020 U.S. Presidential Election* (2020). IZA DP No. 13862, pp 1-38. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3730460 [Accessed 19 January 2021].

Blaise, André, Lapp, Miriam, Young, Robert. 2000. *The Calculus of Voting: An Empirical Test.* European Journal of Political Research, Volume 37, pp 181-2001. Available from: https://www.researchgate.net/publication/316674755_The_Calculus_of_Voting_An_Empirical_Test/link/590b90f5a6fdcc5d421ed5eb/download [Accessed 13 February 2021].

Flanders, W. Dana, Flanders, William D., Goodman, Michael. 2020. *The association of voter turnout with county-level coronavirus disease 2019 occurrence early in the pandemic*. Annals of Epidemiology, Volume 49, September 2020, pp 42-49. Available from: https://www.sciencedirect.com/science/article/abs/pii/S1047279720302088 [Accessed 19 January 2021].

Gamio, Lazaro, Love, Juliette, Stevens, Matt. 2020. *Where Americans Can Vote by Mail in the 2020 Elections.* New York Times. Available from: https://www.nytimes.com/interactive/2020/08/11/us/politics/vote-by-mail-us-states.html [Accessed 13 May 2021].

Jakobsen, Tor Georg, Mehmet, Mehmetoglu. 2016. *Applied Statistics Using Stata - A Guide for the Social Sciences*. Sage publications inc.

Johnson, Andrew F., Pollock, Wendi, Rauhaus, Beth. 2020. *Mass casualty event scenarios and political shifts: 2020 election outcomes and the U.S. COVID-19 pandemic*. Administrative

Theory & Praxis, Volume 42, Issue 2, pp 249-264. Available from: https://www.tandfonline.com/doi/full/10.1080/10841806.2020.1752978 [Accessed 19 January 2021].

Landman, Todd, Splendore, Luca Di Gennaro. 2020. *Pandemic democracy: elections and COVID-19*. Journal of Risk Research, Volume 23, Issue 7-8: COVID-19 Special Issue. Available from: https://www.tandfonline.com/doi/full/10.1080/13669877.2020.1765003 [Accessed 19 January 2021].

Maslakci, Ahmet, Sürücü, Lütfi. 2020. *Validity and Reliability in Quantitative Research.* Business & Management Studies: An International Journal, Vol. 8, Issue 3, pp 2694-2726. Available from: https://www.bmij.org/index.php/1/article/view/1540/1365 [Accessed 16 May 2021].

Riker, William H., Ordershook, Peter C. 1968. *A Theory of the Calculus of Voting.* The American Political Science Review, Volume 62, No. 1, pp 25-42. Available from: https://www.jstor.org/stable/1953324?seq=1 [Accessed 16 February 2021].

Reinfeldt, Fredrik. 2020. *3 november 2020 ödesvalet: om en demokrati i fara och en världsordning i upplösning*. [Stockholm]: Albert Bonniers förlag.

Smets, Kaat, van Ham, Carolien. 2013. *The embarrassment of riches? A meta-analysis of individual-level research on voter turnout.* Electoral Studies, Volume 32, Issue 2, pp 344-359. Available from: https://www.sciencedirect.com/science/article/abs/pii/S0261379412001527 [Accessed 27 January 2021].

Sullivan, Kate. 2020. *Impact of COVID-19 on the 2020 US presidential election*. International Institute for Democracy and Electoral Assistance. Available from: https://www.idea.int/sites/default/files/impact-of-covid-19-on-the-2020-us-presidential-election.pdf [Accessed 10 May 2021].

Online Resources

Centers for Disease Control and Prevention. 2021. *United States COVID-19 Cases and Deaths by State over Time.* Available from: https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36 [Accessed 10 April 2021].

Centers for Disease Control and Prevention. 2020. *United States COVID-19 Cases in the Last 7 Days By State and Jurisdiction.* Available from: https://twitter.com/CDCgov/status/1313943466150367238/photo/1 [Accessed 15 May 2021].

CNN Politics. 2021. *Presidential Results.* Available from: https://edition.cnn.com/election/2020/results/president [Accessed 26 March 2021].

Council on Foreign Relations. 2021. *Major Epidemics of the Modern Era 1899-2021.* Available from: https://www.cfr.org/timeline/major-epidemics-modern-era [Accessed 15 May 2021].

Council on Foreign Relations. 2021. *The 2020 Election by Numbers.* Available from: https://www.cfr.org/blog/2020-election-numbers [Accessed 1 April 2021].

Hayes, Adam, Scott, Gordon. 2021. *Confidence Interval.* Investopedia. Available from: https://www.investopedia.com/terms/c/confidenceinterval.asp [Accessed 19 May 2021].

Moving. 2021. *The 10 Largest States by Population.* Available from: https://www.moving.com/tips/the-10-largest-states-by-population/ [Accessed 29 March 2021].

National Geographic. 2021. *Coronavirus in the U.S.: Where the cases are growing and declining.* Available from: https://www.nationalgeographic.com/science/graphics/graphic-tracking-coronavirus-infections-us [Accessed 20 March 2021].

National Conference of State Legislatures. 2020. *COVID-19 and 2020 Primary Elections.* Available from: https://www.ncsl.org/research/elections-and-campaigns/state-action-on-covid-19-and-elections.aspx#absentee/mail [Accessed 15 May 2021].

NBC News. 2020. *Decision 2016: Presidential Results.* Available from: https://www.nbcnews.com/politics/2016-election/president/ [Accessed 26 March 2021].

NBC News. 2021. *U.S. Presidential Election Results 2020: Biden wins.* Available from: https://www.nbcnews.com/politics/2020-elections/president-results [Accessed 27 March 2021].

Statista. 2021. *Preliminary voter turnout as a share of eligible voters in the 2016 United States presidential election, by state.* Available from: https://www.statista.com/statistics/632113/2016-us-presidential-election-voter-turnout-by-state/ [Accessed 15 May 2021].

Statista. 2021. *Real Gross Domestic Product (GDP) of the United States in the Q3 2020, by state*. Available from: https://www.statista.com/statistics/248053/us-real-gross-domestic-product-gdp-by-state/ [Accessed 20 April 2021].

Statista. 2020. *Resident population of the U.S. in 2020, by state (including the District of Columbia).* Available from: https://www.statista.com/statistics/183497/population-in-the-federal-states-of-the-us/ [Accessed 1 May 2021].

The American Journal of Managed Care. 2021. *A Timeline of COVID-19 Developments in 2020.* Available from: https://www.ajmc.com/view/a-timeline-of-covid19-developments-in-2020 [Accessed 9 February 2021].

World Health Organization. 2021. *Archived: WHO-Timeline – COVID-19*. Available from: https://www.who.int/news/item/27-04-2020-who-timeline---covid-19 [Accessed 8 February 2021].

Worldometer. 2021. *United States Population*. Available from: https://www.worldometers.info/world-population/us-population/ [Accessed 14 March 2021].

World Population Review. 2021. *Alaska Population 2021.* Available from: https://worldpopulationreview.com/states/alaska-population [Accessed 29 March 2021].

Appendix

Scatterplot

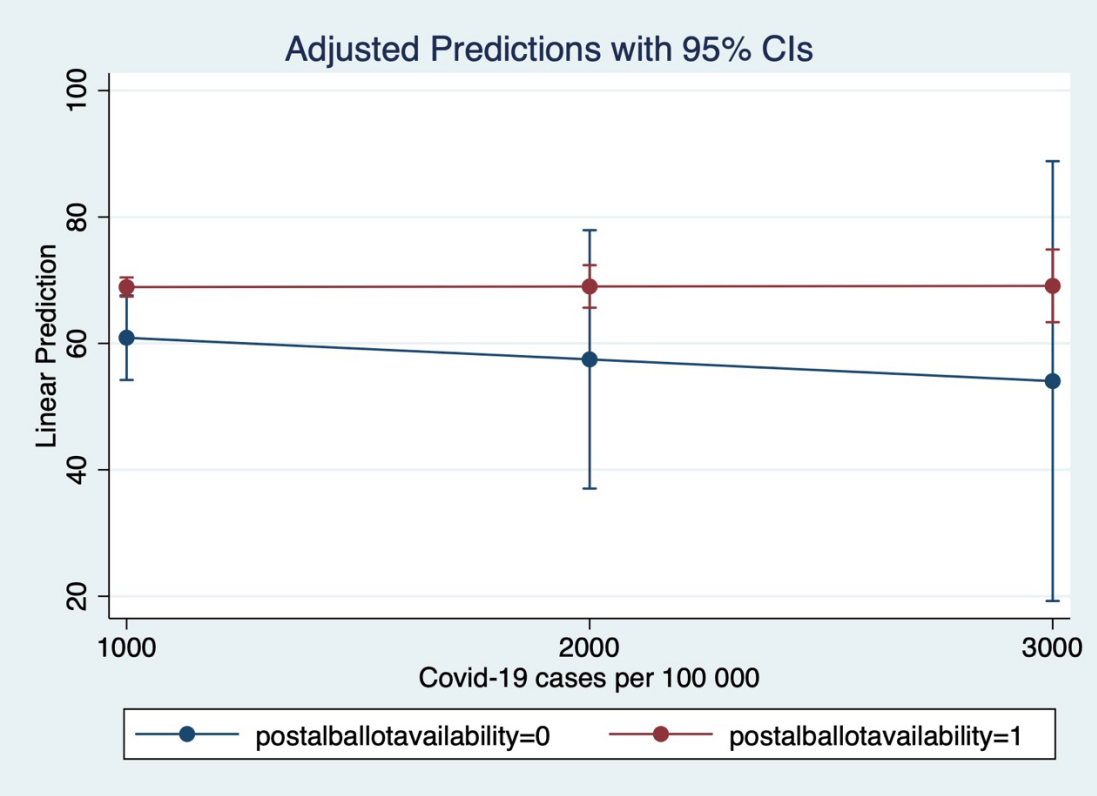

Figure 1. *Marginsplot of the moderating variable, 'Postal Ballot Availability'.*

Histograms

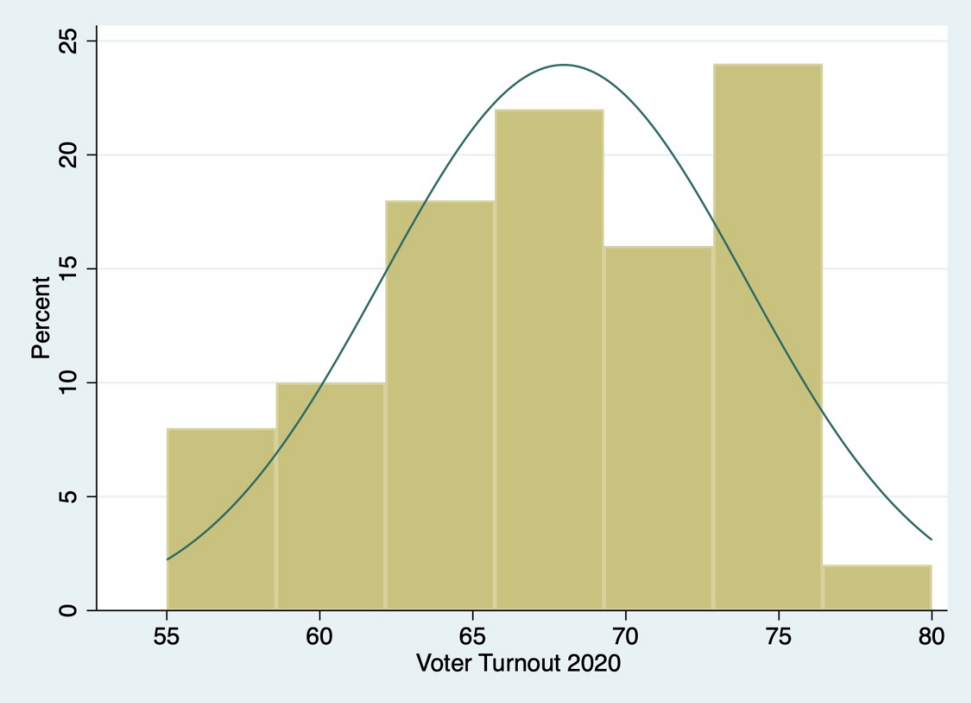

Figure 2. *Histogram of the dependent variable 'Voter Turnout 2020'.*

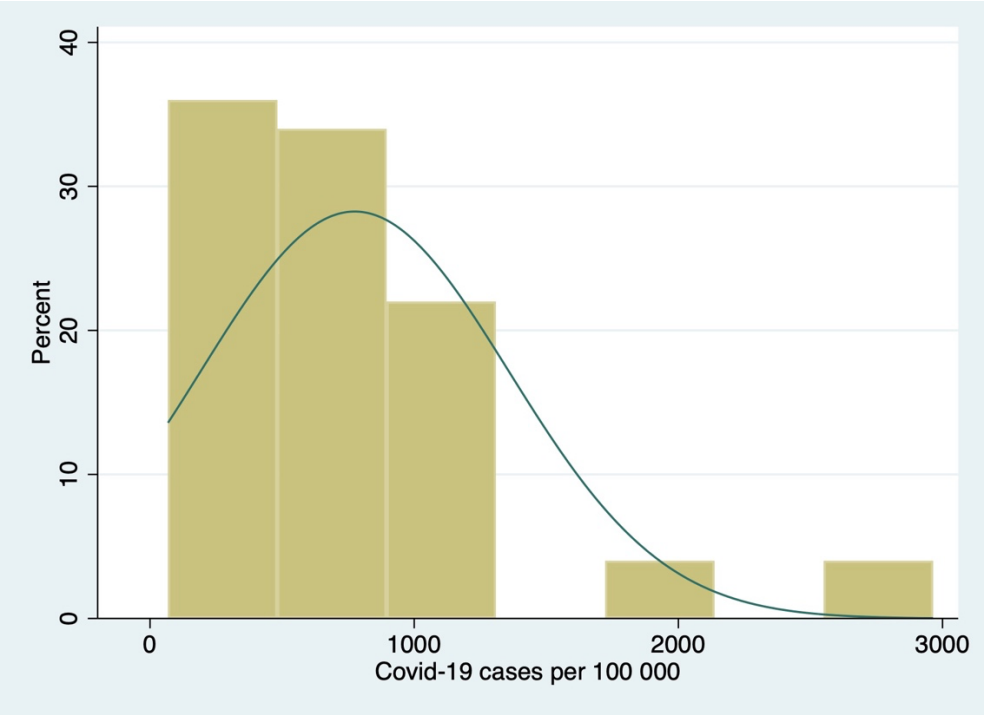

Figure 3. *Histogram of the independent variable 'Covid-19 cases per 100 000 inhabitants.*

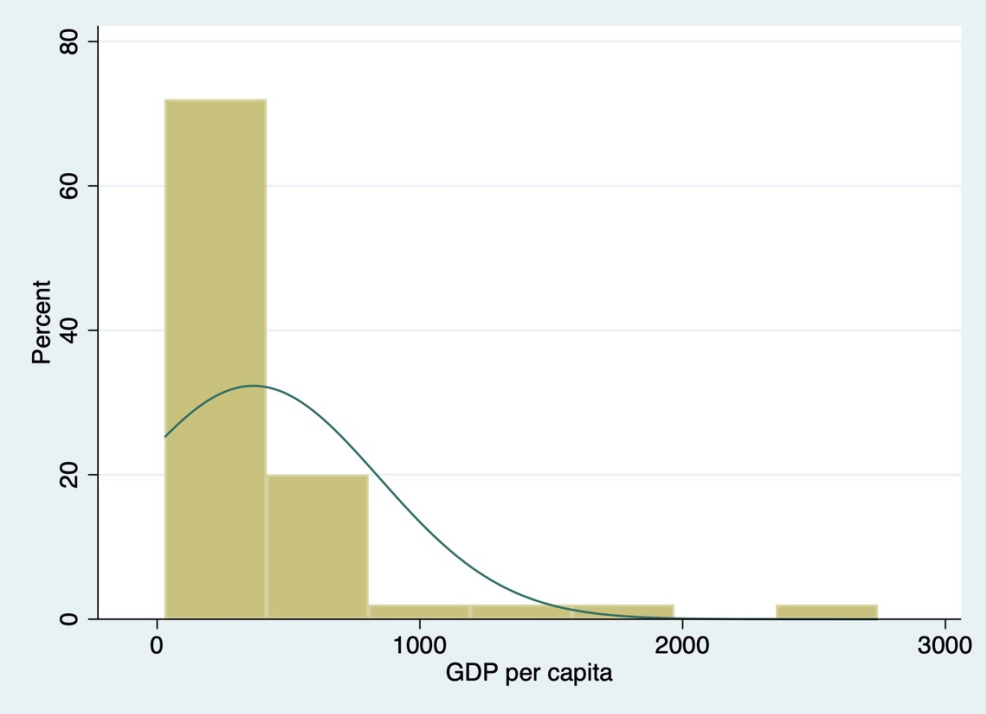

Figure 4. *Histogram of control variable 'GDP per capita', which displays the real gross domestic product of the United States by the third quarter of 2020.*

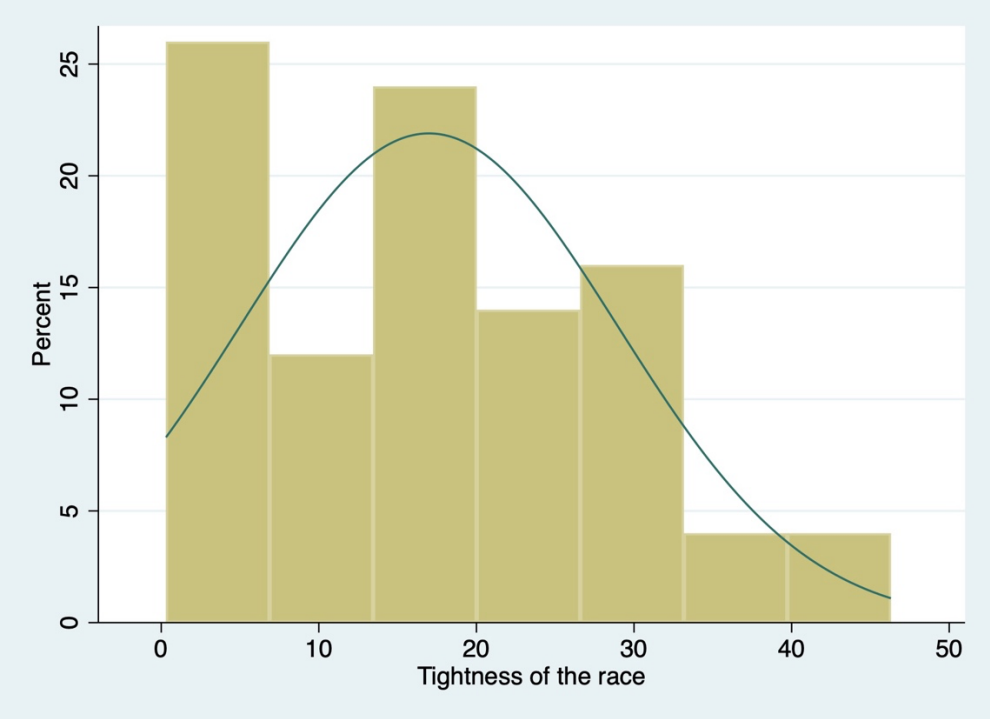

Figure 5. *Histogram of control variable 'Tightness of the race', which displays the vote share of presidential election-2016 from the Democratic and the Republican party in percentage on state-level.*

DFBETA-test

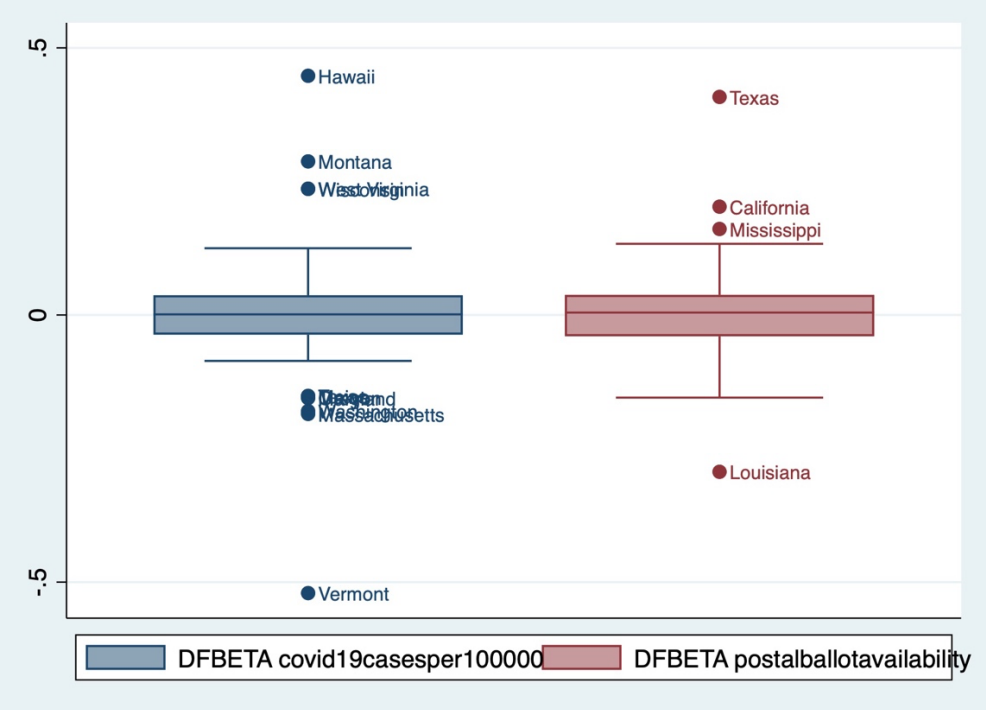

Figure 6. *A figure of DFBETA-test. This figure displays the measurement of the observation effect, where it measures below or above -2 and 2 if it has an effect. Vermont is the only state where the voter turnout is affected by the covid-19 pandemic.*

Heteroscedasticity-test

Breusch-Pagan / Cook-Weisberg test for heteroscedasticity
Ho: Constant variance Variables: fitted values of voterturnout2020 Chi2(1) = 0.06 Prob > chi2 = 0.8054

Table 1. *The heteroscedasticity-test.*

Multicollinearity-test

Variable	VIF	1/VIF
Covid-19	1.21	0.826725
GDP per capita	1.14	0.876619
Tightness of the race	1.11	0.900793
Postal Ballot	1.06	0.943367
Mean VIF	1.13	

Table 2. *The Multicollinearity-test.*