
# MODELING FUTURE OUTBREAKS OF COVID-19 USING TRAFFIC AS LEADING INDICATOR

**Final Report**

**DECEMBER 2023**

**SCOTT PARR[1], SIRISH NAMILAE[2], AND DAHAI LIU[3]**

*[1]Dept. of Civil Engineering, [2]Dept. of Aerospace Engineering,*
*[3]College of Aviation*

*Embry-Riddle Aeronautical University, Daytona Beach, Florida 32114, USA*

**US DEPARTMENT OF TRANSPORTATION GRANT 69A3551747125**

## DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

| | | |
|---|---|---|
| 1. Report No. | 2. Government Accession No. | 3. Recipient's Catalog No. |
| 4. Title and Subtitle<br>Modeling Future Outbreaks of COVID-19 Using Traffic as Leading Indicator | | 5. Report Date<br>December 2023 |
| | | 6. Source Organization Code |
| 7. Author(s)<br>SCOTT PARR, SIRISH NAMILAE, AND DAHAI LIU | | 8. Source Organization Report No.<br>CATM-2023-R7-ERAU |
| 9. Performing Organization Name and Address<br><br>Center for Advanced Transportation Mobility<br>Transportation Institute<br>1601 E. Market Street<br>Greensboro, NC 27411 | | 10. Work Unit No. (TRAIS)<br><br>11. Contract or Grant No.<br>69A3551747125 |
| 12. Sponsoring Agency Name and Address<br><br>University Transportation Centers Program<br>Office of the Secretary of Transportation–Research<br>U.S. Department of Transportation<br>1200 New Jersey Avenue, SE<br>Washington, DC 20590-0001 | | 13. Type of Report and Period Covered<br>Final Report:<br>NOV 2020-APRIL 2023<br><br>14. Sponsoring Agency Code<br>USDOT/OST-R/CATM |
| 15. Supplementary Notes: | | |

16. Abstract

The purpose of this research was to determine the relationship between traffic movements and COVID-19 infections, and ultimately hospitalizations and deaths, throughout various U.S. States using the infection curve and equations from the Susceptible-Infected-Recovered (SIR) model. As a result of state and national governmental restrictions and public perception of the virus, traffic patterns were severely altered throughout the peak of the pandemic in 2020 and 2021. Traffic volumes experienced the greatest reduction when governmental restrictions were first enforced at the beginning of the pandemic and began to approach pre-pandemic values during 2021 as facilities throughout the country reopened. The prediction model applies the traffic volume conditions during the initial stage of the pandemic to the entire study period to determine the effect traffic volumes have on COVID-19 infections. Once the observed infection data were modeled, the adjusted, predicted model was determined using a series of modified SIR equations that reflect changes in traffic, and the findings suggest infection numbers may have been reduced compared to the observed data for each U.S. state studied. The number of hospitalizations and deaths that may be reduced during the second peak given the traffic conditions from the beginning of the pandemic were calculated based on the predicted model results for each state. The findings suggested by the predicted model (i.e., a reduction in infections, hospitalizations, and deaths) can benefit health service facilities by limiting overcrowding and the shortage of ventilators, which can result in fewer deaths caused by COVID-19. This research provides insights for practitioners, researchers, and government entities developing and accessing plans for future pandemics. It is also expected that the findings of this study can be built upon by future researchers who continue to study various aspects of the COVID-19 pandemic and assess the public response to governmental actions.

| | |
|---|---|
| 17. Key Words<br><br>Mobility, pandemic, infection, Susceptible-Infected-Recovered model | 18. Distribution Statement<br><br>Unrestricted; Document is available to the public through the National Technical Information Service; Springfield, VT. |

| | | | |
|---|---|---|---|
| 19. Security Classif. (of this report)<br><br>Unclassified | 20. Security Classif. (of this page)<br><br>Unclassified | 21. No. of Pages<br>95 | 22. Price<br><br>… |

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

## TABLE OF CONTENTS

## EXECUTIVE SUMMARY

The purpose of this research was to determine the relationship between traffic movements and COVID-19 infections, and ultimately hospitalizations and deaths, throughout various U.S. States using the infection curve and equations from the Susceptible-Infected-Recovered (SIR) model. As a result of state and national governmental restrictions and public perception of the virus, traffic patterns were severely altered throughout the peak of the pandemic in 2020 and 2021. Traffic volumes experienced the greatest reduction when governmental restrictions were first enforced at the beginning of the pandemic and began to approach pre-pandemic values during 2021 as facilities throughout the country reopened. The prediction model applies the traffic volume conditions during the initial stage of the pandemic to the entire study period to determine the effect traffic volumes have on COVID-19 infections. Once the observed infection data were modeled, the adjusted, predicted model was determined using a series of modified SIR equations that reflect changes in traffic, and the findings suggest infection numbers may have been reduced compared to the observed data for each U.S. state studied. The number of hospitalizations and deaths that may be reduced during the second peak given the traffic conditions from the beginning of the pandemic were calculated based on the predicted model results for each state.

The findings suggested by the predicted model (e.g., a reduction in infections, hospitalizations, and deaths) can benefit health service facilities by limiting overcrowding and the shortage of ventilators, which can result in fewer deaths caused by COVID-19. This research provides insights for practitioners, researchers, and government entities developing and accessing plans for future pandemics. It is also expected that the findings of this study can be built upon by future researchers who continue to study various aspects of the COVID-19 pandemic and assess the public response to governmental actions.

## 1.0 INTRODUCTION

The SARS-CoV-2 virus, the cause of coronavirus disease (COVID), led to a global pandemic following its emergence in 2019. The COVID-19 pandemic brought unprecedented levels of economic and social disruption. In response to the pandemic, public officials throughout the world issued health directives to limit person-to-person contact. COVID-19 was associated with 704,000 deaths in the U.S. throughout 2020 and 2021, (CDC 2022, CDC, 2023). However, despite the nearly ubiquitous adoption of social distancing measures, the length and severity of governmental actions varied from country to country, and within the United States, from state to state. Transportation, the movement of people and goods, has been used as a surrogate measure of societal interaction. Typically, for members of separate households to meet and intermingle, one or more trips must first be taken. This suggests that traffic counts may be correlated with exposure within a population. It is therefore no surprise that stay-at-home orders led to drastic decreases in traffic volumes throughout the United States (Parr et al. 2020).

This research investigated the relationship between traffic reductions through the various phases of COVID-19 outbreaks across the U.S. in 2020 and 2021, and the subsequent effect on COVID-19 cases on a state basis. The initial wave of COVID-19 outbreaks began in the U.S. in March of 2020. During this phase of the pandemic, compliance with stay-at-home orders and other social distancing guidance was the highest, as measured through significant and sustained traffic reductions. However, as the first wave of COVID-19 infections began to wane in the spring and summer of 2020, phased re-openings of the economy occurred; as such, traffic slowly crept toward pre-pandemic levels (Kristiansson 2021). As cases began to increase again during the fall and winter of 2020 and early into 2021, traffic data suggests that social distancing and stay-at-home orders were not as effective at reducing travel in this phase. Infection rates,

especially within a viral epidemic, depend upon the contact patterns of people in the infected population. A realistic model for the spread of infection takes into account the pattern of mixing within a population, the virulence of the infection, the probability of transmission per contact, and the changes in behavior in the affected population as a result of the epidemic (De Valle et al. 2013). An objective of this research is to investigate if the mixing of individuals, and the corresponding infection rates, within an affected population can be observed and predicted using traffic counts. Trips taken by individuals represent individuals from different households interacting, whether that be through recreational or essential trips, the extent of interaction depends on the type of trip. Human interaction, on a basic level, within a population during an epidemic may be synonymous with traffic volume data. This goal is achieved by developing Susceptible-Infected-Recovered (SIR) models that can predict infections based on observed data.

In response to the global pandemic that made its way to the U.S., state governments issued State of Emergencies in early March of 2020, all within a week of each other, when cases first started to be tested for and confirmed within the Johns Hopkins COVID-19 database (Kristiansson 2021). These mandatory lockdown restrictions included stay-at-home, social distancing, and mask mandates that were implemented at the discretion of the Center for Disease Control (CDC) to control the spread of the virus throughout the country. As cases began to relatively decline in the following months, the spring and summer of 2020, state governments began to reopen the economy in a phased system, with certain mandates, including social distancing and wearing masks indoors, remaining enforced. There were a variety of factors that affected the phased reopening of States, including the severity of COVID-19 infections within certain States and the political makeup of the governments within each state, both of which vary among the ten States chosen for this study. The earliest instance of a phased reopening occurred

in Vermont on April 17, 2020; meanwhile, the latest reopening of the economy was witnessed in New York, which occurred on June 8, 2020 (Kristiansson 2021). The remaining states had initial phased re-openings in mid to early May, with more restrictions being lifted a month later.

Daily traffic volume counts are used in this study to determine the amount of contact and social mixing that exists within a certain area. Higher levels of daily traffic volume relate to more trips being taken by individuals within a household, which are more likely to interact with an individual from a different household. Traffic volumes were observed to be lower throughout the study period compared to the baseline, same-day volume counts of 2019. The infections and traffic counts were broken up into two separate study periods, the first being defined as the initial peak when cases began increasing in the U.S., March to June of 2020. The second peak is defined as the resurgence of cases that occurred from September 2020 to March 2021 for most states. The average reduction of traffic during the first peak was a 41.3 percent decrease, while the second peak only experienced a 15.8 percent decrease in traffic. The reopening of the economy and lifting of governmental mandates is the most likely cause of the increase in traffic between the two peaks identified.

As seen in Figure 1, the difference between traffic volumes in 2020 compared to the same-day volume in 2019 for New York is greater during the first peak of COVID-19 cases in early 2020 (53.2 percent reduction) than the second peak in late 2020 (20.1 percent reduction). The number of COVID-19 cases in the second peak in New York is more than twice that of the initial peak; however, traffic volumes during this time are approaching the values of 2019. This phenomenon can be seen throughout all of the U.S. States identified. This trend describes the public's willingness to adhere to the lockdown mandates of their respective state began to decrease as time passed; similarly, the severity and enforcement of governmental mandates,

especially in terms of individual businesses willingness to enforce mask mandates, also began to wane as the months within the pandemic passed by.

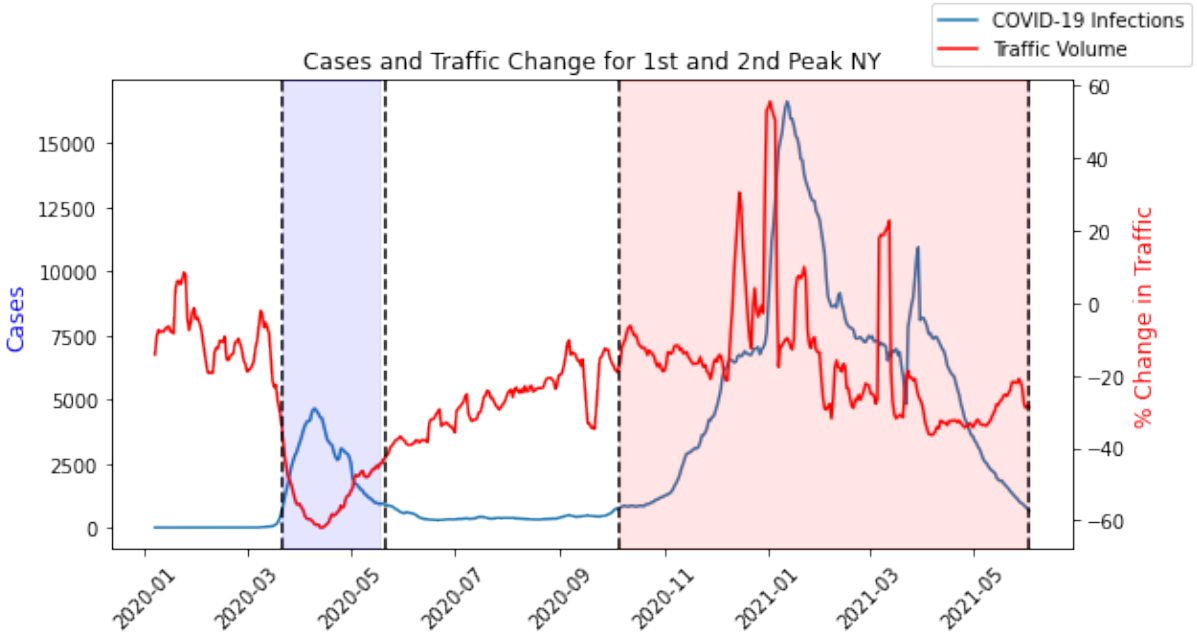


***Figure 1: Seven-day rolling average COVID-19 cases and difference between same-day traffic volume in study period and 2019, including first and second peak, in New York.***

A crucial objective of this research is to determine if traffic counts can be used as a measure that can effectively flatten the curve on local, regional, and global infections. Reducing the maximum peak of infection curves, and total number of infections throughout the duration of a pandemic, can help to minimize the total number of deaths within a population. There is an inherent risk of health services overcrowding during an epidemic, which can further lead to a higher mortality rate of those infected. Almost one in four COVID-19 deaths could be contributed to hospitals being strained by a surging caseload (Sameer et al. 2021). A flattening of the infection curves in each of the states studied can help to mitigate the rise in mortality rate of infected individuals and potentially result in a reduction in the number of deaths throughout the

epidemic surge. The model created within this study will then be utilized to predict the reductions in total infections due to potential policy adaptions and retractions, school closure and phased reopening, for example. This modeling approach analyzes the counterfactuals, as to how the infections could potentially change if lockdown procedures and people's behavior, in terms of travel movements, were the same in late 2020 and early 2021 as they were when COVID-19 initially broke out in the U.S. This method results in the predicted infected curves that peak below the observed COVID-19 infection data for each state. The reduced infection curves describe the extent to which infections may have been reduced in each state, resulting in less hospitalizations overall, which reduces the strain on health services, further reducing the number of deaths caused by the COVID-19 virus throughout the U.S. The conclusions based on the results of the model created will then be used to predict the effectiveness of polices instilled by state government officials throughout the COVID-19 epidemic.

## 2.0 LITERATURE REVIEW

Modeling outbreaks of pandemics has been an area of study long before the outbreak of COVID-19 in the U.S. and around the world. The widespread, global nature of the most recent virus has led to an increase in popularity and necessity for modeling the propagation of COVID-19, and future instances of viruses, in many populations. There are many methods that can be utilized to model infection outbreaks, the sections that follow focus on the classical Susceptible-Infected-Recovered (SIR) model and the Nelder-Mead Estimation approach that is used to determine the parameters that allow for the SIR model to be effective for multiple populations. The unprecedented nature of the virus and its effects on countries throughout the world led to many changes in societies, especially in the U.S. The main change identified in this study is the reduction of traffic volumes in the U.S., as well as the status of hospitals and health services as a response to the growing number of infection cases. The onset of infected cases as a result of increases in traffic volume, or increased interaction amongst those within a population, does not happen in the same day. The offset of positive COVID-19 tests in relation to traffic volumes was studied, determining the time lag in which infection cases show as a result of previous activity. The viral characteristics of COVID-19 are studied as seasons changed, as well as when the dominate variant changed, which both affected the infection rate of the virus.

### 2.1 Epidemiological Models and Modeling COVID-19

Several predictive epidemiological models were developed for countries around the world as a result of health and economic effects caused by this widespread COVID-19 virus. To predict the effects of an ongoing epidemic, the parameters that are applied to the equations that produce the final model must be estimated. The observed COVID-19 infection numbers for a

country, or state, are compiled using daily reported numbers. For the country of Algeria, a country with a population of 43 million, inputs for the parameter estimation for a predicted infection curve were the daily reported case numbers and the population according to United Nations reports (Lounis 2020). The parameters within this study are estimated utilizing the differential equations that make up the classic SIR model. The results of the study determined the date of the peak of infection cases and estimated 800,000 infected individuals during the study period; the reproduction number, R_o, is predicted to be 1.23 (Lounis 2020). A study that looked at populations within Pakistan created an estimation of the COVID-19 epidemic by creating a numerical simulation using the Non-Standard Finite Difference (NFDS) scheme, resulting in figures that can be interpreted and used by government agencies (Din 2021). There are several factors that cause changes in the number of infections within a pandemic, including changes in global population demographics and distribution, human behavior change, and the breakdown of public health systems (Church 2004). Local governmental lockdowns have a major influence on infections, directly causing changes in the transmission rate of individuals within a population. Governmental directives were the leading factor for the epidemiological simulation of the population in India, producing parameters that led to a flattening of the infection curve, reducing the epidemic spread of COVID-19 (Bagal et al. 2020).

The well-studied classical compartmental epidemic models used in most of the above studies such as SIS (Susceptible-Infected-Susceptible), SIR (Susceptible-Infected-Recovered) and SIRS (Susceptible-Infected-Recovered-Susceptible) divide the host population into susceptible, infected, and recovered compartments with a set of differential equations describing dynamics between these different compartments (Brauer et al. 2010). Classical and fractional order SEIR (susceptible, exposed, infections, removed) Ebola epidemic models were used to

analyze and estimate epidemic parameters to create predictive models for future Ebola epidemics (Gendreau 2015). The use of epidemiological models was also used to counteract contagions. Contagions are also a part of the socio-economic issues, such as governmental intervention, where the use of mathematical analysis can bode favorable results in important decision-making processes for effective mitigation of epidemic. In the conventional SIR dynamics model, the population is divided into three compartments using first order differential equations (Brauer et al. 2010), mainly S – Susceptible, I – Infected and R - Recovered population. The parameters β (transmission rate per capita) and γ (recovery rate) govern the rate at which populations move from one compartment to another, respectively. The β rate parameter is the rate at which susceptible population becomes infected during an epidemic. The γ rate parameter determines the rate at which infected individuals recover from the infection causing epidemic. Unlike the SIS, SEIR and SIRS dynamics models, the SIR model operates with the assumption that recovered individuals gain immunity from the contagion and do not get infected again. It is one of the simplest forms of epidemiological model with only two constant rate parameters governing the dynamics of the model. (Islam 2019).

### 2.2 Control Methods in Epidemiological Models

The least-squares method is used for the estimation of parameters with the highest probability, or maximum likelihood, of being correct given some critical assumptions. The Nelder-Mead Estimation method is a method that is described for the minimization of a function with certain variables, which depends on the comparison of function values at the vertices of a general simplex, followed by the replacement of the vertex with the highest value by another point (Nelder 1965). The simplex created adapts itself to the local landscape and contracts onto

the final minimum; the method is effective and computationally compact (Nelder 1965). The Nelder-Mead Estimation, a least-squares method, can accurately predict parameters for SIR model equations based on observed infection data. The COVID-19 epidemic progression in Cameroon was modeled through the early stages of 2020 by using Nelder-Mead Estimation to find the parameters, leading to the conclusion that the peak of infections could have occurred at the end of May, with 7.7 percent of the population being infected out of the 25 million population; this conclusion and usage of this estimation method does not include the intervention of governmental mandates (Nguemdjo 2020). There are a few coding languages that utilize this method of least-squares, and more specifically Nelder-Mead Estimation, a number of which were utilized for the aforementioned COVID-19 modeling studies.

### 2.3 Traffic Mobility as a Result of COVID-19

COVID-19, closures, and other measures have had significant impacts on general mobility. In the Netherlands, 80 percent of people had an overall trip reduction of 55 percent due to a reduction in outdoor activities (de Hass et al. 2020). A study in Australia showed that household trips were reduced by more than 50 percent across all modes of travel. Transit trips were reduced from 14 percent pre-lockdown to only seven percent (Beck and Hensher 2020). Traffic volumes in Florida dropped by 47.5 percent looking at roadway detectors across the state (Parr et al. 2020).

Other studies examined the association between COVID-19 spread and mobility reductions. Carteni et al. (2020) conducted a study in Italy using a multiple linear regression model to show the similarity between positive COVID-19 cases and transportation accessibility in an area. Accessibility contributed about 40 percent in weight to new COVID-19 cases and

weighted heavier than the other variables. The greater the transport accessibility is of an area, the easier it is for the virus to reach the population. The authors stated that the study shows that a more sustainable policy for restrictions and lockdowns to containing social interactions, could be to look more closely to the proportionality of transport accessibility in the area of interest. The greater the accessibility is, the more restrictive policies on mobility should be implemented (Carteni et al. 2020). Another study in the United Kingdom, described how mobility reductions caused significant decreases in COVID-19 cases (Hadjidemetriou et al. 2020). A study on the relationship between daily trips in the U.S. and the COVID-19 infections in the near future used time-series forecast models to project future trends from November 2020 to February 2021 (Truong & Truong 2021). The study discovered a closed loop scenario, where people's travel behavior dynamically changes depending on their risk perception of COVID-19 in an infinite loop. This loop can only be broken if proper and prompt mitigation strategies are put in place to reduce the burden on hospitals and healthcare systems, thus saving more lives.

Mobility reduction has an impact on the spread of COVID-19. Mobility data from Google were applied to the effective reproduction rate, R_t, a measure of viral infectiousness (Noland 2021), to understand the impact from reducing six different trips and activates. The study shows that "Staying at home" is effective in lowering the R_t value. "Activities at parks" appear to not have a significant effect on increase R_t. A return to baseline levels of activity for transit, workplaces, and retail, will increase R_t. 20–40 percent of mobility reductions are needed to attain an R_t below 1.0. A reproduction number below 1.0 is the point at which a disease will perish within a population, above 1.0 and the disease will spread more quickly. The World Health Organization initially estimated the reproduction number for COVID-19 to be 1.4-2.4

(Achaiah 2020). Noland (2021) cautions policy makers about encouraging people to return to normal mobility behavior, particularly when it comes to, transit, workplaces, and retail locations.

A study in Japan showed that during the initial stage of the pandemic, no strong restrictions, such as lockdowns, were put in place by the government (Hara & Yamaguchi 2021). Even though there were no major restrictions, the study detected nation-wide behavioral change using mobile phone network mobility data. During the "state-of-emergency" in Japan, results showed a significant reduction in inter-prefectural travel and trips without strong restrictions from the policymakers. Another interesting finding was that the population density index decreased by 20 percent as people avoided traveling to these densely populated areas. The study showed that after the state of emergency was lifted, people's behaviors did not immediately return to pre-pandemic levels, but instead, recovered slowly. A study in Poland looked at the overall reduction in travel time after the Polish government introduced administrative measures to slow down the spread of COVID-19. A significant decrease in travel times was observed, with no difference between age groups and gender. The more a respondent experienced a fear of COVID-19, the more he or she shortened their daily travel time and stayed closer to home. (Borkowski 2021).

### 2.4 Governmental Directives as a Result of COVID-19

The impact of policymakers and governmental stringency has a major impact on people's actions and adherence to social distancing. Wang (2021) assessed the impact of national culture and government policies from major economies, on social distancing to lower the spread of COVID-19. Government enforcement has a much larger impact on social distancing than what national culture does. There is clear proof that social distancing increases with government

stringency. There are two cultural dimensions that matter when it comes to social distance: it decreases with 'Long-term Orientation'; and increases for 'Indulgence'. The results show that it is necessary for policymakers to act decisively instead of blaming the culture (Wang 2021). Lower COVID-19 infection and mortality rates have been shown to be linked to stricter enforcement policies and more severe penalties for violating stay-at-home orders (Mahmoudi et al. 2021). Policies that allow gradual relaxation of travel restrictions, social distancing and facemask usage, are connected to lower COVID-19 infection and mortality rates (Wang 2021).

A study on China, constructed a city-based epidemic and mobility model (CEMM) to stimulate the spatiotemporal of COVID-19, using multi-agent technology and big data on population migration (Wei et al. 2021). The urban network perspective model emphasizes the important role of high-speed transportation networks and intercity population mobility. The model was able to simulate the initial stage of the inter-city spread of COVID-19 with high precision. The simulation showed that the total number of infectious cases in China would have been 4.46 times higher, 138,824 cases in February 2020, if the city lockdown decreasing population mobility did not occur (Wei et al. 2021).

### 2.5 Heath Service Overcrowding

The quick and widespread onset of the novel Coronavirus in the United States caused strain upon health services in many U.S. States. Hospitals experienced emergency department and intensive care unit overcrowding, as well as a shortage of ventilators since COVID-19 affects the respiratory system of the infected individual. Of 625 hospitals throughout 29 U.S. States, 63 percent of hospitals experienced at least one of the aforementioned alerts from March 7, 2020, to April 30, 2021. Of the hospitals that experienced at least one instance of lack of

emergency services, 63 percent experienced emergency department overcrowding, 61 percent experienced ICU overcrowding, and 12 percent experienced ventilator shortages (Sandhu et al., 2022). The strain on health service centers throughout the country limited the amount and quality of care individuals were able to obtain once admitted to a hospital. During July 2020 to July 2021, which included the SARS-CoV-2 B.1.617.2 (Delta) variant, it was predicted that as intensive care unit beds were at 75 percent capacity, this resulted in 12,000 excess deaths two weeks later. This is compared to when hospitals experienced 100 percent capacity of intensive care unit beds, a predicted 80,000 deaths were expected in the two weeks following (French 2021). Reducing the number of infections in critical areas can lead to the reduction of admittance into hospitals as a result of COVID-19, ultimately leading to a reduction of deaths in the weeks following.

### 2.6 Reduced Mobility Impact on Traffic Crashes and Fatalities

An investigation in Connecticut studied the impact of COVID-19 stay at home orders on daily vehicle miles traveled (VMT) and motor vehicle crashes (Doucette et al. 2021). Looking at the crash severity and number of vehicles involved in crashes from January through April, the study found that the daily VMT decreased by 43 percent. A decrease in daily counts of crashes was noted, but the single vehicle crash rate increased 2.29 times and the single vehicle fatality rate increased 4.10 times. The study concluded that the potential role of reduced police presence, less congested roads, and speeding could contribute to these results. The high speed-related fatal crash rate in Japan during the COVID-19 lockdown were higher than pre-lockdown (Inada 2021). The authors reviewed police data on crash fatalities between January 2010 and February 2020 in which motor vehicle drivers were at fault and found that speeding, speed enforcement by

police, and driver behavior during lockdown were leading causes for the increase in crash fatalities. Traffic crash patterns before and after the outbreak of COVID-19 in Southern Florida for the first half of the years of 2019 and 2020 (Lee & Abdel-Aty 2021), shows a considerable reduction during March to June 2020. The total numbers of crashes decreased by 21 percent, with the most significant reductions occurring during morning peak-hour (33.3 percent), crashes involving alcohol/drug (58.0 percent), and pedestrian crashes (38.3 percent). Another study in Florida (Pierre et al., 2021), looked at the impacts of the COVID-19 pandemic on traffic crashes on freeway (I-10, I-75, and I-95). The paper showed that since the first confirmed COVID-19 case in Florida, there was a decrease in the total traffic crashes, dropping significantly by up to 45.3 percent. A decrease in the rear-end crashes and an increase in the run-off-road were observed. Calderon-Anyosa & Kaufman (2021) conducted studies on external causes of death such as suicide, homicide, and car crashes during COVID-19. The authors wanted to understand how violent and accidental deaths were impacted by the COVID-19 lockdowns. After the lockdown, all forms of deaths suddenly dropped. The largest decrease was seen in traffic related accident deaths, with a reduction of 12.22 and 3.55 deaths per million per month men and women, respectively. Homicide and suicide presented a similar decrease in the initial stages of the lockdown for women while homicide in men increased by 6.66 deaths per million men per year (Calderon-Anyosa & Kaufman 2021).

### 2.7 Reduced Mobility Impact on Air Quality

Prior research has investigated the environmental impact of the COVID-19 global pandemic, social distancing, and subsequent changes in human behavior on air quality. Local lockdowns within states, cities, or whole countries, helped in improving the air quality (Carteni

et al. 2020). Changes in traffic volumes had a direct impact on air quality within the United States. For a study conducted in Florida, traffic volumes and air pollutant concentrations were collected for the span of the pandemic throughout 2020, including the time periods encompassing the lockdown, pre-lockdown, and post-lockdown. During this lockdown period, when the percent change in traffic volumes was the largest compared to 2019, Ozone levels experienced a decrease of 21.3 percent for a monthly average, CO had a reduction of 13 percent and NO2 concentrations decreased by an average of 14 percent for this same time period. PM2.5 concentrations lagged behind the other pollutants, decreasing by 26 percent in the post-lockdown period throughout May (El-Sayed 2021). The reduction of the NO2 pollutant was observed due to the reduction of vehicular emissions, while other factors could have caused decreases in the other pollutants. Light vehicles experienced higher changes in volume compared to heavy vehicles, most likely due to the importance of trips taken by light vehicles being less critical when governmental mandates were being enforced throughout Florida.

In Italy, an analysis of its carbon footprint indicator found the country's carbon footprint shrank by approximately 20 percent (Cartenì et al. 2020). Lockdown and social distancing created a decrease in traffic movements, which correlated to a direct decline in PM2.5 concertation. (Chauhan & Singh 2020). After observing air quality, meteorological parameters, and mobility in six major Italian cities, the authors found that road traffic was reduced by 48-60 percent, NO2 reduced by 24-59.1 percent, and PM2.5 by 17 percent. O3 levels remained essentially unchanged or showed a slight increase of up to 11.4-14.7 percent (Gualtieri et al. 2020). Most studies focused on the impact on PM2.5, NO2, and ozone since these are criteria pollutants, and they can harm people's health and the environment and cause property damage.

An analysis of 50 capital cities worldwide found a 12 percent decrease in particulate matter emissions (PM2.5) (Rodríguez-Urrego 2020). An analysis in Wuhan City, China found the average monthly air quality index, improved by 33.9 percent during the lockdown and PM2.5 decreased by 36.5 percent. Nitrogen dioxide (NO2) decreased by in the city by 53.3 percent. However, Ozone (O3) increased by 116.6 percent (Lian et al. 2020). An analysis conducted in Rio de Janeiro, Brazil also found increased levels of O3 while showing decreased levels of NO2 and carbon monoxide (CO) (Siciliano et al. 2020). An investigation in the continental United States found NO2 reductions of 25.5 percent and statistically significant reductions of PM2.5 (Berman & Ebisu 2020).

### 2.8 COVID-19 Case Date Offset

When exposed to COVID-19, symptoms usually take about five days to appear in a newly infected person. Some people experience symptoms as soon as two days after being exposed. The majority of people infected show COVID-19 symptoms after 12 days, and most people were sick by day 14 (Nazario 2020). Another study on 181 confirmed cases shows an incubation time of 5.1 days for COVID-19 and 97.5 percent of the study group experienced symptoms within 11.5 days of being infected by the virus. In some cases, people develop symptoms after 14 days of being exposed (Lauer et al. 2020). A study conducted on COVID-19 reporting in New York City estimated a mean delay in reporting as five days, with 15 percent of cases reported after ten or more days (Harris 2020). Two other studies conducted by Parr et al. (2020, 2021) were temporally offset from the traffic data by two weeks (that is, the COVID-19 case data is reported for the date two weeks prior to its posting). Kristiansson (2021) explains that health data was temporally offset from traffic data to associate the extent of COVID-19

cases to traffic conditions during the approximate time when the infection occurred. A correlation analysis between the number of daily cases and the decrease in traffic from 2019 to 2020 was conducted and found that, for the 13 regions identified, the average days offset was five (Kristiansson, 2021). The Centers for Disease Control and Prevention (CDC) recommend this period to account for viral incubation and testing time (CDC, 2024).

### 2.8. COVID-19 during Winter Months

Contaminations caused by many respiratory viruses, including influenza and some coronaviruses, increase during winter and decrease during summer (Mallapaty 2020). An increased risk of transmission occurs when people interact indoors and in places with poor ventilation. Studies show that the COVID-19 virus favors dry, cold conditions. The virus degrades faster on surfaces in more humid and warmer environments. During the winter period, people will usually heat their houses and the air is dry and not well ventilated. The Director for Centers for Disease Control and Prevention (CDC) Robert Redfield predicted that the COVID-19 pandemic would take a severe turn for the worse during the winter months (McEvoy 2020). Redfield stated that January and February will be the hardest time for the US, in all history of public health. He also predicted that the death toll rate would increase by 50,000 every two months, resulting in 450,000 deaths by the end of February. CDC released warnings and advised against traveling for Thanksgiving and the holidays. Nine million airport travelers were reported during the holiday period (Newburger 2020). A coronavirus task force coordinator from the White House, Dr. Deborah Birx, made a statement saying that the U.S. will be seeing a large increase in new COVID-19 cases, deaths, and hospitalization in the whole country after the holidays.

A review of the literature mentioned shows that there is a gap in COVID-19 infection models for U.S. States. A vast amount of research has been done in modeling the early stages of COVID-19 in many countries, however, there exists a need to show how infections interact with other factors in a society. Traffic volumes can be a leading indicator for infections as it represents the behavior of the population and how it is interacting with itself. There exists a need for an analysis of how governmental directives play a role in the onset of infections, many of the literature reviewed mentions that governmental directives have a role in the number of infections; however, each scenario is not thoroughly addressed and modeled, providing numerical results. This study seeks to provide data that can be utilized at the state or national level as insight on the effect that mandates and restrictions can have on hospitalizations and deaths in a population for future cases of viral outbreaks. The research presented seeks to build upon the prior knowledge and expand the scientific understanding of SIR modelling and the many factors and social interactions that exist between infections and traffic.

# 3.0 METHODOLOGY

Traffic volumes from 2019-2021 and COVID-19 case infection rates between initial and second wave outbreaks were analyzed for ten U.S. States. This data was used to obtain the rate parameters within the SIR model for the respective states identified for the two phases separately. The parameters are utilized within the classical SIR model equations to be able to model the propagation of cases during a certain period of time. These rate parameters are then correlated to traffic reductions compared to the 2019 baseline. For all the states identified throughout the U.S., the second peak of COVID-19 cases was much greater than the original peak in March and April of 2020, as the data will show in the remainder of this section. However, the traffic reduction was much higher in the first phase, indicating more effective social distancing and reduced exposure. The increased number of cases in the second peaks were likely due in part to the reopening of states, resulting in a higher number of vehicle counts and trips being made throughout the country, leading to more exposure of those within the population.

The objectives of the data and methods used in this research are to develop a model that can be utilized to predict the number of infections in ten U.S. States. The number of infections can then predict the extent of hospitalizations and deaths that may be reduced from these populations. This section provides details on the type of data that was collected and manipulated for the analysis, as well as the methodology for the model formulation that was applied throughout the study.

## 3.1 Data Sources

### 3.1.1 Traffic Data and Statistics

This study includes traffic data collected from the Department of Transportation (DOT) of various U.S. States. The Federal Highway Administration (FHWA) in the United States mandates that every state DOT submit annual traffic statistics as part of the National Highway Performance Monitoring System, or HPMS (USDOT 2016). Transportation agencies in every state construct, operate, and maintain permanent traffic monitoring stations with the purpose of collecting traffic count information, among other measures. These stations are referred to as continuous count stations. To meet the federal requirements outlined by the HPMS, the traffic count detectors report hourly traffic counts continuously throughout the year, every year. The ten states identified in this study publish their traffic count data as publicly available through data requests (FL), through data files on public websites (NY), or have permitted a third-party vendor to share HPMS data publicly online (IL, IN, MA, MI, MT, NH, OH, VT). Illinois, Indiana, Michigan, and Ohio are in the Midwest region of the U.S., and the states have 91, 56, 73 and 182 traffic detectors, respectively. Florida is in the southeast region of the U.S. and has 276 traffic detectors. Montana is in the northwestern part of the U.S. and has 91 traffic detectors. New York, New Hampshire, Massachusetts, and Vermont are located in the northeastern region of the U.S., and they have 137, 52, 82, and 32 traffic detectors, respectively. The distribution of the number of urban and rural detectors can be seen in Table 1. The states chosen for analysis in this research were based on availability of traffic volume data, as well as geographic location. A diverse array of geographic regions within the U.S. are represented by the states chosen to provide a variety of political and geographic variables. The diversity of states also displays general traffic differences

based on existing transportation infrastructure that can influence the adherence to governmental mandates and on overall infections within the predicted model. Same-day equivalent traffic volumes were analyzed throughout this study to avoid normal changes in traffic due to the day of the week. The traffic data from 2019 was offset accordingly, depending on the time of the year, to compare same-day traffic volumes, a Monday in 2019 with a Monday in 2020, etc.

***Table 1: U.S. State quantity and type of continuous count stations***

| State | Urban Detectors | Rural Detectors |
|---|---|---|
| New York | 17 | 118 |
| Florida | 120 | 95 |
| Massachusetts | 43 | 7 |
| New Hampshire | 22 | 26 |
| Illinois | 54 | 25 |
| Indiana | 14 | 30 |
| Ohio | 105 | 65 |
| Michigan | 31 | 34 |
| Montana | 14 | 73 |

Due to a majority of U.S. States having HPMS data publicly available across the same web application, the mass amount of data that needed to be compiled for this research study was able to be done in an efficient and automated manner using code script within MATLAB. The cloud-based software MS2: Modern Traffic Analytics publishes the HPMS data per count station per month for each of the state DOTs. The traffic volume data compiled for this study had no distinction of vehicle class and the counts were aggregated for 24-hour periods for the date periods identified. Traffic count data was compiled from January 1, 2019, to May 30, 2021, for a total of 1,072 count stations throughout the ten states identified in this study. More than 940,000 daily traffic volumes were collected for this research topic. The programming and numeric computing platform, MATLAB, was utilized to aide in the accumulation and concatenation of all

the traffic volumes and the corresponding count station and date. A code script was written that parsed the corresponding DOT data website, downloading the appropriate count station for the date period identified. The MATLAB script can be seen in Appendix A. Another script was written that compiled the downloaded CSV files for each count station for each month into a single array from a matrix that could be transposed into a central Excel spreadsheet, also found in Appendix A. This automated process allowed for more data to be included in the study as it became readily available on the MS2 database, allowing for more robust and accurate information to be processed.

### 3.1.2 Traffic Data Irregularities

A common error found in the traffic data was missing data which was reported as a zero value for the volume from the station. The reason for the zero value may be due to road closures because of scheduled maintenance, incidents, or malfunctioning sensors. When missing data occurred in the dataset, the sensor was removed from consideration for that day, but still included in the analysis for other days when data was available. Some days were shown as "NaN" or "Null" values. These were similarly removed for the aforementioned case; however, the sensor was considered in the analysis when the data was available for other days. This resulted in the daily number of observations varying for different stations.

### 3.1.3 COVID-19 Data

The COVID-19 data for this research was compiled from the Johns Hopkins COVID-19 dashboard from January 1, 2020 to May 30, 2021. The early stages of the pandemic were characterized by low initial infections, as uncertainty about the virus was common and the

population generally adhered to stay at home mandates throughout the country. As cases began to spike in early March for most states, emergency mandates were issued throughout all the U.S. States studied. The dates of the declaration of these initial emergency mandates can be seen in Table 2. At the onset of these State of Emergencies, or shortly thereafter, public schools and restaurants were closed on a grand scale throughout the U.S. The volume of traffic and number of infection cases during this initial lockdown period is described as the first peak throughout this study. The modeling portion of this study required the number of cases for the first and second infection peaks, the first peak for all states occurred in the early part of 2020. The second peak of infections occurred after states began to lift restrictions and mandates; this second peak occurred in the winter of 2020 and most instances lasted into the beginning of 2021, only Florida had the second peak of infections end before 2021.

***Table 2: Governmental directives as a response to the COVID-19 pandemic in the U.S (Kristiansson, 2021).***

| State | State of Emergency | Schools Closed | Restaurants Closed | Phase 1 | Phase 2 | Phase 3 | Phase 4 | Phase 5 |
|---|---|---|---|---|---|---|---|---|
| New York | 3/7 | 3/16 | 3/16 | 6/8 | 6/19 | 7/6 | 7/20 | N/A |
| Florida | 3/9 | 3/13 | 3/20 | 5/18 | 6/5 | 9/25 | N/A | N/A |
| Massachusetts | 3/10 | 3/17 | 3/17 | 5/18 | 6/8 | 7/6 | N/A | N/A |
| New Hampshire | 3/13 | 3/16 | 3/16 | 5/18 | 6/15 | 6/29 | 8/24 | N/A |
| Illinois | 3/9 | 3/13 | 3/17 | 5/1 | 5/29 | 6/26 | N/A | N/A |
| Indiana | 3/6 | 3/19 | 3/16 | 5/4 | 5/22 | 6/11 | 9/26 | N/A |
| Vermont | 3/13 | 3/18 | 3/17 | 4/17 | 5/22 | 6/8 | 7/10 | 9/18 |
| Ohio | 3/9 | 3/16 | 3/16 | 5/12 | 6/22 | 8/25 | N/A | N/A |
| Michigan | 3/10 | 3/16 | 3/16 | 5/7 | 6/1 | 6/10 | N/A | N/A |
| Montana | 3/12 | 3/16 | 3/21 | 4/27 | 6/1 | N/A | N/A | N/A |

### 3.2 Data Analysis

Observed COVID-19 infection data was plotted for each of the ten states to determine the time period that the infection rates peaked throughout 2020 and 2021. After the initial mandatory lockdowns, when COVID-19 cases first started developing in the U.S., there was a peak in infection numbers in early April of 2020 that lasted until late June, and in some instances, early July. This peak is categorized as the "first" peak of cases. When lockdown restrictions began to loosen in the following months, and phased reopening occurred, infection numbers peaked again in late 2020, or early 2021, in some instances; this peak is categorized as the "second" peak and involved more overall cases than the first peak in all the states studied.

Governmental directives changed occasionally throughout the duration of the pandemic, most drastically in the early stages of the virus's development in early 2020. The response of state governments was similar throughout the country, aligning with national mandates, during the onset of the rise in COVID-19 cases in the U.S. in early March of 2020. It was during this time when states of emergency were declared for all the U.S. States. The flux of infections during this time period throughout the country is determined as the first peak in this study, both in terms of the number of infections, as well as the behavior of the population in response to governmental directives and groupthink. The main difference between the behavior of different states was the way in which the phased reopening of the economy was initiated in the months following the first peak of cases. Based on executive orders and official governmental websites, many states provided a written approach on how and when which parts of the economy were reopened and to what extent (Kristiansson 2021). The second peak referred to in this study is the corresponding number of infections and loosened governmental restrictions that occurred during this phased reopening period that all states experienced. The key dates from these phased

reopening announcements were collected to provide a roadmap of how governmental directives changed as a response to the pandemic and are summarized for each state in Table 2. The description and goal of each reopening phase varied from state to state and were published by each state's government as a public health initiative.

### 3.3 Model Formulation

SIR models divide the host population into susceptible, infected, and recovered compartments with a set of differential equations describing dynamics between these different compartments (Allen 2008). The well-studied classical compartmental SIR epidemic model is reformulated to incorporate the effect of the changes in the traffic volume on infectious disease spread. In terms of differential equations, a simple form of the SIR model can be specified as follows:

$$\frac{dS(t)}{dt} = \frac{\beta * S(t) * I(t)}{N} \tag{1.1}$$

$$\frac{dI(t)}{dt} = \frac{\beta * S(t) * I(t)}{N} - \gamma * I(t) \tag{1.2}$$

$$\frac{dR(t)}{dt} = \gamma * I(t) \tag{1.3}$$

Here S, I, and R denote the susceptible, infected, and recovered populations respectively, and the parameters $\beta$ and $\gamma$ represent the transmission rate and recovery rate, respectively. The classical SIR equations, equations 1.1, 1.2, and 1.3, were used to model the observed infection data, which produced an estimation for the transmission rate and recovery rate parameters. These

equations and parameters were then used as the baseline for the infection curves and represented the observed data.

There are several studies that have utilized SIR models for COVID-19 spread in Europe and the United States; these studies can provide the details of the dynamic parameters. Further, the transmission term, $\beta$, is often qualified in models to account for factors, like social separation policies induced by governmental mandates, that reduce the transmission. In order to determine the relationship between traffic counts and COVID-19 infections, the SIR differential equations have been modified to incorporate the effect of traffic, and reducing the transmission rate, $\beta_2$, as follows:

$$I_2(t) = I(t) - \Delta I \tag{1.4}$$

$$\frac{dS(t)}{dt} = \frac{\beta_2 * S(t) * I_2(t)}{N} \tag{1.5}$$

$$\frac{dI_2(t)}{dt} = \frac{\beta_2 * S(t) * I_2(t)}{N} - \gamma * I_2(t) \tag{1.6}$$

$$\frac{dR(t)}{dt} = \gamma * I_2(t) \tag{1.7}$$

$\beta$, or the infection rate, is multiplied by a traffic value to achieve an SIR curve that produces a total number of infections for the given time period, calculated through integration of the infection curve, which is lower than the observed infection curve. The differential of infections from the observed data to the predicted data is the variable $\Delta I$, as seen in equations 1.4 and 1.11. The $\Delta I$ value is calculated by determining how many vehicles would have been absent from the roadway in the second peak, given the percent change in volumes from the first peak, $\Delta T$, and multiplied by the regression value, as shown in equation 1.11:

$$T_2 = T_b * (1 - p_2) \tag{1.8}$$

$$T_1 = T_b * (1 - p_1) \tag{1.9}$$

$$\Delta T = T_2 - T_1 \tag{1.10}$$

$$\Delta I = \Delta T * n \tag{1.11}$$

where $T_2$ is the observed traffic count during the second peak and $T_1$ is the predicted traffic count for the second peak based on first peak conditions. $p_2$ is the percent decrease in traffic in the second peak and $p_1$ is the percent decrease in traffic of the first peak. The total reduction in traffic volumes is denoted by $\Delta T$ in equation 1.10. The value resulting from the integration of the observed infection curve was reduced by this $\Delta I$ value to result in the reduced infection curve. The resulting parameters, notably the new $\beta$ rate, found in equation 1.6, were determined based on the reduction in overall cases. The predicted infection curves were plotted using equations 1.5, 1.6 and 1.7.

Once the general timeframe of both the first and second peaks were determined for the U.S. States, an infection curve from the classical SIR model could be fit to the observed COVID-19 case data gathered from the Johns Hopkins database for each of the ten U.S. States. The Nelder-Mead Estimation method was used to fit a predicted infection curve to the data for each of the two peaks, using the least squares estimation technique. The date range for each of the state's second peak COVID-19 infections were determined based on the mean absolute error of the predicted model analysis. The extent of the second peak for each U.S. state was determined, then the Nelder-Mead Estimation model was run varying the initial and end dates alternatively by ten days in each direction, determining the mean absolute error for each iteration. The date range that encapsulated observed second peak COVID-19 infections and produced the minimum mean

absolute error were used for the predicted model. The result estimated by the model based on the observed data produced a beta and gamma value, which are the parameters within the SIR model equations; these parameters, along with the number of susceptible individuals in each state, could then be used to model infection curves independent from the observed data. The estimation of contact and recovery rate parameters that make up the infection curve within an SIR model included the use of Fisher information matrices and negative log likelihood functions to achieve a curve that fit the observed COVID-19 data for the U.S. States studied. The process was run with a Python script in a Google Colaboratory notebook, as seen in Appendix B, and various iterations of the model were tested to create the predicted infection curves for the second peaks for all ten U.S. States. The original code that utilizes Nelder-Mead Estimation to predict parameters for COVID-19 was collected from GitHub, called param-estimation-SIR by Marisa Eisenburg (2017). The code needed an overhaul of changes that could modify the existing SIR equations to reflect equations 1.4-1.7. As all ten U.S. States needed to be modeled, a number of variables were created that reflected each state, including start and end dates and population numbers. A separate script was written in Google Colab that contained equations 1.1-1.3 and 1.5-1.7 and utilized the parameters obtained from this initial calculation, as found in Table 4; this script also provided the final observed and predicted infection curves for all of the ten states, and the script can be seen in Appendix B.

One of the objectives of this study is to determine the relationship between traffic volumes and infection rates within a population. The predicted infection curves that are the end output of the methodology, are shown to reflect the number of infections during the second peak that result from the percent change of volumes being the same as the percent change of traffic volumes in the first peak. This methodology represents a population in the second peak that has similar

characteristics of the population during the initial outbreak of COVID-19 in the U.S. A linear regression analysis was completed for each state during the first peak to determine the relationship between changes in traffic volumes and ensuing COVID-19 cases, which were analyzed during a five-day lag after the traffic volume counts. The five-day offset of daily reported cases is based on a correlation analysis between the reported cases and decrease in traffic from 2019 to 2020 (Kristiansson 2021). The number of days offset varied from zero to 28, and the final offset was determined by the value of the correlation analysis which was closest to -1. Out of the ten U.S. States analyzed, the average days offset is five days, which is the value applied to the linear regression analysis in this study (Kristiansson 2021). To calculate the change in infections due to this reduction in traffic, the value of the slope from the linear regression analysis of the change in traffic volumes to the number of COVID-19 cases, as calculated as a seven-day average with a five-day lag, was multiplied by the difference in actual and calculated 2020 traffic volumes. The results of the regression analysis can be seen in Table 3, the slope of the linear equation between changes in traffic and COVID-19 cases is described by the variable *n*. This relationship is applied to traffic volumes in the second peak to determine the number of infections that may have been reduced in the scenario where traffic volume change, as a percentage, was the same as that of the first peak, $\Delta I$ from equation 1.11. The contact rate, $\beta_2$, in the modified SIR equations represents the transmissivity that reflects the predicted infection curve, as found in equation 1.6.

***Table 3: Regression analysis results for changes in traffic and COVID-19 infections.***

| State | N (millions) | n | $R^2$ |
|---|---|---|---|
| New York | 19 | .00073 | .079 |
| Florida | 21 | .00016 | .258 |
| Massachusetts | 7 | .00125 | .476 |
| New Hampshire | 1.5 | .00004 | .046 |
| Illinois | 12.5 | .00078 | .059 |
| Indiana | 7 | .00036 | .282 |
| Ohio | 12 | .00004 | .012 |
| Vermont | .65 | .00038 | .428 |
| Michigan | 10 | .00067 | .306 |
| Montana | 1 | .00007 | .226 |

The parameters calculated from the predicted SIR model based on the observed infection data and the number of susceptible individuals for each state, taken from population counts provided by U.S. Census data, were used to predict infections within each state. The parameter for number of susceptible persons in each simulation can be found in Table 3 as *N*. The *n* factor, or the slope calculated from the linear regression analysis of the first peak, represents the relationship between traffic volumes and infection rates per state while adherence to governmental mandates was at the highest. Population density data was compiled from Statista (2020) for each of the U.S. states and can be found in Table 5. Population density data is used to identify the relationship between population density and the *n* factor. States with a higher population density might tend to have a higher contact rate amongst people within the population, leading to changes in traffic having a greater impact on the reduction of overall cases and more of a flattening of the infection curve. To determine the effect that the mandatory lockdown had on traffic volumes, and ultimately COVID-19 infections, the beta variable in the

SIR model during the second peak was altered to represent an infection curve that could potentially result from lower traffic volumes. The total percent change in traffic volumes in the first peak of COVID-19 infection, the peak during March and April of 2020, is the baseline value used for each U.S. state, as seen as % traffic reduction first peak in Table 4. As traffic volumes were higher during the second peak due to fewer stay-at-home restrictions, the percent change in traffic volumes from 2019 to 2020 was lower, compared to the first peak, as seen as % traffic reduction second peak in Table 4. The difference between the calculated, predicted volume and the actual 2020 traffic volume is determined as the number of vehicles, seen as $\Delta T$ in equation 1.10, that would be absent from the road if the quarantine restrictions from the beginning of the pandemic were replicated during the dates of the second peak. The reduction in traffic counts are then used to determine how many infections may have been reduced throughout the second peak, seen as $\Delta I$ in equation 1.11. Another objective was to determine the potential hospitalizations and deaths that may have been prevented based on the reduction of infections. The overall flattening of the curve and reduction of overall cases within a population relies on the contact rate, as well as the number of susceptible persons within the population, due to the contact rate in the SIR equations being directly affected by the number of people within the population, *N*. The total number of reduced hospitalizations and deaths from the observed data to the predicted data is based on the corresponding rates for each state for the date period in which the second peak occurred:

$$\Delta H = \frac{\#\ Adult\ patients\ confirmed\ and\ suspected\ with\ COVID{-}19}{\#\ Infections} * \Delta I \tag{1.12}$$

$$\Delta D = \frac{\#\ Deaths\ caused\ by\ COVID{-}19}{\#\ Infections} * \Delta I \tag{1.13}$$

where the hospitalization rate is the total number of adult patients confirmed and suspected of having COVID-19 compared to the total number of infections for a certain state during the

second peak (U.S. Department of Health and Human Services, 2023). The mortality rate is calculated as the number of deaths compared to the total number of infections during the second peak; the number of predicted infections is multiplied by both of these rates to determine the number of hospitalizations, $\Delta H$, and the number of deaths, $\Delta D$ that may have been prevented throughout the second peak of infections.

# 4.0 RESULTS

The findings from the research completed during this study include the final SIR model infection curves that represent the observed infection cases and the adjusted infection cases, which are predicted based on traffic patterns throughout the pandemic for each U.S. state. The resulting reduction in infections, hospitalizations, and deaths that could be a result of these modified infection curves are analyzed on an individual state basis.

The 2020 daily traffic compared to same-day 2019 daily traffic, along with the daily recorded infection numbers, in the ten states in the U.S. are shown in Figure 2 through Figure 11. Figure 2 shows the southeast state (Florida), Figure 3 through Figure 6 show the northeast states (New York, Massachusetts, New Hampshire, and Vermont, respectively). Figure 7 through Figure 10 show the midwestern states (Illinois, Indiana, Ohio, and Michigan, respectively). Figure 11 shows the northwest state (Montana). Each of the U.S. states identified in this study were analyzed separately due to regional changes in traffic volumes and infection rates; hospitalization and death rates also vary depending on the geographic region and the population characteristics of each. Population density, for example, may have an effect on the contact and transmission rate of viruses within a population. A higher population density may result in a higher transmission rate as individuals are closer physically and the chances of interaction are higher than an area with a lower population density.

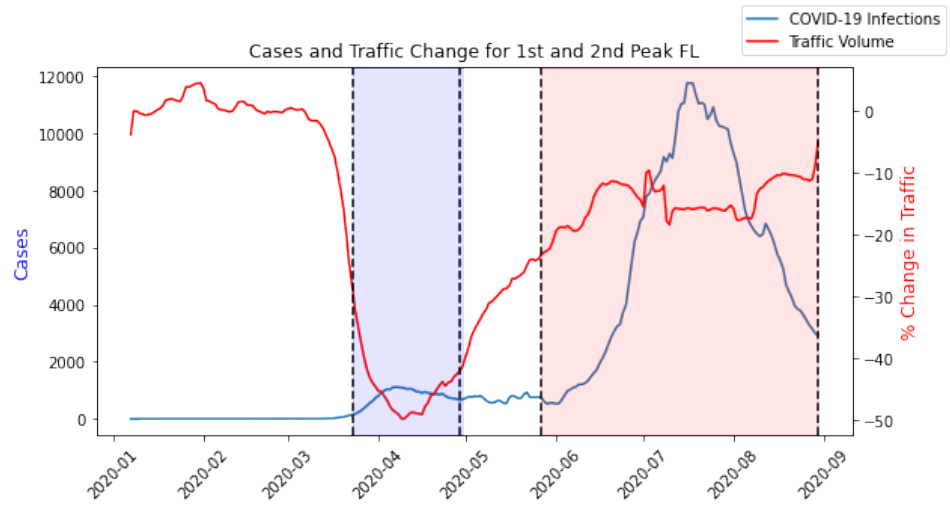


***Figure 2: Seven-day rolling average COVID-19 cases and difference between same-day traffic volume in study period and 2019, including first and second peak, in Florida.***

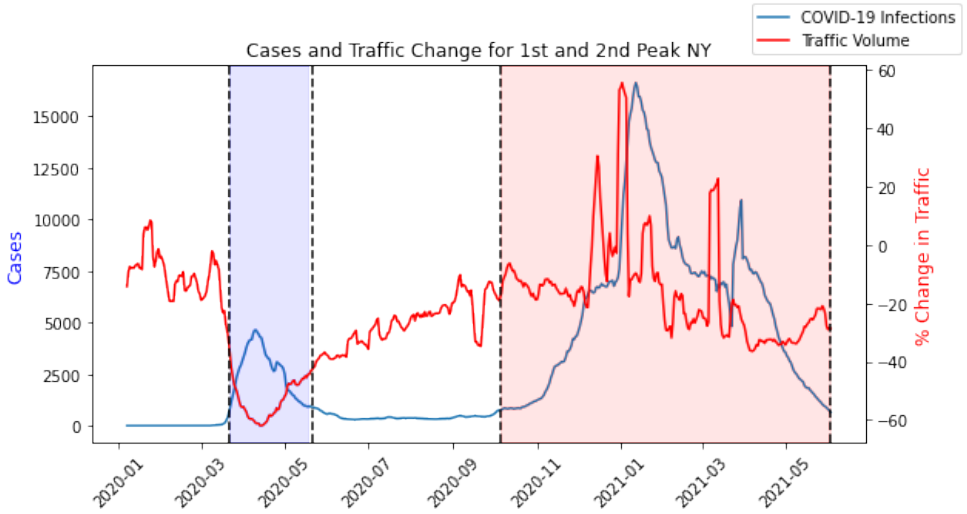


***Figure 3: Seven-day rolling average COVID-19 cases and difference between same-day traffic volume in study period and 2019, including first and second peak, in New York.***

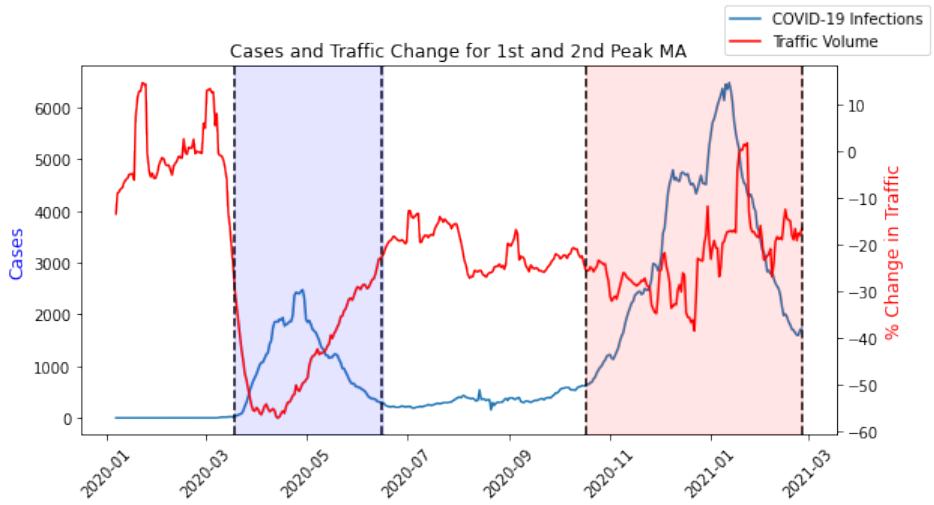


***Figure 4: Seven-day rolling average COVID-19 cases and difference between same-day traffic volume in study period and 2019, including first and second peak, in Massachusetts.***

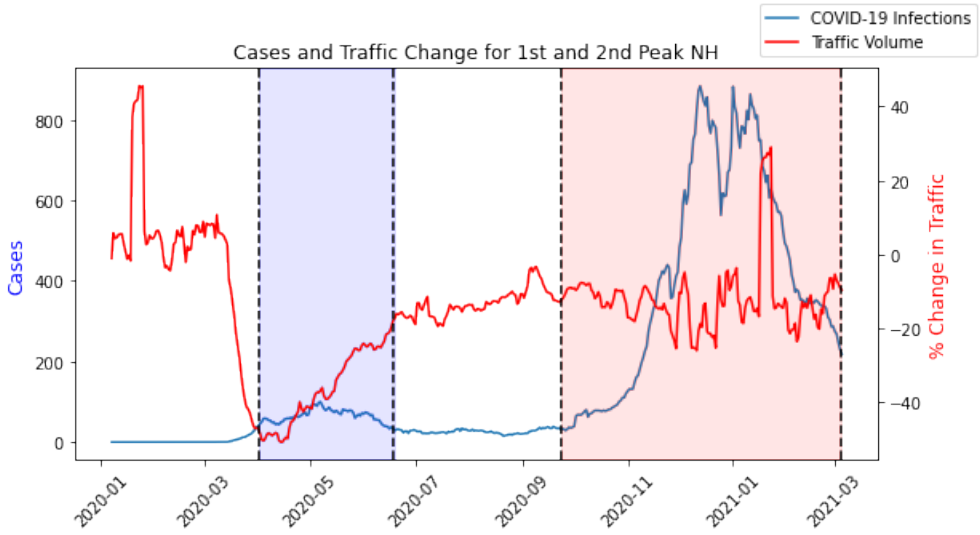


***Figure 5: Seven-day rolling average COVID-19 cases and difference between same-day traffic volume in study period and 2019, including first and second peak, in New Hampshire.***

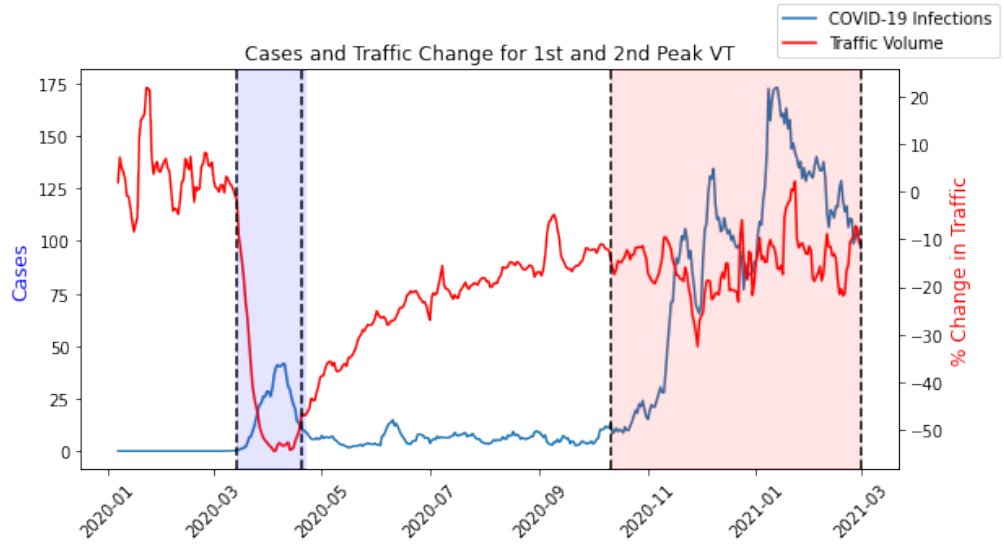


***Figure 6: Seven-day rolling average COVID-19 cases and difference between same-day traffic volume in study period and 2019, including first and second peak, in Vermont.***

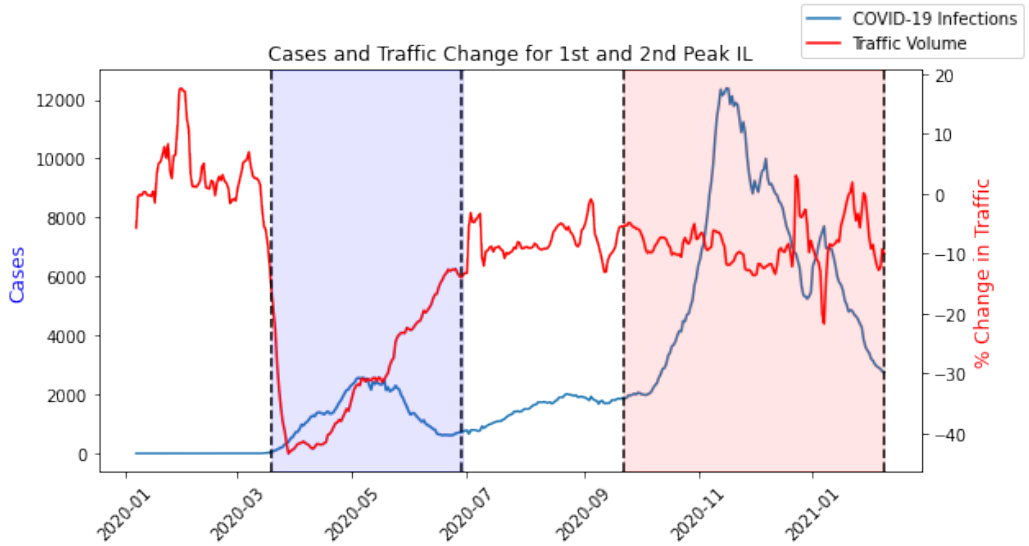


***Figure 7: Seven-day rolling average COVID-19 cases and difference between same-day traffic volume in study period and 2019, including first and second peak, in Illinois.***

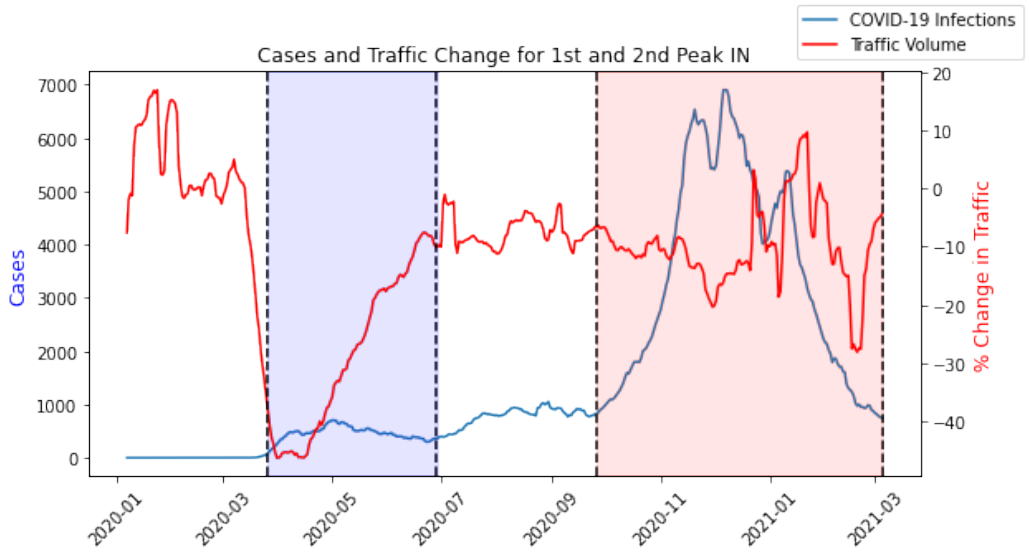


***Figure 8: Seven-day rolling average COVID-19 cases and difference between same-day traffic volume in study period and 2019, including first and second peak, in Indiana.***

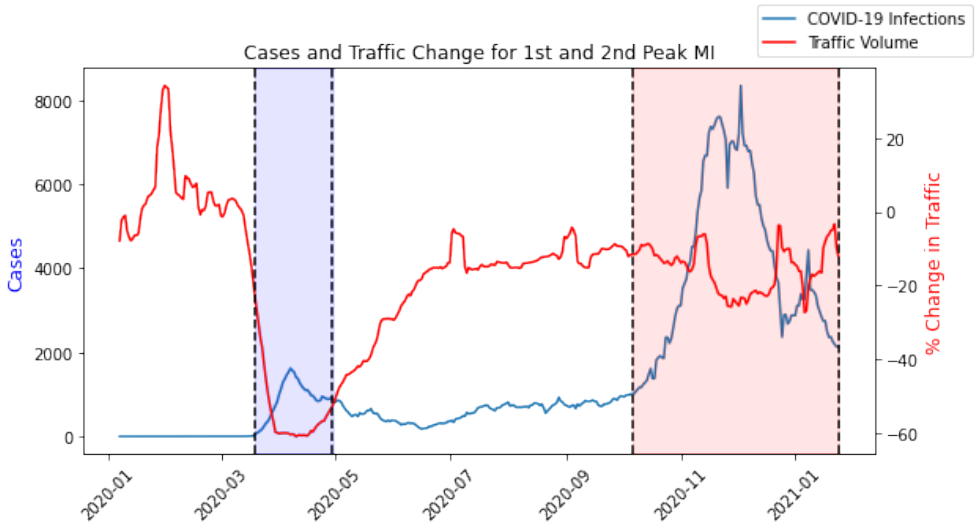


***Figure 9: Seven-day rolling average COVID-19 cases and difference between same-day traffic volume in study period and 2019, including first and second peak, in Michigan.***

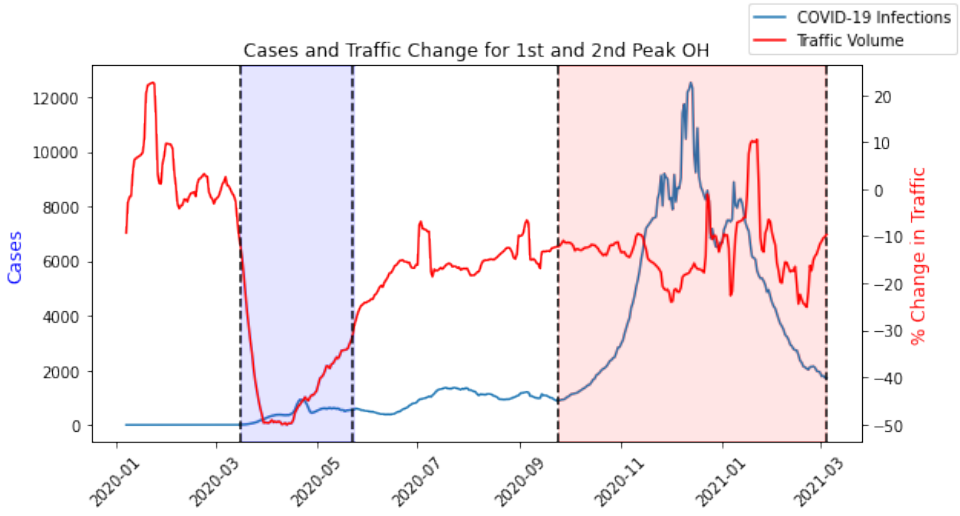


***Figure 10: Seven-day rolling average COVID-19 cases and difference between same-day traffic volume in study period and 2019, including first and second peak, in Ohio.***

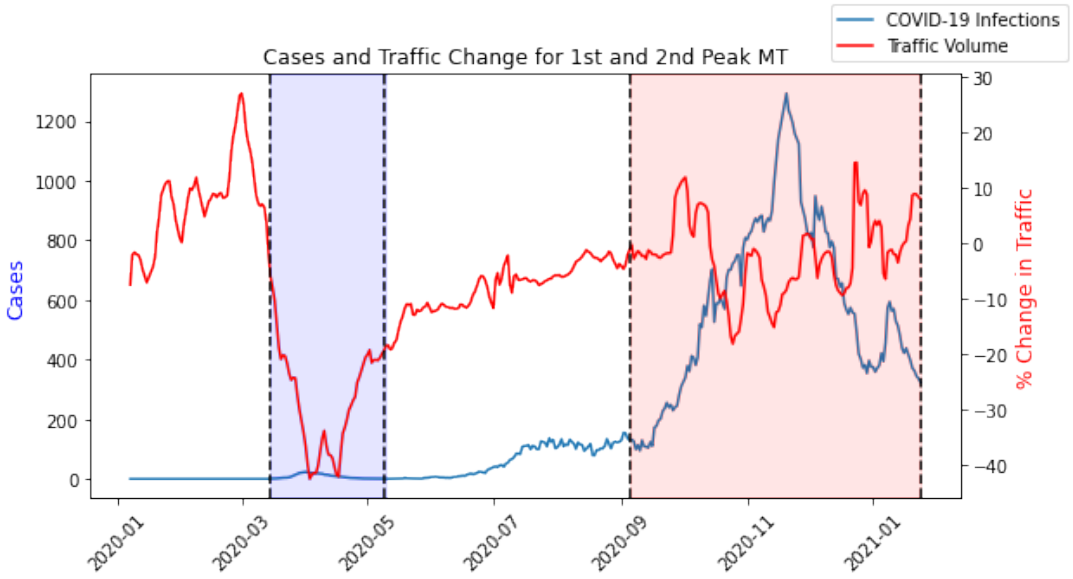


***Figure 11: Seven-day rolling average COVID-19 cases and difference between same-day traffic volume in study period and 2019, including first and second peak, in Montana.***

COVID-19 infection data for all of 2020 and the beginning half of 2021 was plotted as a time series, along with the change in traffic volume for same day equivalents from 2019 to 2020. From the time series plots shown in the figures below, change in traffic volumes during the first

peak of cases (highlighted by the blue region) is much higher than when cases peak for the second time (highlighted by the red region) in late 2020 and early 2021, as traffic volumes creeped back towards pre-pandemic levels. Every state experienced higher traffic volumes during the second peak compared to the first peak. The percent change of traffic from the base year of 2019 to the study year 2020 is shown in Table 4. All states had a negative change in traffic volumes from 2019 to 2020. Michigan had the largest change in traffic volumes (56.1 percent) while Indiana had the lowest change in traffic volumes (28.1 percent). To model infection curves for the data collected, the parameters for the SIR model are needed to make predictions on the data. The dates that correspond to the second peak of cases from the time series plot were used as the range for the Nelder-Mead Estimation for the observed case data.

The parameter estimation curves for New York, Florida, Massachusetts, New Hampshire, Illinois, Indiana, Vermont, Ohio, Michigan, and Montana can be seen in Figures 12 through 21. These are the plots that are produced by the least squares likelihood analysis using Nelder-Mead Estimation, using the population and case data from each state. The black data points are the daily observed COVID-19 infection data, and the blue curve is the output from the predicted model based on the estimation technique. The corresponding beta and gamma parameters from the estimation can be found in the Table 4. These parameters were then used in conjuncture with the population of each respective state in the traditional SIR model equations to achieve an infection curve that represents the initial, or observed, conditions.

*Table 4: Estimated SIR model parameters based on observed data for second peak date period.*

| State | Beta variable | Gamma variable | Reproduction Number, Ro | % Traffic Reduction First Peak | % Traffic Reduction Second Peak | Date Period First Peak | Date Period Second Peak | MAE |
|---|---|---|---|---|---|---|---|---|
| New York | .929 | .897 | 1.036 | -53.2 | -20.1 | 3/21/20 – 5/19/20 | 10/3/20 – 6/3/21 | 1506 |
| Florida | 2.253 | 2.183 | 1.032 | -44.9 | -14.0 | 3/23/20 – 4/29/20 | 5/27/20 - 8/30/20 | 552 |
| Massachusetts | 1.187 | 1.142 | 1.039 | -41.9 | -24.0 | 3/19/20 – 6/16/20 | 10/17/20 - 2/25/21 | 287 |
| New Hampshire | 1.39 | 1.345 | 1.033 | -34 | -14.1 | 4/1/20 – 6/18/20 | 9/23/20 - 3/5/21 | 51 |
| Illinois | 1.032 | .995 | 1.037 | -29.4 | -9.8 | 3/19/20 – 6/28/20 | 9/22/20 - 2/8/21 | 1086 |
| Indiana | 1.045 | 1.003 | 1.042 | -28.1 | -11.0 | 3/26/20 – 6/28/20 | 9/26/20 - 3/5/21 | 366 |
| Ohio | 1.104 | 1.063 | 1.039 | -41 | -14.3 | 3/16/20 – 5/23/20 | 9/24/20 - 3/5/21 | 634 |
| Vermont | 1.794 | 1.757 | 1.021 | -46.7 | -16.5 | 3/14/20 – 4/20/20 | 10/11/20 - 3/1/21 | 16 |
| Michigan | 1.536 | 1.483 | 1.036 | -56.1 | -16.2 | 3/19/20 – 4/29/20 | 10/6/20 - 1/24/21 | 662 |
| Montana | .978 | .938 | 1.043 | -28.5 | -2.9 | 3/15/20 – 5/9/20 | 9/5/20 - 1/24/21 | 83 |

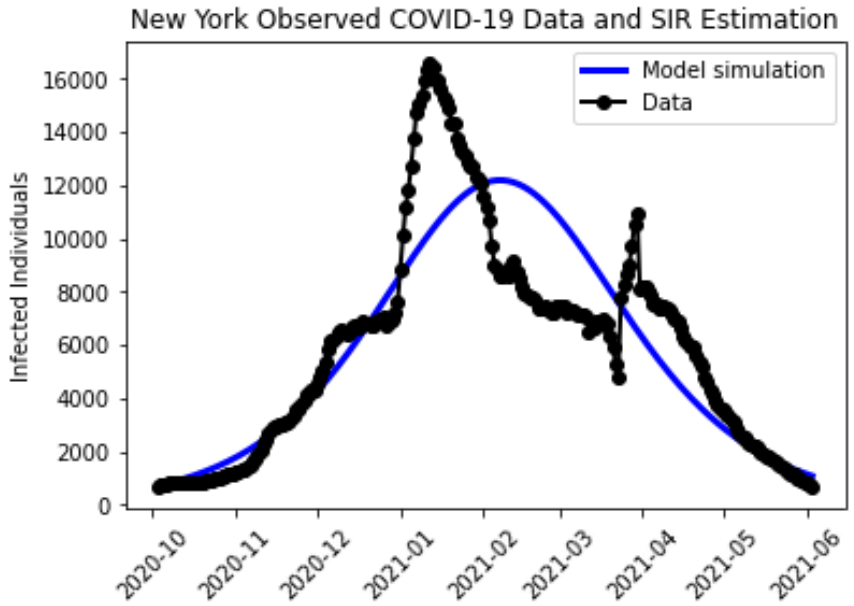


***Figure 12: Observed COVID-19 cases and the fit infected curve based on Nelder-Mead Estimation for New York.***

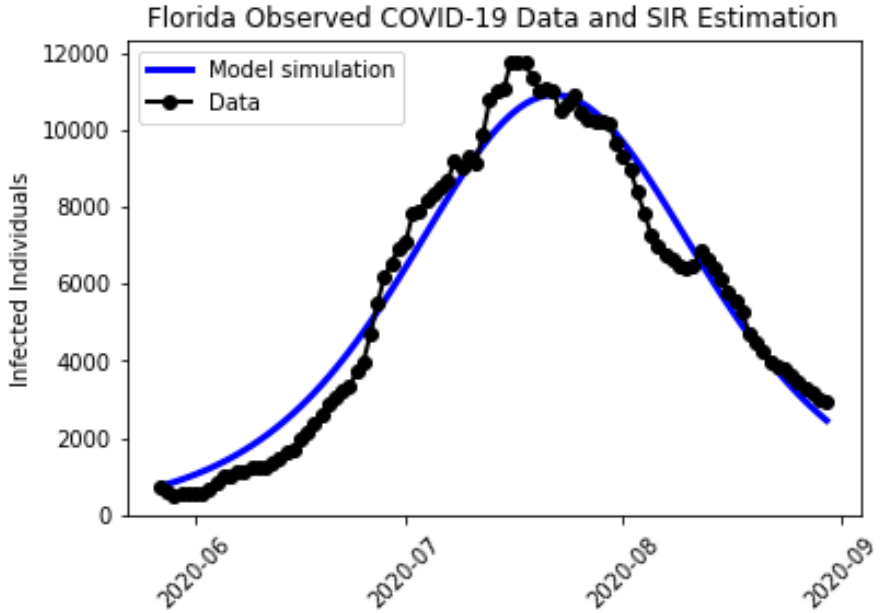


***Figure 13: Observed COVID-19 cases and the fit infected curve based on Nelder-Mead Estimation for Florida.***

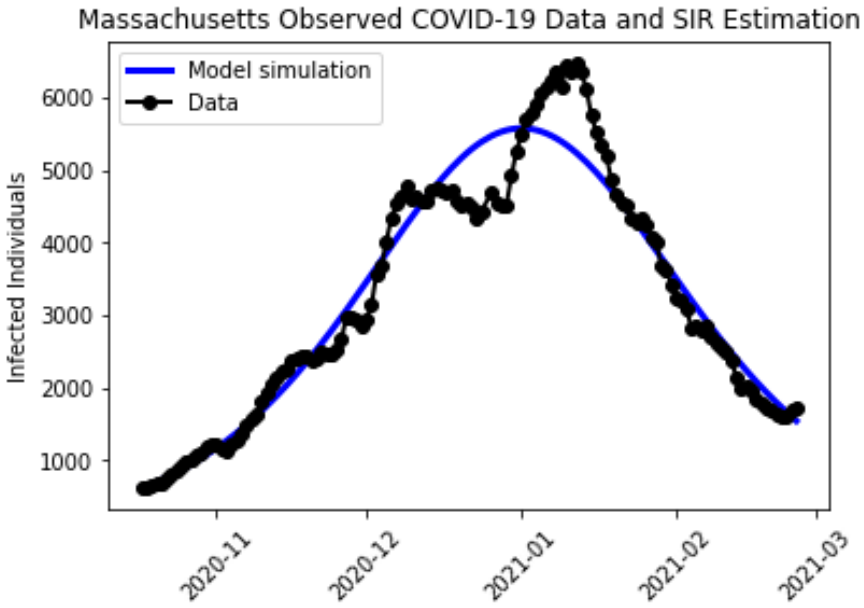


***Figure 14: Observed COVID-19 cases and the fit infected curve based on Nelder-Mead Estimation for Massachusetts.***

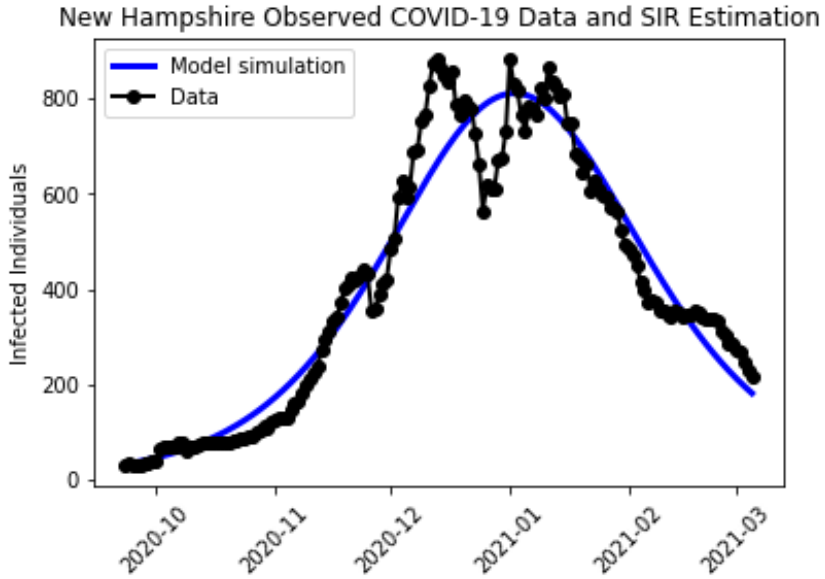


***Figure 15: Observed COVID-19 cases and the fit infected curve based on Nelder-Mead Estimation for New Hampshire.***

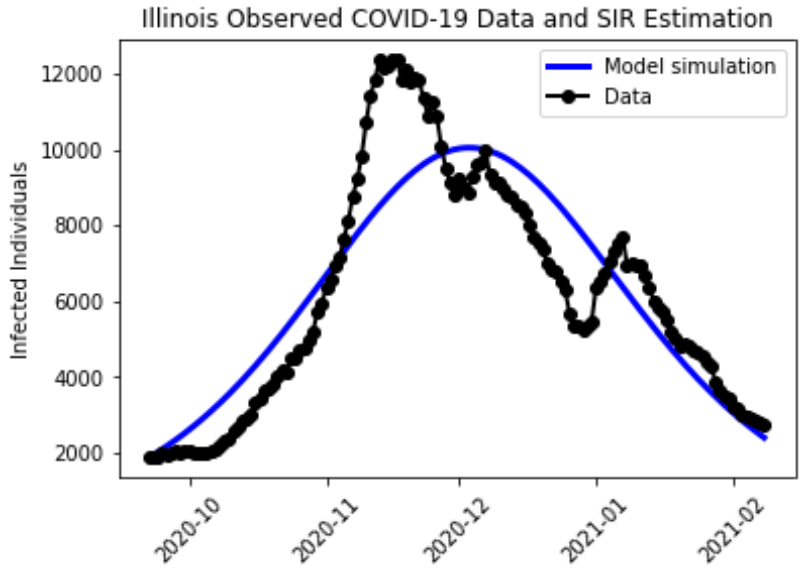


***Figure 16: Observed COVID-19 cases and the fit infected curve based on Nelder-Mead Estimation for Illinois.***

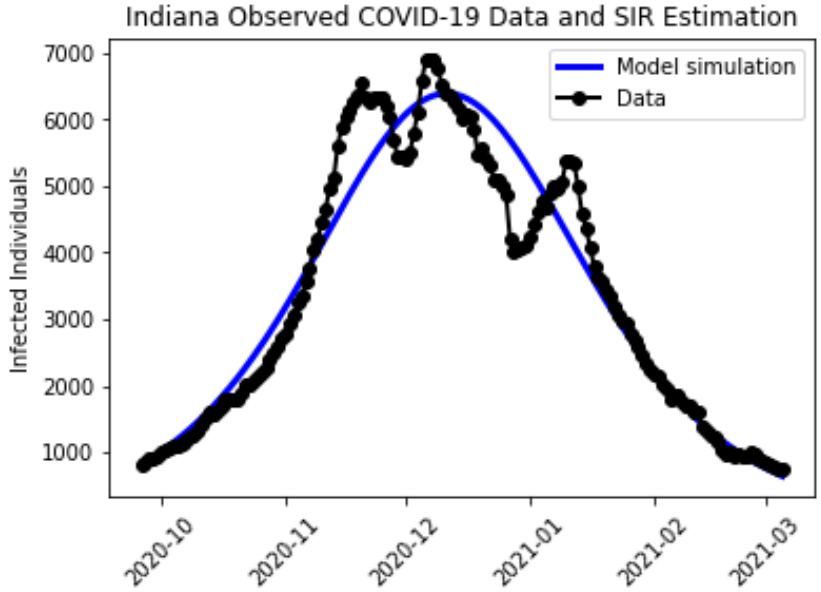


***Figure 17: Observed COVID-19 cases and the fit infected curve based on Nelder-Mead Estimation for Indiana.***

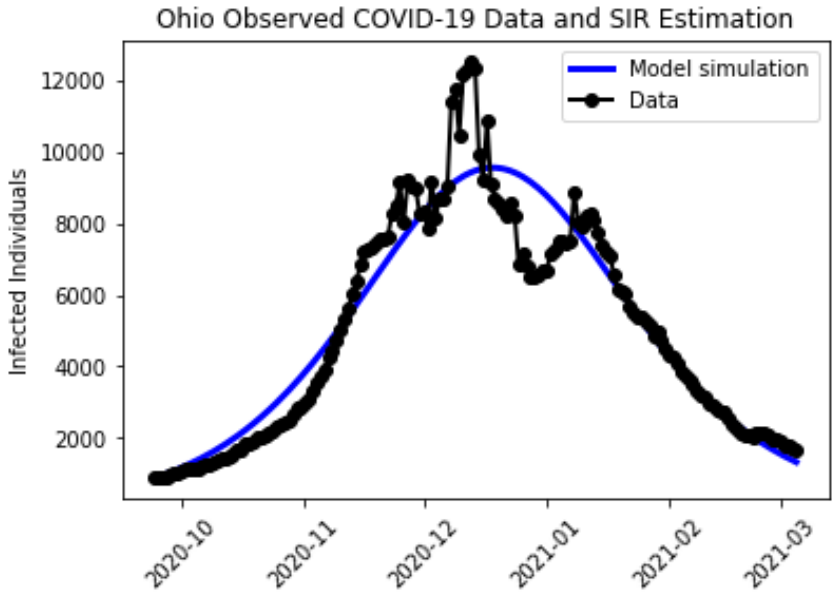


***Figure 18: Observed COVID-19 cases and the fit infected curve based on Nelder-Mead Estimation for Ohio.***

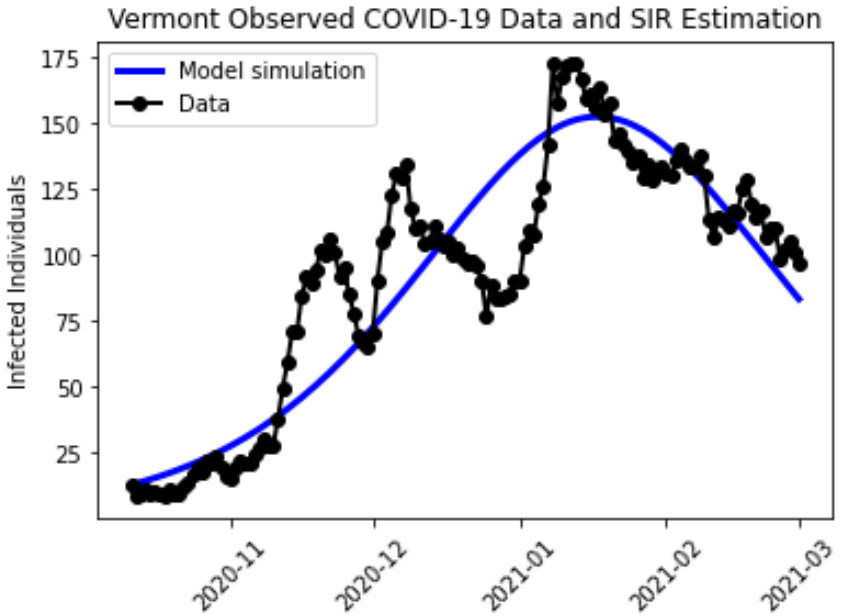


***Figure 19: Observed COVID-19 cases and the fit infected curve based on Nelder-Mead Estimation for Vermont.***

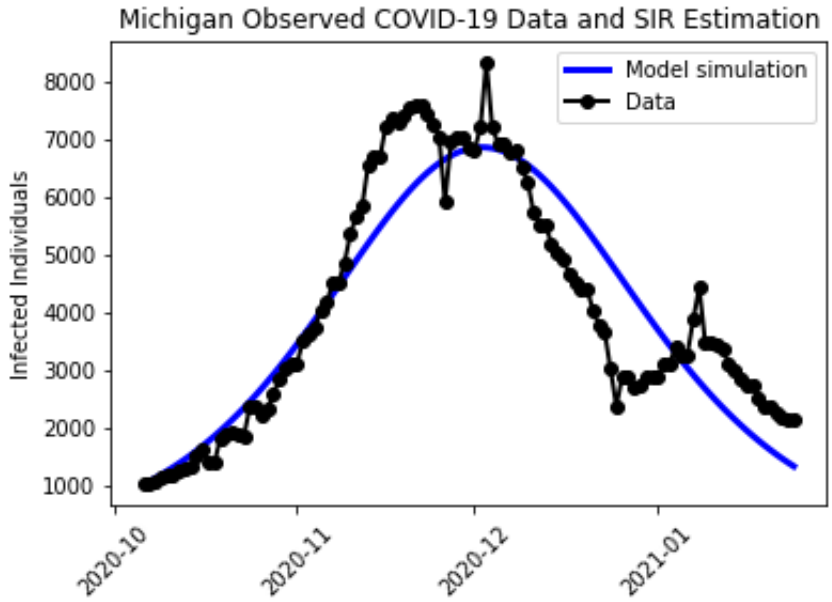


***Figure 20: Observed COVID-19 cases and the fit infected curve based on Nelder-Mead Estimation for Michigan.***

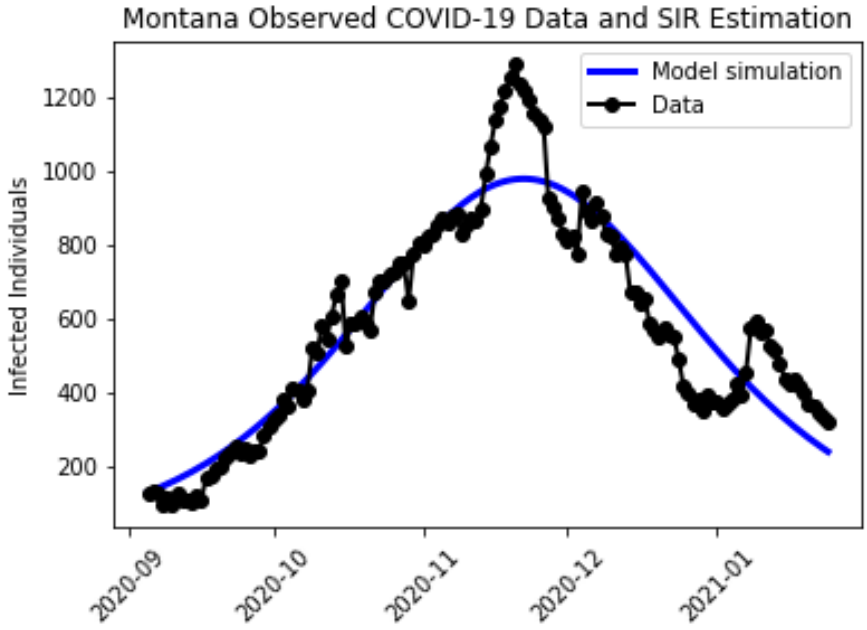


***Figure 21: Observed COVID-19 cases and the fit infected curve based on Nelder-Mead Estimation for Montana.***

The parameters determined through the Nelder-Mead Estimations were then used as the baseline observed values for the second peak of COVID-19 infections that occurred in late 2020 and early into 2021. The red curves in the 2nd Peak Infected Cases Observed and The adjusted plots shown in Figures 22 through 31 represent this data, which applies the model simulation curve parameters to the classic SIR equations. In order to determine the effect consistent government restrictions, including stay-at-home orders, might have had on COVID-19 infections during this second (bigger) peak, conditions and behavior from the initial peak are then applied to the aforementioned peak. Transportation user behavior was much different between the two time periods studied, as evidence by the change in traffic volumes compared to the same-day equivalent in 2019. During the early months, there were fewer users on the roadways, which is a result of state and federal governmental restrictions being more strict at this stage, as well as business incentivizing work-at-home practices. As the months progressed after this first peak, users became more comfortable and restrictions began to lift. This is evidenced by the immediate decrease in change in traffic volumes during this time period. The prediction model found in this study is concerned with the effect that governmental restrictions, that proved to be successful in reducing the number of trips taken throughout households during the initial surge of cases in the U.S., might have on traffic volumes during the second peak, and ultimately the ensuing effect on the second substantial peak of infections. The predicted infections depict a situation where, when conditions in each state predicted a surge of cases, state governments could enact a set of restrictions on household trips taken, reducing the total volume of vehicles within the system,

and flattening the resulting infection curve, which can ultimately reduce the load on health care services throughout the country.

The first step of the analysis of the second peak was to simulate traffic conditions as they were in the first peak. A linear regression analysis was completed for all of the U.S. states studied, determining the relationship between the change in daily traffic volumes from 2019 to 2020/2021 and the number of daily infections, taken from a five-day lag period from the traffic volume data (Kristiansson, 2021). The values from the analysis can be seen in Table 3, denoted by the variable *n.* States with a larger *n* value resulted in a higher reduction in infections during the second peak, meaning that traffic volumes were more of a factor in infection cases. Massachusetts had the highest *n* value, while Ohio and New Hampshire recorded the lowest *n* value for the ten states studied. There are many factors that can contribute to the number of infections during a pandemic within a specific population, this study identified only the effect that traffic volumes may have played.

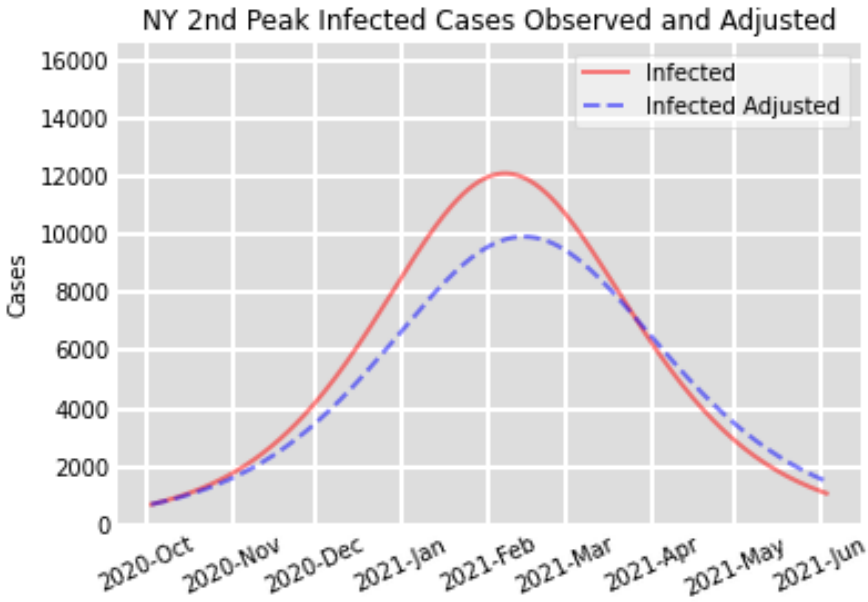


***Figure 22: Second peak observed and adjusted infected curves based on traffic patterns from first peak for New York.***

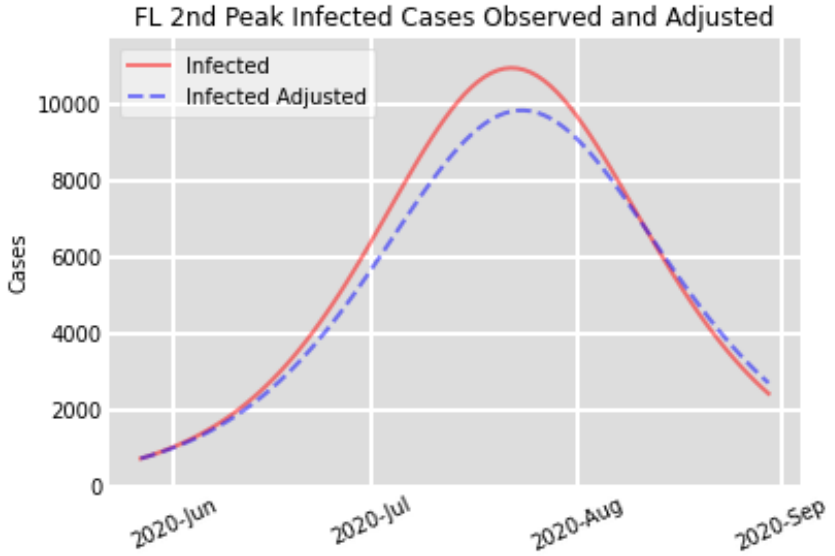


***Figure 23: Second peak observed and adjusted infected curves based on traffic patterns from first peak for Florida.***

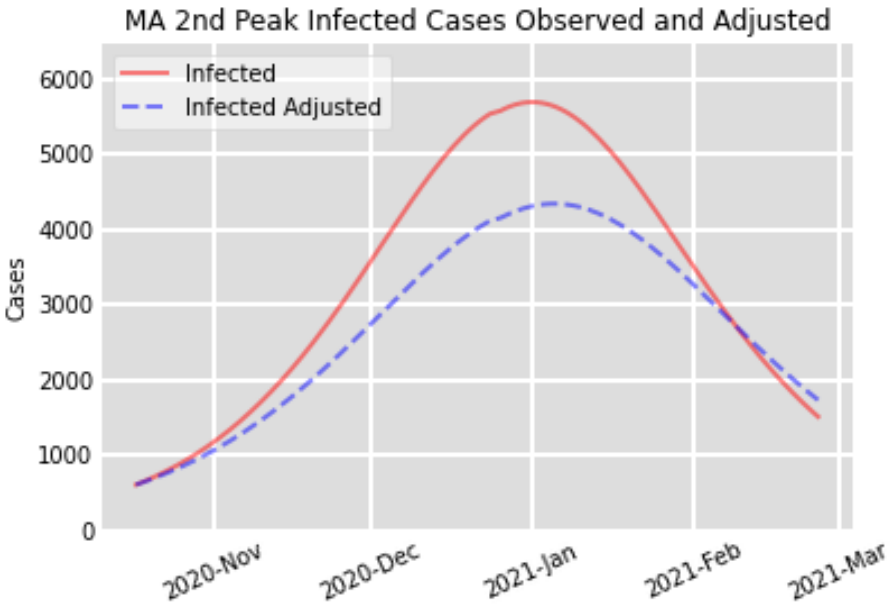


***Figure 24: Second peak observed and adjusted infected curves based on traffic patterns from first peak for Massachusetts.***

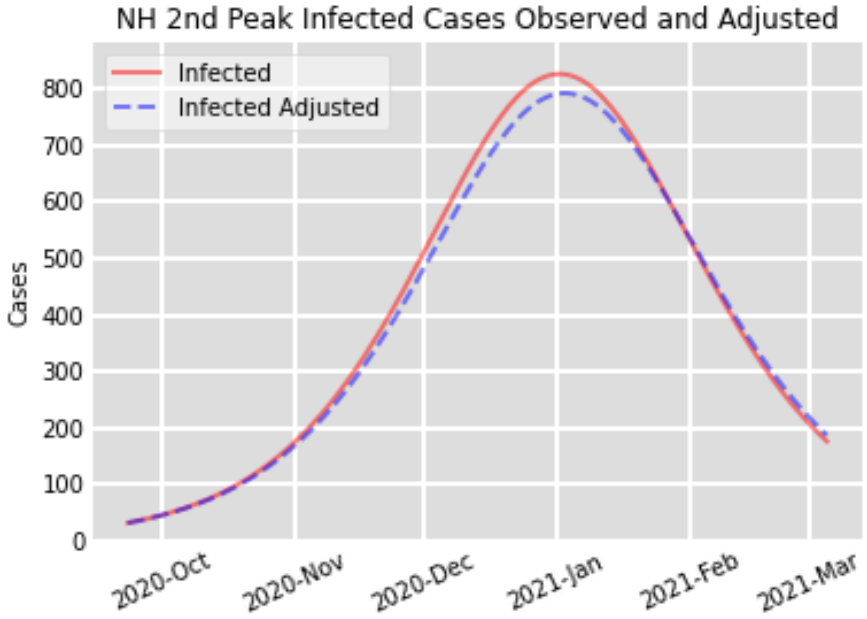


***Figure 25: Second peak observed and adjusted infected curves based on traffic patterns from first peak for New Hampshire.***

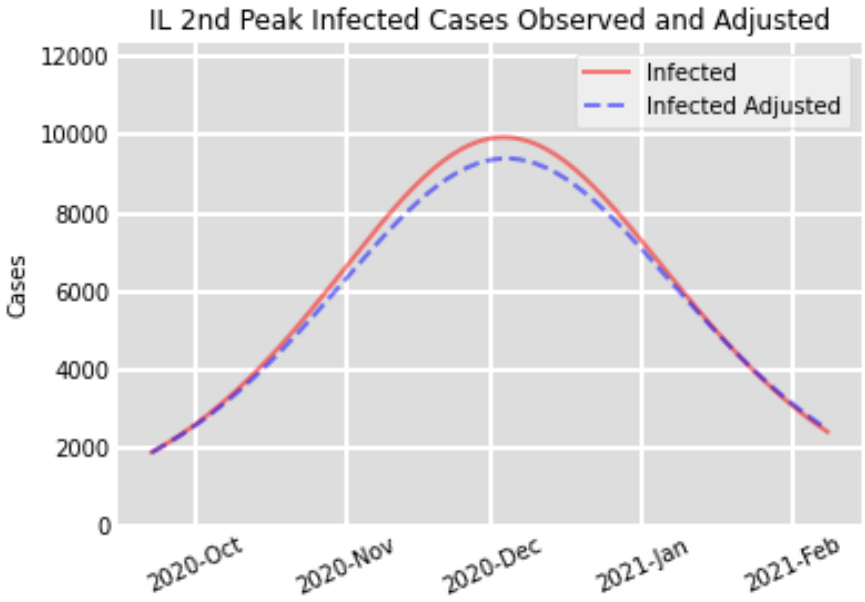


***Figure 26: Second peak observed and adjusted infected curves based on traffic patterns from first peak for Illinois.***

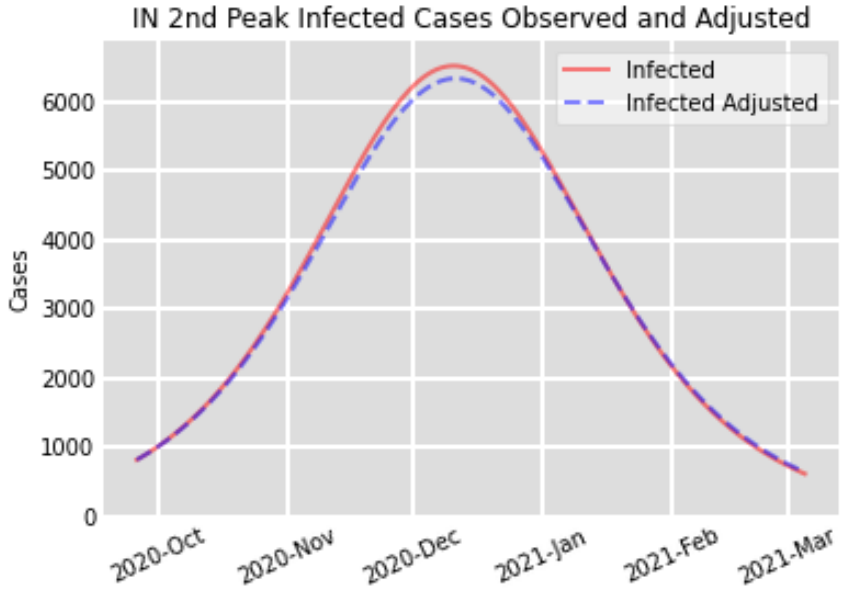


***Figure 27: Second peak observed and adjusted infected curves based on traffic patterns from first peak for Indiana.***

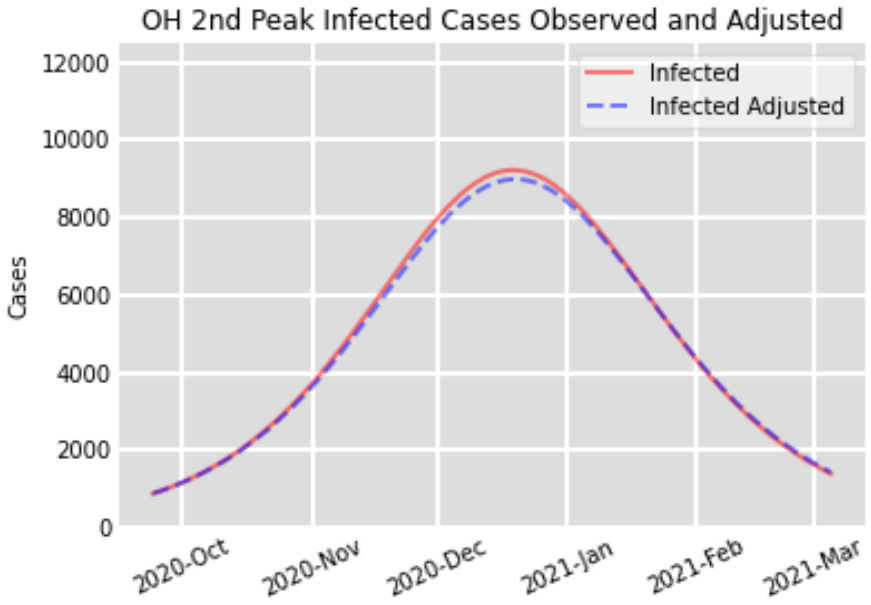


***Figure 28: Second peak observed and adjusted infected curves based on traffic patterns from first peak for Ohio.***

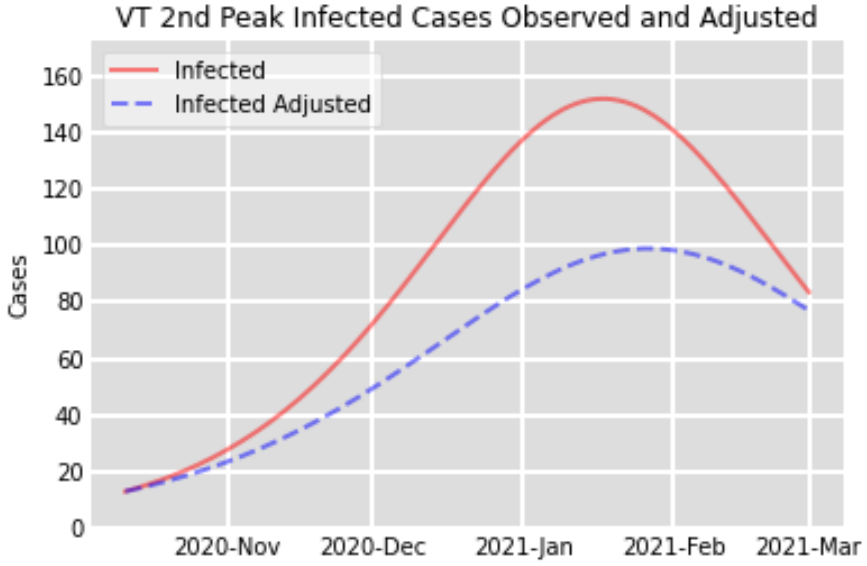


***Figure 29: Second peak observed and adjusted infected curves based on traffic patterns from first peak for Vermont.***

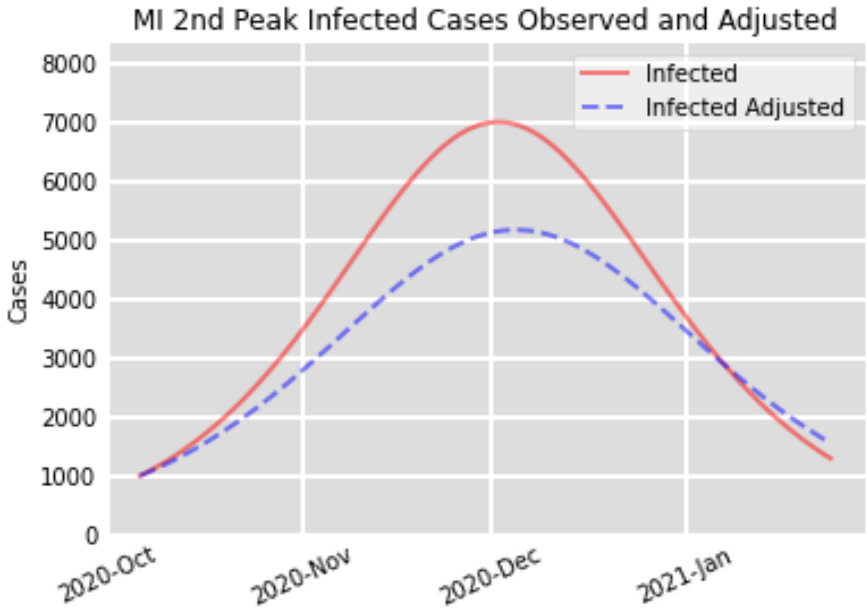


***Figure 30: Second peak observed and adjusted infected curves based on traffic patterns from first peak for Michigan.***

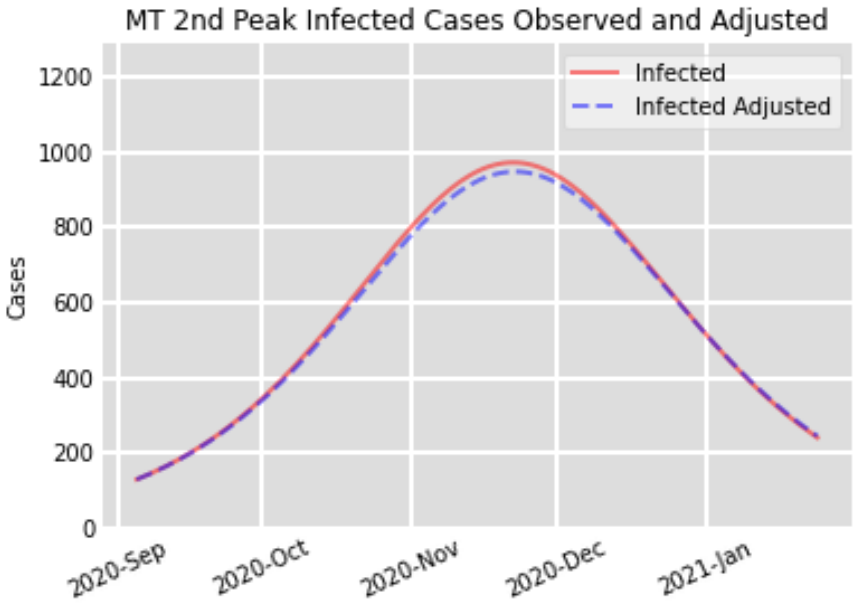


***Figure 31: Second peak observed and adjusted infected curves based on traffic patterns from first peak for Montana.***

The SIR model equation parameters calculated for the observed infection curve, as well as the final adjusted infection curve parameters, can be found in Table 5. The final observed infection curves were plotted along with the reduced infection curves to show the comparison between the two. The total number of COVID-19 infections which may have been reduced, calculated as the difference in the integral of infection curves between the observed (red curve) and predicted equations (blue curve), can be seen in Table 6. Based on the SIR curve models created, the contact rate values to be used in the modified SIR equations were identified for each of the ten U.S. states. This contact rate, which is based on the reduction in traffic during the second peak, $\beta_2$, represents the amount by which traffic conditions within the date period encompassing the first peak affect the reduction of infections in the predicted SIR infected curve models. The results of the calculated contact rate values for each of the states identified can be seen in Table 5.

A single trip taken by an infected individual in a densely populated area would most likely result in more contact of individuals compared to the same single trip by the infected individual in a sparsely populated area. This relationship between population density and the *n* factor depicts why more densely populated states have a higher *n* value. Massachusetts, the state with the highest *n* factor of the states identified, is also the most densely populated state, with a density of 884 people per square mile. Alternatively, Montana has the lowest population density, seven people per square mile, and has the second lowest *n* factor of the states studied. For all ten U.S. states, a total of 7,112 deaths may have been prevented throughout the second peak, with a further 273,062 hospitalizations not being required if the

governmental restrictions and behavior from the first peak lasted throughout 2020 and into early 2021. A total of 407,800 COVID-19 infection cases may have been prevented in these ten U.S. states within the time period that the second peak of infections occurred for each state. Figure 32 illustrates the linear regression plot between the predicted number of deaths reduced from the observed data and population density for each U.S. state.

***Table 5: SIR model parameters for observed and adjusted equations for second peak date periods.***

| State | $\beta$ | $\beta_2$ | $\gamma$ | N (millions) | Pop Density (p/mi^2) | Traffic Reduction |
|---|---|---|---|---|---|---|
| New York | .929 | .9257 | .897 | 19 | 413 | -.201 |
| Florida | 2.253 | 2.249 | 2.183 | 21 | 401 | -.140 |
| Massachusetts | 1.187 | 1.1804 | 1.142 | 7 | 884 | -.240 |
| New Hampshire | 1.39 | 1.389 | 1.345 | 1.5 | 152 | -.141 |
| Illinois | 1.032 | 1.0307 | .995 | 12.5 | 228 | -.098 |
| Indiana | 1.045 | 1.0443 | 1.003 | 7 | 188 | -.110 |
| Ohio | 1.104 | 1.1034 | 1.063 | 12 | 286 | -.143 |
| Vermont | 1.794 | 1.787 | 1.757 | .65 | 68 | -.165 |
| Michigan | 1.536 | 1.527 | 1.483 | 10 | 177 | -.162 |
| Montana | .978 | .9774 | .938 | 1 | 7 | -.029 |

***Table 6: Effect of reduction in travel on health service infrastructure based on predicted infection curves for U.S. states.***

| State | Reduced Infections | % Change in Infections | Hospitalization Rate | Reduced Hospitalizations | Death Rate Infected | Reduced Deaths |
|---|---|---|---|---|---|---|
| New York | 159,000 | 11.1 | 84.9 | 134,991 | 2.07 | 3,292 |
| Florida | 38,000 | 6.86 | 65.2 | 24,776 | 1.60 | 608 |
| Massachusetts | 75,300 | 17.2 | 40.4 | 30,422 | 1.69 | 1,273 |
| New Hampshire | 1,400 | 2.75 | 40.8 | 572 | 1.42 | 20 |
| Illinois | 31,800 | 3.70 | 59.5 | 18,921 | 1.41 | 449 |
| Indiana | 8,000 | 1.66 | 59.8 | 4,784 | 1.60 | 128 |
| Ohio | 11,000 | 1.58 | 62.5 | 6,875 | 1.60 | 176 |
| Vermont | 3,100 | 26.6 | 31.0 | 961 | 1.42 | 43 |
| Michigan | 79,000 | 18.2 | 64.3 | 50,797 | 1.40 | 1,106 |
| Montana | 1,200 | 1.67 | 49.4 | 593 | 1.41 | 17 |
| **Total** | **407,800** | | | **273,062** | | **7,112** |

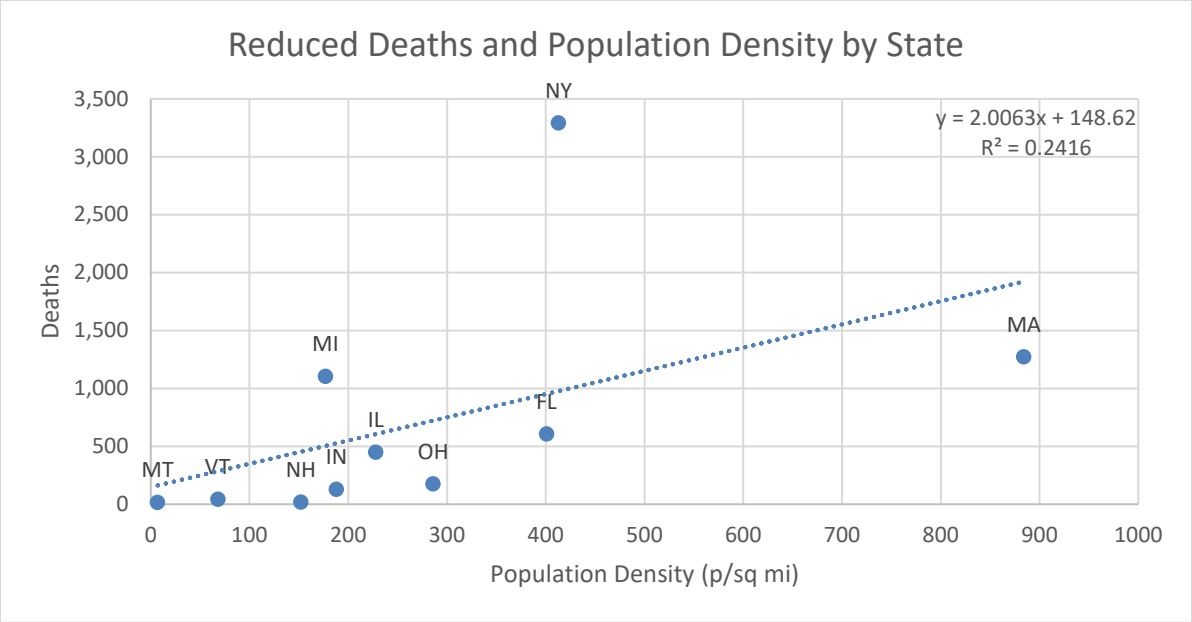


***Figure 32: Linear regression plot between the predicted number of deaths reduced from the observed data and population density for each U.S. state***

The death rate for infected individuals is the highest for New York and Massachusetts, most likely due to the geographic location of these states, being further north and the weather being cold during the second peak of cases could have contributed to higher death rates than states with milder climates. These states also have the highest population densities, relative to the rest of the states identified, which could contribute to higher deaths among those infected. The hospitalization rate of New York was observed as being the highest among the ten states. Vermont is the state that has the highest percent change (26.6 percent) in infections in the predicted model compared to the observed data. This result conveys that implementing stay-at-home mandates over the first 12 months of the pandemic may have been the most effective in Vermont. Even though New York had the largest

number of infections, hospitalizations, and deaths reduced, the largest percentage of the population that observes an effect of these changes is Vermont, with Michigan having the next largest percentage change in infections (18.2 percent).

# 5.0 CONCLUSION

The COVID-19 pandemic that swept across the world in 2020 and throughout 2021 had societal, environmental, and economic impacts that affected every country. The research conducted identified how the epidemic had an effect on the U.S. population, specifically within ten states – Florida, New York, Massachusetts, New Hampshire, Illinois, Indiana, Vermont, Ohio, Michigan, and Montana. The effects of the virus had various impacts on differing populations, dependent on a number of factors. The factors studied in this research topic included the effect that traffic volumes had on COVID-19 infections and the resulting hospitalizations and deaths that may have been prevented. The number of infections during the epidemic, led to a high number of hospitalizations, which induced a strain on health care facilities throughout the U.S. The hospitalization and death rates for the specific states during the second peak were determined and utilized to calculate the number of hospitalizations and deaths that could have been avoided if the initial lockdown environment was prolonged throughout the dates in which the second peak of infections occurred. Two separate peaks were analyzed throughout the study, the initial "first" peak, which generally occurred from March to June 2020, and the "second" peak, which took place in the winter of 2020 and early 2021. Traffic conditions during the first peak, that were a result of governmental mandates and stay-at-home orders, were simulated for the dates of the second peak and the resulting reduction of COVID-19 infections was calculated for each of the U.S. states. The extent to which the observed infection curves were flattened by reducing the number of vehicular

traffic can be utilized by various national and state governments to control the spread of infections during future epidemics.

In general, the results of the research showed that a reduction in traffic during the second peak of infections flattened the resulting infection curve for the same time period. The total number of infections that may have been reduced by maintaining governmental procedures from the initial outbreak of COVID-19 cases in the ten U.S. states was 407,800 cases. As a result of these 407,800 cases, 273,062 individuals may have been removed from hospitals, reducing the overall load on health care facilities. There also may have been 7,112 fewer deaths in the second peak alone, given the traffic conditions of the first peak of infections. A specific example of the general results of the research is illustrated by the state of Massachusetts. To achieve an equivalent percent change in traffic from 2019 to 2020, a reduction of 24 percent in traffic in the second peak volumes was calculated. The resulting number of vehicles reduced from the second peak time period is used to calculate the number of infections that are reduced from the second peak, 75,300 in the case of Massachusetts. From this value, the corresponding number of hospitalizations (30,422) and number of deaths (1,273) that are potentially reduced for the second peak time period are calculated by using the hospitalization and mortality rates, respectively. This method of reducing traffic volumes is effective at flattening the curve of infections within a population, as evidenced by the SIR infection curves for the predicted versus the observed models. Each of the predicted infection curves results in a lower peak and total number of infections than the original infection curves, which are based on the observed COVID-19 data.

All of the states identified in the study had predicted infection curves that were below the observed infection curves, meaning that there is a direct relationship between the number of trips taken and the infection rate. This trend is visualized in Figure 33, as the percent reduction in traffic between 2019 and 2020 increases, the higher the percent of infections reduced in the predicted model compared to the observed model. This relationship indicates that traffic volumes may serve as a leading indicator of infections, and ultimately hospitalizations and deaths, within a population. The state with the highest reduction in traffic during the first peak period, Michigan, resulted in having one of the highest percent reductions in infections, an 18.2 percent reduction of infections in the predicted model compared to the observed infection data. As more people leave their households and interact with others in public spaces, the exposure rate increases, leading to more cases. On the contrary, fewer total vehicle volume counts might result in fewer overall infections for a state or country. The rate at which this relationship develops is due to a multitude of external factors as the dynamic of a pandemic within a society is complex and interconnected. The results from the predicted model created depict a correlation between traffic volumes and infections in which traffic volumes can be an indicator of the amount of interaction between individuals and help to predict future infection counts.

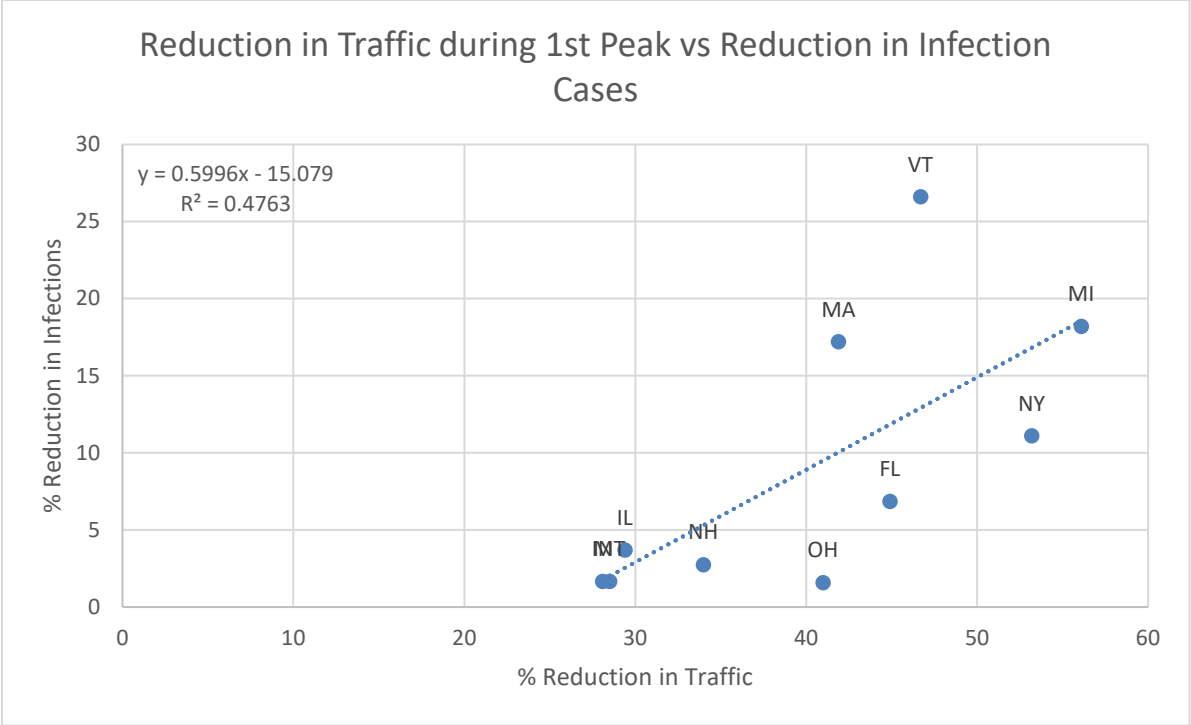


***Figure 33: Percent change in traffic during first peak analyzed against the number of deaths reduced from observed and predicted infection models for each state.***

The overall reduction of peak infections in the second peak, quantified by the percent change in infections, as seen in Table 6, are impacted by a countless number of factors. The relationship between the number of cases and traffic volumes may be determined as being linear, however, it is far from being a direct and independent relationship. Epidemics within a population are a complex and dynamic interaction that relies on a number of factors. Some of these factors that directly relate to traffic could include the amount of public transit utilized, the type of trips being taken within a population, and the behavior of the population during the first peak. A higher number of public transit users might result in more infections as there exists more people in close contact with one another, increasing the contact rate of this

specific population, as opposed to individuals travelling alone in passenger cars. Recreational trips could result in a higher contact rate, due to longer contact time between individuals, than essential trips in which the person is wearing a mask and practicing proper social distancing. The results of the potential reduction of cases in the second peak for this study rely on the behavior of individuals during the first peak; therefore, any changes in behavior amongst individuals in a certain area have a direct impact on results for the reduction of infections.

A surge in hospitalizations during the peak of infections can have a compounding effect on the demand for medical resources and ultimately lead to an increase in the death rate of those infected with the COVID-19 virus. Nearly one in four COVID-19 deaths could be attributed to hospitals strained by surging caseload (Sameer et al., 2021). The surge-mortality relationship was stronger as the pandemic waned on in the summer of 2020, June to August, than it was in the spring, March to May, of 2020. The mortality rate that was calculated for each state studied is dependent on the time period and healthcare infrastructure capacity of the specific region of infections. Any change in the virus itself can also have an impact on the transmission and mortality rates. The primary variants of the COVID-19 virus during the study period were the Alpha, Beta, Gamma, and Delta variants. A study conducted on these four variants from June 2020 to October 2021 concluded that the mortality rates were 2.16, 2.23, 1.50, and 2.08, respectively (Lin 2021). Advances in the virus protection throughout the study period can also have an effect on the contact rate of the virus. The SARS-CoV-2 vaccine was rolled out in the United States in the beginning of 2021 for

essential healthcare workers, and quickly became eligible for all age groups by the summer of 2021. The dynamic of social interaction and spread of the virus was altered once a large proportion of the population were able to be vaccinated against the early variants of the virus. Once the vaccine became readily available, individuals might have resorted to pre-pandemic behaviors, including taking more vehicular trips in the second peak.

Additional limitations of the research conducted relate to the generalization of traffic and infection trends within each state. The relationship between traffic volumes and infection numbers involves data that is compiled at a state level and is only geographically defined by the state in which the values are recorded. Continuous count stations, from which the traffic volumes were collected, include both rural and urban count stations. The contact rate for COVID-19 might differ in each of these regions, with higher densities in urban populations, the contact rate may be higher than the contact rate in a rural population. Geographically specific COVID-19 infection data could allow corresponding traffic detectors to be used to determine the relationship between traffic volumes and infections on a more local level, which could be scaled up to represent the entire population of the state. Homogeneous mixing can lead to an overestimation of the epidemic size and the magnitude of intervention needed to control the outbreak (Stroud, 2006).

Another limitation of this research is that only a fraction of the U.S. states were represented in this study. Due to the lack of available traffic volume data, the number of states that could be studied was limited. While a range of geographic regions were

represented by the states studied, a larger variety of states might be able to provide richer results that could be applied to the remaining U.S. states.

The infection data compiled for this research is based on reported values, which may be underestimated, compared to the actual number of COVID-19 infections around the U.S. An underestimation of infections could also be contributed to the availability of tests during the first peak of infections, compared to the second peak where tests were more readily available for individuals and more accurate at detecting the virus. The reporting of infection numbers occurs when an individual tests positive for the Coronavirus; therefore, the majority of reports come from hospitals and healthcare facilities. Hospitalization rates that were calculated for the predicted infection values may be skewed higher than the actual values, as reported cases are determined at hospitals, which contribute also to the number of patients that are admitted due to COVID-19.

The SIR model is very general and can apply to a multitude of pandemics. COVID-19, and every new virus that becomes a pandemic, has certain characteristics that differentiate itself from other viruses. The model used throughout this study was not specifically tailored towards the COVID-19 virus but instead towards a general virus that causes a pandemic. There are an unconceivable number of factors that result in the transmission of a virus within a pandemic and the SIR model seeks to simplify these factors into a few parameters.

Implementing governmental mandates and directives that reduce the total number of vehicles on the roadways within a population may decrease the total number of infections

that are likely to occur; however, there are additional environmental benefits that result from the reduction of motor vehicle trips. A study conducted by El-Sayed looking at the impact of the COVID-19 pandemic on the air quality in Florida found that a number of air pollutants were reduced during the time period of the lockdown in Florida. Reduction in nitrogen dioxide and carbon monoxide, two of the six EPA criteria pollutants, were observed throughout the state in most major cities, as a result of a reduction of motor vehicles during this time period. Likewise, concentrations of ozone were also measured to be reduced, as a result of the nitrogen oxide concentrations in the atmosphere being reduced. (El-Sayed 2021). An additional benefit of reducing the total volume of cars on the roadway would be that there is a reduction of criteria pollutants in the atmosphere, which has the potential to improve air quality and make for a healthier environment for the corresponding population.

Traffic volume data for the remaining forty U.S. states can be compiled, along with the infection counts, and the prediction models can be created for every U.S. state. A larger database of traffic volumes and infections can provide more refined results and infection predictions that can be used on a federal or state level to provide guidance during future epidemics. The results from this study can be used in comparison with an analysis of more localized traffic and infection data to determine changes in flattening of the infection curve based on geographic region. Data on a county or city level can be used to predict infections within a certain area and governmental directives can be imposed on potential hotspots for infections. Location-based service, or LBS, can be utilized to determine common areas of grouping and the exposure time between individuals in a certain area, which can be used as

input for the transmission rate and development of infection curve models and ultimately infection control.

The findings from this study can be utilized by government officials and communities that are concerned with the mitigation of disease and infection rates within a specific area. Implementing stay-at-home mandates, or even work-at-home incentives for certain employees, can reduce the total traffic volume within a population.

## 6.0 REFERENCES


Achaiah NC, Subbarajasetty SB, Shetty RM. $R_0$ and $R_e$ of COVID-19: Can We Predict When the Pandemic Outbreak will be Contained? Indian J Crit Care Med. 2020 Nov;24(11):1125-1127. doi: 10.5005/jp-journals-10071-23649. PMID: 33384521; PMCID: PMC7751056.

Allen, Linda JS, Fred Brauer, Pauline Van den Driessche, and Jianhong Wu. Mathematical epidemiology. Vol. 1945. Berlin: Springer, 2008.

Bagal DK, Rath A, Barua A, Patnaik D. Estimating the parameters of susceptible-infected-recovered model of COVID-19 cases in India during lockdown periods. Chaos Solitons Fractals. 2020 Nov;140:110154. doi: 10.1016/j.chaos.2020.110154. Epub 2020 Jul 29. PMID: 32834642; PMCID: PMC7388782.

Beck, Matthew J; Hensher, David A. Insights into the impact of COVID-19 on household travel and activities in Australia – The early days under restrictions. Transport Policy, Volume 96, Issue 0, 2020, pp 76-93. https://trid.trb.org/view/1719131

Berman, Jesse D; Ebisu, Keita. Changes in U.S. air pollution during the COVID-19 pandemic. Science of the Total Environment, Volume 739, Issue 0, 2020, 139864. https://trid.trb.org/view/1713548

Borkowski, Przemyslaw; Jazdzewska-Gutta, Magdalena; Szmelter-Jarosz, Agnieszka. Lockdowned: Everyday mobility changes in response to COVID-19. Journal of Transport Geography, Volume 90, Issue 0, 2021 https://trid.trb.org/view/1753543

Brauer, F., Driesseche, P., & Wu, J. Mathematical Epidemiology, 2010.

Calderon-Anyosa, Renzo J C; Kaufman, Jay S. Impact of COVID-19 lockdown policy on homicide, suicide, and motor vehicle deaths in Peru. Preventive Medicine, Volume 143, Issue 0, 2021. https://trid.trb.org/view/1753959

Cartenì, Armando; Di Francesco, Luigi; Martino, Maria. How mobility habits influenced the spread of the COVID-19 pandemic: Results from the Italian case study. Science of the Total Environment, Volume 741, Issue 0, 2020, 140489. https://trid.trb.org/view/1717215

Centers for Disease Control and Prevention (CDC). "Morbidity and Mortality Weekly Report: Provisional Mortality Data – United States, 2021." *CDC*, U.S. Department of Health & Human Services, 29 April. 2022, www.cdc.gov/mmwr/volumes/71/wr/mm7117e1.htm. Accessed 16 Sept. 2024.

Centers for Disease Control and Prevention (CDC). "Morbidity and Mortality Weekly Report: COVID-19 Mortality Update – United States, 2022." *CDC*, U.S. Department of Health & Human Services, 5 May. 2023, www.cdc.gov/mmwr/volumes/72/wr/mm7218a4.htm. Accessed 16 Sept. 2024.

Centers for Disease Control and Prevention (CDC). "COVID-19 Clinical Presentation." *CDC*, U.S. Department of Health & Human Services, 14 June. 2024, www.cdc.gov/covid/hcp/clinical-care/covid19-presentation.html#cdc_generic_section_4-incubation-period. Accessed 16 Sept. 2024.

Chauhan, A., and Singh R. P.. "Decline in PM2.5 concentrations over major cities around the world associated with COVID-19." Environmental Research 187: 109634, 2020.

Church DL. Major factors affecting the emergence and re-emergence of infectious diseases. Clin Lab Med. 2004 Sep;24(3):559-86, v. doi: 10.1016/j.cll.2004.05.008. PMID: 15325056; PMCID: PMC7119055.

Covid-19 Hospitalizations. *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, https://gis.cdc.gov/grasp/covidnet/covid19_3.html.

COVID-19 Pandemic: Statewide Analysis of Social Separation and Activity Restriction. currents. 2010 Feb 22.

COVID-19 United States Cases by County. (2020). Retrieved September 09, 2022, from https://coronavirus.jhu.edu/us-map

de Haas, Mathijs; Faber, Roel; Hamersma, Marije. How COVID-19 and the Dutch 'intelligent lockdown' change activities, work and travel behaviour: Evidence from longitudinal data in the Netherlands. Transportation Research Interdisciplinary Perspectives, Volume 6, Issue 0, 2020, 100150, https://trid.trb.org/view/1714701.

Del Valle SY, Hyman JM, Chitnis N. Mathematical models of contact patterns between age groups for predicting the spread of infectious diseases. Math Biosci Eng. 2013 Oct-Dec;10(5-6):1475-97. doi: 10.3934/mbe.2013.10.1475. PMID: 24245626; PMCID: PMC4002176.

Din RU, Algehyne EA. Mathematical analysis of COVID-19 by using SIR model with convex incidence rate. Results Phys. 2021 Apr;23:103970. doi: 10.1016/j.rinp.2021.103970. Epub 2021 Feb 19. PMID: 33623731; PMCID: PMC7893319.

Doucette, Mitchell L; Tucker, Andrew; Auguste, Marisa E; Watkins, Amy; Green, Christa; Pereira, Flavia E; Borrup, Kevin T; Shapiro, David; Lapidus, Garry. Initial impact of COVID-19's stay-at-home order on motor vehicle traffic and crash patterns in Connecticut: an interrupted time series analysis. Injury Prevention, Volume 27, Issue 1, 2021, n.p. https://trid.trb.org/view/1765118.

Eisenburg, M. "Param-Estimation-Sir/License at Master · Epimath/Param-Estimation-Sir." *GitHub*, https://github.com/epimath/param-estimation-SIR/blob/master/license, 2017.

El-Sayed, M.H., Koehler, K., Elshorbany, Y.F.: "On the impact of COVID-19 pandemic on air quality in Florida", submitted to Environmental Pollution, 2021.

French G, Hulse M, Nguyen D, et al. Impact of Hospital Strain on Excess Deaths During the COVID-19 Pandemic — United States, July 2020–July 2021. MMWR Morb Mortal Wkly Rep 2021;70:1613–1616.
DOI: http://dx.doi.org/10.15585/mmwr.mm7046a5external icon.

Gendreau, M.. "Halting the travels of infectious disease: Lessons from the Ebola epidemic" TR News 299: 40–42, 2015. https://doi.org/10.17226/23612.

Gualtieri, G. Brilli, L. Carotenuto, F. Vagnoli, C. Zaldei, A. Gioli, B. Quantifying road traffic impact on air quality in urban areas: A Covid19-induced lockdown analysis in Italy, 2020. https://doi.org/10.1016/j.envpol.2020.115682.

Hadjidemetriou, Georgios M; Sasidharan, Manu; Kouyialis, Georgia; Parlikad, Ajith K. The impact of government measures and human mobility trend on COVID-19 related

deaths in the UK. Transportation Research Interdisciplinary Perspectives, Volume 6, Issue 0, 2020, 100167. https://trid.trb.org/view/1718474.

Hara, Yusuke; Yamaguchi, Hiromichi. Japanese travel behavior trends and change under COVID-19 state-of-emergency declaration: Nationwide observation by mobile phone location data. Transportation Research Interdisciplinary Perspectives, Volume 9, Issue 0, 2021, 100288 https://trid.trb.org/view/1760520.

Harris, E. Jeffrey. August 4, 2022. Overcoming Reporting Delays Is Critical to Timely Epidemic Monitoring: The Case of COVID-19 in New York City. doi: https://doi.org/10.1101/2020.08.02.20159418.

Inada, Haruhiko; Ashraf, Lamisa; Campbell, Sachalee. COVID-19 lockdown and fatal motor vehicle collisions due to speed-related traffic violations in Japan: A time-series study. Injury Prevention, Volume 27, Issue 1, 2021, pp 98-100 https://trid.trb.org/view/1765117.

Islam, Sabique Ul, "Fuel Shortages During Hurricanes: Epidemiological Modeling and Optimal Control" (2019). *Doctoral Dissertations and Master's Theses*. 488. https://commons.erau.edu/edt/488.

Kristiansson, Fanny Margaretha, "International Analysis on the Traffic Impact of the COVID-19 Pandemic" (2021). *Doctoral Dissertations and Master's Theses*. 578. https://commons.erau.edu/edt/578.

Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., . . . Lessler, J. (2020). The incubation period of Coronavirus DISEASE 2019 (COVID-19) from

publicly REPORTED Confirmed Cases: Estimation and application. Annals of Internal Medicine, 172(9), 577-582. doi:10.7326/m20-0504

Lee, Jaeyoung; Abdel-Aty, Mohamed. Changes in Traffic Crash Patterns: Before and After the Outbreak of COVID-19 in Southern Florida. Transportation Research Board 100th Annual Meeting, Transportation Research Board, 2021, 23p https://trid.trb.org/view/1759808.

Lian, Xinbo; Huang, Jianping; Huang, Rujin; Liu, Chuwei; Wang, Lina; Zhang, Tinghan. (2020) Impact of city lockdown on the air quality of COVID-19-hit of Wuhan city. Science of the Total Environment, Volume 742, Issue 0, 2020, 140556 https://trid.trb.org/view/1718924.

Lin L, Liu Y, Tang X and He D (2021) The Disease Severity and Clinical Outcomes of the SARS-CoV-2 Variants of Concern. *Front. Public Health* 9:775224. doi: 10.3389/fpubh.2021.775224.

Lounis M, Bagal DK. Estimation of SIR model's parameters of COVID-19 in Algeria. Bull Natl Res Cent. 2020;44(1):180. doi: 10.1186/s42269-020-00434-5. Epub 2020 Oct 19. PMID: 33100825; PMCID: PMC7570398.

Mahmoudi, Jina; Xiong, Chenfeng; Zhang, Lei. How Social Distancing, Mobility, and Preventive Policies Affect COVID-19 Outcomes in Urban Areas: Big Data-driven Evidence from the DMV Megaregion. Transportation Research Board 100th Annual Meeting, Transportation Research Board, 2021, 17p https://trid.trb.org/view/1759950.

Mallapaty, S. (2020, October 23). Why COVID outbreaks look set to worsen this winter. Retrieved December 11, 2022, from https://www.nature.com/articles/d41586-020-02972-4.

McEvoy, J. (2020, December 02). CDC director: Winter could Bring 200,000 deaths, 'MOST Difficult' period of US'S public health history. Retrieved March 11, 2021, from https://www.forbes.com/sites/jemimamcevoy/2020/12/02/cdc-director-winter-could-bring200000-deaths-most-difficult-period-of-uss-public-health-history/?sh=105eb5a75d85.

Nazario, B. April 20, 2020. Coronavirus incubation period: How long and when most contagious. Retrieved March 10, 2021, from https://www.webmd.com/lung/coronavirusincubation-period#1

Nelder, J.A. & Mead, R. A simplex method for function minimization. Comput. J. 1965; 7, 308.

Newburger, E. (2020, November 29). Fauci says Christmas and New Year's restrictions will be necessary due to holiday coronavirus wave. Retrieved March 11, 2021, from https://www.cnbc.com/2020/11/29/coronavirus-fauci-says-christmas-and-new-yearsrestrictions-will-be-necessary.htm.

Nguemdjo U, Meno F, Dongfack A, Ventelou B (2020) Simulating the progression of the COVID-19 disease in Cameroon using SIR models. PLOS ONE 15(8): e0237832.

Noland, Robert B. Mobility and the effective reproduction rate of COVID-19. Journal of Transport & Health, Volume 20, Issue 0, 2021, 101016 https://trid.trb.org/view/1768304.

Parr, S., Wolshon, B., Murray-Tuite, P., Renne, J., and K. Kim, (2020). "Traffic Impacts of the COVID-19 Pandemic: Statewide Analysis of Social Separation and Activity Restriction". ASCE Natural Hazards Review. 21(3): 04020025, doi:10.1061/(asce)nh.1527-6996.0000409.

Parr S, Wolshon B, Murray-Tuite P, Lomax T. Multistate assessment of roadway travel, social separation, and COVID-19 cases. Journal of Transportation Engineering, Part A: Systems. 2021 May 1;147(5):04021012.

Pierre, R., Kobelo, D., Kisabanzira, N., Kodi, J., Kitali, A.E., Sando, T. Invited Student Paper - Investigating Changes in Florida Traffic Crash Trends Due to COVID-19 Pandemic. Transportation Research Board 100th Annual 103 Meeting, Transportation Research Board, 2021, 18p https://trid.trb.org/view/1759002.

Rodríguez-Urrego, Daniella; Rodríguez-Urrego, Leonardo. Air quality during the COVID-19: PM2.5 analysis in the 50 most polluted capital cities in the world. Environmental Pollution, Volume 266, Issue 0, 2020, 115042 https://trid.trb.org/view/1721291

Sameer S. Kadri, Junfeng Sun, Alexander Lawandi, et al. Association Between Caseload Surge and COVID-19 Survival in 558 U.S. Hospitals, March to August 2020. Ann Intern Med.2021; 174:1240-1251. [Epub 6 July 2021]. doi:10.7326/M21-1213

Sandhu P, Shah AB, Ahmad FB, Kerr J, Demeke HB, Graeden E, Marks S, Clark H, Bombard JM, Bolduc M, Hatfield-Timajchy K, Tindall E, Neri A, Smith K, Owens C, Martin T, Strona FV. Emergency Department and Intensive Care Unit Overcrowding and Ventilator Shortages in US Hospitals During the COVID-19 Pandemic, 2020-2021. Public Health Rep. 2022 Jul-Aug;137(4):796-802. doi: 10.1177/00333549221091781. Epub 2022 Jun 1. PMID: 35642664; PMCID: PMC9257510.

Siciliano, Bruno; Dantas, Guilherme; da Silva, Cleyton M; Arbilla, Graciela. Increased O3 levels during the COVID-19 lockdown: Analysis for the city of Rio de Janeiro, Brazil. Science of the Total Environment, Volume 737, Issue 0, 2020, 139765 https://trid.trb.org/view/1709705

Statista. "Population density in the U.S. by federal States including the District of Columbia in 2019." Accessed November 14, 2022. https://www.statista.com/statistics/183588/population-density-in-the-federal-States-of-theus

Stroud, P.D., Sydoriak, S.J., Riese, J.M., Smith, J.P., Mniszewski, S.M., Romero, P.R. Semi-empirical power-law scaling of new Infection rate to model epidemic dynamics with inhomogeneos mixing. *Mathematical Biosciences.* 2006; 203:301–318.

Truong, D., Truong, M.D. Projecting Daily Travel Behavior by Distance During the Pandemic and the Spread of COVID-19 Infections – Are We in a Closed Loop

Scenario? Transportation Research Interdisciplinary Perspectives, Volume 9, Issue 0, 2021, 100283 https://trid.trb.org/view/1758690.

U.S. Department of Health & Human Services. "Covid-19 Reported Patient Impact and Hospital Capacity by State Timeseries." *HealthData.gov*, 8 Feb. 2023, https://healthdata.gov/Hospital/COVID-19-Reported-Patient-Impact-and-Hospital-Capa/g62h-syeh.

U.S. Department of Transportation, Federal Highway Administration. Traffic Monitoring Guide. Washington D.C. 2016.

Wei, Ye; Wang, Jiaoe; Song, Wei; Xiu, Chunliang; Ma, Li; Pei, Tao. Spread of COVID-19 in China: analysis from a city-based epidemic and mobility model. Cities, Volume 110, Issue 0, 2021 https://trid.trb.org/view/1747799.

Yan, Wang. Government Policies, National Culture and Social Distancing during the First Wave of the COVID-19 Pandemic: International Evidence. Safety Science, Volume 135, Issue 0, 2021 https://trid.trb.org/view/1761338.

# APPENDIX A

2/24/23 10:33 AM C:\Users\tateg\OneDri...\downloaddata.m 1 of 2

```
clc;
clear;
close all;

% % read station values
% [~,~,raw] = xlsread('tcds_list (1).xlsx');
% %
% locid = raw([3:end],1);
% locid([1:38,41:43,45:47,49:52,54:56,62:75,79:104,106,109,111:114,117:185]) = [];
%location numbers
%x =↙
{'001121','001125','001170','001180','001190','001200','001210','001220','001230','0012↙
60','001270','001280','001290'};
t = 1;

% x = inputdlg('Enter LocationID: ','s');
% locid = {x{t}};
% while length(locid) == t %loop stops when user presses cancel
%     x = inputdlg('Enter LocationID(if no more locations, press Cancel): ','s');
%     locid = [locid,x];
% t = t+1;
% end
%read locid from excel
[~,~,raw] = xlsread('Data_MT.xlsx');
locid = raw([3:93],1);


for r = 1:length(locid)
%    for u = %5:9
%jan values
%%s and locid{r} if using strings for locid
web(sprintf('https://mdt.public.ms2soft.com/tcds/file_download.aspx?↙
agency_id=1670&tcds_user_type=&format=excel&type=MonthlyVolume&abn=1&local_id=%↙
s&sdate=2021-01-08',locid{r}),'-browser');
%feb values
web(sprintf('https://mdt.public.ms2soft.com/tcds/file_download.aspx?↙
agency_id=1670&tcds_user_type=&format=excel&type=MonthlyVolume&abn=1&local_id=%↙
s&sdate=2021-02-08',locid{r}),'-browser');
%feb values
web(sprintf('https://mdt.public.ms2soft.com/tcds/file_download.aspx?↙
agency_id=1670&tcds_user_type=&format=excel&type=MonthlyVolume&abn=1&local_id=%↙
s&sdate=2021-03-08',locid{r}),'-browser');
%mar values
web(sprintf('https://mdt.public.ms2soft.com/tcds/file_download.aspx?↙
agency_id=1670&tcds_user_type=&format=excel&type=MonthlyVolume&abn=1&local_id=%↙
s&sdate=2021-04-08',locid{r}),'-browser');
%ap value
web(sprintf('https://mdt.public.ms2soft.com/tcds/file_download.aspx?↙
agency_id=1670&tcds_user_type=&format=excel&type=MonthlyVolume&abn=1&local_id=%↙
s&sdate=2021-05-08',locid{r}),'-browser');
```

```
%may values
web(sprintf('https://mdt.public.ms2soft.com/tcds/file_download.aspx?↙
agency_id=1670&tcds_user_type=&format=excel&type=MonthlyVolume&abn=1&local_id=%↙
s&sdate=2021-06-08',locid{r}),'-browser');
%june values
web(sprintf('https://mdt.public.ms2soft.com/tcds/file_download.aspx?↙
agency_id=1670&tcds_user_type=&format=excel&type=MonthlyVolume&abn=1&local_id=%↙
s&sdate=2021-07-08',locid{r}),'-browser');
   pause(2);
   end
   pause(3);
% end
```

```
clc;
clear;
close all;

% read station values
% [~,~,raw] = xlsread('Data_IL.xlsx',2);
finaly = [];
% locid = raw([3:end],1);

%prompt user for location ids
t = 1;

% x = inputdlg('Enter LocationID: ','s');
% locid = {x{t}};
% while length(locid) == t %loop stops when user presses cancel
%     x = inputdlg('Enter LocationID(if no more locations, press Cancel): ','s');
%     locid = [locid,x];
% t = t+1;
% end

%read excel file to get locid in string format
[~,~,raw] = xlsread('Data_MT.xlsx');
locid = raw([3:93],1);


tables = cell(1,length(locid));
for r = 1:length(locid)
% if download in format with parenthesis
% jan = xlsread(sprintf('MonthlyVolumeReport_1_2019 (%d).xlsx',x));
% feb = xlsread(sprintf('MonthlyVolumeReport_2_2019 (%d).xlsx',x));
% mar = xlsread(sprintf('MonthlyVolumeReport_3_2019 (%d).xlsx',x));
% ap = xlsread(sprintf('MonthlyVolumeReport_4_2019 (%d).xlsx',x));
% may = xlsread(sprintf('MonthlyVolumeReport_5_2019 (%d).xlsx',x));
% june = xlsread(sprintf('MonthlyVolumeReport_6_2019 (%d).xlsx',x));
% febvalues = feb([2:29],26);
%2020
% jan = xlsread(sprintf('MonthlyVolumeReport_1_2020 (%d).xlsx',x));
% feb = xlsread(sprintf('MonthlyVolumeReport_2_2020 (%d).xlsx',x));
% mar = xlsread(sprintf('MonthlyVolumeReport_3_2020 (%d).xlsx',x));
% ap = xlsread(sprintf('MonthlyVolumeReport_4_2020 (%d).xlsx',x));
% may = xlsread(sprintf('MonthlyVolumeReport_5_2020 (%d).xlsx',x));
% june = xlsread(sprintf('MonthlyVolumeReport_6_2020 (%d).xlsx',x));

%for loop for each year 2015-2019
% for y = 5:9
%if download in format where locid is present in file name
%if exist(sprintf('MonthlyVolumeReport_%s_6_2019.xlsx',locid{r})) == 2
    %jun = xlsread(sprintf('MonthlyVolumeReport_%s_6_2019.xlsx',locid{r}));
    %junvalues = jun([2:31],26);
%     junv = [junvalues]';
```

```
% else
%     junv = zeros(1,30);
% end
if exist(sprintf('MonthlyVolumeReport_%s_1_2021.xlsx',locid{r})) == 2
    jan = xlsread(sprintf('MonthlyVolumeReport_%s_1_2021.xlsx',locid{r}));
    janvalues = jan([2:32],26);
    janv = [janvalues]';
else
    janv = zeros(1,31);
end
%     if exist(sprintf('MonthlyVolumeReport_%s_2_2020.xlsx',locid{r})) == 2 %&& y ~= 6
%     feb = xlsread(sprintf('MonthlyVolumeReport_%s_2_2020.xlsx',locid{r}));
%     febvalues = feb([2:29],26);
%     febv = [febvalues]';
    if exist(sprintf('MonthlyVolumeReport_%s_2_2021.xlsx',locid{r})) == 2 %&& y == 6; %↙
2016 leap year
        feb = xlsread(sprintf('MonthlyVolumeReport_%s_2_2021.xlsx',locid{r}));
        febvalues = feb([2:29],26);
        febv = [febvalues]';
%     elseif exist(sprintf('MonthlyVolumeReport_%s_2_2021.xlsx',locid{r})) ~= 2 %&& y↙
== 6
%         febv = zeros(1,31);
    else
        febv = zeros(1,28);
    end

    if exist(sprintf('MonthlyVolumeReport_%s_3_2021.xlsx',locid{r})) == 2
    mar = xlsread(sprintf('MonthlyVolumeReport_%s_3_2021.xlsx',locid{r}));
    marvalues = mar([2:32],26);
    marv = [marvalues]';
        else
        marv = zeros(1,31);
    end
    if exist(sprintf('MonthlyVolumeReport_%s_4_2021.xlsx',locid{r})) == 2
    ap = xlsread(sprintf('MonthlyVolumeReport_%s_4_2021.xlsx',locid{r}));
    apvalues = ap([2:31],26);
    apv = [apvalues]';
        else
        apv = zeros(1,30);
    end
    if exist(sprintf('MonthlyVolumeReport_%s_5_2021.xlsx',locid{r})) == 2
    may = xlsread(sprintf('MonthlyVolumeReport_%s_5_2021.xlsx',locid{r}));
    mayvalues = may([2:32],26);
    mayv = [mayvalues]';
        else
        mayv = zeros(1,31);
    end
    if exist(sprintf('MonthlyVolumeReport_%s_6_2021.xlsx',locid{r})) == 2
    june = xlsread(sprintf('MonthlyVolumeReport_%s_6_2021.xlsx',locid{r}));
    junevalues = june([2:31],26);
```

```
    junev = [junevalues]';
    else
    junev = zeros(1,30);
    end
finalm = [janv,febv,marv,apv,mayv,junev];
finaly = [finaly,finalm];

% % tabl = table(finalm);
% if y == 9
tables{r} = finaly;
finaly = [];
% else
% end

end
% end
```

# APPENDIX B



```
import numpy as np
from scipy.integrate import odeint
import matplotlib.pyplot as plt
import matplotlib
import pandas as pd


import os, sys
from google.colab import drive
drive.mount('data')
%cd data/MyDrive/CIV542_DATA_SHARE/
nb_path = '/content/notebooks'
os.symlink('/content/data/MyDrive/CIV542_DATA_SHARE/', nb_path)
sys.path.insert(0, nb_path)


df = pd.read_excel("Summary US Sweden 2019_2020.xlsx", header=5)

statenum = 9 #NY=0,FL=1,MA=2,NH=3,IL=4,IN=5,OH=6,VT=7,MI=8,MT=9
peak = 2
dateindex = [13,25,37,49,61,73,85,97,109,121]
casesindex = [7,19,31,43,55,67,79,91,103,115]
date = df.iloc[dateindex[statenum], :] # 13-NY 25-FL 37-MA 49-NH 61-IL 73-IN 85-OH 97-VT 109-MI 121-MT
cases = df.iloc[casesindex[statenum],:] #19-NY 31-FL 43-MA 55-NH 67-IL 79-IN 91-OH 103-VT 115-MI 127-MT
dateplot = date
traffic2020index = [8,20,32,44,56,68,80,92,104,116]
traffic2020 = df.iloc[traffic2020index[statenum],:] #26-FL 38-MA 50-NH 62-IL 74-IN 98-VT 110-MI
traffic2019 = df.iloc[traffic2020index[statenum]+1,:] #27-FL 39-MA 51-NH 63-IL 75-IN 99-VT 111-MI

Nlist = [19000000,21000000,7000000,1500000,12500000,7000000,12000000,650000,10000000,1000000]
N = Nlist[statenum]

d = {'Date':cases, 'Cases':date, 'DatePlot':dateplot, '2019 Traffic':traffic2019, '2020 Traffic':traffic2020}
dfnew = pd.DataFrame(d)
dfnew = dfnew.drop(index=dfnew.index[0])
dfnew = dfnew.dropna(subset=['Date','Cases'], how='any') #new for FL

dfnew['7-day Average'] = dfnew.iloc[:,1].rolling(window=7).mean()

dfnew.DatePlot = dfnew.DatePlot.apply(pd.to_datetime) # convert to datetime
dfnew['DatePlot'] = matplotlib.dates.date2num(dfnew['Date'])

if peak == 1:
  firstdate = [80,82,78,91,78,85,75,73,78,74]
  seconddate = [140,120,168,170,180,180,144,111,120,130]
if peak == 2:
  firstdate =[276,147,290,266,265,269,267,284,279,248]
  seconddate = [520,243,421,430,405,430,430,426,390,390]

firstdatepeak1 = [80,82,78,91,78,85,75,73,78,74]
seconddatepeak1 = [140,120,168,170,180,180,144,111,120,130]



zero2019 = len(dfnew.loc[dfnew['2019 Traffic'] == 0])
zero2020 = len(dfnew.loc[dfnew['2020 Traffic'] == 0])
dfnew = dfnew.iloc[0:seconddate[statenum],:]
dfnew = dfnew[(dfnew['2019 Traffic'] != 0)]
dfnew = dfnew[(dfnew['2020 Traffic'] != 0)] # remove zeros from 2020 and 2019 traffic volumes to create percent change graph
dfnew['Diff Traffic'] = dfnew['2020 Traffic'] - dfnew['2019 Traffic']
dfnew['Traffic Average'] = dfnew.iloc[:,6].rolling(window=7).mean()
dfnew['% Change Traffic'] = dfnew['Diff Traffic']/(dfnew['2019 Traffic'])*100
dfnew['% Change Average'] = dfnew.iloc[:,8].rolling(window=7).mean()
plt.plot(dfnew['Date'], dfnew['7-day Average']);
plt.ylabel('Cases')
plt.xticks(rotation=45);


# create figure and axis objects with subplots()
fig,ax = plt.subplots(figsize=(9,5))
# make a plot
ax.plot(dfnew['Date'], dfnew['7-day Average'], label='COVID-19 Infections')
# set y-axis label
ax.set_ylabel("Cases",color="blue",fontsize=12)
plt.xticks(rotation=45);
ax.set_title('Cases and Traffic Change for 1st and 2nd Peak MT')
# twin object for two different y-axis on the sample plot
```

```
ax2=ax.twinx()
# make a plot with different y-axis using second axis object
ax2.plot(dfnew['Date'], dfnew['% Change Average'], 'r', label='Traffic Volume')
ax2.set_ylabel("% Change in Traffic",color="red",fontsize=12)
plt.axvline(x = dfnew['Date'].iloc[firstdatepeak1[statenum]], color = 'k', linestyle='--')
plt.axvline(x = dfnew['Date'].iloc[seconddatepeak1[statenum]-1], color = 'k', linestyle='--')
plt.axvline(x = dfnew['Date'].iloc[firstdate[statenum]], color = 'k', linestyle='--')
plt.axvline(x = dfnew['Date'].iloc[seconddate[statenum]-(zero2019+zero2020+1)], color = 'k', linestyle='--')
plt.axvspan(dfnew['Date'].iloc[firstdatepeak1[statenum]], dfnew['Date'].iloc[seconddatepeak1[statenum]-(zero2019+zero2020)], alpha=0.1, color
plt.axvspan(dfnew['Date'].iloc[firstdate[statenum]], dfnew['Date'].iloc[seconddate[statenum]-(zero2019+zero2020+1)], alpha=0.1, color='red')
plt.subplots_adjust(bottom=.2)
fig.legend()
plt.savefig('1stand2ndPeakMTv2.png', bbox_inches='tight',transparent=True)
plt.show()


# create figure and axis objects with subplots()
# make a plot

plt.ylabel("Traffic Volume",color="black",fontsize=12)
plt.xticks(rotation=45);
plt.title('Difference in Traffic Volumes from 2019 to 2020')

# make a plot with different y-axis using second axis object
plt.plot(dfnew['Date'], dfnew['Traffic Average'], 'r')

plt.axvline(x = 737503.0, color = 'k', linestyle='--',label = 'axvline - full height')
plt.axvline(x = 737592.0, color = 'k', linestyle='--',label = 'axvline - full height')
plt.axvline(x =   737711.0, color = 'k', linestyle='--',label = 'axvline - full height')
plt.axvline(x = 737866.0, color = 'k', linestyle='--',label = 'axvline - full height')
plt.subplots_adjust(bottom=.2)
plt.savefig('DifferenceinTraffic.png')
plt.show()



import numpy as np

def model(ini, time_step, params):
    Y = np.zeros(3) #column vector for the state variables
    X = ini
    mu = 0
    beta = params[0]
    gamma = params[1]

    Y[0] = mu - beta*X[0]*X[1] - mu*X[0] #S
    Y[1] = beta*X[0]*X[1] - gamma*X[1] - mu*X[1] #I
    Y[2] = gamma*X[1] - mu*X[2] #R

    return Y

def x0fcn(params, data):
    S0 = 1.0 - (data[0]/N)
    I0 = data[0]/N
    R0 = 0.0
    X0 = [S0, I0, R0]

    return X0


def yfcn(res, params):
    return res[:,1]*N



import numpy as np
from scipy.stats import poisson
from scipy.stats import norm

from scipy.integrate import odeint as ode

def NLL(params, data, times): #negative log likelihood
    params = np.abs(params)
    data = np.array(data)
    res = ode(model, x0fcn(params,data), times, args=(params,))
    y = yfcn(res, params)
    nll = sum(y) - sum(data*np.log(y))
    #nll = -sum(np.log(poisson.pmf(np.round(data),np.round(y)))) # the round is b/c Poisson is for (integer) count data
```

```
        # note this is a slightly shortened version--there's an additive constant term missing but it
        # makes calculation faster and won't alter the threshold. Alternatively, can do:
        # nll = -sum(np.log(poisson.pmf(np.round(data),np.round(y)))) # the round is b/c Poisson is for (integer) count data
        # this can also barf if data and y are too far apart because the dpois will be ~0, which makes the log angry

        # ML using normally distributed measurement error (least squares)
        #nll = -sum(np.log(norm.pdf(data,y,0.1*np.mean(data)))) # example WLS assuming sigma = 0.1*mean(data)
        # nll = sum((y - data)**2)  # alternatively can do OLS but note this will mess with the thresholds
        #                           for the profile! This version of OLS is off by a scaling factor from
        #                           actual LL units.
        return nll

    import numpy as np


    from scipy.integrate import odeint as ode

    def minifisher (times, params, data, delta = 0.001):
        #params = np.array(params)
        listX = []
        params_1 = np.array (params)
        params_2 = np.array (params)
        for i in range(len(params)):
            params_1[i] = params[i] * (1+delta)
            params_2[i]= params[i] * (1-delta)

            res_1 = ode(model, x0fcn(params_1,data), times, args=(params_1,))
            res_2 = ode(model, x0fcn(params_2,data), times, args=(params_2,))
            subX = (yfcn(res_1, params_1) - yfcn(res_2, params_2)) / (2 * delta * params[i])
            listX.append(subX.tolist())
        X = np.matrix(listX)
        FIM = np.dot(X, X.transpose())
        return FIM


    import numpy as np
    import scipy.optimize as optimize
    import copy


    def proflike (params, profindex, cost_func, times, data, perrange = 0.5, numpoints = 10):
        profrangedown = np.linspace(params[profindex], params[profindex] * (1 - perrange), numpoints).tolist()
        profrangeup = np.linspace(params[profindex], params[profindex] * (1 + perrange), numpoints).tolist()[1:] #skip the duplicated values
        profrange = [profrangedown, profrangeup]
        currvals = []
        currparams = []
        currflags = []

        def profcost (fit_params, profparam, profindex, data, times, cost_func):
            paramstest = fit_params.tolist()
            paramstest.insert(profindex, profparam)
            return cost_func (paramstest, data, times)

        fit_params = params.tolist() #make a copy of params so we won't change the origianl list
        fit_params.pop(profindex)
        print('Starting profile...')
        for i in range(len(profrange)):
            for j in profrange[i]:
                print(i, j)
                optimizer = optimize.minimize(profcost, fit_params, args=(j, profindex, data, times, cost_func), method='Nelder-Mead')
                fit_params = np.abs(optimizer.x).tolist() #save current fitted params as starting values for next round
                #print optimizer.fun
                currvals.append(optimizer.fun)
                currflags.append(optimizer.success)
                currparams.append(np.abs(optimizer.x).tolist())

        #structure the return output
        profrangedown.reverse()
        out_profparam = profrangedown+profrangeup
        temp_ind = range(len(profrangedown))
        #temp_ind.reverse()
        out_params = [currparams[i] for i in temp_ind]+currparams[len(profrangedown):]
        out_fvals = [currvals[i] for i in temp_ind]+currvals[len(profrangedown):]
        out_flags = [currflags[i] for i in temp_ind]+currflags[len(profrangedown):]
        output = {'profparam': out_profparam, 'fitparam': np.array(out_params), 'fcnvals': out_fvals, 'convergence': out_flags}
        return output
```

```
import matplotlib.dates as mdates
import scipy.optimize as optimize
import numpy as np
import matplotlib.pyplot as plt
import scipy.stats as stats

from scipy.integrate import odeint as ode
#### Load Data ####
# loop to iterate different starting positions
estimations = []
#for y in range(10):
    #for x in range(5):

dfnew2 = dfnew.iloc[firstdate[statenum]:seconddate[statenum]]

times = dfnew2['DatePlot'].values.tolist()
data = dfnew2['7-day Average'].values.tolist()

#### Set initial parameter values and initial states ####
params = [2, 2] #make sure all the params and inition states are float
# NY - 2,2,19000000
# FL - 2,2,21000000 2nd - 2,2,21000000
# MA - 2,2,7000000 2nd - 2,2,7000000
# NH - 2, 2, 1500000 2nd - 2,2,1500000
# IL - 2, 2, 12500000 2nd - 2,2,12500000
# IN - 2, 2, 7000000 2nd - 2,2,7000000
# OH - 2, 2, 12000000 2nd - 2,2,12000000
# VT - 2,2,650000 2nd - 2,2,650000
# MI - 12,12,10000000 2nd - 2,2,10000000
# MT - 12,12,1000000 2nd - 2,2,1000000
paramnames = ['beta', 'gamma']
ini = x0fcn(params,data)
print(ini)

#### Simulate and plot the model ####
#fig, (ax1,ax2) = plt.subplots(nrows=1, ncols=2, figsize=(12,10), sharey=True)
res = ode(model, ini, times, args=(params,))
#print(res)
sim_measure = yfcn(res, params)
#print(sim_measure)
plt.plot(dfnew2['Date'], sim_measure, 'b-', linewidth=3, label='Model simulation')
plt.plot(dfnew2['Date'], data, 'k-o', linewidth=2, label='Data')
# set monthly locator
plt.gca().xaxis.set_major_locator(mdates.MonthLocator(interval=1))
plt.ylabel('Infected Individuals')
plt.legend()
plt.text(dfnew2['DatePlot'].iloc[int(len(dfnew2['Date'])/2)], min(data), f" Beta = {np.round(params[0],3)} Gamma = {np.round(params[1],3)}")
plt.savefig('InitialParam(NY63-210).png')
plt.show()


#### Parameter estimation ####
optimizer = optimize.minimize(NLL, params, args=(data, times), method='Nelder-Mead')
paramests = np.abs(optimizer.x)
iniests = x0fcn(paramests, data)

# ### Traffic Reduced parameters
reducedparamests = [paramests[0]*.881,paramests[1]]
#### Re-simulate and plot the model with the final parameter estimates ####
xest = ode(model, iniests, times, args=(paramests,))
xestreduced = ode(model, iniests, times, args=(reducedparamests,))
est_measure = yfcn(xest, paramests)
est_measure_reduced = yfcn(xestreduced, reducedparamests)

plt.plot(dfnew2['Date'], est_measure, 'b-', linewidth=3, label='Model simulation')
#plt.plot(dfnew2['DatePlot'], est_measure_reduced, 'r-', linewidth=3, label='Model simulation reduced')
plt.plot(dfnew2['Date'], data, 'k-o', linewidth=2, label='Data')
plt.gca().xaxis.set_major_locator(mdates.MonthLocator(interval=1))
plt.ylabel('Infected Individuals')
plt.xticks(rotation=45)
plt.legend()
#plt.text(dfnew2['Date'].iloc[int(len(dfnew2['Date'])/2)-10], min(data), f" Beta = {np.round(paramests[0],3)} Gamma = {np.round(paramests[1],
plt.title('New York Observed COVID-19 Data and SIR Estimation')
plt.savefig('ObservedSIREstimationPeak2(NY).png', bbox_inches='tight')
```

```
plt.show()

#### Calculate the simplified Fisher Information Matrix (FIM) ####
FIM = minifisher(times, params, data, delta = 0.001)
#print(np.linalg.matrix_rank(FIM)) #calculate rank of FIM
#print(FIM)

#### Generate profile likelihoods and confidence bounds ####
threshold = stats.chi2.ppf(0.95,len(paramests))/2.0 + optimizer.fun
perrange = 0.25 #percent range for profile to run across

profiles={}
for i in range(len(paramests)):
  profiles[paramnames[i]] = proflike(paramests, i, NLL, times, data, perrange=perrange)
  # plt.figure()
  # plt.scatter(paramests[i], optimizer.fun, marker='*',label='True value', color='k',s=150, facecolors='w', edgecolors='k')
  # plt.plot(profiles[paramnames[i]]['profparam'], profiles[paramnames[i]]['fcnvals'], 'k-', linewidth=2, label='Profile likelihood')
  # plt.axhline(y=threshold, ls='--',linewidth=1.0, label='Threshold', color='k')
  # plt.xlabel(paramnames[i])
  # plt.ylabel('Negative log likelihood')
  # plt.legend(scatterpoints = 1)
  paramnames_fit = [n for n in paramnames if n not in [paramnames[i]]]
  paramests_fit = [v for v in paramests if v not in [paramests[i]]]
  #print(paramnames_fit)
  #print(paramests_fit)
  #plot parameter relationships
  #for j in range(profiles[paramnames[i]]['fitparam'].shape[1]):
  # plt.figure()
  # plt.plot(profiles[paramnames[i]]['profparam'],profiles[paramnames[i]]['fitparam'][:,j],'k-', linewidth=2, label=paramnames_fit[j])
  # plt.scatter(paramests[i], paramests_fit[j], marker='*',label='True value', color='k',s=150, facecolors='w', edgecolors='k')
  # plt.xlabel(paramnames[i])
  # plt.ylabel(paramnames_fit[j])
  # plt.legend(scatterpoints = 1)
  #print(profiles)
estimations.append([np.round(paramests[0],3), np.round(paramests[1],3)])
;
print(estimations)


Diff = abs(data-est_measure)
MAE = sum(Diff)/len(Diff)
MAE


dfnew2['Date']


print(f"Beta parameter is {np.round(paramests[0],3)}. Gamma parameter is {np.round(paramests[1],3)}. ")


(sum(dfnew['2020 Traffic'].iloc[80:140]) - sum(dfnew['2019 Traffic'].iloc[80:140]))/ sum(dfnew['2019 Traffic'].iloc[80:140])


dfnew['2019 Traffic'].iloc[75:144].mean()
```

```
import numpy as np
from numpy import nan
from scipy.stats import pearsonr
import sympy as sy
from scipy.integrate import odeint
import matplotlib.pyplot as plt
import matplotlib
import pandas as pd

import seaborn as sns
from sklearn.linear_model import LinearRegression


import os, sys
from google.colab import drive
drive.mount('data')
%cd data/MyDrive/CIV542_DATA_SHARE/
nb_path = '/content/notebooks'
os.symlink('/content/data/MyDrive/CIV542_DATA_SHARE/', nb_path)
sys.path.insert(0, nb_path)


df = pd.read_excel("Summary US Sweden 2019_2020.xlsx", header=5)
statenum = 9  #NY=0,FL=1,MA=2,NH=3,IL=4,IN=5,OH=6,VT=7,MI=8,MT=9
peak = 2
dateindex = [13,25,37,49,61,73,85,97,109,121]
casesindex = [7,19,31,43,55,67,79,91,103,115]
date = df.iloc[dateindex[statenum], :] # 13-NY 25-FL 37-MA 49-NH 61-IL 73-IN 85-OH 97-VT 109-MI 121-MT
cases = df.iloc[casesindex[statenum],:] #19-NY 31-FL 43-MA 55-NH 67-IL 79-IN 91-OH 103-VT 115-MI 127-MT

traffic2020index = [8,20,32,44,56,68,80,92,104,116]

traffic2020 = df.iloc[traffic2020index[statenum],:] #26-FL 38-MA 50-NH 62-IL 74-IN 98-VT 110-MI
traffic2019 = df.iloc[traffic2020index[statenum]+1,:] #27-FL 39-MA 51-NH 63-IL 75-IN 99-VT 111-MI
d = {'Date':cases, 'Cases':date, '2019 Traffic':traffic2019, '2020 Traffic':traffic2020}



dfnew = pd.DataFrame(d)
dfnew = dfnew.drop(index=dfnew.index[0])
dfnew = dfnew.dropna(subset=['Date','Cases'], how='any') #new for FL
dfnew['7-day Average'] = dfnew.iloc[:,1].rolling(window=7).mean()
dfnew['7-day Avg 5-day Lag'] = dfnew['7-day Average'].shift(-5)
dfnew['Date'] = matplotlib.dates.date2num(dfnew['Date'])


dfnew['t'] = range(len(dfnew))
if peak == 1:
    firstdate = [80,82,78,91,78,85,75,73,78,74]
    seconddate = [140,120,168,170,180,180,144,111,120,130]
if peak == 2:
  firstdate =[276,147,290,266,265,269,267,284,279,248]
  seconddate = [520,243,421,430,405,430,430,426,390,390]

dfnew2 = dfnew.iloc[firstdate[statenum]:seconddate[statenum]]

maxcases = max(dfnew2['7-day Average'])
maxdifference2019 = max(dfnew2['2019 Traffic'])
maxdifference2020 = max(dfnew2['2020 Traffic'])
#print(dfnew2.loc[peak])



model_list = []
xx = np.array(dfnew2['2019 Traffic']-dfnew2['2020 Traffic']).reshape(-1,1)
yy = np.array(dfnew2['7-day Avg 5-day Lag'])
model = LinearRegression().fit(xx, yy)
print(f"The data from set has a linear regression line equation of {str(np.around(model.coef_[0],5)) + 'x ' + '+ ' + str(np.around(model.inte
model_list.append(model)


d = 0
x_line = np.linspace(-2500000,2500000,10).reshape(-1,1)
y_vals = [model.coef_[0] * i + model.intercept_ for i in x_line]
xx = np.array(dfnew2['2019 Traffic']-dfnew2['2020 Traffic'])
yy = np.array(dfnew2['7-day Avg 5-day Lag'])
plt.scatter(xx, yy)
```

```
plt.xlim(-3000000,maxdifference2019)
plt.ylim(0,maxcases)
plt.plot(x_line, y_vals, '--')
plt.text(15,2,f"R2 = {np.around(model_list[0].score(xx.reshape(-1,1), yy),3)}")
d += 1

matplotlib.pyplot.plot_date(dfnew['Date'], dfnew['7-day Average'])


plt.plot(dfnew['Date'], dfnew['7-day Average']);


if peak==1:
  percenttraffic = (sum(dfnew2['2019 Traffic'])-sum(dfnew2['2020 Traffic']))/(sum(dfnew2['2019 Traffic']))
# % change in traffic (do for first peak then mulitply 2020 traffic volume from second peak by 1-result)
print('Did you change peak=1 for new state?')


percenttraffic


trafficred = -sum(dfnew2['2019 Traffic'])*(percenttraffic) + sum(dfnew2['2019 Traffic']) # traffic volume in 2020 for second peak so percent
#print(sum(dfnew2['2019 Traffic']),sum(dfnew2['2020 Traffic']))
#print(trafficred)
trafficdiff = sum(dfnew2['2020 Traffic'])-trafficred
peak1change = (sum(dfnew2['2020 Traffic'])-sum(dfnew2['2019 Traffic']))/sum(dfnew2['2019 Traffic'])*100
print(peak1change)


nlist = [.00073,.00016,.00125,.00004,.00078,.00036,.00004,.00038,.00067,.00007]
trafficdiff*nlist[statenum]


# Total population, N.
Nlist = [19000000,21000000,7000000,1500000,12500000,7000000,12000000,650000,10000000,1000000]
N = Nlist[statenum]
N2 = Nlist[statenum]
# Initial number of infected and recovered individuals, I0 and R0. R0 = I0*gamma
I0, R0 = dfnew2['7-day Average'].iloc[0], 0
# Everyone else, S0, is susceptible to infection initially.
S0 = N - I0 - R0
S02 = N2 - I0 - R0
# Contact rate, beta, and mean recovery rate, gamma, (in 1/days).
if peak == 1:
  betalist= [5.795,4.646,4.61,4.422,4.351,4.707,10.742,21.011,10.521,30.586]
  gammalist = [5.675,4.495,4.48,4.389,4.256,4.644,10.609,20.755,10.325,30.386]
else:
  betalist = [.929,2.253,1.187,1.39,1.032,1.045,1.104,1.794,1.536,.978]
  gammalist = [.897,2.183,1.142,1.345,.995,1.003,1.063,1.757,1.483,.938]

beta2list = [.9257,2.249,1.1804,1.389,1.0307,1.0443,1.1034,1.787,1.527,.9774]
beta, gamma = betalist[statenum], gammalist[statenum]
beta2, gamma2 = beta2list[statenum], gamma
# A grid of time points (in days)
t = np.linspace(0, 300, 300)
t = [i for i in range(301)]

# The SIR model differential equations.
def deriv(y, t, N, beta, gamma):
    S, I, R = y
    dSdt = -beta * S * I / N
    dIdt = beta * S * I / N  - gamma * I
    dRdt = gamma * I
    return dSdt, dIdt, dRdt

def deriv(y2, t, N2, beta2, gamma2):
    S2, I2, R2 = y2
    dSdt2 = -beta2 * S2 * I2 / N2
    dIdt2 = beta2 * S2 * I2 / N2   - gamma2 * I2
    dRdt2 = gamma2 * I2
    return dSdt2, dIdt2, dRdt2

# Initial conditions vector
y0 = S0, I0, R0
# Integrate the SIR equations over the time grid, t.
ret = odeint(deriv, y0, t, args=(N, beta, gamma))
S, I, R = ret.T

y02 = S02, I0, R0
```

```
ret2 = odeint(deriv, y02, t, args=(N2, beta2, gamma2))
S2, I2, R2 = ret2.T


# Plot the data on three separate curves for S(t), I(t) and R(t)
fig = plt.figure(facecolor='w')
ax = fig.add_subplot(111, facecolor='#dddddd', axisbelow=True)
#ax.plot(t, S, 'b', alpha=0.5, lw=2, label='Susceptible')
ax.plot_date(dfnew2['Date'], I[:len(dfnew2['Date'])], 'r', alpha=0.5, lw=2, label='Infected')
ax.plot_date(dfnew2['Date'], I2[:len(dfnew2['Date'])], 'b', linestyle='dashed', alpha=.5, lw=2, label='Infected Adjusted')
#ax.plot(t, R, 'g', alpha=0.5, lw=2, label='Recovered with immunity')
#ax.plot(t,dfnew2['7-day Average'], alpha=0.5, lw=2, label='NY 7-day Average Cases')
statename = ['NY','FL','MA','NH','IL','IN','OH','VT','MI','MT']
ax.set_ylabel('Cases')
ax.set_title('NY 2nd Peak Infected Cases Observed and Adjusted')
ax.set_ylim(0,maxcases)
ax.yaxis.set_tick_params(length=0)
ax.xaxis.set_tick_params(length=0, rotation=25)
ax.xaxis.set_major_formatter(matplotlib.dates.DateFormatter('%Y-%b'))
ax.xaxis.set_major_locator(matplotlib.dates.MonthLocator())
ax.grid(b=True, which='major', c='w', lw=2, ls='-')
ymax = round(max(I),1)
#ax.annotate(ymax, xy=(25, ymax), xytext=(25, ymax+5), arrowprops=dict(facecolor='black', shrink=0.05),)
ymax2 = round(max(I2),1)
#ax.annotate(ymax2, xy=(75, ymax2), xytext=(75, ymax2+5), arrowprops=dict(facecolor='black', shrink=0.05),)
legend = ax.legend()
legend.get_frame().set_alpha(0.5)
for spine in ('top', 'right', 'bottom', 'left'):
    ax.spines[spine].set_visible(False)

plt.savefig('SIRModelAdjustedNY2ndPeakMethod4.png')
plt.show()


from numpy import trapz
area = trapz(I[:len(dfnew2['Date'])], dx=1)
area2 = trapz(I2[:len(dfnew2['Date'])], dx=1)
#print(round(area),round(area2))
redinfections = round(sum(I[:len(dfnew2['Date'])]))-round(sum(I2[:len(dfnew2['Date'])])) #total number of infections reduced
changeinf = redinfections/sum(I[:len(dfnew2['Date'])])*100
print(redinfections)
print(changeinf)
```