Eli Corpron
12/12/22
Data 512

El Paso County Research Paper

Introduction

The Covid pandemic hit the entire country, and El Paso County in Colorado was no different. They, like everyone else in the country, were blindsided by an unexpected illness, found themselves in a series of mask mandates and quarantine restrictions, and to top it off was an intensifying political situation.

In addition, crime and the perception of crime exploded in the public consciousness over the course of the pandemic. On one hand people were staying at home a lot more, not going out, and the government covid relief meant for a while many people didn't have to immediately worry about losing where they live or running out of money.On the other hand we had extended periods of isolation from quarantine rules. On top of that, while immediate worries about money may have been paused for many with the Covid relief bills, the worries didn't go away, especially with rising inflation. Understanding how these forces interact would be a powerful tool in approaching future pandemic responses.

We decided to approach this in two parts. We first focused on the interaction on masking mandates and covid deaths. This was a time series analysis to attempt and see if there was a relation in the covid death rate and when masking mandates were put into effect. We did this by looking at changepoints in the time series of covid deaths, and comparing those to when masking mandates were in effect. The second part was an analysis on the interaction between covid deaths, masking mandates, and the number of crimes committed. We checked this by running multiple linear regression with number

of offenses as the response variable and covid deaths and if a masking mandate was in effect as the predicting variables. Both of these were done focusing on El Paso county. For the time series analysis we found that there wasn't a strong relation between changes in masking policy and covid deaths. In addition we didn't find strong evidence in our model that there was a relation between covid, masking mandates, and crimes committed.

Background and Related Work

There is a growing body of research into the effectiveness of masking and masking mandates and their impact on the spread of Covid. Research currently finds that masking mandates do slow down the spread of Covid (Huang et al., 2020). If masking mandates are effective then we would expect to see that when masking mandates go into effect we would see a lowered infection rate. We also expect that the infection rate of covid is pretty linked with the death rate for covid related deaths. If infections go up, more people have covid, and more people are likely to die from it. When infection rates go down, the number of people with covid decreases, and the number of people who die covid related deaths then go down. We decided to use the Covid death rate as an indicator for infection rate for our study.

For crime we found a lot of interesting and seemingly conflicting information. We saw reports that crime in 2020 and 2021 roughly maintained similar levels and year-end homicides were little over half what they were 25 years ago (Maxine, 2022). We also saw a lot of articles making claims similar to VOA News, with a title of "Why Homicide Rates in US Spiked 30% During COVID Pandemic" (Mekouar, 2022). With these two

different ideas of what crime was like, exploding or keeping to expected trends, we wanted to see what crime was like in El Paso County.

Methodology

To see how masking mandates impacted covid deaths in El Paso County I used the COVID-19 data from John Hopkins University. This is a daily updated data set, actively maintained by John Hopkins. The information for masking mandates comes from the CDC. This dataset contains masking mandate information for each county in the US to September 2021, when masking mandates were no longer required by the Federal government. The process of joining this data gave me information for the data ranges from January 22, 2020 to October 15, 2021.

The Covid data gives a cumulative total of deaths from Covid. We calculated the number of people who died on a given day to get the death rate from covid. We then created a time series and calculated the change points to see where there are significant changes in the Covid death rate. We decided to use the Pelt method for calculating change points because it executes quickly and is exact. We would then compare the change points to when masking mandates were put into effect to try and see if they had an impact or not.

I decided on answering this question by using change points because it is intuitive and easy to understand. We aren't trying to answer the question of is masking effective since many studies already explore that. What we want to see is if it had an impact in El Paso County. An easy way to see that is if there is a change in the rate people are dying from covid when masking policies change. We are essentially approaching this as a Consequentialist. We are seeing if the harms caused by masking

policies are worth it by seeing if they impacted death rates. We also wanted to keep the results easy to understand not just in what they are, but also in how we find them so readers can easily judge for themselves if our results make sense. This was through our goal of being human-centered.

To check how Covid deaths and masking mandates related to crimes committed in El Paso County we used crime data from the Colorado Crime Statistics. Colorado Crime Statistics is funded by the Colorado Automobile Theft Prevention Authority, which is a part of the Colorado State Patrol. The data source allowed me to specify the number of offenses committed, restricted to El Paso County, with the data given per day. This way I was able to compare the number of offenses to the Covid and masking mandate day to day from the date ranges of January 22, 2020 to October 15, 2021.

The hypothesis we are exploring here is does the number of Covid related deaths and whether a masking mandate was in effect impact the number of offenses that occured in the El Paso County, with a null hypothesis is that there is not a relation between covid deaths and masking mandates and the number of incidents that occur. We decided to use the standard significance value of .05. Since we are predicting a numeric response variable with numeric and categorical variables we decided to use multiple linear regression. Just as with checking if masking mandates impacted covid death rates, we picked the model method here with human-centered front and center. Multiple linear regression is an easy to understand model that will allow for most of our readers to easily do sanity checks on our results and understand why we made the conclusions we did.

Findings

Figure 1 is the result of plotting the change points in the number of Covid deaths in El Paso County against if a masking mandate was in effect. One change point is in early March, when Covid started being reported, and another is October 15th, when the data ends. The middle two change points occur on roughly November 1st, 2020 and January 1st, 2021.

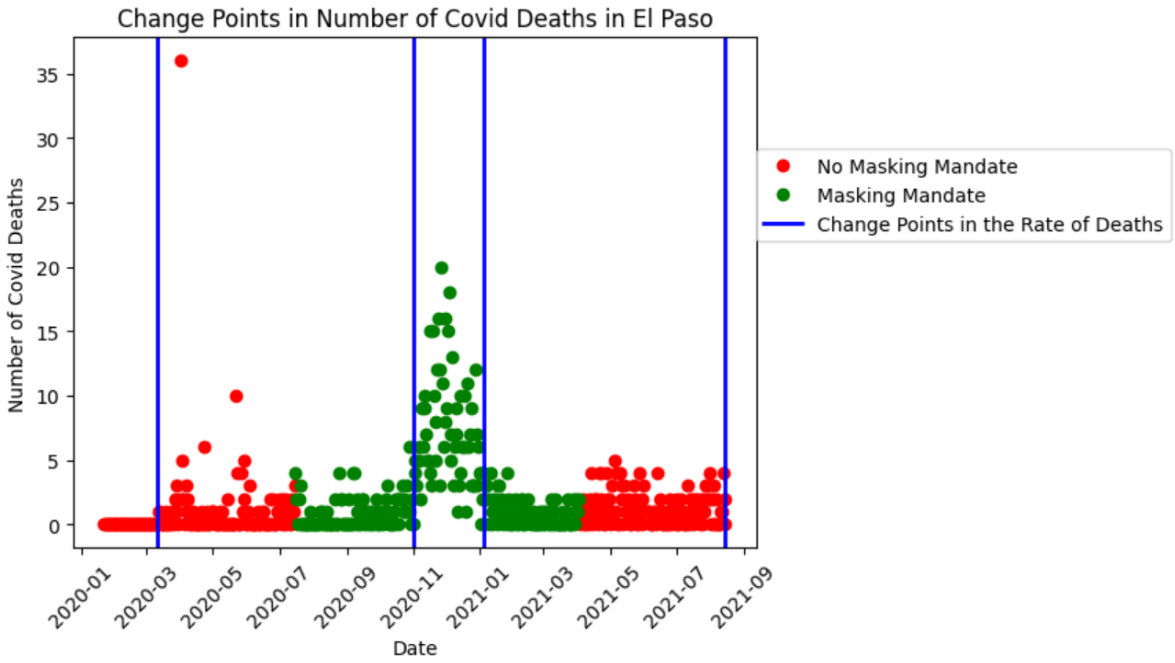

Figure 1: Time series of the number of Covid deaths, with bars denoting change points where there is a significant change in the rate of Covid deaths

The results for the multiple linear regression analysis on offsenses in El Paso County were not incredible. We found significant p-values for both covid deaths ($p = .019$) and masking mandates ($p < .001$). However the R-squared value was only .136 and we had a skew value of 8.248, indicating significant skew in the model.

When we look at the residuals we can see that they are not normally distributed and do not have constant variance. This violates several assumptions necessary for multiple linear regression.

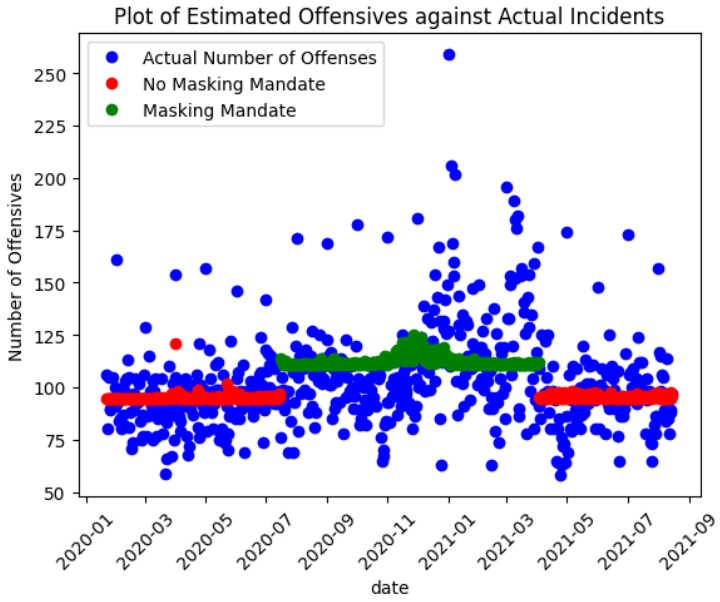

Figure 2: Plot of the predicted number offenses committed against the actual values

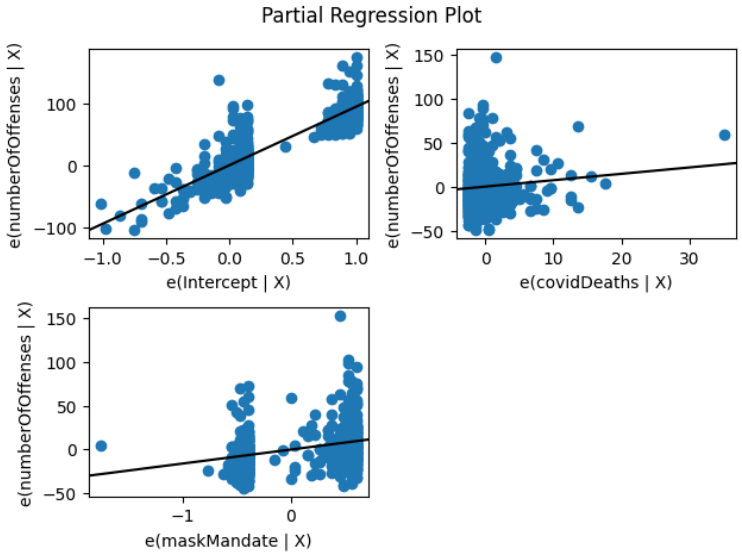

Figure 3: Plot of the residuals from the multiple linear regression model.

Discussion and Implications

The change point research has an interesting implication about the effectiveness of mask mandates in El Paso County. The time it takes from when someone becomes sick from Covid to showing symptoms ranges a bit, but on average it takes about 5.6 days (WebMD, 2020). We would then expect that, if the masking mandates were effective, around one week after being implemented there would be a significant shift in infection rates, which would then lead to a fall in Covid deaths. What we see instead is that the first significant change point is roughly four months after the implementation of the masking mandate, and it is due to a spike up in Covid deaths. The next change point is roughly two months later, and is when the masking mandate is still in effect. It is interesting that there is no change point close to where masking mandates are enacted or removed.

I think this can be explained by the fact that the infection rates of covid were not constant. This is hinted at the fact that there is an explosion in covid deaths months after the masking mandate was put in place, and happens to coincide when the Omicron variant began spreading in the US. I think this points to the fact that further research is necessary to completely understand when and why change points occur.

The results for the multiple linear regression were also not incredible. The significant p-values for covid deaths and mask mandate variables are undermined by assumptions for constant variance and being normally distributed. On top of that the model has a very low R-squared value at .136. This makes sense when comparing the predicted values against values. It almost appears to plot the average number of offenses. I think this points to the fact that most, if not all of the variables are heavily

skewed. I think that the model can be improved on by transforming the variables to be more normally distributed.

Human centered data science had a heavy influence on how we conducted our research. Our goal in every step was to have a model that was both understandable and explainable. We did not want a model that performed well but we didn't understand or couldn't explain. We also focused on using general and collective data about a broad geographic area, without identifying information about people. Lastly, the point of this research is to better improve our response to future pandemics. It is for these reasons we feel that our research was human centered.

Limitations

There are a number of limitations with our research. It only applies to El Paso County from the dates of January 22, 2020 to October 15, 2021. The change point detection research implicitly assumes that the infection rates for covid are constant, and that all masking mandates are the same. Neither of these are true; different Covid variants behave differently and there are a wide range of masking mandates with a wide range of requirements. The multiple linear regression model has a number of key assumptions not held, making the results not reliable. On top of that, the model has a very low R-squared value, meaning it does a poor job actually modeling the data. For all of these reasons our results and conclusions that can be drawn from them are extremely limited and unreliable.

Conclusion

For the entire project our goal was to keep it human centered. To do so we focused on easily understood and explainable models, data that protected people's privacy, and our goal with the research was to improve future responses to pandemics.

Our research question came in two parts. The first was what was the impact of masking mandates on Covid deaths in El Paso County. The second was the number of Covid related deaths and whether a masking mandate was in effect impacted the number of offenses that occured in El Paso County. Our hypothesis for the first part was that there would be a change point around when masking mandates are enacted and removed, and our null hypothesis for the second part is that there is no relation between covid deaths and masking mandates.

For the first part we found that change points appeared to appear regardless of when the masking mandates situation changed. However we feel that our model makes a number of untrue assumptions that make possible conclusions here unreliable. On top of that our model for the second part violated several key assumptions. Unfortunately this makes possible conclusions from the model unreliable as well.

References

Huang, J., Fisher, B. T., Tam, V., Wang, Z., Song, L., Shi, J., La Rochelle, C., Wang, X., Morris, J. S., Coffin, S. E., & Rubin, D. M. (2022). The effectiveness of government masking mandates on COVID-19 county-level case incidence across the United States, 2020. *Health Affairs*, *41*(3), 445–453. https://doi.org/10.1377/hlthaff.2021.01072

Maxine. (2022, March 3). *Impact report: Covid-19 and crime*. Council on Criminal Justice. Retrieved December 11, 2022, from https://counciloncj.org/impact-report-covid-19-and-crime/

Mekouar, D. (2022, February 2). *Why homicide rates in US spiked 30% during COVID*

pandemic. VOA. Retrieved December 11, 2022, from https://www.voanews.com/a/why-homicide-rates-spiked-30-during-the-pandemic-/6420391.html

WebMD. (2020, February 24). *Coronavirus incubation period: How long and when most contagious*. WebMD. Retrieved December 12, 2022, from https://www.webmd.com/lung/coronavirus-incubation-period

Data Sources

John Hopkins Covid-19 Data - https://www.kaggle.com/datasets/antgoldbloom/covid19-data-from-john-hopkins-university

CDC Masking Data - https://data.cdc.gov/Policy-Surveillance/U-S-State-and-Territorial-Public-Mask-Mandates-Fro/62d6-pm5i

El Paso County Crime Data - https://coloradocrimestats.state.co.us/tops