Measuring Causality between Mask & Closure policies and COVID-19 deaths at the state level

Dmitry Baron, Leena Bhai, Daniel Ryu

Importance and Context

The COVID-19 pandemic in 2020 presented the United States with an unprecedented set of challenges. One of the toughest: each state needed to take decisive action to contain the virus and save lives, while minimizing impact on its residents socially, emotionally, and financially. This dilemma underscored the cultural differences between states and their leadership - while most states implemented some combination of face mask requirements and business closures, the duration of their implementation and severity of their enforcement varied. As waves of increased COVID cases came and went, different states removed and re-implemented their policies, while others remained conservative. The cost of these measures and death toll of the virus alike caused great harm on the residents of every state, as criticisms abounded about each state governments' responses (or lack thereof).

Before clinically approved vaccines are widely available, there was no better way to protect the residents from SARS-CoV-2 than personal preventive behaviors such as social distancing and wearing masks, and public health measures, including active testing, case tracing and restrictions on social *gatherings*[1]. The states issued guidance and orders to temporarily close businesses where social distancing was not possible. The business where social distancing is generally difficult by design are restaurants, indoor group fitness places, movie theatres, bars to name a few.

Government mandated business closures were broad and exceptions were granted to operate certain low contact business activities and essential business. This study focuses on describing the relationship between a state's public policies - public mask mandates and business closures - and COVID-related death rates. We narrowed down to firms in exposed industries like restaurants, gyms, and movie theatres. On reviewing existing literature on impact of COVID related closures on business, we found that restaurant industry was particularly negatively impacted. The results from a survey conducted in April 2020, published by the United States (U.S.) National Bureau of Economic Research (NBER), indicated that the restaurant industry is likely susceptible to a long crisis. Restaurant operators, who responded to the NBER survey believe that they have a 72% chance of survival if the crisis lasts one-month, but if the crisis lasts four-months, then they give themselves only a 30% chance of survival. If the crisis lasts for six-months, the expectations of survival shrink to 15% (Bartik et al., 2020). In the light of this research, we narrowed our reserach Question to ** "Did public policy of mask mandates and closing restaurants for dine-in business cause a decrease in COVID related death rates at the state level?" **

We define death rates as COVID-related deaths (as per the New York Times public dataset) over total recorded population (as per the US Census Bureau, 2019) per state. We define public policy of mask mandate as an order mandated at the state level to use mask at public places. The public policy to close restaurants is defined as a public order to restrict dining on premise of restaurants.

The study takes into account the fifty states as well as District of Columbia (51 total). The dates included are from March 1st (the first COVID death in the US) to December 14th (the first COVID vaccines administered in the US), or "exploration period". For the remainer of the paper, we define "public policies" for the purposes of this study as a combination of public policy of mask mandate, and the public policy to restrict restaurants.

Description of Data

Restaurant Closures

To operationalize the public policy to close restaurants for dine-in business, we first looked at closure dates. We noticed that different states closed on different days. The order applied to entire state. We then looked at the date when states reopened restaurants (for indoor and/or outdoor dining) statewide for the first time. We calculated the difference between the reopening and closing dates to calculate the number of days for which the restaurants remained closed. We understand when states issued reopening orders, the states also limited capacities of the restaurant.

The states closed and reopened their restaurants in waves. To calculate the total days of restaurant closure, we added up the difference between reopening and closure dates of each set of closure orders by state. All the states except South Dakota closed and reopened the restaurants in the first set of closures. For second closure, one state closed after our exploration end date, so we imputed it to bring it to the exploration end date. In the second reopening, two states reopened after the study end date, so we imputed them to our exploration end date. In the third set of closures, only one state started and closed before our exploration end date, so we did not do any modification there. The below chart shows the timeline of closures and reopening and our need to impute some of the dates.

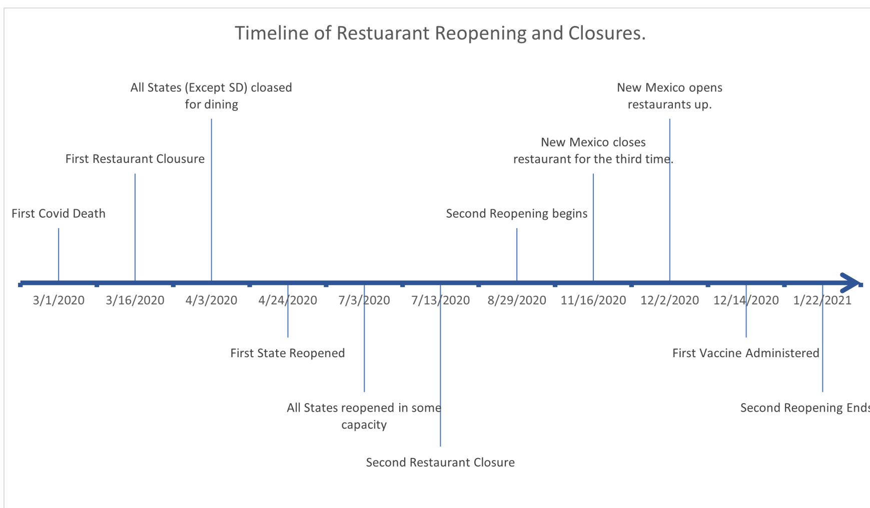


Figure 1: Timeline of Restaurant Reopenings and Closures

The following USA map shows that the number of days that the restaurants remained closed by state.

From the above map we can see that restaurants remained closed in some states a lot longer than in others at the state level. We quickly notice that the highest number of days for the closure was 131 days and lowest number of days was 0. Washington state closed by 109 days, while California closed its restaurants for 63 days for dine-in services. We also notice there is no clear pattern geographically in the days of closures.

On investigation we find, some states that removed statewide mandates created a blueprint for reopening the state by creating a county-wise reopening guide dependent on the number of infections in that county. This dataset does not contain information about such county specific closures.

Next, let us look at the histogram of the number of days that the restaurants remained closed.

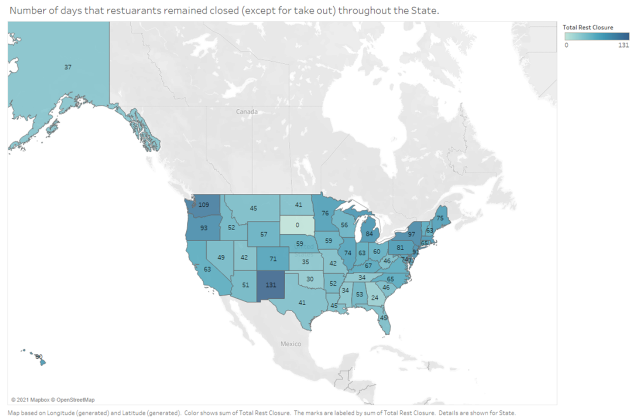


Figure 2: Number of Days Restaurants Remained Closed by State

The above histogram gives us an idea of the distribution of the variable. Most states fall between 25 to 90 days. Two states closed their restaurants for less than 25 days and two states closed their restaurants for more than 100 days. The spread is critical to the survival of restaurant industry as mentioned in the motivation section.

Masks

For the face mask state policies we focused on the public face mask mandate only. We selected the public mandate only (as opposed to Business) as Public face mandate is general and had variance: 38 out of 50 states and Washington DC, vs. 46 states, which have implemented the business mask mandate.

Focus on the public mask mandate allowed for more differentiation between the states. The first public mask mandate was introduced by the state of NJ on 4/08/2020, while the last one - by NH on 11/20/2020. The majority of public mask policy introduction falls at the period of approximately mid-June to early August of 2020.

We calculated the following variables associated with the public face policy: number of days it took to implement the mask policy since the March 1, 2020 and number of days during which the mask mandate was effective. The mean value for the number of days it took the states to implement the Public mask policy is 83.25. The mean value for the number of days during which the mask mandate was effective is 131.3 days.

We also attempted to assess seriousness of the states to implement the public mask policies by looking at the number of states, which had fines associated with not complying with the mask policy (18 states or approx. 35%) and any associated criminal charges or citations (14 states or approx. 27.5%). We used these variable as binary variables.

After the data exploration stage, we decided to use only the binary variable if a state had the public mask mandate ("TRUE") or did not have one ("FALSE").

During the initial data exploration we didn't find any significant difference in death rate in the states with mask mandate vs. the states without the mask mandate.

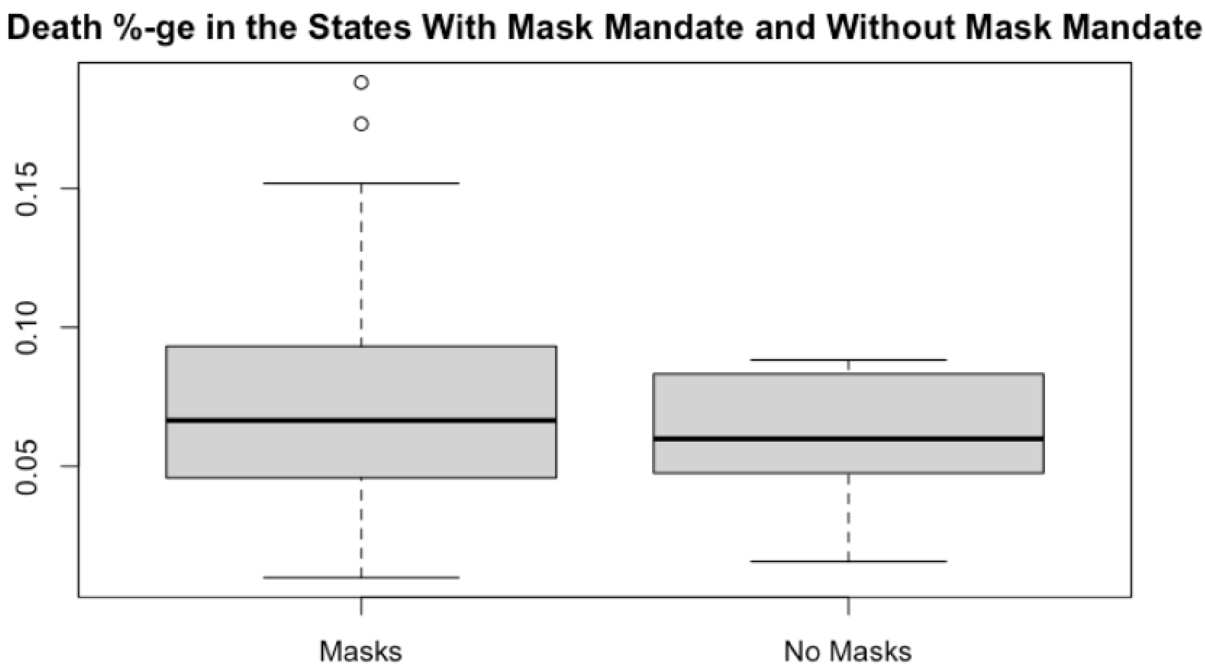


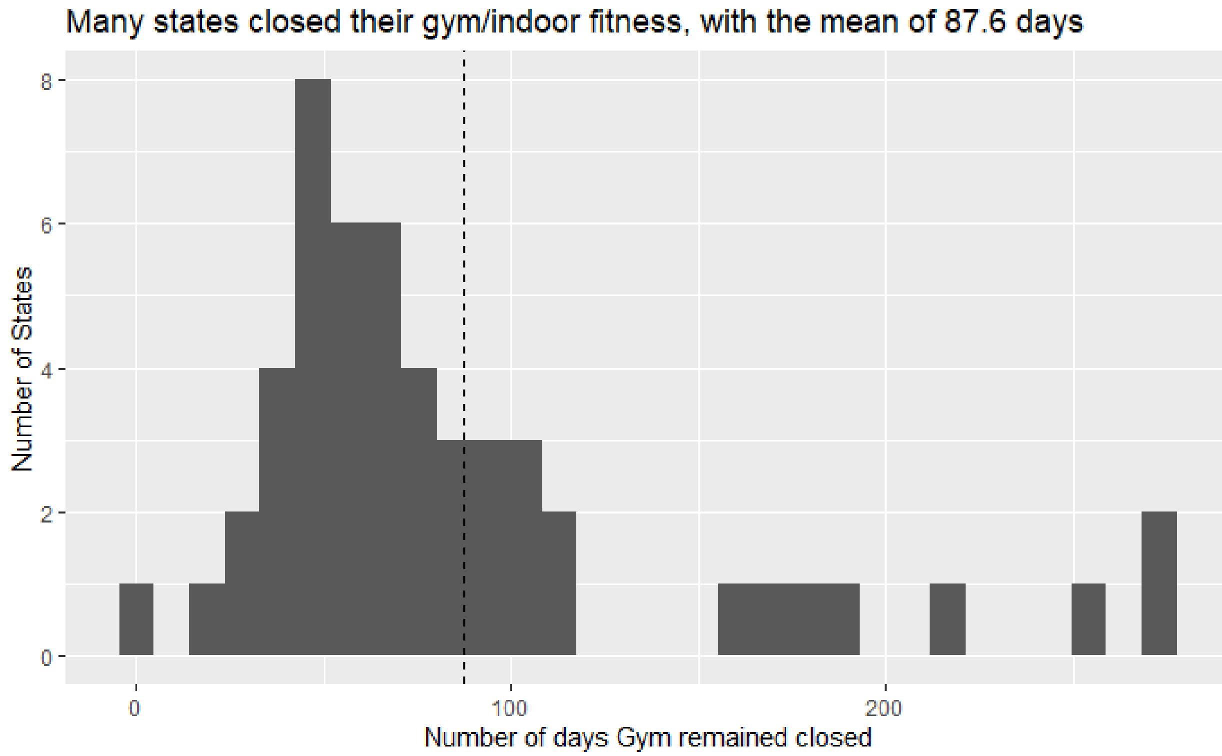
Figure 3: Boxplot of death rate in states with and without the public face mask mandate

Through the box plot we see that the states with no mask mandate have lower death percentage than in the states with mask mandate. There are two states with mask mandates that are outliers. 50% of the states with mask mandates have death rates between 0.05% and 0.1%.

Gym Closures

We also looked at Gym Closures in the third model (try everything approach) to see if it has any effect on death rates. Just like restaurants, all states except for SD closed their gyms and indoor fitness activities starting from 16th of March, 2020 to 3rd of April, 2020. Some states started reopening their indoor gyms/fitness centers statewide at April 24, 2020. 47 States reopened their gym statewide, while 3 states did not and their gyms remained closed for the remainder of the year. In these cases, their reopening date was imputed as end of the study date, so that we can calculate the total days for which the gyms and indoor fitness remained closed by state. Pennsylvania and Rhode Island closed their gyms on 12/12/2020 and 11/30/2020 and reopened on 1/4/2021 and 12/21/2020, respectively. Since the reopening dates are after our study end date, we imputed the reopening dates to the study end date.

Let us look at the histogram of the number of days that the gyms/indoor fitness remained closed.



The above histogram gives us an idea of the distribution of this variable. The above histogram shows us a stark contrast in the response of the states. Some states were stricter than others in keeping their gyms closed. The above histogram has a positive skew or right skew which indicates than mode and median of the histogram are greater than the mean of 87.6 days.

A Model Building Process

To iterate, our research Question is "Did public policy of mask mandates and closing restaurants for dine-in business cause a decrease in COVID related death rates at the state level?"

Our Null Hypothesis is that "Public policy of mask mandates and closing restaurants for dine-in business" have no effect on "COVID related death rates" at the state level.

Independent covariates are:

1. Public face mask mandate in effect (True) or not in effect (False)
2. Number of days for which the restaurants remained closed (except takeout) during the exploration period
3. Number of days for which the gym and indoor fitness remained closed during the exploration period

Control Variables are:

1. Total population by state
2. Proportion of people aged 65 and older. Based on our EDA, the only variable, which we decided to be worth log-transforming was the total state population.

Dependent variable is: The depended or outcome variable is the Death Rate, which is measures as the cumulative number of the deaths per state as of 12/14/2020 divided by the state population.

We have identified three models from our EDA:

- Model 1: Limited Model
  This model include only a few key variables such as Death Rate as caused by the public mask mandate (True or False) + Number of days for which the restaurants remained closed during the exploration period.

$$DeathsPerCapita = \beta_0 + \beta_1 MaskMandate + \beta_2 RestaurantClosureDays$$

β_0 will be the states with no mask mandate and 0 days of restaurant closures. β_1 will be decrease in the number of deaths if the mask mandate is in place. β_3 will be resulting decrease in deaths due to increase in restaurant closure by 1 day.

```
Call:
lm(formula = deathsPerCapita * 1e+05 ~ mask_mandate + TotalRestClosure,
    data = StateDeathsClosureMasks)

Residuals:
    Min      1Q  Median      3Q     Max
-70.670 -30.144  -3.051  15.240 112.343

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)      84.79670   18.23000   4.651 2.61e-05 ***
mask_mandateTRUE -3.04541   17.59334  -0.173    0.863
TotalRestClosure  0.06621    0.31183   0.212    0.833
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 43.55 on 48 degrees of freedom
Multiple R-squared:  0.00107,	Adjusted R-squared:  -0.04055
F-statistic: 0.02572 on 2 and 48 DF,  p-value: 0.9746
```

Figure 4: Model1

I have mulitiplied the deaths per capita by 100K to make the numbers interpretable. Model1 does not do a good job of explaining the variances in death. We notice that multiple R-squared is close to 0 and adjusted R-squared is negative. The model is not useful and does not explain anything. If mask mandate is implemented, it would result in 3 less deaths. Similarly increasing the number of days that restaurants remaned closed, it would increase deaths by 0.06621. Both the mask mandate and restaurant closures are not statistically significant, so we cannot reliably interpret the coefficients.

- Model 2: More Expanded Model

$$DeathsPerCapita = \beta_0 + \beta_1 MaskMandate + \beta_2 RestaurantClosureDays + \beta_3 StatePopulation$$

```
Call:
lm(formula = deathsPerCapita * 1e+05 ~ mask_mandate + TotalRestClosure +
    log(total_population), data = StateDeathsClosureMasks)

Residuals:
    Min      1Q  Median      3Q     Max
-58.799 -31.509  -5.315  16.561 107.687

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)           -34.52913   91.50656  -0.377    0.708
mask_mandateTRUE        1.17676   17.74022   0.066    0.947
TotalRestClosure       -0.02518    0.31690  -0.079    0.937
log(total_population)   8.00498    6.01763   1.330    0.190

Residual standard error: 43.21 on 47 degrees of freedom
Multiple R-squared:  0.03732,   Adjusted R-squared:  -0.02413
F-statistic: 0.6073 on 3 and 47 DF,  p-value: 0.6136
```

Figure 5: Model2

Model2 is very slightly better than model1 with R-squared of 0.0373. Adjusted R-squared still remains negative. β_0 will be the number of deaths for a state with no population, no mask mandate and no restuarant closures. Such a state will have -34 deaths. β_1 will be incremental deaths with mask mandate. β_2 will be decrease in deaths for increase in each day of restaurant closures. We added state population as the control variable. Increasing state population by 1% provides an additional 8 deaths on average. All the features are not statistically significant, so we cannot reliabily interpret the coefficients.

- Model 3: The Most Inclusive Model

This model includes all of the above variables. We added to this model the proportion of people aged 65 and older, who are most vulnerable to the COVID19. We also added another business gyms, which were closed due to COVID restrictions of the state as a kitchen sink approach to see if that helps.

$$DeathsPerCapita = \beta_0 + \beta_1 MaskMandate + \beta_2 RestaurantClosureDays + \beta_3 StatePopulation + \beta_4 SixtyFivePlusPercentag$$

Model3 is slightly better than model2 with R-squared of 0.1204. Adjusted R-squared is still close to zero. β_0 will be the number of deaths for a state with no features. Such a state will have -89 deaths. β_1 will be incremental deaths with mask mandate. β_2 is decrease in deaths for increase in each day of restaurant closures. We added total population and sixty_five_plus population as the control variables. Increasing state population by 1% provides an additional 8 deaths on average. However since none of the coeffecients are significant, it is not reliable to interpret the coefficients. Some of the casual relation between gym closure days and restuarnat closure days may be absorbed due to a relationship between the two.

for a cluster of states. there is a linear relationship between number of days that restaurants and gyms remained closed.

```
Call:
lm(formula = deathsPerCapita * 1e+05 ~ mask_mandate + TotalRestClosure +
    log(total_population) + sixtyfive_plus_percentage + TotalGymClosure,
    data = StateDeathsClosureMasks)

Residuals:
    Min      1Q  Median      3Q     Max
-62.913 -29.311  -3.433  18.229 105.901

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               -14.15178  120.36201  -0.118    0.907
mask_mandateTRUE            3.08489   18.84466   0.164    0.871
TotalRestClosure           -0.15325    0.40725  -0.376    0.708
log(total_population)       6.69277    6.57493   1.018    0.314
sixtyfive_plus_percentage  -0.06772    3.28113  -0.021    0.984
TotalGymClosure             0.07855    0.13086   0.600    0.551

Residual standard error: 43.98 on 45 degrees of freedom
Multiple R-squared:  0.04518,   Adjusted R-squared:  -0.06091
F-statistic: 0.4258 on 5 and 45 DF,  p-value: 0.8282
```

Figure 6: Model3

Most appropriate test

The relationship between states' public policies and COVID deaths is summarized by this causal graphical model:

This study focuses on how states can prevent transmission, which is a couple steps removed from COVID deaths; our focus is on deaths because this is ultimately the goal in a public health emergency.

There are other causes that prevent deaths by means of preventing transmissions or helping recovery, but these are not studied - these are discussed in *Omitted Variables*.

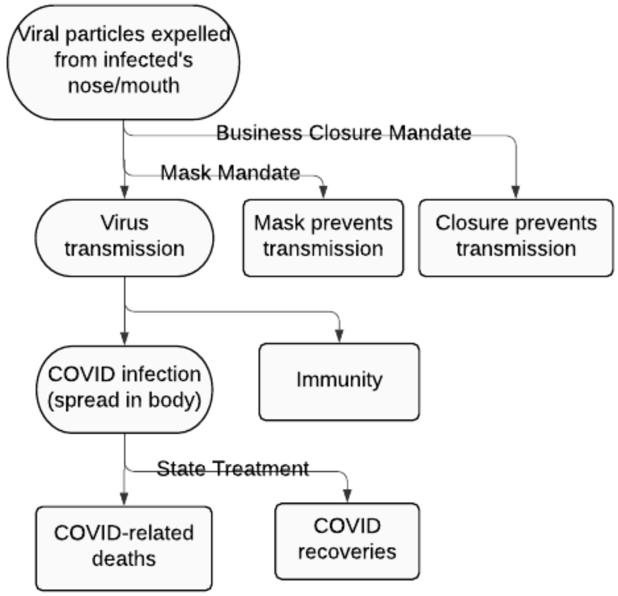


Figure 7: Graphical Model: causation of COVID deaths

Regression Table

	Dependent variable:		
	deathsPerCapita * 1e+05		
	(1)	(2)	(3)
mask_mandate	-3.045	1.177	3.085
	(17.593)	(17.740)	(18.845)
TotalRestClosure	0.066	-0.025	-0.153
	(0.312)	(0.317)	(0.407)
log(total_population)		8.005	6.693
		(6.018)	(6.575)
sixtyfive_plus_percentage			-0.068
			(3.281)
TotalGymClosure			0.079
			(0.131)
Constant	84.797***	-34.529	-14.152
	(18.230)	(91.507)	(120.362)
Observations	51	51	51
R2	0.001	0.037	0.045
Adjusted R2	-0.041	-0.024	-0.061
Residual Std. Error	43.554 (df = 48)	43.209 (df = 47)	43.978 (df = 45)
F Statistic	0.026 (df = 2; 48)	0.607 (df = 3; 47)	0.426 (df = 5; 45)
Note:			*p<0.1; **p<0.05; ***p<0.01

The above table compares the three models. All the three models do not have any significant features. So it difficult to interpret any coefficients. The three models explain very little of what is going on in the covid world. We are unable to explain the variances in death rate by using the independent variables.

We cannot talk about practical significance as none of the factors have a statistically Significant effect. It is interesting to see that the factors that caused economic pain like restuarant closures and civic debates like masks do not have any effect on COVID related death rates in state. We would examine some omitted variables and model issues in the subsequent sections.

Results

]We conclude that there is no effect of “Public policy of mask mandates and closing restaurants for dine-in business” on “COVID related death rates” at the state level

Test Assumptions for the Classical Linear Model

We test assumptions on Model2.

One risky assumption this model relies on is that the mask and closure policies of one state are independent from the next. This challenges our I.I.D assumption of the data. During the beginning of the pandemic in the United States, there were likely inter-state effects as COVID cases and panic alike spread across the country. There could be a clustering effect between start dates and end dates of restrictions, potentially reducing standard errors of the model. However, we believe that these effects are partially neutralized by the complexity of the relationship between increasing cases and implementing restrictions. The same per-capita increase of cases would potentially incite a different response by different state governments; different levels of case increases at different times must have incited even more varied responses.

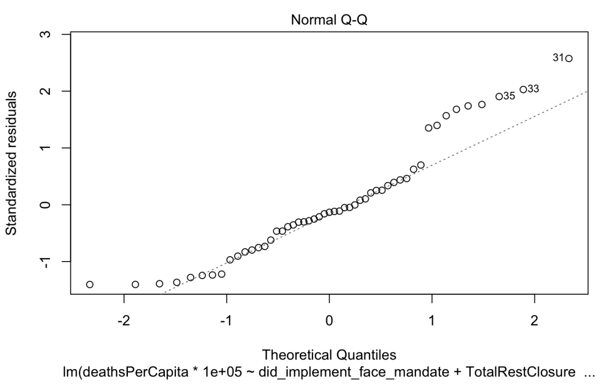
The collinearity within our main model (model 2) is quantified by the following VIF values:

Table 1: Low VIF values for second model

did_implement_face_mandate	TotalRestClosure	population_density)
1.355	1.376	1.056

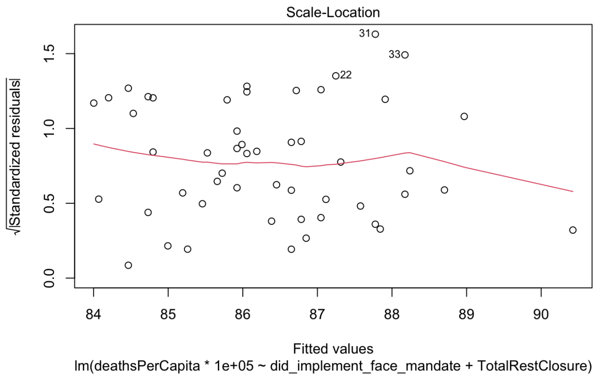
These VIF values around 1.3 indicate a low level of collinearity, which we find acceptable; VIF values around 10 or greater would indicate the coefficients underrepresent the effect of closure lengths on the COVID-related death rate in our model.

The assumption of normally distributed errors presents some major issues. The Q-Q plot reveals a right skew and heavy tails, with residuals past the first quantile diverging from the normal distribution:



However, to assess if the skewness of residuals is significantly different from a normal distribution, we employed the Shapiro-Wilk test which produced a p-value of 0.008, successfully rejecting the null hypothesis of normality. We see that there is strong evidence of non-normal residuals. In most situations, a transformation of the dependent variable would be appropriate; however, because the dependent variable is already a rate, an additional log transformation would further obscure the practicality of this model. We should examine our model more skeptically considering there may be an uncaptured relationship.

Heteroskedasticity is the final potentially problematic assumption. We assess the spread of the standardized residuals across fitted values in the Scale-Location plot:



The small quantity of points makes it difficult to distinguish signal from noise in assessing homoskedasticity. The variance of the residuals seems to increase with the value of the fitted death rate variable, dropping due to the outlier to the right of the graph. To assess the potential of heteroskedasticity with a statistical test, we employ the studentized Breusch-Pagan test, which generates a p-value of 0.34 - insufficient to reject the null hypothesis of homoskedasticity. We fail to find evidence of a linear relationship between death rate and the error variance; due to the small sample size of 51 "states", the lack of a linear relationship is sufficient considering practical constraints to make the assumption of homoskedasticity.

Test Limitations

Several limitations of the model design also may have affected this study.

Perhaps the greatest impact is the limitation of granularity. The only available data within the scope of the study was aggregated at the state level. Certainly governors issued mandates statewide, but individual counties often enforced stricter supplementary mandates or were much less stringent than at the state level. We consider larger and more heterogenous states such as California, where political attitudes and socioeconomic environments vary greatly; summarizing both policies and deaths statewide surely ignores the many microcosms within the state.

An additional limitation is the inability of broad terms to capture the variety of applications between states. For example, the New York Times database, from which we pull our death counts per state, delineates that on April 5th the Council of State and Territorial Epidemiologists advised states to report additional "probable case" counts based on criteria for symptoms and exposure. Some states took this recommendation and reported probable cases, which are reflected in the death rate in our model; others continued to report only confirmed cases. Likewise, different states applied and interpreted mask and closure policies differently; for example, New York allowed restaurants to set up partitioned tents to seat customers outdoors amidst winter conditions; a warmer state like California wouldn't have interpreted an edge case so generously. The model is based on variable definitions that cannot account for a broad range of outcomes - the design generalizes inefficiently in this regard.

The third limitation in the design is the measurement of policy effectiveness with a static date range. The model design relies on comparing the aggregated application of policies with the aggregated number of deaths; the broad stroke of this method unfortunately ignores the many time-based phenomena that occurred at certain ranges within the year. Our study also measures number of days relative to a fixed range of dates for all states; we logically assert the first COVID death in the country should have incited all state governments to consider action, but considering the diverse vulnerabilities of each state, a more specific reference date from which to count may have been more precise to quantify how long a state had a policy in place.

Omitted Variables

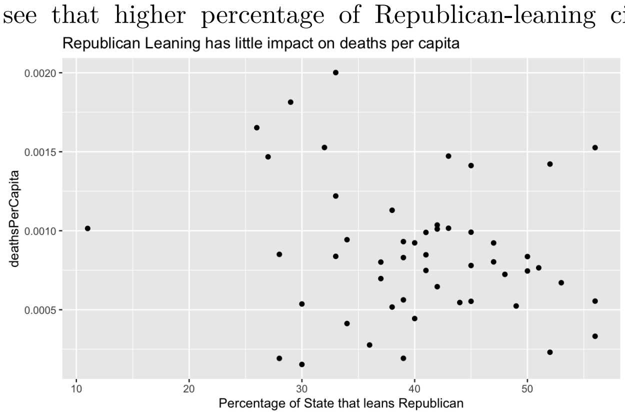
As discussed in *Most Appropriate Test*, this model does not capture all possible causes at play in this situation. The relative lack of explanatory power of the model can be attributed to an inability to access several potential causes. We've identified a few omitted variables that may have causal effect on our covariates as well as per-capita COVID death rates:

- *Mask/Closure adherence*: The regression measures the causality between the policies and COVID deaths, however the adherence of a state's citizens to the policies is omitted. Because adherence should have a positive relationship with policy length/punishment and a negative relationship with COVID deaths per capita, the bias of omitting adherence should be away from zero. The effect of mask/closure policy adherence is partially captured by placing the policies, resulting in a possible overestimation of their effect. Because of the omission of this variable, we recommend caution in interpreting the causality of our model.

- *First mover disadvantage*: Certain states were quicker to implement mask & closure mandates. Some of these states also experienced a higher rate of COVID deaths earlier in the time range. Our model assumes each state's closure and mask policies are equal, however later implemented policies will have been slightly more effective in increasing adherence and reducing deaths. This omitted "learnings" variable has a positive correlation with the number of days taken to implement mask/closure policies and a negative correlation with COVID death rates, therefore the direction of the overall omitted variable bias should be towards zero, which adds credibility to our model coefficients' significance. The closest proxy variable available is how many days after the first mask/closure mandate a state enacts its own mask/closure mandate, but this is insufficient to account for how effective this learning period was.

- *Amount of inbound and intrastate travel*: COVID-19 did not start in the U.S., but was brought in by international travelers. Likewise, states vary in terms of how many travelers enter from other states, and travel among counties within the state. The mobility of citizens and travelers within a state has a positive correlation with COVID death rate, and should have a positive correlation with the length and quickness of mask/closure policies (since higher mobility states will be more cautious), therefore we would expect the omitted mobility variable to bias coefficients away from zero.

- *Political affiliation (Republican)*: Governors of more conservative-leaning states were politically incentivized to more conservatively apply mask and closure mandates, due to the party's emphasis on less-regulative government. One might also expect a more conservative state to also have higher death rates, resulting in an omitted variable bias toward zero; however, on inspection we see that higher percentage of Republican-leaning citizens and death rates have little correlation.

  Since the effect of political affiliation on COVID deaths per capita is small, we can also expect any bias to be insignificant, although away from zero.

- *COVID case treatment capability*: One important variable omitted from the model is a state's ability to prevent COVID cases from becoming deaths. Whether this be hospital capacity, hospital availability, or the citizens' attitude toward going to the hospital, many factors contribute to a negative correlation between case treatment and COVID deaths. We also expect that better case treatment has a negative correlation with mask/closure policies, since the state is more willing to take risks. Therefore, the omission of case treatment ability should bias the model away from zero, since the modeled effect of implementing mask/closure mandates captures part of the case treatment capability.

Reverse Causality

One additional issue that arises is that of reverse causality: would states experiencing more deaths result in extending its public mask and/or closure policies? The decision-making of state leadership during this time was exceedingly complex and confusingly opaque, perhaps even to the leadership themselves. Our model is based on an absence of this factor; however, we believe this reverse causality played a big part in the extension or termination of these policies per state. Because we expect negative causality from policies to deaths, and positive causality from deaths to policies, we believe that the coefficients understate the influence of policies on deaths, since these two causal relationships are at odds with one another. In addition to the aforementioned omitted variables, this reverse causality phenomenon could affect the lack of explanatory power in our model.

Conclusion

The purpose of this study was to examine the relationship between a state's public mask mandate and business closures and COVID-related death rates. Intuitively, many would assume that states that implemented longer mandates prevented a greater per-capita rate of COVID-related deaths. However, the results of this study indicate that we cannot reject the null hypothesis - there are no proven causal relationships between a state's public mask & closure policies, and the COVID-related death rate between March 1st and December 14th.

Tellingly, the model also doesn't indicate that high percentages of 65 plus citizens or high population densities cause higher death rates. The lack of a relationship between these control variables and death rates reveals that there may have been substantive design issues in this study, apart from a lack of effectiveness of state policies. We identified and explored several omitted variables that may have affected our model, as well as potential reverse causality, with higher deaths causing longer policy duration. More broadly, the lack of granularity in the state-level data causes major issues in terms of small sample size and a loss of information regarding the diverse applications and reactions to policies within the different counties of a state.

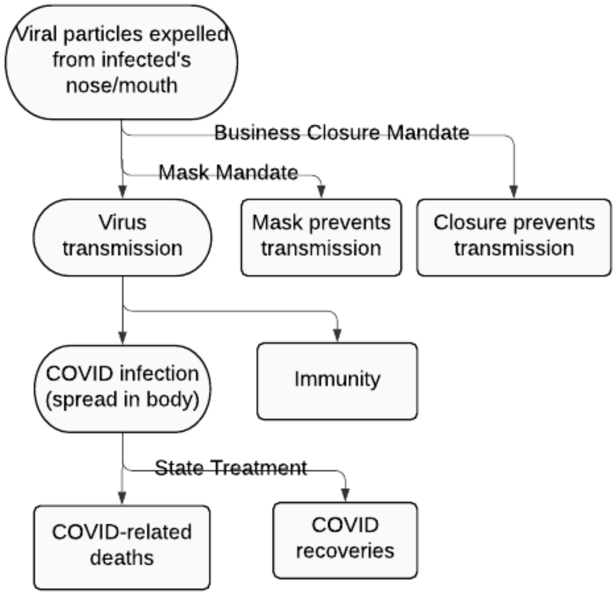
An additional reason for the lack of explanatory power could be the high number of steps between the implementation of a public policy and the COVID deaths it prevents; we studied the correlation between policies and infection rates rather than death rates, and found more evidence of a relationship (see addendum). As a whole, future researchers would do well to consider how to better capture these effects in future models. The complexity of the COVID situation is challenging to disentangle, but with careful reasoning and future availability of data unavailable during this study, subsequent studies will progressively resolve these issues and more successfully distinguish the effectiveness of the public mask and business closure policies in protecting American citizens from future pandemics.

References:

1 https://www.nature.com/articles/s41579-020-00459-7
2. https://www.sciencedirect.com/science/article/pii/S0278431920303182

Addendum-1

Our proposed graphical model traces causality from the mask and business closure policies of a state, to its eventual COVID death rates:



However, the high number of results in the chain between policies and deaths may explain the lack of relationship between policy and deaths per capita in the data. The more events between the modeled cause and effect, the more opportunities there are for unforeseen omitted variables or other factors to obscure the true relationship. In reviewing the results of our model, our team changed the dependent variable of our model from deaths per capita to infections per capita, to see if the effect of policies could be seen in this "shorter" causal model.

We re-examine restaurant closures and mask policy with COVID infection rates rather than death rates:

Table 2: Significant relationship between restaurant closure length and infection rate

	Dependent variable:
	infectionsPerCapita $*1e+05$
did_implement_face_mandate	-159.728
	(734.087)
TotalRestClosure	-43.946^{***}
	(13.011)
Constant	$8,143.663^{***}$
	(760.652)
Observations	51
R^2	0.250
Adjusted R^2	0.218
Residual Std. Error	1,817.309 (df = 48)
F Statistic	7.983^{***} (df = 2; 48)
Note:	*p<0.1; **p<0.05; ***p<0.01

We see a significant relationship between restaurant closures and infections per capita: for one additional day a state had restaurants closed, there was an additional 0.04% decrease in infections per capita between March 1st and December 14th.

Then, we examined if in line with intuition, there is a strong correlation between COVID infections and COVID deaths:

Table 3: Medium correlation between infections and death rates

	Dependent variable:
	deathsPerCapita
infectionsPerCapita	0.008^{***}
	(0.003)
Constant	0.0004^{***}
	(0.0002)
Observations	51
R^2	0.154
Adjusted R^2	0.137
Residual Std. Error	0.0004 (df = 49)
F Statistic	8.936^{***} (df = 1; 49)
Note:	*p<0.1; **p<0.05; ***p<0.01

Surprisingly, the relationship between COVID infection and death rates is not as strong as we would expect, with only an R^2 value of 15.4%. However, on second thought, this is in line with what we understand scientifically about the disease: mortality rates are low, with asymptomatic or mild symptomed carriers comprising the majority of infections.

Considering the variation from the model between restaurant closures and infections, and that between infections and deaths, the explanatory power of a model between various public policies and deaths may be quite affected by random variation alone, not to mention the various omitted variables and reverse causality discussed in *Test Limitations*. We advise future research to focus on infection rates, in order to offset against these other confounding factors as much as possible.