

Impacts of Stay at Home Orders on Individual Mobility

Elaine Chang

Dom Dillingham

Jesse Miller

Michael Wang

1. An Introduction

Context

In Spring of 2020 as the novel COVID-19 pandemic made waves across headlines, United States (US) state-level governments attempted to slow community spread of COVID through Stay at Home (SAH) orders. By the end of March 2020, 39 states had an active SAH order that limited its citizens' movement patterns, impacting an estimated **75 million Americans**, roughly one-fifth of the US population. This represented one of the strongest public policy interventions in modern American history, whereby mandates leveraged by state governments greatly impacted constituent freedom of choice.

State leadership such as Governor Newsom of California and Governor Cuomo of New York were praised for early action. However, the ramifications of the state-issued SAH orders revealed to be immense and devastating in scope. The SAH orders spurred a complimentary set of closure orders resulting in job loss, delays in healthcare treatment, and an economic halt. This was all done in the name of public health.

More than one year later, SAH orders are still re-implemented by local governments as a response to surges in COVID cases. Importantly, SAH orders will likely be implemented during future health disasters as well.

Research Focus

We believe it is imperative to understand the effectiveness of SAH policies so that policymakers can better weigh the benefits and drawbacks of such directives in the future, and it is through this lens that we frame our research.

We examine the effectiveness of SAH orders in reducing citizen mobility, and we aim to understand whether Americans in states with SAH orders imposed adhered to the orders and limited their non-essential activities. Formally, our research seeks to find a correlation between the length of a state's SAH order and the reduction in its constituent mobility during lockdown. If we establish that such a relationship exists, policymakers will have a baseline from which they can quantify the efficacy of their policy tools. A natural next step beyond this research would be to evaluate the effect of reduced citizen mobility on the spread of COVID, thereby completing a causal chain. However, this paper will be limited in scope to the former question.

Research Operationalization

We will focus our study during a research window that spans from 3/15/20 to 6/15/20, a period of 92 days. This time period coincides with the first COVID wave in the US (Figure 1), and is the time period during which states differed in their SAH order policies. Prior to 3/15/2020, no states had implemented SAH orders, and after 6/15/2020, only New York, California, Oregon, and New Hampshire had a SAH order in place. Thus, our research window serves as a standardized time period that captures most states' SAH orders.

It is with respect to this first US COVID wave that we examine changes in citizen mobility, as opposed to subsequent spikes in cases, in order to control for confounding factors. Due to the US's decision to delegate COVID measures to the state level, differences in each state's approach becomes compounded due to unique county and state -wide conditions. Additionally, since our research focuses on citizen mobility, the Spring season helps to control for the impacts of extreme seasonality (e.g., Winter, Summer) on COVID spread.

The dependent variable we capture is the change in citizen mobility, which serves as a proxy for the degree of adherence citizens have toward the SAH order imposed in their states. We leverage **Google mobility reports**, an anonymized dataset that aggregates the location data of users of Google products, to evaluate the degree to which citizens adhered to SAH orders. Mobility metrics are calculated by Google against a baseline set as the median mobility value from the 5-week period of 1/3/2020 – 2/6/2020, and are reported as a percent change with respect to this baseline. Established baselines values vary in time and location. Full documentation of Google's datasets can be found [here](#).

While there are many mobility measures that Google reports on including "transit stations", "parks", and the "workplace", our research will focus on the change in mobility for the "retail and recreation" category.

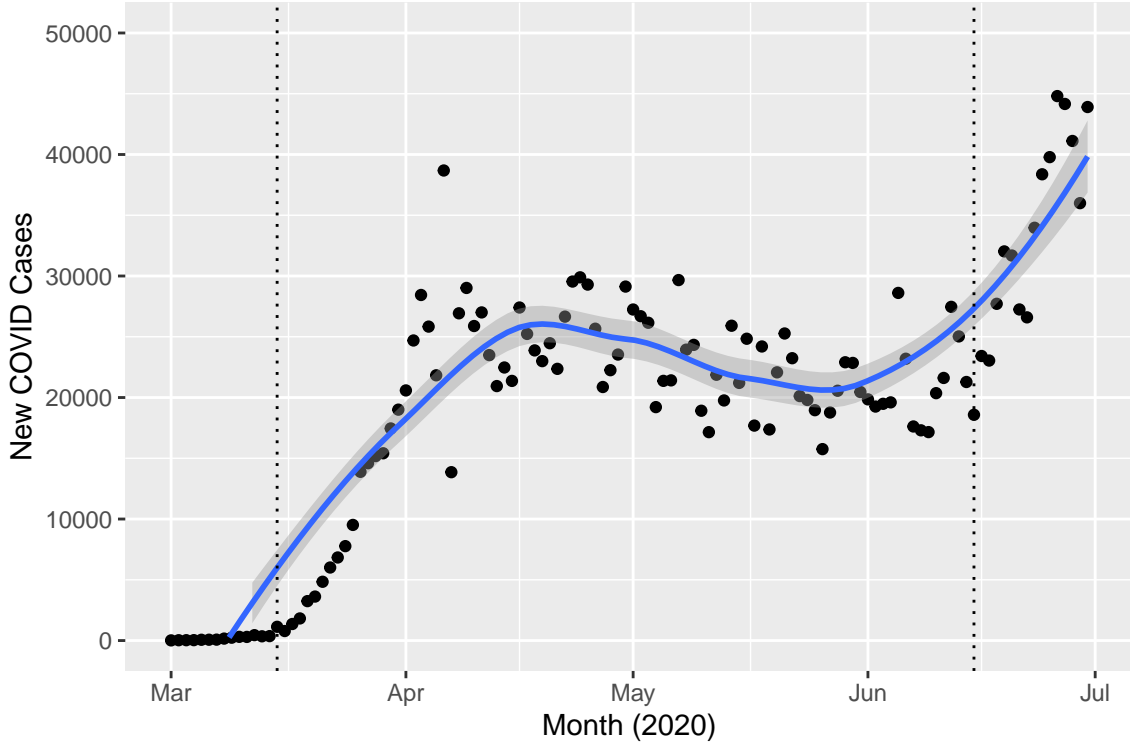


Figure 1: First US COVID Wave (dotted lines represent research window from 3/15/2020 to 6/15/2020)

Unlike the other categories, we believe that “retail and recreation” best captures the type of non-essential activity SAH policies intend to reduce.

Limitations of Data

We acknowledge that there are limitations to the datasets and we make assumptions that surround our research. Notwithstanding, we believe it is appropriate to proceed with our research. Although these caveats may increase noise and decrease the precision of our results, they do not systematically bias our findings in any meaningful way:

- State-imposed SAH orders have varying restrictions on movement. North Carolina, for example, issued a SAH order that **exempts places of worship** on the basis that they operate as an essential activity. There is a lack of normalization in what a SAH order may mean between state lines.
- Google mobility reports are a good proxy for citizen movement, but they are not perfect. Google implements a **differential privacy algorithm** to intentionally add noise to data in order to mask individual user behavior. Although this preserves user privacy, this hurts our model accuracy, as the dataset deviates from reality. Additionally, individuals who do not use Google products or who have Location History turned off will not have their mobility data recorded, which makes the dataset unrepresentative of reality. However, we believe this will not systematically bias Google mobility reports in a way that will skew the results of our research. (Interestingly, other companies such as **Apple** have also released their versions of mobility reports that detail their user bases. In the interest of preserving scope, we have not examined these reports in detail, although this could be an extension of our research for future studies.)

2. A Model Building Process

The models we create aim to describe the effect of independent variables on the dependent variable. In subsequent sections, we construct three models, each of which adds complexity in the form of introducing additional independent variables.

Formally, the dependent variable is the median percent change in “retail and recreational” (R&R) mobility in all 50 states, as well as the District of Columbia, during the research window that spans from 3/15/2020 to 6/15/2020. A lower, more negative value for the dependent variable corresponds to a greater median reduction in mobility during the research window. For example, Alabama and Pennsylvania observed median mobility reductions of -15% and -34%, respectively. Thus, we can conclude that Pennsylvania had a greater median reduction in mobility than Alabama during the research window, as compared to the states’ respective pre-COVID, baseline movement patterns.

We choose to examine the median percent change in mobility, instead of another aggregate such as mean, for two benefits. Firstly, using the median aligns with Google’s methodology of setting the baseline with which all daily comparisons are made. Secondly, the median will be less sensitive to anomalous fluctuations in mobility throughout the research window as compared to the mean.

For the duration of this analysis, we will use $\alpha = 0.05$ to determine statistical significance in accordance with common industry practices.

Model 1

Variables

In our most limited model, we examine the relationship that a state's SAH order duration has on the dependent variable. To calculate the length of a SAH order, we rely on the **COVID-19 US State Policy Database**, a repository maintained by the Boston University School of Public Health. This dataset contains start and end dates for each state's SAH order, which we convert to a single integer representing the number of days a SAH order is imposed for a given state. This calculation is dependent on a state having imposed a SAH order that explicitly restricts movement of the general public. Certain states, such as Connecticut, imposed SAH orders that did not limit movement. We treat these states as having a SAH order length of 0 days.

Modeling Goal

Our modeling goal is to build an explanatory model to determine if a relationship exists between SAH order length and citizen mobility.

EDA

Our initial EDA shows a strong, linear relationship between our SAH order length and the dependent variable (Figure 2).

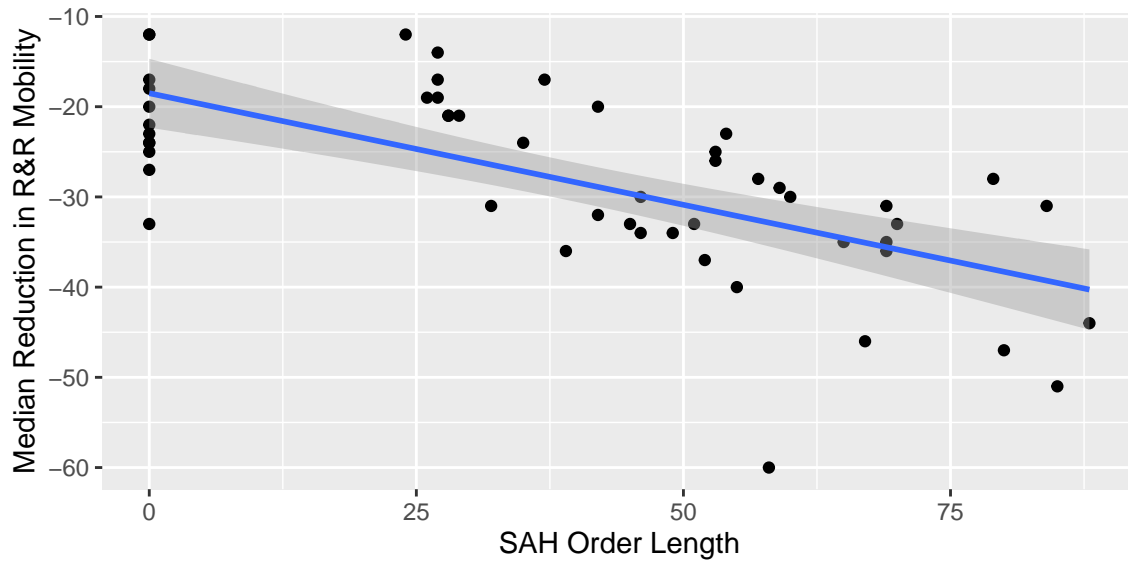


Figure 2: Model 1 EDA, Linear Relationship Exhibited When States Without SAH Order Included

However, we notice the cluster of states that did not implement a mobility-restricting SAH order and the horizontal gap in which no states implemented a SAH order between 1 to 23 days in length. We observe these attributes will impact how a linear predictor fits the data. If we removed the data points representing states that did not implement a mobility-restricting SAH order, we observe that a non-linear relationship may describe the relationship more meaningfully (Figure 3). This matches our intuitive sense, as we would expect the effects of a SAH order on mobility to taper off at the extremes.

In Figure 4, we observe the distribution of our dependent variable does appear to exhibit a slightly heavier left tail than a normal distribution as a result of a few states having far lower reductions in R&R mobility than other states

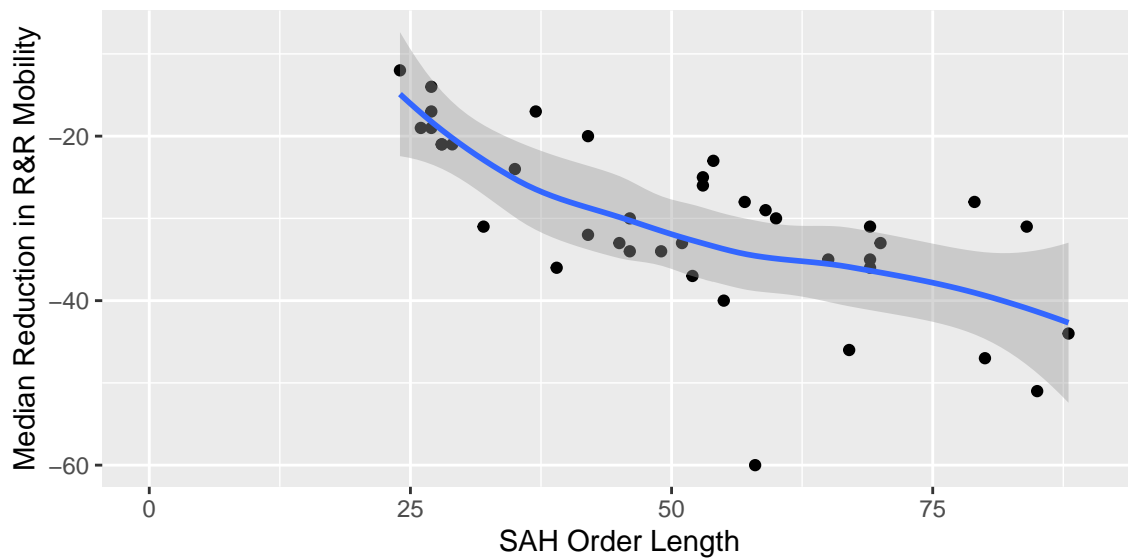


Figure 3: Model 1 EDA, Non-Linear Relationship Exhibited When States Without SAH Order Excluded

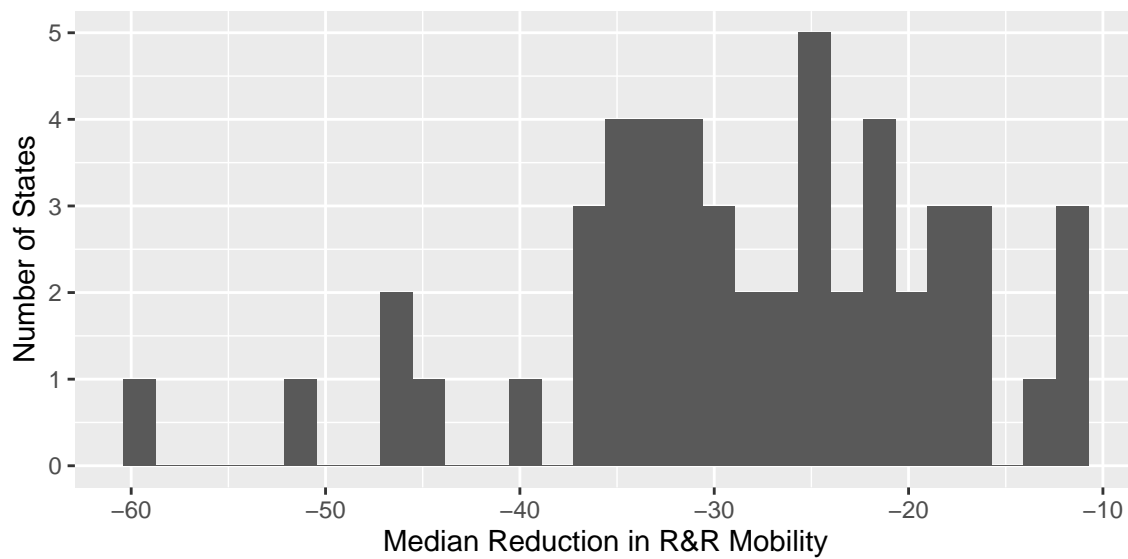


Figure 4: Model 1 EDA, Distribution of the Dependent Variable

Transformations

Based on the EDA, we may consider excluding the states that did not implement a mobility-restricting SAH order. However, this causes our model to be less interpretable. Given that our goal is not to perfectly describe the relationship between SAH order length and the dependent variable in this first, most limited model, but rather to build an explanatory model showing that changes in the independent variable may cause changes of magnitude in the dependent variable, we decided against applying the aforementioned transformation.

Further, although we do notice a deviation from normality for our dependent variable, the deviation from normality is not strong enough to suggest using a transformation that may hurt the interpretability of this study. Our aim with Model 1 is to generalize a trend while preserving model interpretability.

Model Construction

Model 1 is reported as: $Y = -18.5 + -0.25 * X_0$, where:

- Y : median percent change in R&R mobility

- X_0 : Length of SAH order during the 92 day research window from 3/15/20 to 6/15/20, in days

```
model1 = lm(y_var~1+length,data=df)
summary(model1)$coefficients[1:2, 1:2]
```

```
##              Estimate Std. Error
## (Intercept) -18.50698  1.89976195
## length      -0.24733  0.03978374
```

Evaluation

```
coeftest(model1, vcov = vcovHC(model1))
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error  t value  Pr(>|t|)
## (Intercept) -18.506979    1.779985 -10.3973 5.436e-14 ***
## length      -0.247330    0.040557  -6.0983 1.651e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# summary(model1) # Adjusted R-squared:  0.4295
```

The p-values, using robust standard errors, for the intercept and independent variable, SAH length, are 5.436e-14 and 1.651e-07, respectively, indicating statistically significant relationships. We also find an adjusted R-squared value of 0.4295. An ANOVA against the base model, which excludes SAH length as a variable, reports $\Pr(>F) = 1.08e-07$.

```
model_base <- lm(y_var~1,data=df)
anova(model1,model_base,test='F')
```

```
## Analysis of Variance Table
##
## Model 1: y_var ~ 1 + length
## Model 2: y_var ~ 1
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1      49 2899.9
## 2      50 5187.2 -1   -2287.3 38.649 1.083e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We observe that changes in X_0 have a significant relationship with changes in Y .

However, as our modeling and analysis goal is explanatory, we need to evaluate the likelihood of causality. There are two potential factors that could undermine the establishment of causality:

1. *Reverse causality*: Reverse causality would imply that reductions in mobility would cause longer SAH orders, which violates intuition and is therefore unlikely. SAH orders were issued by local governments with the explicit intent to reduce citizen movement as a community mitigation measure to slow the spread of COVID. We would expect states with reduced mobility to potentially lift SAH orders sooner as the mandate would appear unnecessary.
2. *Omitted variable bias*: There is substantial risk of omitted variable bias undermining the explanatory power of Model 1 (Figure 5). There are several variables not included in Model 1 that may cause both longer SAH orders and decreases in mobility, creating a strong bias away from zero. We explore these independent variables in Model 2.

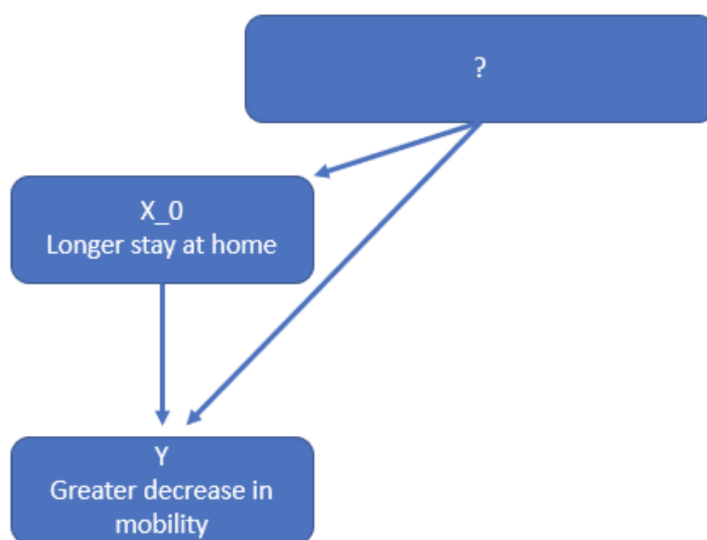


Figure 5: Schematic of Model 1 Omitted Variable Bias

Model 2

Variables

In our second model, we include two additional variables in our analysis: the political party of the state's governor and new COVID case count per 1000 residents during the research window. The political party of the state's governor serves as a proxy for the political beliefs of the state's residents, and the COVID case count per 1000 residents serves as a normalized comparison of COVID severity between states.

Modeling Goal

The goal of our second model is again explanatory, but with a focus on covariates that could have a relationship both with the dependent variable, Y , and the SAH length, X_0 . If we observe that these covariates are significant, it could undermine confidence in the explanatory power of Model 1.

EDA

Based on our EDA (Figure 6), it appears there exists a relationship between COVID cases per 1000 residents and the median reduction in mobility. The more COVID cases a state records per 1000 residents, the greater the median reduction in mobility. Additionally, there appears to exist a relationship between political beliefs and the median reduction in mobility. In states with Democratic Governors, the median reduction in mobility was greater in magnitude.

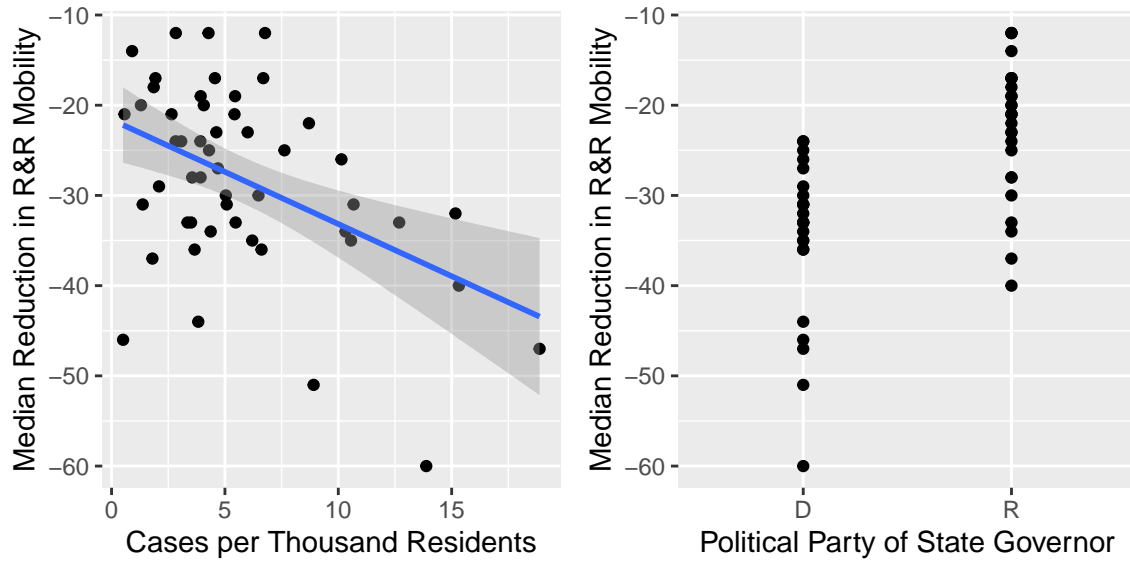


Figure 6: Model 2 EDA, Effects of X_1 and X_2 on Y

This potentially undermines the explanatory power of Model 1, as both of these independent variables may also explain positive changes in the X_0 variable, SAH Order Length, thus biasing Model 1's beta coefficient away from zero. To further assess this possibility, we explore the relationship between our independent variables.

We observe in Figure 7 that more COVID cases per 1000 residents appears to be correlated with longer SAH orders. Similarly, states with Democratic Governors correlated with longer SAH orders.

Transformations

Based on our EDA, no transformations appear to be called for.

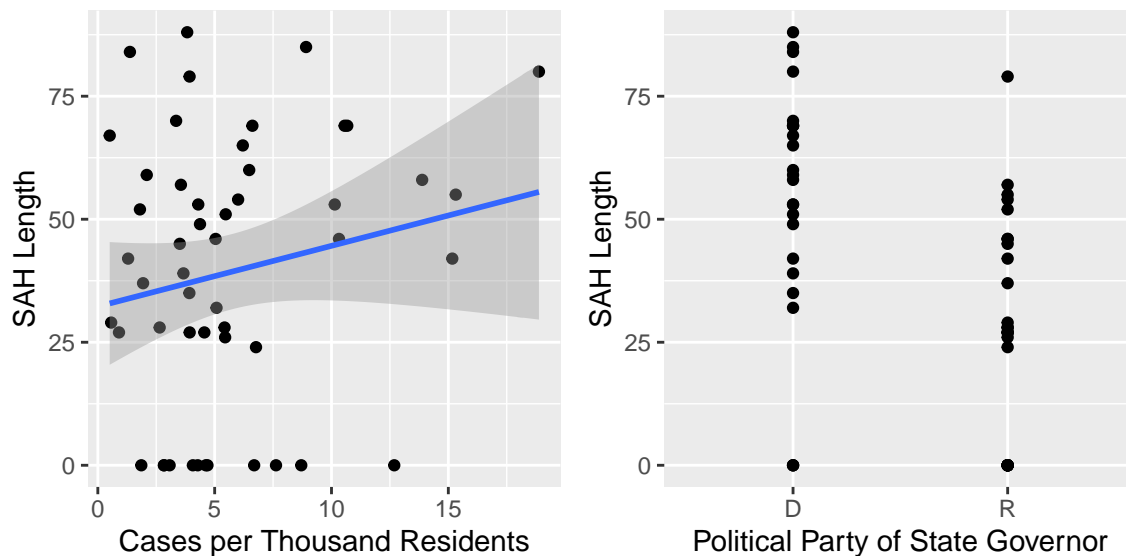


Figure 7: Model 2 EDA, Effect of X_1 and X_2 on X_0

Model Construction

Model 2 is reported as: $Y = -20.73 - 0.17 * X_0 - 0.74 * X_1 + 6.06 * X_2$, where:

- Y : median percent change in R&R mobility
- X_0 : Length of SAH order during the 92 day research window from 3/15/20 to 6/15/20, days
- X_1 : Party of Governor (0 = Democrat, 1 = Republican)
- X_2 : Total new COVID cases per 1000 residents during the research window

```
model2 = lm(y_var~1+length+new_case_per_thousand+Party,data=df)
summary(model2)$coefficients[, 1:2]
```

```
##              Estimate Std. Error
## (Intercept)  -20.3998210  2.81351654
## length       -0.1729392  0.03792942
## new_case_per_thousand -0.7397533  0.23067868
## PartyR        6.0596136  2.10165535
```

Evaluation

```
coeftest(model2, vcov = vcovHC(model2))
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -20.399821  2.242532 -9.0968 6.202e-12 ***
## length       -0.172939  0.031091 -5.5623 1.229e-06 ***
## new_case_per_thousand -0.739753  0.297339 -2.4879 0.016453 *
## PartyR        6.059614  1.871777  3.2374 0.002215 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(model2,model1,test='F')
```

```
## Analysis of Variance Table
##
## Model 1: y_var ~ 1 + length + new_case_per_thousand + Party
## Model 2: y_var ~ 1 + length
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1      47 1930.5
## 2      49 2899.9 -2   -969.39 11.801 7.034e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

When covariates X_1 and X_2 are introduced, we observe that both the magnitude and p-value for X_0 decreases substantially. Additionally, X_1 and X_2 both have a statistically significant relationship with Y .

This calls into question the explanatory power of the X_0 variable, SAH order length, as the causal chain may resemble Figure 8:

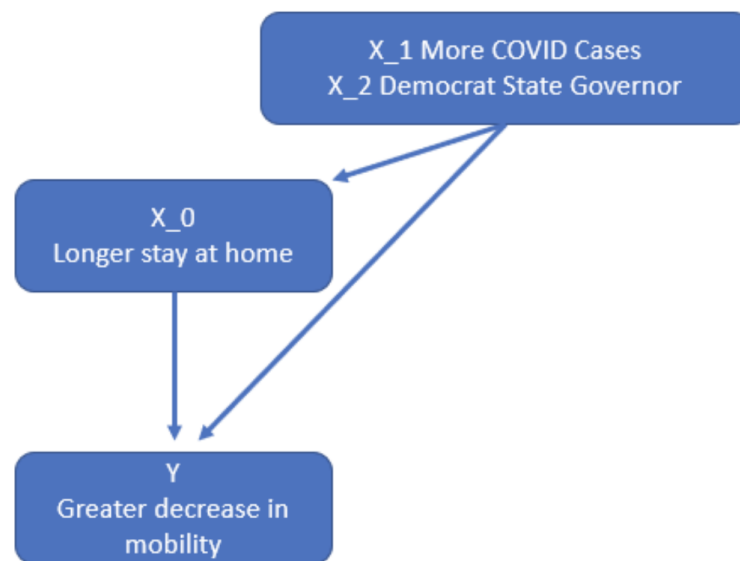


Figure 8: Schematic of Model 2 Omitted Variable Bias

Calculating the omitted variable bias in Model 1 against models that include X_1 and X_2 separately, we observe:

$$Y = -18.51 + -0.25 * X_0$$

$$Y = -14.41 + -0.22 * X_0 + -0.88 * X_1$$

Omitted variable bias = $B1 - B1 = -0.03$

$$Y = -18.50 + -0.25 * X_0$$

$$Y = -25.10 + -0.18 * X_0 + 7.49 * X_2$$

Omitted variable bias = $B1 - B1 = -0.07$

```
model2_1 <- lm(y_var~1+length+new_case_per_thousand,data=df)
model2_2 <- lm(y_var~1+length+Party,data=df)

coeftest(model2_1, vcov = vcovHC(model2_1))
```

```
##
## t test of coefficients:
##
```

```
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -14.411371   1.992834 -7.2316 3.249e-09 ***
## length        -0.222997   0.036362 -6.1327 1.572e-07 ***
## new_case_per_thousand -0.881291   0.306620 -2.8742 0.006019 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
coeftest(model2_2, vcov = vcovHC(model2_2))
```

```
##
## t test of coefficients:
##
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -25.099377   2.287880 -10.9706 1.119e-14 ***
## length       -0.180591   0.037332  -4.8374 1.399e-05 ***
## PartyR        7.493858   2.205106   3.3984 0.001372 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We observe that the omitted variable bias is away from zero in both cases.

Although Model 2 still shows that SAH order length has the most significant relationship with the dependent variable, it raises concerns about the explanatory power of SAH order length. As we are working with limited data, our models may not fulfill all assumptions of the Classical Linear Model. This will be addressed in further detail in the *Limitations of Model* section.

Model 3

Variables

In our third model, we include two additional variables in addition to previous variables: the median age of residents and the percentage of residents at risk for serious COVID infection. Regarding the former variable, we hypothesize younger residents, who will have a greater tolerance for risk relative to older residents, may be less inclined to adhere to SAH orders. Regarding the latter variable, we hypothesize that residents at greater relative risk for serious COVID infection may be less inclined to adhere to SAH orders.

Modeling Goal

The goal of our third model is again explanatory, as we are seeking to determine if additional covariates that may naturally cause citizens to reduce their mobility may also explain the reduction in median R&R mobility observed across states.

EDA

In our EDA, we observe that median age and the percent of residents at risk for serious COVID illness do not appear to show a strong relationship with the Y variable (Figure 9). Furthermore, the percent of residents at risk for serious COVID illness actually appears to show the opposite relationship with the Y variable than intuition would suggest, indicating that states with a higher percent of people at risk for serious COVID illness experienced a lower median reduction in mobility during the research window.

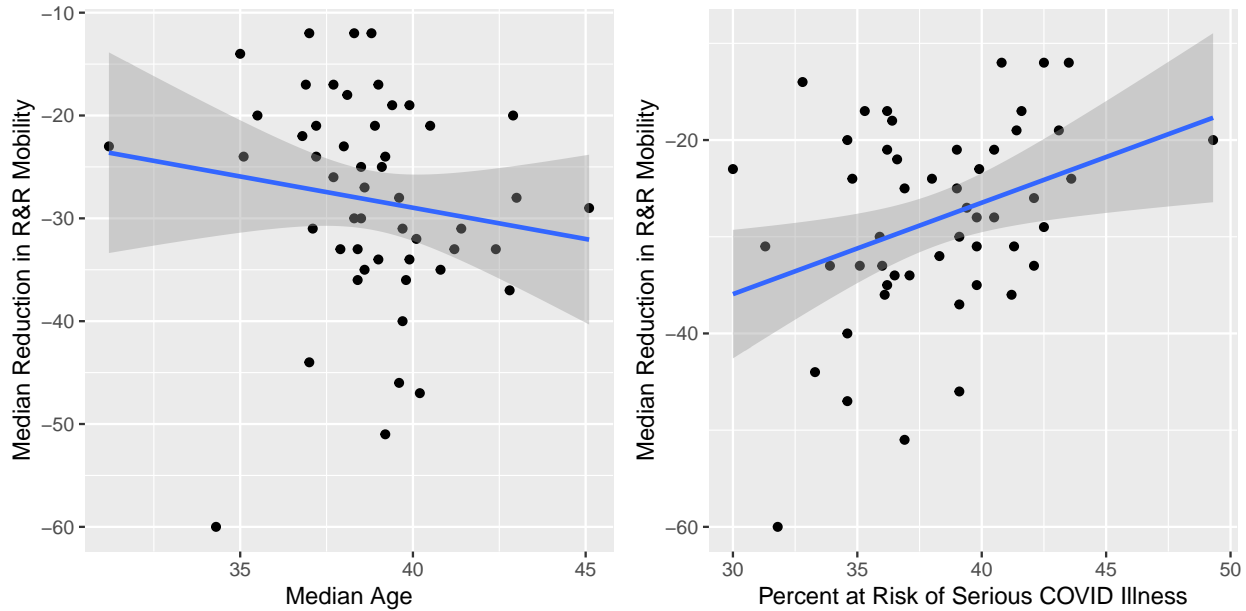


Figure 9: Model 3 EDA, Effect of X_3 and X_4 on Y

Transformations

Based on our EDA, no transformations appear to be called for.

Model Construction

Model 3 is reported as $Y = -30.40 - 0.16 * X_0 - 0.55 * X_1 + 5.13 * X_2 - 0.77 * X_3 + 1.02 * X_4$, where:

- Y : median percent change in R&R mobility
- X_0 : Length of SAH order during the 92 day research window from 3/15/20 to 6/15/20, days

- X_1 : Party of Governor (0 = Democrat, 1 = Republican)
- X_2 : Total new COVID cases per 1000 people in the period
- X_3 : Median Age of State Residents
- X_4 : Percent of Residents at Risk for Serious COVID Infection

```
model3 <- lm(y_var~1+length+new_case_per_thousand+Party+MedianAge+RISKCOV,data=df)
summary(model3)$coefficients[, 1:2]
```

```
##              Estimate Std. Error
## (Intercept)  -30.3973039 14.66807994
## length       -0.1637986  0.03812169
## new_case_per_thousand -0.5485156  0.21950896
## PartyR        5.1310066  1.94902815
## MedianAge    -0.7718604  0.52802378
## RISKCOV       1.0218033  0.32489341
```

Evaluation

```
coeftest(model3, vcov = vcovHC(model3))
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -30.397304  22.014331 -1.3808 0.174161
## length       -0.163799   0.037297 -4.3918 6.75e-05 ***
## new_case_per_thousand -0.548516  0.299251 -1.8330 0.073429 .
## PartyR        5.131007   1.999429  2.5662 0.013680 *
## MedianAge    -0.771860   0.724634 -1.0652 0.292480
## RISKCOV       1.021803   0.314725  3.2467 0.002208 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(model3,model2,test='F')
```

```
## Analysis of Variance Table
##
## Model 1: y_var ~ 1 + length + new_case_per_thousand + Party + MedianAge +
##           RISKCOV
## Model 2: y_var ~ 1 + length + new_case_per_thousand + Party
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1      45 1556.3
## 2      47 1930.5 -2   -374.21 5.4102 0.007842 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

After including covariates X_3 and X_4 , we observe that the p-value for X_3 is high, and that the direction of X_4 is actually the opposite of what intuition would suggest, which might be caused by this essentially being an outcome variable on the right side caused by an omitted variable such as lack of medical knowledge. This outcome leads us to believe covariates X_3 and X_4 do not carry strong explanatory power for changes in the Y variable.

Regression Table

We summarize the variables from the 3 regression models in the regression table below.

```
se.model1 = coeftest(model1, vcov = vcovHC(model1))[ , "Std. Error"]
se.model2 = coeftest(model2, vcov = vcovHC(model2))[ , "Std. Error"]
se.model3 = coeftest(model3, vcov = vcovHC(model3))[ , "Std. Error"]
```

```
stargazer(model1, model2, model3, type = "text",
  se = list(se.model3),
  star.cutoffs = c(0.05, 0.01, 0.001),
  title = paste("Table 1: A comparison of models that describe ",
    "the effect of SAH mandates on constituent mobility"))
```

```
##
## Table 1: A comparison of models that describe the effect of SAH mandates on constituent mobility
## =====
##                               Dependent variable:
##                               -----
##                               y_var
##                               (1)      (2)      (3)
## -----
## length                -0.247***      -0.173***      -0.164***
##                        (0.037)        (0.038)        (0.038)
##
## new_case_per_thousand                -0.740**      -0.549*
##                        (0.231)        (0.220)
##
## PartyR                6.060**      5.131*
##                        (2.102)        (1.949)
##
## MedianAge                -0.772
##                        (0.528)
##
## RISKCOV                1.022**
##                        (0.325)
##
## Constant                -18.507      -20.400***      -30.397*
##                        (22.014)        (2.814)        (14.668)
## -----
## Observations                51      51      51
## R2                0.441      0.628      0.700
## Adjusted R2                0.430      0.604      0.667
## Residual Std. Error      7.693 (df = 49)      6.409 (df = 47)      5.881 (df = 45)
## F Statistic      38.649*** (df = 1; 49) 26.430*** (df = 3; 47) 20.998*** (df = 5; 45)
## =====
## Note:                                *p<0.05; **p<0.01; ***p<0.001
```

Limitations of Model

We evaluated our model against the five Classical Linear Model (CLM) assumptions and find that Model 2 most reasonably satisfies the CLM requirements (Figure 10). To establish that our estimators in Model 2 are unbiased, we must establish that the data are independent and identically distributed (IID), that there is no perfect collinearity, and that a linear conditional expectation is met. While Model 2 shows evidence that there is no perfect collinearity (as evidenced by R not dropping a coefficient in the model as well as by applying the ocular test to the “Residuals vs Fitted” figure below), there are some IID considerations to note. There is likely a time dependence in our observations, as SAH orders are imposed in response to COVID spread. We attempt to mitigate this impact by focusing our analysis on the first wave of COVID in the US. These considerations of IID for Model 2 are noted as a limitation of our model.

Next, we evaluate Model 2 for homoskedastic errors. While there is a mostly flat trendline in the below “Scale-Location” subplot, we detect a slight downward slope to the line plotted. A Breusch-Pagan test results in a p-value of 0.28, which fails to reject the null hypothesis and suggests there is a lack of evidence for heteroskedastic errors. To remain conservative in our practice, and because of the relatively small sample size of our analysis, we opt to default the model to robust standard errors.

The final CLM requirement of normally distributed errors is met through the ocular test of the “Normal QQ” plot as well as a statistical evaluation via the Shapiro-Wilk normality test. The Shapiro-Wilk normality test yields a p-value of 0.14, which leads us to fail to reject the null hypothesis that we have normally distributed errors.

```
par(mfrow=c(2,2))
plot(model2)
```

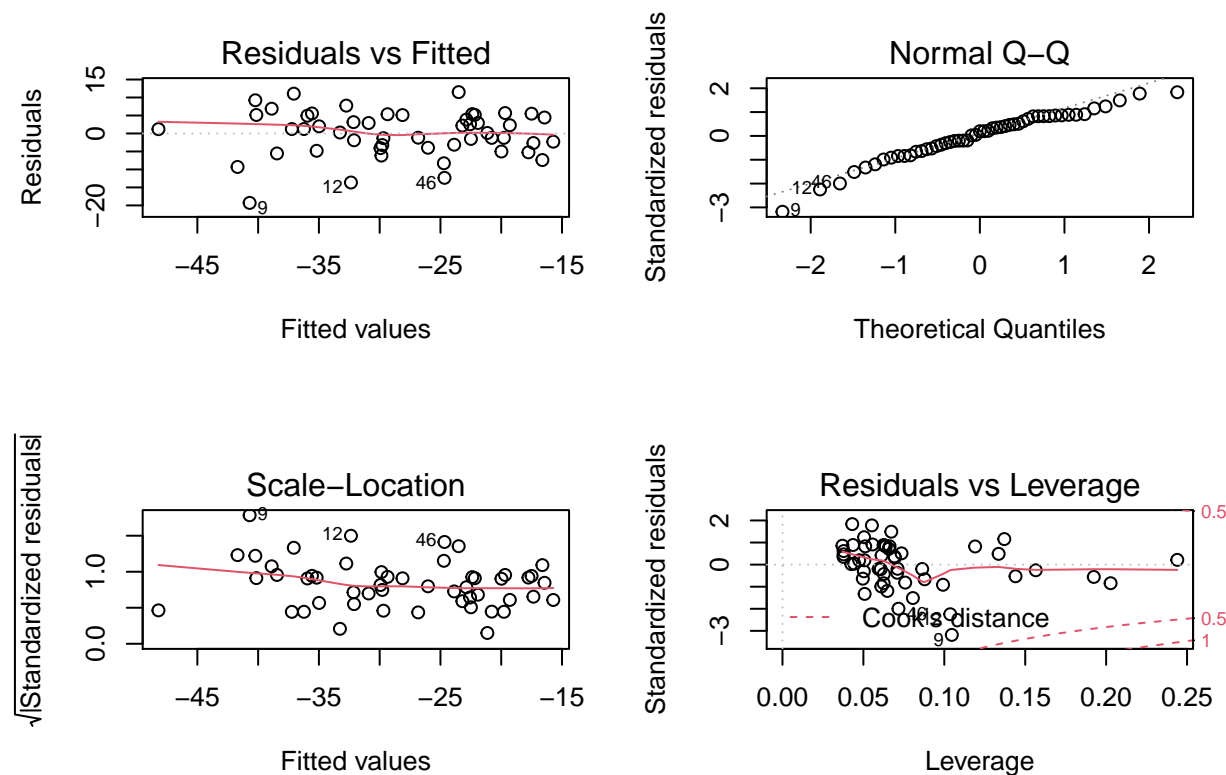


Figure 10: Model 2 Summary Statistics


```

# Assessing homoskedastic errors
bptest(model2)

##
## studentized Breusch-Pagan test
##
## data: model2
## BP = 3.813, df = 3, p-value = 0.2824
# Assessing for normally distributed error
shapiro.test(model2$residuals)

##
## Shapiro-Wilk normality test
##
## data: model2$residuals
## W = 0.96537, p-value = 0.1412

```

Discussion of Omitted Variables

As part of the evaluation of Model 1, we noted its susceptibility to omitted variable bias and we identified two variables that could cause such biases: X_1 (COVID Cases per 1000 Residents) and X_2 (Political Party of State Governor). We then included these variables in Model 2. The Political Party of State Governor is a proxy for the political beliefs of the state's residents, which is an omitted variable discussed below. Additionally, we examine several omitted variables that may impact Models 2 and 3, below.

Omitted Variable 1: Political Beliefs of State Residents

Model 2 includes X_2 , Political Party of State Governor, which we observe was an omitted variable that had biased Model 1 away from zero (Figure 11). However, X_2 (Political Party of State Governor) is really just a rough proxy for the political beliefs of the state's residents. Our concern is that if political beliefs of the state's residents were included as a variable, it might be shown to have an even stronger relationship with SAH order length and the dependent variable of mobility, thus biasing Models 1 and 2 away from zero.

For this reason we remain hesitant about the explanatory power of the SAH order length.

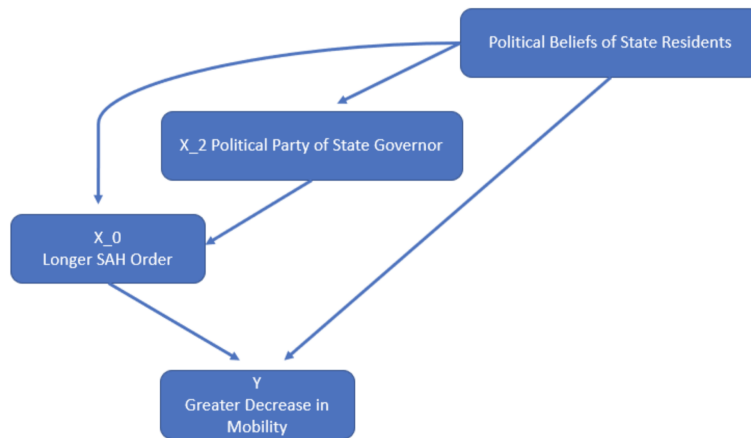


Figure 11: Schematic of Omitted Variable 1, Political Beliefs of State Residents

Omitted Variable 2: Percentage of Urban Residents

The percent of urban residents in a state is an omitted variable that could bias multiple coefficients in Model 2 away from zero (Figure 12). Residents in urban areas come in more frequent contact with other residents, increasing the risk of COVID spread. We would expect states with a higher percentage of urban residents to implement longer SAH Orders, experience a higher number of COVID cases, and observe residents in urban areas who are more inclined to reduce their R&R mobility due to higher perceived risk. This highlights the difficulty we face in ascribing clear explanatory power to SAH order length.

Omitted Variable 3: New Unemployment Claims

New unemployment claims is an omitted variable that could affect SAH order length the dependent variable, and it would bias the SAH order length beta coefficient away from zero. We would expect states with a greater number of new unemployment claims to try to end their SAH order quickly because of negative economic impact or a SAH order. In addition, we would expect states with a relatively large number of new unemployment claims to observe a smaller median decrease in R&R mobility, as unemployed citizens will have extra time on their hands and may find it difficult to reduce their R&R mobility.

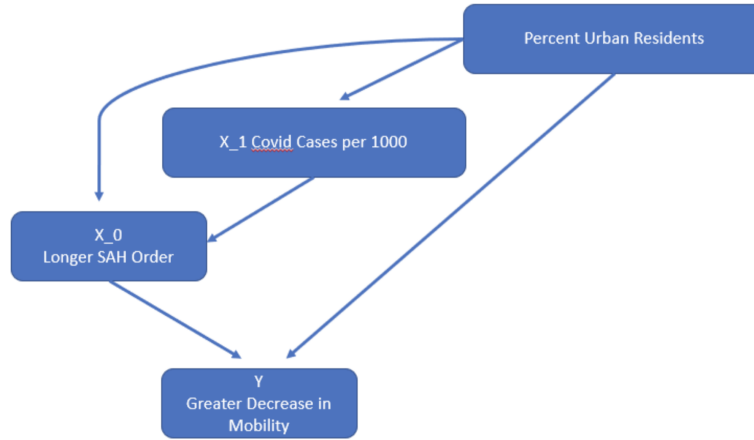


Figure 12: Schematic of Omitted Variable 2, Percentage of Urban Residents

Omitted Variable 4: Medical Knowledge

Medical knowledge is an omitted variable that may have a relationship with both COVID Cases per 1000 Residents and the dependent variable. We would expect that in states with lower average medical knowledge, there may be a greater number of COVID cases per 1000 residents and a smaller median decrease in R&R mobility. This would bias our beta coefficient associated with COVID Cases per 1000 Residents toward zero.

In addition, medical knowledge as an omitted variable may aid in explaining why the the Percent of Residents at Risk for Serious COVID Infection variable has a beta coefficient with a sign that is opposite of what intuition would suggest. The Percent at Risk of Serious COVID Illness could be an outcome variable of the omitted variable, medical knowledge. In Model 3, we would have put this outcome variable on the right side. Lack of medical knowledge would cause both a greater number of citizens to be at risk of serious COVID illness due to pre-existing conditions such as tobacco use, obesity, and diabetes, as well as cause these same individuals to not limit their R&R mobility during COVID.

Omitted Variable 5: Risk Tolerance

In Model 2, we observed that states with a greater number of COVID cases per 1000 residents saw a greater median decrease in R&R mobility. Risk tolerance is an omitted variable that could have a relationship with both COVID Cases per 1000 Residents and the dependent variable. However, it would bias our beta coefficient toward zero. We could imagine that states with a greater average risk tolerance would have more COVID cases per 1000 residents and a smaller median decrease in R&R mobility.

Conclusion

Our exploratory analysis shows a statistically significant relationship between increasing lengths of implemented SAH orders and reductions in resident mobility for non-essential activities during the first wave of COVID in the United States, even when accounting for normalized COVID case count during the research window as well as the political party of the governor. These findings add to the body of literature that support CDC's guidance to local governments in issuing SAH orders as a measure to reduce mobility to limit COVID spread. As more data becomes available, we can begin to answer the question of whether these actions were worthwhile from a social and economic standpoint. We look to future research to continue this exploration to test the causal relationship between SAH order length and mobility reduction, as well as examine the impact that SAH orders may have had on reducing COVID infections and corresponding deaths.