

# MODEL FITTING

Ben Chang, Nicolas Min, Emmanuel Rodriguez

4/1/2021

## 1. Data

We have created our final data set (**covid\_econ.csv**) by combining and cleaning data from:

- Bureau of Labor Statistics
- Google: COVID-19 Community Mobility Reports
- Oxford: COVID-19 Government Response Tracker

To reproduce plots and analysis in this document, read in **covid\_econ.csv** and run codes starting from *Data Manipulation* at the end of this section.

### Database creation

```
#####
##### Employment data #####
#####

# Source: Bureau of Labor Statistics
# Link: https://download.bls.gov/pub/time.series/sm/sm.data.1.AllData
# Series description: The original series are the SA thousand of employees at each month
# by sector and state

#Read data
all_data <- read.delim('https://download.bls.gov/pub/time.series/sm/sm.data.1.AllData')

#Create index of series
#(note: other services' not included)
index <- read.delim('https://download.bls.gov/pub/time.series/sm/sm.series') %>%
  filter(seasonal=='S', area_code == 00000, data_type_code == 01,
         industry_code == 1000000*supersector_code, state_code <= 56) %>%
  select(c(series_id,state_code,supersector_code)) %>%
  filter(supersector_code %in% c(10, 20, 30, 41, 42, 43, 50, 55, 60, 65, 70, 90))
states_index <- read.delim('https://download.bls.gov/pub/time.series/sm/sm.state')
supersector_index <- read.delim('https://download.bls.gov/pub/time.series/sm/sm.supersector')

index <-
  index %>% merge(states_index, by = 'state_code') %>%
  merge(supersector_index, by = 'supersector_code') %>%
  select(series_id, state_name, supersector_name)
```

```

##Filter data
series <- unique(index[,1])
filtered_data <- all_data %>% filter(series_id %in% series) %>%
  merge(index, by='series_id') %>%
  mutate(month = substr(period,2,3)) %>%
  filter(year >= 2019) %>%
  select(value, month, year, state_name, supersector_name)

##Transform data

#Obtain monthly change and drop off 2019 (jan-nov)
employment_data <- filtered_data %>%
  mutate(emp_monthly_change = as.numeric(value)/lag(as.numeric(value))-1) %>%
  mutate(year_month = paste0(year,month)) %>%
  filter(year_month >= 201912) %>%
  select(value, emp_monthly_change, month, year, state_name, supersector_name)

#Built date variable
dates <-
  as.Date(as.yearmon(paste(
    employment_data$year, employment_data$month, sep = '-'
  )), frac = 1)
employment_data <- employment_data %>%
  mutate(date = dates) %>%
  select(date, value, emp_monthly_change, state_name, supersector_name)

colnames(employment_data)[2] <- 'employees'

#####
##### Mobility data #####
#####

# Source: Google

#Read data
mobility_original <- read.csv('Data/2020_US_Region_Mobility_Report.csv', header = T, na.strings = "") %>%
  mutate(date = as.Date(date)) %>%
  filter(is.na(sub_region_2) & !is.na(sub_region_1)) %>%
  select(c(date, sub_region_1, retail_and_recreation_percent_change_from_baseline,
           grocery_and_pharmacy_percent_change_from_baseline, parks_percent_change_from_baseline,
           transit_stations_percent_change_from_baseline, workplaces_percent_change_from_baseline,
           residential_percent_change_from_baseline)) %>%
  'colnames<-'(c('date','state_name','Mob_ret_rec','Mob_gro_pharm', 'Mob_parks','Mob_transit','Mob_work', 

#Seasonal adjusted daily data (only taking residential mobility)

mobility <- mobility_original %>%
  select(date, state_name, Mob_res) %>% mutate(Mob_res_SA = NA)

```

```

states <- unique(mobility$state_name)
for(i in 1:length(states)){
  state <- states[i]

  state_data <- mobility %>% filter(state_name == state) %>% select(date, Mob_res)
  ts <-
    ts(state_data$Mob_res,
        frequency = 7,
        start = c(2020, as.numeric(format(
          as.Date('2020-02-15'), "%j"
        ))))
  decompose_ts <- decompose(ts, 'additive')
  adjust_ts <- ts - decompose_ts$seasonal

  mobility[mobility$state_name == state,]$Mob_res_SA <- adjust_ts
}

#Computing 30 days change (to obtain the monthly change)
mobility <- mobility %>% mutate(delta_Mob_res_SA = NA)
states <- unique(mobility$state_name)
for(i in 1:length(states)){
  state <- states[i]

  mob <- mobility %>% filter(state_name == state) %>% select(Mob_res_SA)
  delta_mob <- unname(unlist(mob/lag(mob, n=30)-1))

  mobility[mobility$state_name == state,]$delta_Mob_res_SA <- delta_mob
}

#####
#### Stringency data #####
#####

#load data
stringency_us <- read.csv("https://raw.githubusercontent.com/0xCGRT/USA-covid-policy/master/data/0xCGRT")
stringency_us <- stringency_us %>%
  select(Date, RegionName, StringencyIndex) %>% filter(RegionName != '')

#Modify name of one state
stringency_us[stringency_us$RegionName == 'Washington DC',]$RegionName <- rep('District of Columbia', length(stringency_us))

#Convert dates
stringency_us <- stringency_us %>% mutate(date = as.Date(as.character(Date), '%Y%m%d')) %>%
  select(date, RegionName, StringencyIndex) %>%
  `colnames<-`c('date', 'state_name', 'stringency_index')

#Compute monthly change of stringency index
stringency_us <- stringency_us %>% mutate(delta_stringency_index = NA)
states <- unique(stringency_us$state_name)
for(i in 1:length(states)){
  state <- states[i]
}

```

```

str <- stringency_us %>% filter(state_name == state) %>%
  select(stringency_index)
delta_str <- unname(unlist(str - lag(str, n=30)))

stringency_us[stringency_us$state_name == state,]$delta_stringency_index <- delta_str
}

#####
##### Merge all data #####
#####

#Merge employment and mobility

data <- merge(employment_data, mobility, by.x=c('date','state_name'))
data <- merge(data, stringency_us, by.x = c('date','state_name'))

modified_data <- data
colnames(modified_data)[4] <- c("change_in_emp")
colnames(modified_data)[5] <- c("Sector")

#####

##### Write csv #####
#####

write.csv(modified_data, file = "covid_econ.csv")

```

### Note: Seasonal adjustment of Google Mobility Data

Since February of 2020, Google began publishing local mobility reports, which “show movement trends over time ordered by geographical areas and classified in various categories of places, such as shops and leisure spaces, supermarkets and pharmacies, parks , transportation stations, workplaces and residential areas”.

We chose to work with the Residential Mobility index, considering it to be the most representative index that represents the effect of Stay-at-home orders.

Residential Mobility index is a non-seasonally adjusted daily data. So, in order to make it comparable with the seasonally adjusted employment data, we used an additive time series model to eliminate the weekly seasonality factor from the series. Below is one example for Alabama.

```

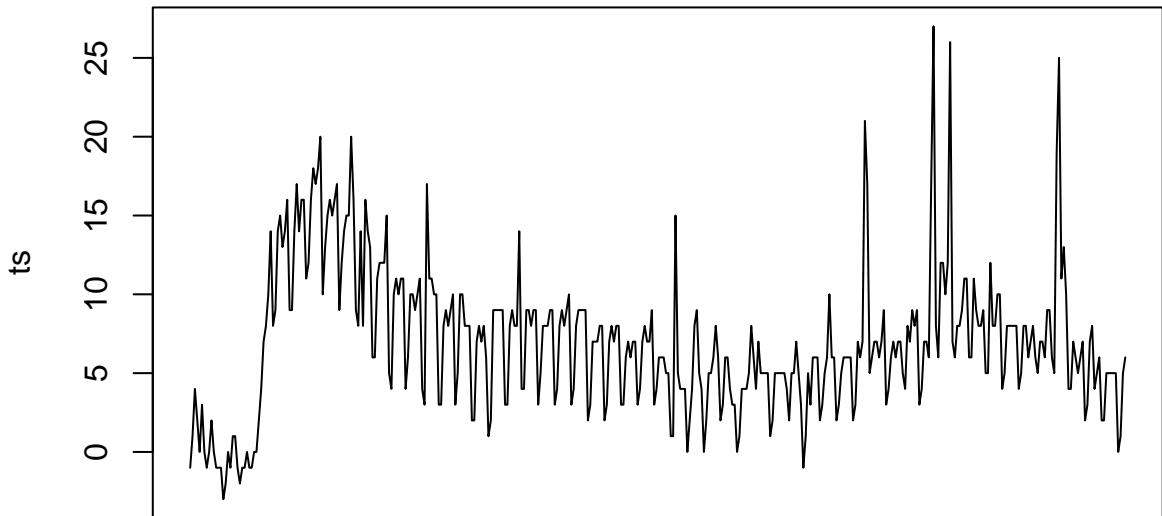
## Example for Alabama (weekly seasonal adjusted)
example <- mobility %>% filter(state_name == 'Alabama') %>% select(date, Mob_res)

ts <- ts(example$Mob_res, frequency = 7, start = c(2020,as.numeric(format(as.Date('2020-02-15'), "%j"))))

plot(ts, xaxt='n', main='Alabama Residential Mobility Index (Original Time Series)')

```

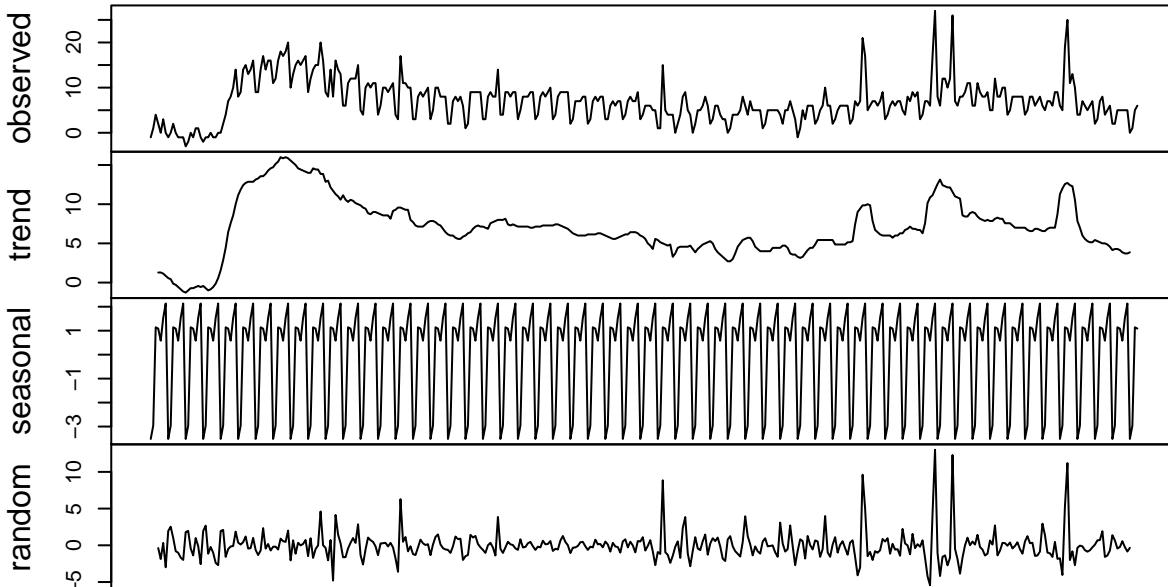
## Alabama Residential Mobility Index (Original Time Series)



Time

```
decompose_ts <- decompose(ts, 'additive')
plot(decompose_ts, xaxt='n')
```

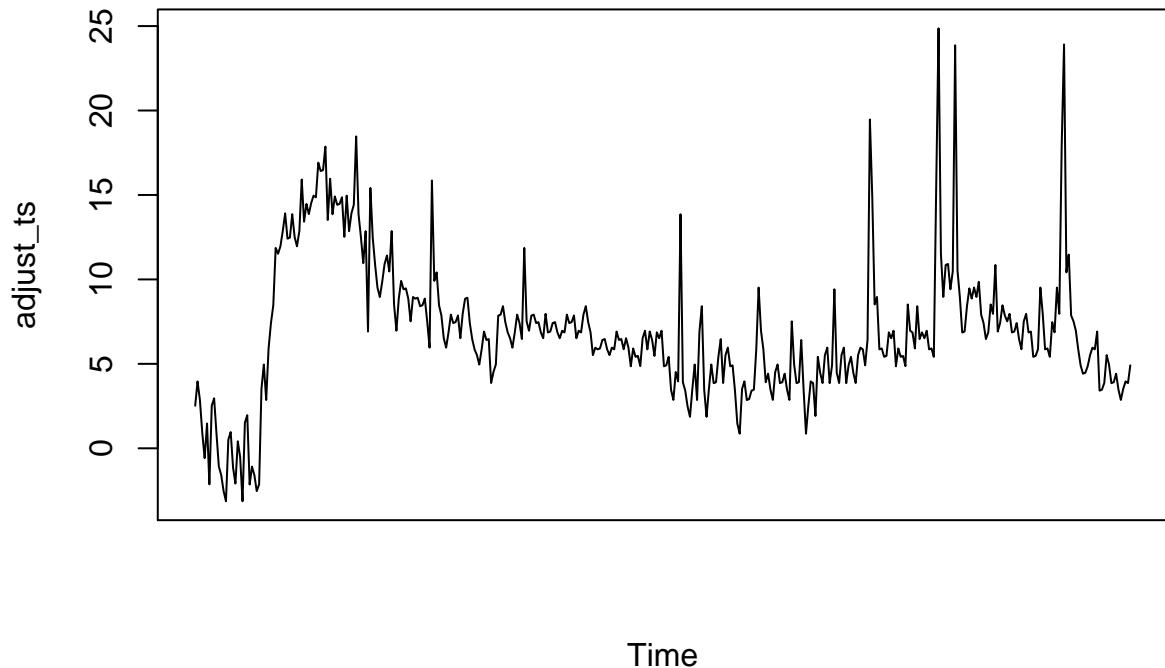
### Decomposition of additive time series



Time

```
adjust_ts <- ts - decompose_ts$seasonal
plot(adjust_ts, xaxt='n', main='Alabama Residential Mobility Index (Seasonal Adjusted Time Series)')
```

## Alabama Residential Mobility Index (Seasonal Adjusted Time Series)



## Data Manipulation

```
#=====
#data manipulation
=====

modified_data_ols <- read.csv("covid_econ.csv")

#drop state=DC from data set
modified_data_ols <- modified_data_ols[!(modified_data_ols$state_name=="District of Columbia"),]

#add squared term
modified_data_ols$Mob_res_SA_sq <- (modified_data_ols$Mob_res_SA)^2

#add exponential term
modified_data_ols$Mob_res_SA_exp <- exp(modified_data_ols$Mob_res_SA)

#sector as categorical variable
modified_data_ols$Sector <- factor(modified_data_ols$Sector)

#dates when mobility_exp is greater than the 95th percentile
unique(modified_data_ols$date[modified_data_ols$Mob_res_SA_exp > quantile(modified_data_ols$Mob_res_SA, 0.95)]))

## [1] "2020-03-31" "2020-04-30" "2021-01-31"

extreme_dates <-
  unique(modified_data_ols$date[modified_data_ols$Mob_res_SA_exp > quantile(modified_data_ols$Mob_res_SA, 0.95)])

#exponentiate mobility for just the leisure&hospitality sector
#square mobility for the rest
modified_data_ols$f2_Mob_res_SA <- modified_data_ols$Mob_res_SA_sq

modified_data_ols$f2_Mob_res_SA[modified_data_ols$Sector == "Leisure and Hospitality"] <-
  modified_data_ols$Mob_res_SA_exp[modified_data_ols$Sector == "Leisure and Hospitality"]
```

## 2. Plots

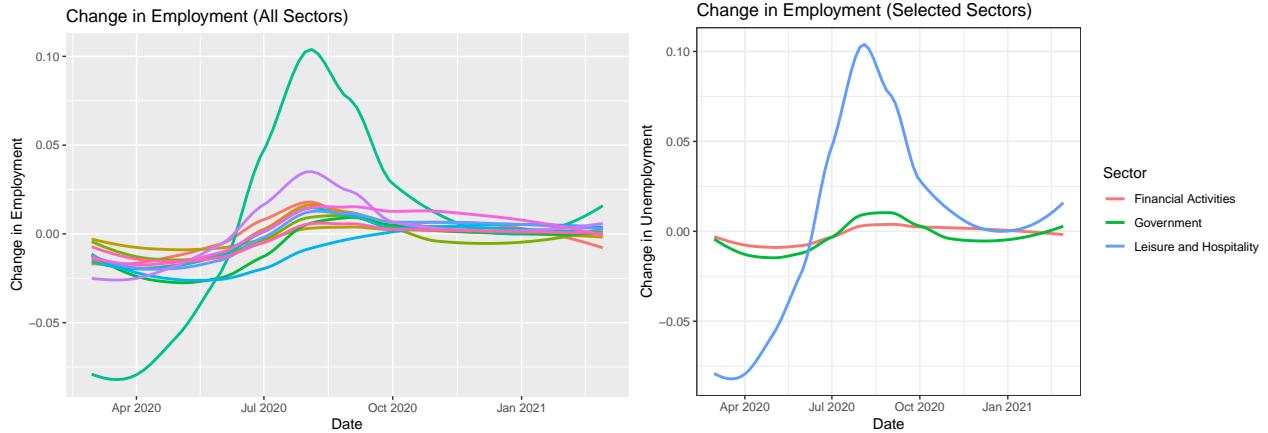
### Change in Employment by Sector

We see how employments in different sectors / different states were affected differently by COVID-19.

```
ggplot(data = modified_data,
       aes(x = date, y = change_in_emp, color = Sector)) +
  geom_smooth(se = FALSE) +
  theme(legend.position = "none") +
  labs(x = "Date", y = "Change in Employment", title = "Change in Employment (All Sectors)")

modified_data2 <-
  modified_data[modified_data$Sector == "Leisure and Hospitality" |
    modified_data$Sector == "Financial Activities" |
    modified_data$Sector == "Government", ]

ggplot(data = modified_data2,
       aes(x = date, y = change_in_emp, color = Sector)) +
  geom_smooth(se = FALSE) + theme(legend.position = "bottom") +
  theme_bw() +
  labs(x = "Date", y = "Change in Unemployment", title = "Change in Employment (Selected Sectors)")
```



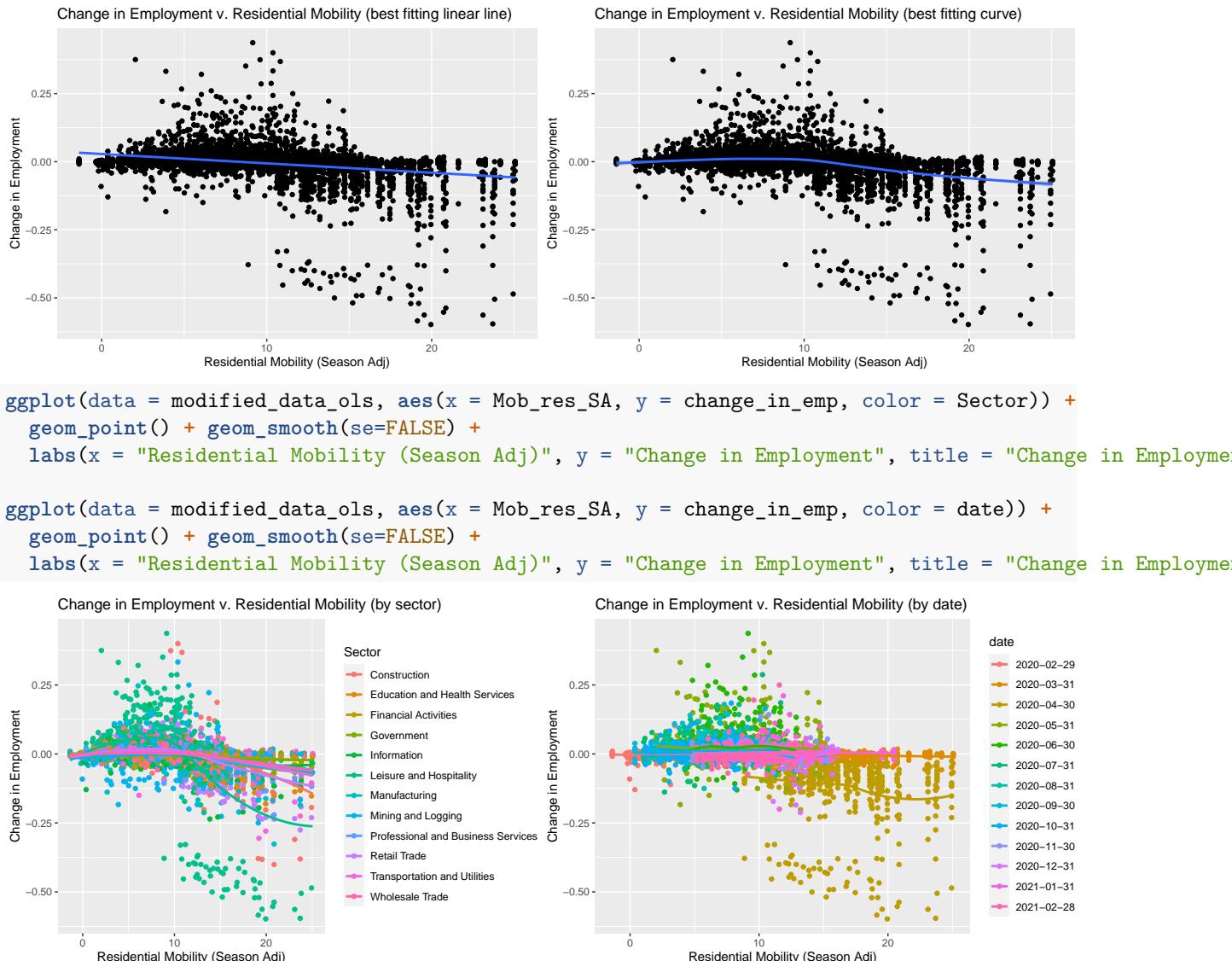
### Scatterplots - Mobility

Though we see a negative relationship between *change in employment* and *residential mobility*, we observe that a linear model would not be sufficient to explain the relationship between mobility and employment. The linear model cannot capture the significant negative change in employment that takes place when residential mobility is high.

By color coding, we noticed an interesting pattern. While for all other dates and all other sectors showed somewhat linear relationship (or slightly quadratic relationship) between mobility and change in employment, 2020-04-30 and Leisure&Hospitality Sector data showed a completely different relationship.

```
#Mobility Seasonally Adj
ggplot(data = modified_data_ols, aes(x = Mob_res_SA, y = change_in_emp)) +
  geom_point() + geom_smooth(method = lm) +
  labs(x = "Residential Mobility (Season Adj)", y = "Change in Employment", title = "Change in Employment")

ggplot(data = modified_data_ols, aes(x = Mob_res_SA, y = change_in_emp)) +
  geom_point() + geom_smooth() +
  labs(x = "Residential Mobility (Season Adj)", y = "Change in Employment", title = "Change in Employment")
```

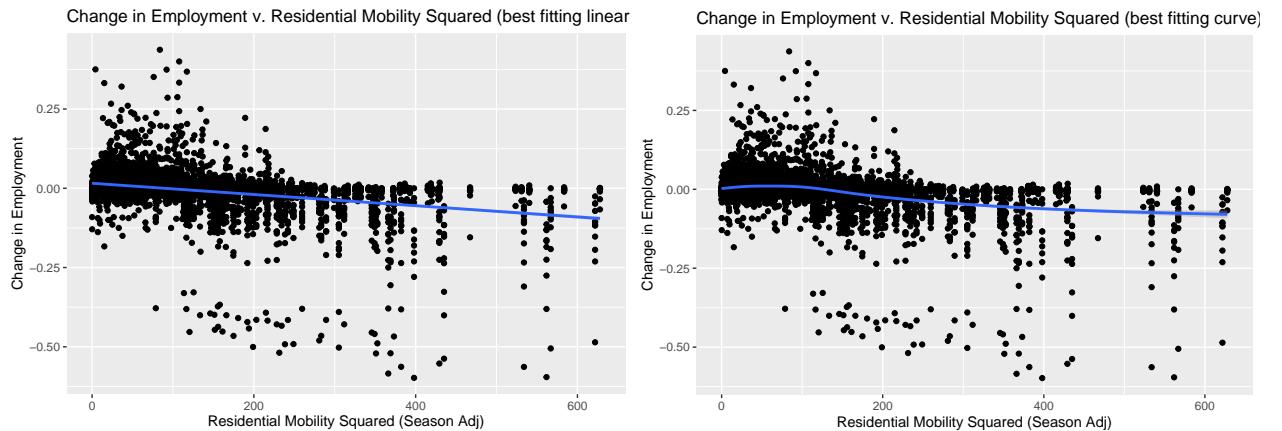


To account for this, we first tried applying mobility squared for all the data points. The best-fitting curve became a bit more linear, but the linear model still could not capture the significant negative change in employment that takes place when residential mobility is high.

By color coding, we could identify the nature of these data points with significantly negative *change in employment*. As predicted, these points were an intersection of Leisure and Hospitality Sector, and date=2020-04-30. Notice the curves for Leisure and date=2020-04-30 are still curved. This suggests a linear fit would not be sufficient for explaining the relationship between mobility squared and change in employment.

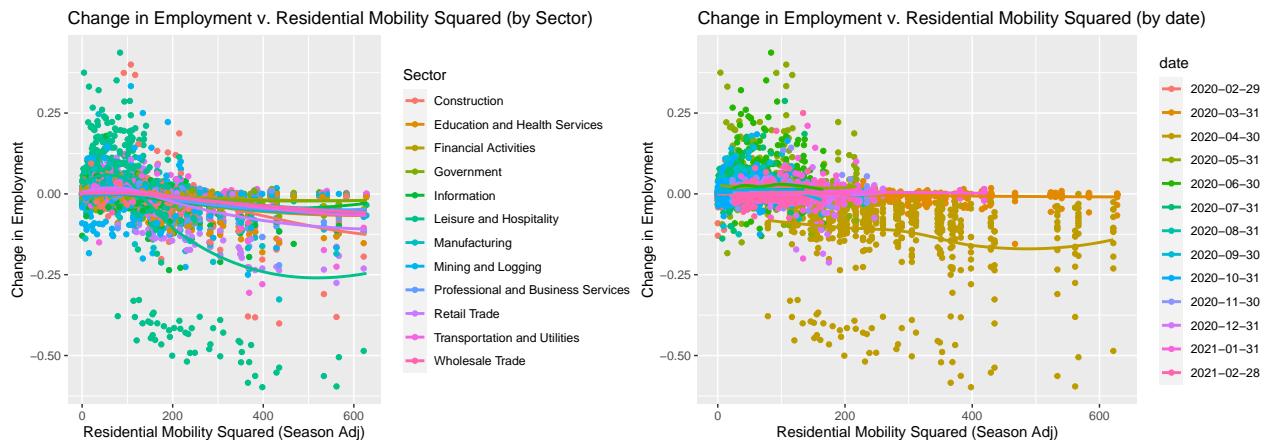
```
#Mobility Seasonally Adj Squared
ggplot(data = modified_data_ols, aes(x = Mob_res_SA_sq, y = change_in_emp)) + geom_point() + geom_smooth()
  labs(x = "Residential Mobility Squared (Season Adj)", y = "Change in Employment", title = "Change in Employment v. Residential Mobility Squared (Season Adj)")

ggplot(data = modified_data_ols, aes(x = Mob_res_SA_sq, y = change_in_emp)) + geom_point() + geom_smooth()
  labs(x = "Residential Mobility Squared (Season Adj)", y = "Change in Employment", title = "Change in Employment v. Residential Mobility Squared (Season Adj) (Best Fitting Curve)")
```



```
ggplot(data = modified_data_ols, aes(x = Mob_res_SA_sq, y = change_in_emp, color = Sector)) + geom_point()
  labs(x = "Residential Mobility Squared (Season Adj)", y = "Change in Employment", title = "Change in Employment v. Residential Mobility Squared (by Sector)")

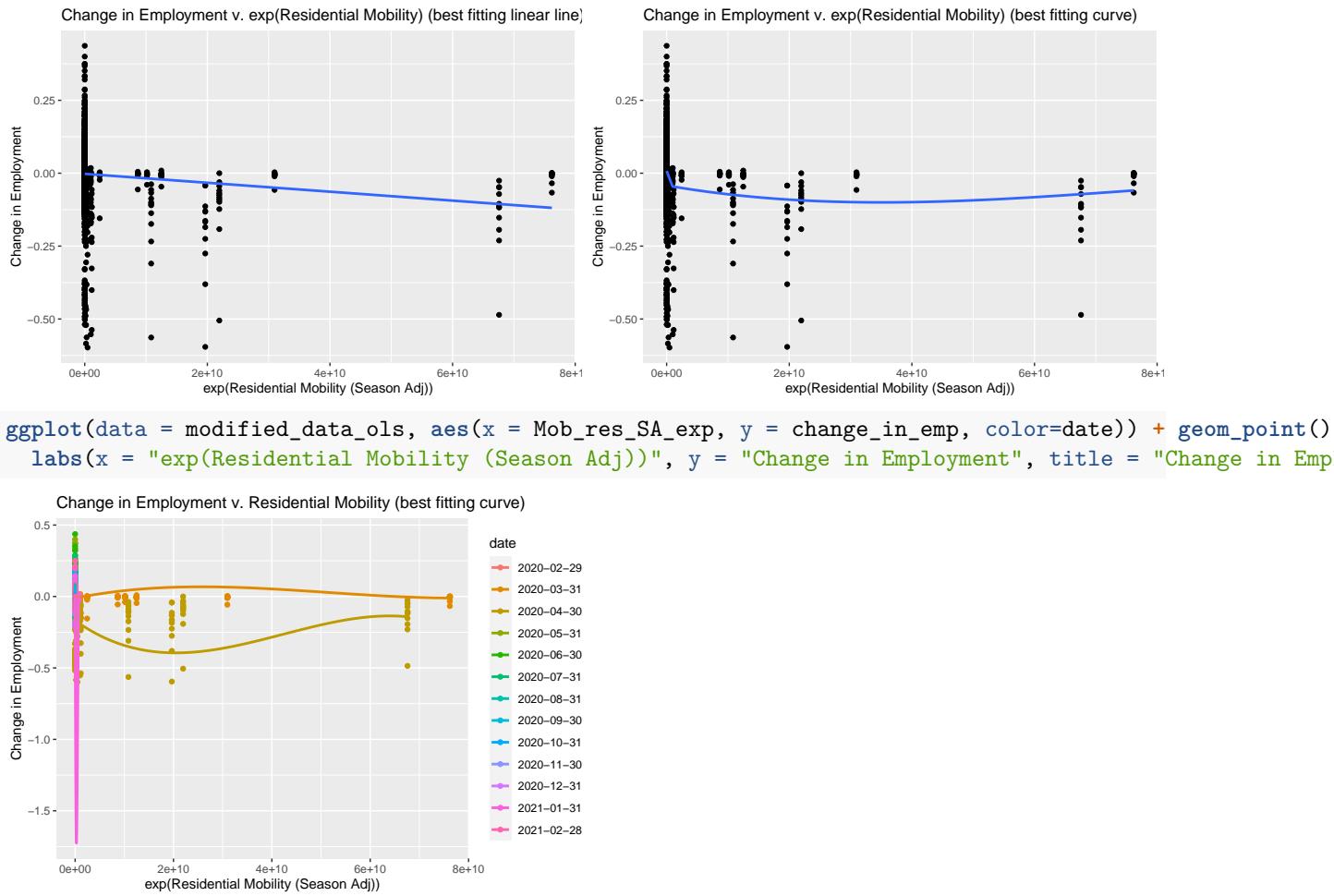
ggplot(data = modified_data_ols, aes(x = Mob_res_SA_sq, y = change_in_emp, color = date)) + geom_point()
  labs(x = "Residential Mobility Squared (Season Adj)", y = "Change in Employment", title = "Change in Employment v. Residential Mobility Squared (by date)")
```



So, we tried observing different functional forms of Mobility (i.e.,  $f(\text{mobility})$ ). Here we look at ' $\exp(\text{mobility})$ '. We noticed, by looking at the steep kink of the curve, that the performance of a linear model was worse here than the squared mobility above. Also, we noticed what's driving the curve were the outliers at the far right. Again, by color coding, we could identify the nature of these outliers - they were an intersection of state=New York and date=2020-04-30.

```
#Mobility Seasonally Adj Squared
ggplot(data = modified_data_ols, aes(x = Mob_res_SA_exp, y = change_in_emp)) + geom_point() + geom_smooth()
  labs(x = "exp(Residential Mobility (Season Adj))", y = "Change in Employment", title = "Change in Employment v. exp(Residential Mobility (Season Adj))")

ggplot(data = modified_data_ols, aes(x = Mob_res_SA_exp, y = change_in_emp)) + geom_point() + geom_smooth()
  labs(x = "exp(Residential Mobility (Season Adj))", y = "Change in Employment", title = "Change in Employment v. exp(Residential Mobility (Season Adj))")
```



## Scatterplots - Stringency

We see a fairly linear relationship between stringency and change in employment. We could have very well tried polynomial terms of stringency, but that seemed unnecessary given how flat the regression line is.

```
#Stringency
ggplot(data = modified_data_ols, aes(x = stringency_index, y = change_in_emp)) + geom_point() + geom_smooth()
  labs(x = "Stringency Index", y = "Change in Employment", title = "Change in Employment v. Stringency")
```

```
#Stringency
ggplot(data = modified_data_ols, aes(x = stringency_index, y = change_in_emp)) + geom_point() + geom_smooth()
  labs(x = "Stringency Index", y = "Change in Employment", title = "Change in Employment v. Stringency")
```



### Scatterplots - Relationship between explanatory variables (Stringency and Mobility)

We see a fairly linear correlation between stringency and mobility (correlation coefficient = 0.67).

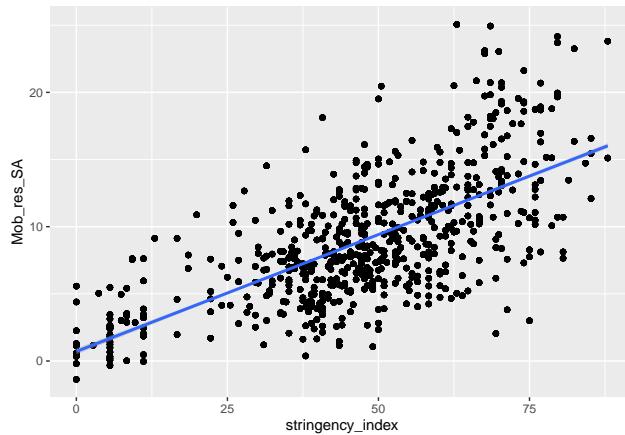
*#Stringency and Mobility*

```
ggplot(data = modified_data_ols, aes(x = stringency_index, y = Mob_res_SA)) + geom_point() + geom_smooth()

#drop NAs to calc correlation
modified_data_ols_no_na <- modified_data_ols[!is.na(modified_data_ols$stringency_index),]

#correlation
cor(modified_data_ols_no_na$stringency_index, modified_data_ols_no_na$Mob_res_SA)

## [1] 0.6744251
```



### 3. Model Fitting

Based on our data analysis, we believe that the change in employment for a given state, time, and sector can be predicted using mobility, stringency, and sector composition in the following formulation:

$$\Delta Emp_{s,t,sec} = \alpha + \mu_{sec} + \beta_1 f(Mob_{s,t}) + \beta_2 String_{s,t} + \beta_3 f(Mob_{s,t}) * Sector_{s,t} + \beta_4 String_{s,t} * Sector_{s,t} + X_{s,t} * \Gamma + \epsilon_{s,t,sec}$$

where

- $\Delta Emp_{s,t,sec}$  is the change in employment for a given state (in the U.S.), at a given time (limited to the pandemic era; 02/01/2020-06/01/2021), for a given sector (one of 12)
- $\alpha$  is the global intercept and  $\mu_{sec}$  is the local intercept that varies for each sector.
- $Mob_{s,t}$  is the seasonally adjusted residential mobility of a given state at a given time (so higher when people move less). We think different sector takes a different functional form of Mobility. In this report, we tried:
  - ‘mobility’ for all the data set
  - ‘mobility squared’ for all the data set
  - ‘exp(mobility)’ for all the data set
  - ‘exp(mobility)’ for Leisure&Hospitality Sector, but ‘mobility squared’ for the rest of the sector.
- $String_{s,t}$  is the stringency level of a given state at a given time (higher index means more stringent)
- $Sector_{s,t}$  is the categorical variable.
- $X_{s,t}$  encapsulates vectors of control variables such as the number of new COVID-19 cases in a given state at a given time. Note that, in this report, control variables are not yet included.

Using data from 2/29/2020 to 1/31/2021, we will derive coefficients in our model using k-fold cross validation ( $k=10$ ; the choice of  $k$  here is arbitrary). Then, using these coefficients, we plan to predict the 2021 April/May change in unemployment for a given state for a given sector. Finally, we will compute the RMSE between the predicted values and the actual values to assess the performance of our model.

In this project, we do not anticipate our model to necessarily produce the *lowest* RMSE. To minimize RMSE and let our model have the most predictive power, we should include as many features as possible that have any relevance to unemployment. However, taking this approach, the coefficient for the variables of our interest would lose most of their interpretability. Since we are interested in the interpretation of these coefficients as well as the predictive power of our model, we will instead take the approach of including only a handful of control variables in our model.

In this report, instead of running the cross validation, we ran a linear model using the entire data set (using the formulation above) to first assess the fit of our model and the significance of the coefficient of each variable of our interest.

## 4. OLS Table for the model

OLS1: With  $f(\text{mobility}) = \text{mobility}$ , for all sectors

```
ols1 <- lm(change_in_emp ~ Mob_res_SA*Sector + stringency_index*Sector , data = modified_data_ols)
summary(ols1)

##
## Call:
## lm(formula = change_in_emp ~ Mob_res_SA * Sector + stringency_index *
##     Sector, data = modified_data_ols)
##
## Residuals:
##      Min        1Q    Median        3Q       Max
## -0.44465 -0.00930  0.00300  0.01451  0.43501
##
## Coefficients:
##                               Estimate Std. Error
## (Intercept)                1.903e-02 5.451e-03
## Mob_res_SA                 -4.324e-03 5.537e-04
## SectorEducation and Health Services 2.343e-03 7.630e-03
## SectorFinancial Activities -1.077e-02 7.630e-03
## SectorGovernment            -8.932e-03 7.630e-03
## SectorInformation           -3.466e-03 7.659e-03
## SectorLeisure and Hospitality 8.214e-02 7.630e-03
## SectorManufacturing          2.106e-03 7.630e-03
## SectorMining and Logging    -1.459e-02 7.761e-03
## SectorProfessional and Business Services 1.251e-03 7.630e-03
## SectorRetail Trade           1.326e-02 7.630e-03
## SectorTransportation and Utilities -1.879e-03 7.630e-03
## SectorWholesale Trade        -6.762e-03 7.630e-03
## stringency_index              3.865e-04 1.443e-04
## Mob_res_SA:SectorEducation and Health Services 9.761e-04 7.760e-04
## Mob_res_SA:SectorFinancial Activities 3.385e-03 7.760e-04
## Mob_res_SA:SectorGovernment   3.135e-03 7.760e-04
## Mob_res_SA:SectorInformation 2.155e-03 7.797e-04
## Mob_res_SA:SectorLeisure and Hospitality -1.319e-02 7.760e-04
## Mob_res_SA:SectorManufacturing 1.105e-03 7.760e-04
## Mob_res_SA:SectorMining and Logging 3.372e-03 7.849e-04
## Mob_res_SA:SectorProfessional and Business Services 1.431e-03 7.760e-04
## Mob_res_SA:SectorRetail Trade -1.153e-03 7.760e-04
## Mob_res_SA:SectorTransportation and Utilities 1.353e-03 7.760e-04
## Mob_res_SA:SectorWholesale Trade 2.394e-03 7.760e-04
## SectorEducation and Health Services:stringency_index -2.648e-04 2.010e-04
## SectorFinancial Activities:stringency_index -4.215e-04 2.010e-04
## SectorGovernment:stringency_index -4.545e-04 2.010e-04
## SectorInformation:stringency_index -4.536e-04 2.023e-04
## SectorLeisure and Hospitality:stringency_index 7.711e-04 2.010e-04
## SectorManufacturing:stringency_index -2.778e-04 2.010e-04
## SectorMining and Logging:stringency_index -4.637e-04 2.050e-04
## SectorProfessional and Business Services:stringency_index -3.065e-04 2.010e-04
## SectorRetail Trade:stringency_index -2.342e-05 2.010e-04
## SectorTransportation and Utilities:stringency_index -1.647e-04 2.010e-04
## SectorWholesale Trade:stringency_index -3.521e-04 2.010e-04
```

```

##                                     t value Pr(>|t|)
## (Intercept)                      3.491 0.000484 ***
## Mob_res_SA                       -7.810 6.50e-15 ***
## SectorEducation and Health Services          0.307 0.758762
## SectorFinancial Activities          -1.411 0.158177
## SectorGovernment                   -1.171 0.241804
## SectorInformation                  -0.453 0.650921
## SectorLeisure and Hospitality      10.765 < 2e-16 ***
## SectorManufacturing                 0.276 0.782557
## SectorMining and Logging           -1.880 0.060206 .
## SectorProfessional and Business Services   0.164 0.869770
## SectorRetail Trade                  1.738 0.082195 .
## SectorTransportation and Utilities    -0.246 0.805467
## SectorWholesale Trade                -0.886 0.375563
## stringency_index                   2.679 0.007402 **
## Mob_res_SA:SectorEducation and Health Services 1.258 0.208496
## Mob_res_SA:SectorFinancial Activities 4.362 1.31e-05 ***
## Mob_res_SA:SectorGovernment        4.040 5.41e-05 ***
## Mob_res_SA:SectorInformation       2.763 0.005735 **
## Mob_res_SA:SectorLeisure and Hospitality -16.999 < 2e-16 ***
## Mob_res_SA:SectorManufacturing     1.424 0.154356
## Mob_res_SA:SectorMining and Logging 4.296 1.76e-05 ***
## Mob_res_SA:SectorProfessional and Business Services 1.844 0.065270 .
## Mob_res_SA:SectorRetail Trade      -1.486 0.137442
## Mob_res_SA:SectorTransportation and Utilities 1.744 0.081284 .
## Mob_res_SA:SectorWholesale Trade   3.086 0.002039 **
## SectorEducation and Health Services:stringency_index -1.317 0.187877
## SectorFinancial Activities:stringency_index -2.097 0.036063 *
## SectorGovernment:stringency_index     -2.261 0.023793 *
## SectorInformation:stringency_index    -2.242 0.024964 *
## SectorLeisure and Hospitality:stringency_index 3.836 0.000126 ***
## SectorManufacturing:stringency_index   -1.382 0.167012
## SectorMining and Logging:stringency_index -2.263 0.023694 *
## SectorProfessional and Business Services:stringency_index -1.525 0.127422
## SectorRetail Trade:stringency_index     -0.116 0.907265
## SectorTransportation and Utilities:stringency_index -0.819 0.412594
## SectorWholesale Trade:stringency_index   -1.751 0.079901 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05 on 7674 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared: 0.1962, Adjusted R-squared: 0.1925
## F-statistic: 53.51 on 35 and 7674 DF, p-value: < 2.2e-16

```

### OLS2: With $f(\text{mobility}) = \text{mobility squared}$ , for all sectors

We see that R squared improves quite significantly!

```

ols2 <- lm(change_in_emp ~ Mob_res_SA_sq*Sector + stringency_index*Sector , data = modified_data_ols)
summary(ols2)

```

```

##
## Call:
## lm(formula = change_in_emp ~ Mob_res_SA_sq * Sector + stringency_index *
## 
```

```

##      Sector, data = modified_data_ols)
##
## Residuals:
##      Min       1Q    Median      3Q     Max
## -0.45807 -0.00908  0.00130  0.01264  0.41719
##
## Coefficients:
##                               Estimate Std. Error
## (Intercept)                3.773e-03  5.419e-03
## Mob_res_SA_sq              -2.282e-04 2.222e-05
## SectorEducation and Health Services 6.544e-03 7.581e-03
## SectorFinancial Activities 1.372e-03 7.581e-03
## SectorGovernment            2.714e-03 7.581e-03
## SectorInformation           4.598e-03 7.610e-03
## SectorLeisure and Hospitality 4.165e-02 7.581e-03
## SectorManufacturing         6.564e-03 7.581e-03
## SectorMining and Logging   -3.675e-03 7.715e-03
## SectorProfessional and Business Services 7.008e-03 7.581e-03
## SectorRetail Trade          1.014e-02 7.581e-03
## SectorTransportation and Utilities 2.970e-03 7.581e-03
## SectorWholesale Trade      1.762e-03 7.581e-03
## stringency_index            3.924e-04 1.277e-04
## Mob_res_SA_sq:SectorEducation and Health Services 6.667e-05 3.104e-05
## Mob_res_SA_sq:SectorFinancial Activities 1.826e-04 3.104e-05
## Mob_res_SA_sq:SectorGovernment 1.767e-04 3.104e-05
## Mob_res_SA_sq:SectorInformation 1.223e-04 3.123e-05
## Mob_res_SA_sq:SectorLeisure and Hospitality -5.778e-04 3.104e-05
## Mob_res_SA_sq:SectorManufacturing 6.973e-05 3.104e-05
## Mob_res_SA_sq:SectorMining and Logging 1.595e-04 3.152e-05
## Mob_res_SA_sq:SectorProfessional and Business Services 8.962e-05 3.104e-05
## Mob_res_SA_sq:SectorRetail Trade -4.152e-05 3.104e-05
## Mob_res_SA_sq:SectorTransportation and Utilities 7.370e-05 3.104e-05
## Mob_res_SA_sq:SectorWholesale Trade 1.283e-04 3.104e-05
## SectorEducation and Health Services:stringency_index -3.155e-04 1.776e-04
## SectorFinancial Activities:stringency_index -4.390e-04 1.776e-04
## SectorGovernment:stringency_index -4.960e-04 1.776e-04
## SectorInformation:stringency_index -4.853e-04 1.788e-04
## SectorLeisure and Hospitality:stringency_index 4.008e-04 1.776e-04
## SectorManufacturing:stringency_index -3.163e-04 1.776e-04
## SectorMining and Logging:stringency_index -4.071e-04 1.814e-04
## SectorProfessional and Business Services:stringency_index -3.545e-04 1.776e-04
## SectorRetail Trade:stringency_index -8.443e-05 1.776e-04
## SectorTransportation and Utilities:stringency_index -1.733e-04 1.776e-04
## SectorWholesale Trade:stringency_index -3.614e-04 1.776e-04
## 
## (Intercept)                      0.696  0.48627
## Mob_res_SA_sq                     -10.270 < 2e-16 ***
## SectorEducation and Health Services 0.863  0.38805
## SectorFinancial Activities        0.181  0.85642
## SectorGovernment                  0.358  0.72033
## SectorInformation                 0.604  0.54575
## SectorLeisure and Hospitality    5.493  4.07e-08 ***
## SectorManufacturing               0.866  0.38662
## SectorMining and Logging          -0.476  0.63386

```

```

## SectorProfessional and Business Services          0.924  0.35535
## SectorRetail Trade                            1.338  0.18103
## SectorTransportation and Utilities            0.392  0.69527
## SectorWholesale Trade                         0.232  0.81626
## stringency_index                             3.074  0.00212 **
## Mob_res_SA_sq:SectorEducation and Health Services 2.148  0.03176 *
## Mob_res_SA_sq:SectorFinancial Activities      5.883  4.19e-09 ***
## Mob_res_SA_sq:SectorGovernment                 5.694  1.29e-08 ***
## Mob_res_SA_sq:SectorInformation                3.917  9.04e-05 ***
## Mob_res_SA_sq:SectorLeisure and Hospitality   -18.615 < 2e-16 ***
## Mob_res_SA_sq:SectorManufacturing              2.247  0.02470 *
## Mob_res_SA_sq:SectorMining and Logging         5.062  4.24e-07 ***
## Mob_res_SA_sq:SectorProfessional and Business Services 2.887  0.00390 **
## Mob_res_SA_sq:SectorRetail Trade                -1.338  0.18105
## Mob_res_SA_sq:SectorTransportation and Utilities 2.375  0.01760 *
## Mob_res_SA_sq:SectorWholesale Trade             4.135  3.59e-05 ***
## SectorEducation and Health Services:stringency_index -1.776  0.07578 .
## SectorFinancial Activities:stringency_index      -2.471  0.01348 *
## SectorGovernment:stringency_index                 -2.792  0.00525 **
## SectorInformation:stringency_index                -2.714  0.00666 **
## SectorLeisure and Hospitality:stringency_index    2.256  0.02409 *
## SectorManufacturing:stringency_index               -1.780  0.07506 .
## SectorMining and Logging:stringency_index          -2.243  0.02490 *
## SectorProfessional and Business Services:stringency_index -1.996  0.04602 *
## SectorRetail Trade:stringency_index                  -0.475  0.63458
## SectorTransportation and Utilities:stringency_index -0.976  0.32923
## SectorWholesale Trade:stringency_index              -2.034  0.04195 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04858 on 7674 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.2413, Adjusted R-squared:  0.2378
## F-statistic: 69.73 on 35 and 7674 DF,  p-value: < 2.2e-16

```

### OLS3: With f(mobility) = exp(mobility), for all sectors

As expected, exp(mobility) is a poor linear predictor of change in employment - R squared is much worse than mobility squared!

```
ols3 <- lm(change_in_emp ~ Mob_res_SA_exp*Sector + stringency_index*Sector , data = modified_data_ols)
summary(ols3)
```

```

##
## Call:
## lm(formula = change_in_emp ~ Mob_res_SA_exp * Sector + stringency_index *
##     Sector, data = modified_data_ols)
##
## Residuals:
##      Min        1Q        Median       3Q        Max
## -0.54384 -0.00640  0.00389  0.01244  0.44543
##
## Coefficients:
##   Estimate Std. Error
## (Intercept) 1.424e-02  5.863e-03
```

```

## Mob_res_SA_exp           -1.986e-12  4.835e-13
## SectorEducation and Health Services      3.633e-03  8.204e-03
## SectorFinancial Activities          -6.959e-03  8.204e-03
## SectorGovernment            -5.212e-03  8.204e-03
## SectorInformation           -4.552e-04  8.237e-03
## SectorLeisure and Hospitality        7.019e-02  8.204e-03
## SectorManufacturing          3.712e-03  8.204e-03
## SectorMining and Logging         -1.130e-02  8.349e-03
## SectorProfessional and Business Services 3.191e-03  8.204e-03
## SectorRetail Trade             1.229e-02  8.204e-03
## SectorTransportation and Utilities     -2.014e-04  8.204e-03
## SectorWholesale Trade          -4.176e-03  8.204e-03
## stringency_index              -3.119e-04  1.157e-04
## Mob_res_SA_exp:SectorEducation and Health Services 7.509e-13  6.834e-13
## Mob_res_SA_exp:SectorFinancial Activities 1.640e-12  6.834e-13
## Mob_res_SA_exp:SectorGovernment       1.742e-12  6.834e-13
## Mob_res_SA_exp:SectorInformation      1.738e-12  6.835e-13
## Mob_res_SA_exp:SectorLeisure and Hospitality   -2.674e-12  6.834e-13
## Mob_res_SA_exp:SectorManufacturing      1.005e-12  6.834e-13
## Mob_res_SA_exp:SectorMining and Logging      1.090e-12  6.838e-13
## Mob_res_SA_exp:SectorProfessional and Business Services 1.118e-12  6.834e-13
## Mob_res_SA_exp:SectorRetail Trade        -7.869e-14  6.834e-13
## Mob_res_SA_exp:SectorTransportation and Utilities 8.833e-13  6.834e-13
## Mob_res_SA_exp:SectorWholesale Trade      1.062e-12  6.834e-13
## SectorEducation and Health Services:stringency_index -1.148e-04  1.609e-04
## SectorFinancial Activities:stringency_index 1.231e-04  1.609e-04
## SectorGovernment:stringency_index        4.384e-05  1.609e-04
## SectorInformation:stringency_index       -1.263e-04  1.620e-04
## SectorLeisure and Hospitality:stringency_index -1.449e-03  1.609e-04
## SectorManufacturing:stringency_index      -1.123e-04  1.609e-04
## SectorMining and Logging:stringency_index 9.497e-05  1.647e-04
## SectorProfessional and Business Services:stringency_index -8.742e-05  1.609e-04
## SectorRetail Trade:stringency_index        -2.210e-04  1.609e-04
## SectorTransportation and Utilities:stringency_index 4.717e-05  1.609e-04
## SectorWholesale Trade:stringency_index      3.600e-05  1.609e-04
## t value Pr(>|t|)
## (Intercept)                         2.428  0.01520 *
## Mob_res_SA_exp                      -4.108  4.03e-05 ***
## SectorEducation and Health Services  0.443  0.65794
## SectorFinancial Activities          -0.848  0.39636
## SectorGovernment                    -0.635  0.52525
## SectorInformation                  -0.055  0.95593
## SectorLeisure and Hospitality      8.555  < 2e-16 ***
## SectorManufacturing                 0.452  0.65094
## SectorMining and Logging          -1.353  0.17604
## SectorProfessional and Business Services 0.389  0.69730
## SectorRetail Trade                  1.498  0.13428
## SectorTransportation and Utilities -0.025  0.98041
## SectorWholesale Trade              -0.509  0.61081
## stringency_index                   -2.695  0.00704 **
## Mob_res_SA_exp:SectorEducation and Health Services 1.099  0.27189
## Mob_res_SA_exp:SectorFinancial Activities 2.400  0.01642 *
## Mob_res_SA_exp:SectorGovernment      2.549  0.01081 *
## Mob_res_SA_exp:SectorInformation     2.543  0.01100 *

```

```

## Mob_res_SA_exp:SectorLeisure and Hospitality      -3.913 9.21e-05 ***
## Mob_res_SA_exp:SectorManufacturing               1.471  0.14142
## Mob_res_SA_exp:SectorMining and Logging          1.594  0.11098
## Mob_res_SA_exp:SectorProfessional and Business Services 1.636  0.10195
## Mob_res_SA_exp:SectorRetail Trade                -0.115  0.90833
## Mob_res_SA_exp:SectorTransportation and Utilities 1.293  0.19621
## Mob_res_SA_exp:SectorWholesale Trade             1.553  0.12037
## SectorEducation and Health Services:stringency_index -0.713  0.47556
## SectorFinancial Activities:stringency_index       0.765  0.44407
## SectorGovernment:stringency_index                 0.272  0.78526
## SectorInformation:stringency_index                -0.780  0.43564
## SectorLeisure and Hospitality:stringency_index     -9.006 < 2e-16 ***
## SectorManufacturing:stringency_index              -0.698  0.48529
## SectorMining and Logging:stringency_index          0.577  0.56422
## SectorProfessional and Business Services:stringency_index -0.543  0.58686
## SectorRetail Trade:stringency_index                -1.374  0.16952
## SectorTransportation and Utilities:stringency_index 0.293  0.76936
## SectorWholesale Trade:stringency_index             0.224  0.82295
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05374 on 7674 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.07149,   Adjusted R-squared:  0.06726
## F-statistic: 16.88 on 35 and 7674 DF,  p-value: < 2.2e-16

```

#### OLS4: With $f(\text{mobility}) = \exp(\text{mobility})$ , for Leisure&Hospitality while $f(\text{mobility}) = \text{mobility squared}$ , for all other sectors

The R-squared is worse than ' $f(\text{mobility}) = \text{mobility squared}$ , for all sectors'. It seems like  $\exp(\text{mobility})$  is not the correct functional form after all. (this is somewhat expected based on the sine-like curve that fits Leisure and Mobility plot)

```
ols4 <- lm(change_in_emp ~ f2_Mob_res_SA*Sector + stringency_index*Sector , data = modified_data_ols)
summary(ols4)
```

```

##
## Call:
## lm(formula = change_in_emp ~ f2_Mob_res_SA * Sector + stringency_index *
##   Sector, data = modified_data_ols)
##
## Residuals:
##      Min        1Q        Median         3Q        Max
## -0.54384 -0.00786  0.00187  0.01275  0.44543
##
## Coefficients: (1 not defined because of singularities)
##                               Estimate Std. Error
## (Intercept)                  1.615e-02 5.737e-03
## f2_Mob_res_SA                -4.660e-12 4.740e-13
## SectorEducation and Health Services -5.829e-03 8.127e-03
## SectorFinancial Activities    -1.100e-02 8.127e-03
## SectorGovernment              -9.659e-03 8.127e-03
## SectorInformation              -7.775e-03 8.159e-03
## SectorLeisure and Hospitality  6.828e-02 8.040e-03
## SectorManufacturing            -5.809e-03 8.127e-03

```

## SectorMining and Logging	-1.605e-02	8.274e-03
## SectorProfessional and Business Services	-5.366e-03	8.127e-03
## SectorRetail Trade	-2.232e-03	8.127e-03
## SectorTransportation and Utilities	-9.403e-03	8.127e-03
## SectorWholesale Trade	-1.061e-02	8.127e-03
## stringency_index	-3.704e-04	1.127e-04
## f2_Mob_res_SA:SectorEducation and Health Services	-1.615e-04	2.353e-05
## f2_Mob_res_SA:SectorFinancial Activities	-4.558e-05	2.353e-05
## f2_Mob_res_SA:SectorGovernment	-5.145e-05	2.353e-05
## f2_Mob_res_SA:SectorInformation	-1.059e-04	2.383e-05
## f2_Mob_res_SA:SectorLeisure and Hospitality	NA	NA
## f2_Mob_res_SA:SectorManufacturing	-1.585e-04	2.353e-05
## f2_Mob_res_SA:SectorMining and Logging	-6.865e-05	2.427e-05
## f2_Mob_res_SA:SectorProfessional and Business Services	-1.386e-04	2.353e-05
## f2_Mob_res_SA:SectorRetail Trade	-2.697e-04	2.353e-05
## f2_Mob_res_SA:SectorTransportation and Utilities	-1.545e-04	2.353e-05
## f2_Mob_res_SA:SectorWholesale Trade	-9.985e-05	2.353e-05
## SectorEducation and Health Services:stringency_index	4.474e-04	1.752e-04
## SectorFinancial Activities:stringency_index	3.238e-04	1.752e-04
## SectorGovernment:stringency_index	2.668e-04	1.752e-04
## SectorInformation:stringency_index	2.775e-04	1.766e-04
## SectorLeisure and Hospitality:stringency_index	-1.390e-03	1.573e-04
## SectorManufacturing:stringency_index	4.466e-04	1.752e-04
## SectorMining and Logging:stringency_index	3.558e-04	1.797e-04
## SectorProfessional and Business Services:stringency_index	4.084e-04	1.752e-04
## SectorRetail Trade:stringency_index	6.784e-04	1.752e-04
## SectorTransportation and Utilities:stringency_index	5.895e-04	1.752e-04
## SectorWholesale Trade:stringency_index	4.015e-04	1.752e-04
##	t value Pr(> t )	
## (Intercept)	2.815	0.004897 **
## f2_Mob_res_SA	-9.830	< 2e-16 ***
## SectorEducation and Health Services	-0.717	0.473242
## SectorFinancial Activities	-1.354	0.175875
## SectorGovernment	-1.188	0.234674
## SectorInformation	-0.953	0.340628
## SectorLeisure and Hospitality	8.493	< 2e-16 ***
## SectorManufacturing	-0.715	0.474741
## SectorMining and Logging	-1.940	0.052462 .
## SectorProfessional and Business Services	-0.660	0.509126
## SectorRetail Trade	-0.275	0.783633
## SectorTransportation and Utilities	-1.157	0.247285
## SectorWholesale Trade	-1.306	0.191686
## stringency_index	-3.286	0.001019 **
## f2_Mob_res_SA:SectorEducation and Health Services	-6.864	7.23e-12 ***
## f2_Mob_res_SA:SectorFinancial Activities	-1.937	0.052815 .
## f2_Mob_res_SA:SectorGovernment	-2.186	0.028818 *
## f2_Mob_res_SA:SectorInformation	-4.443	9.02e-06 ***
## f2_Mob_res_SA:SectorLeisure and Hospitality	NA	NA
## f2_Mob_res_SA:SectorManufacturing	-6.733	1.78e-11 ***
## f2_Mob_res_SA:SectorMining and Logging	-2.829	0.004684 **
## f2_Mob_res_SA:SectorProfessional and Business Services	-5.888	4.06e-09 ***
## f2_Mob_res_SA:SectorRetail Trade	-11.461	< 2e-16 ***
## f2_Mob_res_SA:SectorTransportation and Utilities	-6.565	5.55e-11 ***
## f2_Mob_res_SA:SectorWholesale Trade	-4.243	2.23e-05 ***

```

## SectorEducation and Health Services:stringency_index      2.553 0.010685 *
## SectorFinancial Activities:stringency_index            1.848 0.064599 .
## SectorGovernment:stringency_index                     1.523 0.127795
## SectorInformation:stringency_index                  1.572 0.116084
## SectorLeisure and Hospitality:stringency_index     -8.840 < 2e-16 ***
## SectorManufacturing:stringency_index                2.549 0.010822 *
## SectorMining and Logging:stringency_index           1.980 0.047789 *
## SectorProfessional and Business Services:stringency_index 2.331 0.019790 *
## SectorRetail Trade:stringency_index                 3.872 0.000109 ***
## SectorTransportation and Utilities:stringency_index 3.365 0.000770 ***
## SectorWholesale Trade:stringency_index              2.291 0.021962 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05274 on 7675 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.1054, Adjusted R-squared:  0.1015
## F-statistic:  26.6 on 34 and 7675 DF,  p-value: < 2.2e-16

```

### OLS5: To check, dropping mobility terms

Without the mobility terms, R-squared drops significantly. This verifies the observation from our plot above (where the regression line for change in ‘employment v. stringency’ was flat). Namely, stringency by itself is not a sufficient predictor of change in employment.

```

##
## Call:
## lm(formula = change_in_emp ~ stringency_index * Sector, data = modified_data_ols)
##
## Residuals:
##       Min        1Q        Median        3Q        Max
## -0.54000 -0.00642  0.00423  0.01284  0.44800
##
## Coefficients:
##                               Estimate Std. Error
## (Intercept)                   1.615e-02 5.899e-03
## stringency_index               -3.704e-04 1.159e-04
## SectorEducation and Health Services 2.812e-03 8.256e-03
## SectorFinancial Activities      -8.563e-03 8.256e-03
## SectorGovernment                -6.907e-03 8.256e-03
## SectorInformation                -2.138e-03 8.288e-03
## SectorLeisure and Hospitality    7.239e-02 8.256e-03
## SectorManufacturing                2.667e-03 8.256e-03
## SectorMining and Logging          -1.232e-02 8.400e-03
## SectorProfessional and Business Services 2.047e-03 8.256e-03
## SectorRetail Trade                  1.220e-02 8.256e-03
## SectorTransportation and Utilities -1.139e-03 8.256e-03
## SectorWholesale Trade                -5.270e-03 8.256e-03
## stringency_index:SectorEducation and Health Services -8.979e-05 1.612e-04
## stringency_index:SectorFinancial Activities         1.723e-04 1.612e-04
## stringency_index:SectorGovernment                  9.573e-05 1.612e-04
## stringency_index:SectorInformation                -7.473e-05 1.623e-04
## stringency_index:SectorLeisure and Hospitality     -1.517e-03 1.612e-04
## stringency_index:SectorManufacturing             -8.038e-05 1.612e-04
## stringency_index:SectorMining and Logging          1.265e-04 1.650e-04

```

```

## stringency_index:SectorProfessional and Business Services -5.247e-05 1.612e-04
## stringency_index:SectorRetail Trade -2.185e-04 1.612e-04
## stringency_index:SectorTransportation and Utilities 7.576e-05 1.612e-04
## stringency_index:SectorWholesale Trade 6.942e-05 1.612e-04
##
## t value Pr(>|t|)
## (Intercept) 2.737 0.00621 ***
## stringency_index -3.196 0.00140 ***
## SectorEducation and Health Services 0.341 0.73346
## SectorFinancial Activities -1.037 0.29968
## SectorGovernment -0.837 0.40289
## SectorInformation -0.258 0.79642
## SectorLeisure and Hospitality 8.768 < 2e-16 ***
## SectorManufacturing 0.323 0.74665
## SectorMining and Logging -1.467 0.14238
## SectorProfessional and Business Services 0.248 0.80418
## SectorRetail Trade 1.477 0.13966
## SectorTransportation and Utilities -0.138 0.89026
## SectorWholesale Trade -0.638 0.52327
## stringency_index:SectorEducation and Health Services -0.557 0.57751
## stringency_index:SectorFinancial Activities 1.069 0.28527
## stringency_index:SectorGovernment 0.594 0.55261
## stringency_index:SectorInformation -0.461 0.64514
## stringency_index:SectorLeisure and Hospitality -9.410 < 2e-16 ***
## stringency_index:SectorManufacturing -0.499 0.61804
## stringency_index:SectorMining and Logging 0.767 0.44323
## stringency_index:SectorProfessional and Business Services -0.326 0.74478
## stringency_index:SectorRetail Trade -1.356 0.17521
## stringency_index:SectorTransportation and Utilities 0.470 0.63837
## stringency_index:SectorWholesale Trade 0.431 0.66672
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05424 on 7686 degrees of freedom
## (12 observations deleted due to missingness)
## Multiple R-squared: 0.05268, Adjusted R-squared: 0.04985
## F-statistic: 18.58 on 23 and 7686 DF, p-value: < 2.2e-16

```

## 5. Residual analysis

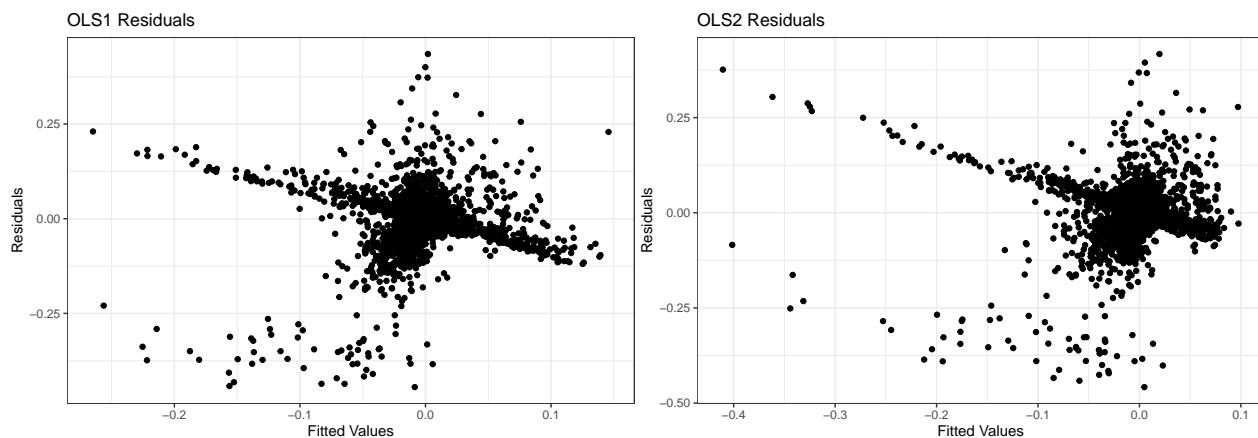
For all residual plots, we noticed a decreasing trend in residuals (as opposed to it being random). That is our models are systematically underestimating high positive (or negative) change in employment values.

These suggest that we use some sine-like function for our  $f(\text{mobility})$ !

```
#Source: https://otexts.com/fpp2/residuals.html
```

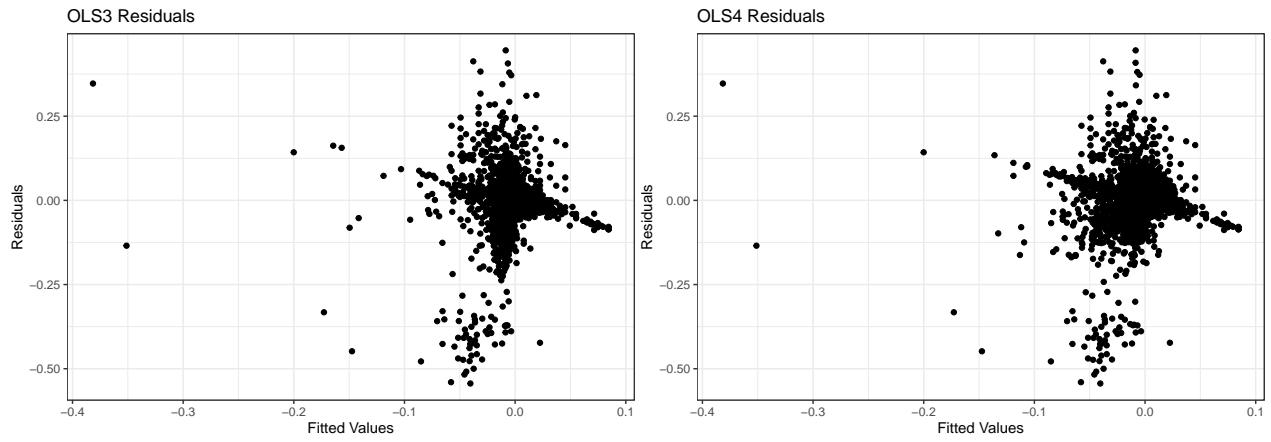
```
#OLS1
## Fitted values vs residuals
mod_ols1 <- fortify(ols1)
ggplot(mod_ols1, aes(x = .fitted, y = .resid)) + geom_point() +
  theme_bw()+
  labs(x = "Fitted Values", y = "Residuals", title = "OLS1 Residuals")

#OLS2
## Fitted values vs residuals
mod_ols2 <- fortify(ols2)
ggplot(mod_ols2, aes(x = .fitted, y = .resid)) + geom_point() +
  theme_bw()+
  labs(x = "Fitted Values", y = "Residuals", title = "OLS2 Residuals")
```



```
#OLS3
## Fitted values vs residuals
mod_ols3 <- fortify(ols3)
ggplot(mod_ols3, aes(x = .fitted, y = .resid)) + geom_point() +
  theme_bw()+
  labs(x = "Fitted Values", y = "Residuals", title = "OLS3 Residuals")

#OLS4
## Fitted values vs residuals
mod_ols4 <- fortify(ols4)
ggplot(mod_ols4, aes(x = .fitted, y = .resid)) + geom_point() +
  theme_bw()+
  labs(x = "Fitted Values", y = "Residuals", title = "OLS4 Residuals")
```



## 6. Conclusion

The project started from the idea that mobility, stringency, and sector composition (and some additional control variables) would sufficiently predict the change in unemployment rate of a given state, for a given sector, at a certain month-end date during the pandemic era.

From this report, we notice a couple of things (note that this is a preliminary analysis of our model; we expect the coefficients will be adjusted once we add more control variables in our model going forward):

- The coefficient for the squared Mobility was statistically significant.
- The coefficient for the stringency was statistically significant. However, stringency itself cannot well predict change in employment.
- The local intercept for the Leisure and hospitality sector, which depends heavily on person-to-person contact, showed statistical significance .
- Most of the coefficients for the interaction term,  $f(Mob_{s,t}) * Sector_{s,t}$  were statistically significant, while only a handful of the coefficients for the interaction term,  $f(Strings_{s,t}) * Sector_{s,t}$  were statistically significant.

Going forward, we plan to enhance our model by following the below steps.

1. Find a functional form of mobility that can capture the deep fluctuation in change in employment that we see (especially from Leisure&Hospitality sector).
2. Investigate the possibility of the impact that lead/lag of f(mobility) has on change in employment.
3. Add more control variables to our model. Some candidates include: new COVID cases, share of elderly people, 2016 Trump Vote share, and Bartik-style (method of isolating local labor demand changes) predicted job loss.

As discussed, after we enhance our model, we plan to:

1. Run a cross validation to contrive a RMSE-minimizing set of coefficients.
2. Using these coefficients, predict April/May 2021 change in unemployment data once the data for independent variables become available.
3. Assess our prediction by computing RMSE between the predicted and the real values.

Finally, it is worth noting that we will not be able to make a causal conclusion based on the coefficient we will obtain at the end of the day. This is because we face the endogeneity problem. Namely, the direction of the causal relationship could go both ways when it comes to the relationship between mobility and unemployment. To make a causal claim, we would need to limit our time frame and use additional econometric tools so that we can treat mobility as a random assignment.