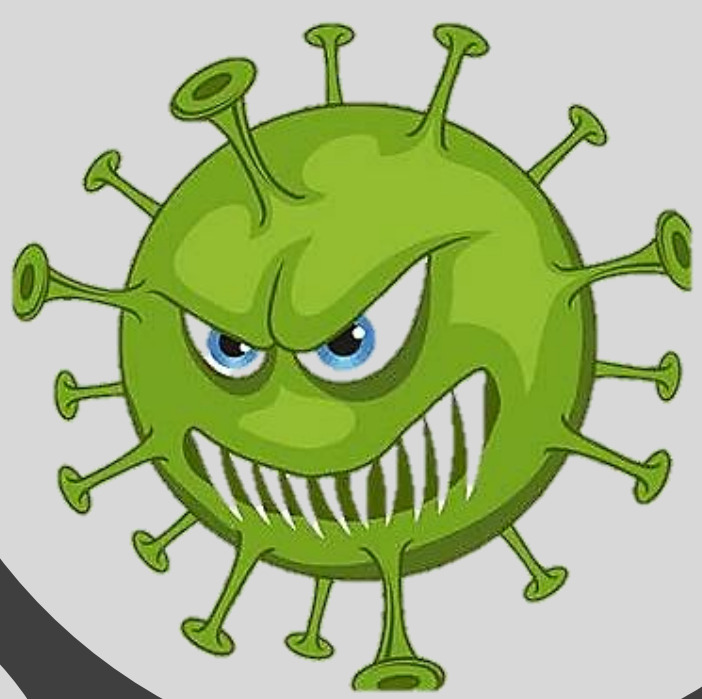
Data analysis of policies impact on COVID-19 infection increasing rate

2020 DataFest: COVID -19 Virtual Data Challenge (TEAM: PiazzzzaHut)



DATA SUMMARY:

Date: March 14, 2020 to June 04, 2020

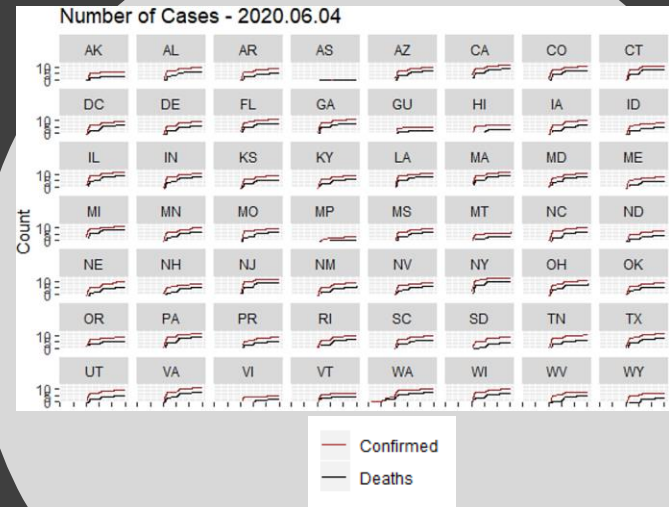
Dependent variable: Changing rate of COVID-19 infection numbers

Independent variables: Policies were implemented in a state

--stay-at-home order --mask mandates

--travel restriction --restriction on restaurants

Note: All of independent variables are dummy variables (if implement policy count 1 ; otherwise count 0)



Panel Data

```
f = change_infection~stay_at_home_order+mask_mandates+travel_restrictions+restaurant_limit
```

```
Oneway (individual) effect Random Effect Model
   (Swamy-Arora's transformation)

Call:
plm(formula = f, data = pdata, model = "random", index = c("date",
    "state"))

Unbalanced Panel: n = 83, T = 48-50, N = 4144

Effects:
                     var   std.dev share
idiosyncratic 0.0303269 0.1741463 0.971
individual    0.0008964 0.0299402 0.029
theta:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.3570  0.3647  0.3647  0.3645  0.3647  0.3647

Residuals:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-0.2708 -0.0614 -0.0323 -0.0001  0.0058  4.7502

Coefficients:
                   Estimate Std. Error  z-value  Pr(>|z|)
(Intercept)         1.1255603  0.0073023 154.1368 < 2.2e-16 ***
stay_at_home_order -0.0357323  0.0071748  -4.9803 6.349e-07 ***
mask_mandates      -0.0352082  0.0093395  -3.7698 0.0001634 ***
travel_restrictions -0.0314276  0.0062457  -5.0319 4.856e-07 ***
restaurant_limit    0.0165912  0.0078462   2.1145 0.0344684 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    150.83
Residual Sum of Squares: 147.58
R-Squared:      0.021708
Adj. R-Squared: 0.020763
Chisq: 91.1741 on 4 DF, p-value: < 2.22e-16
```

Random effect model

model selection(OLS VS Random)

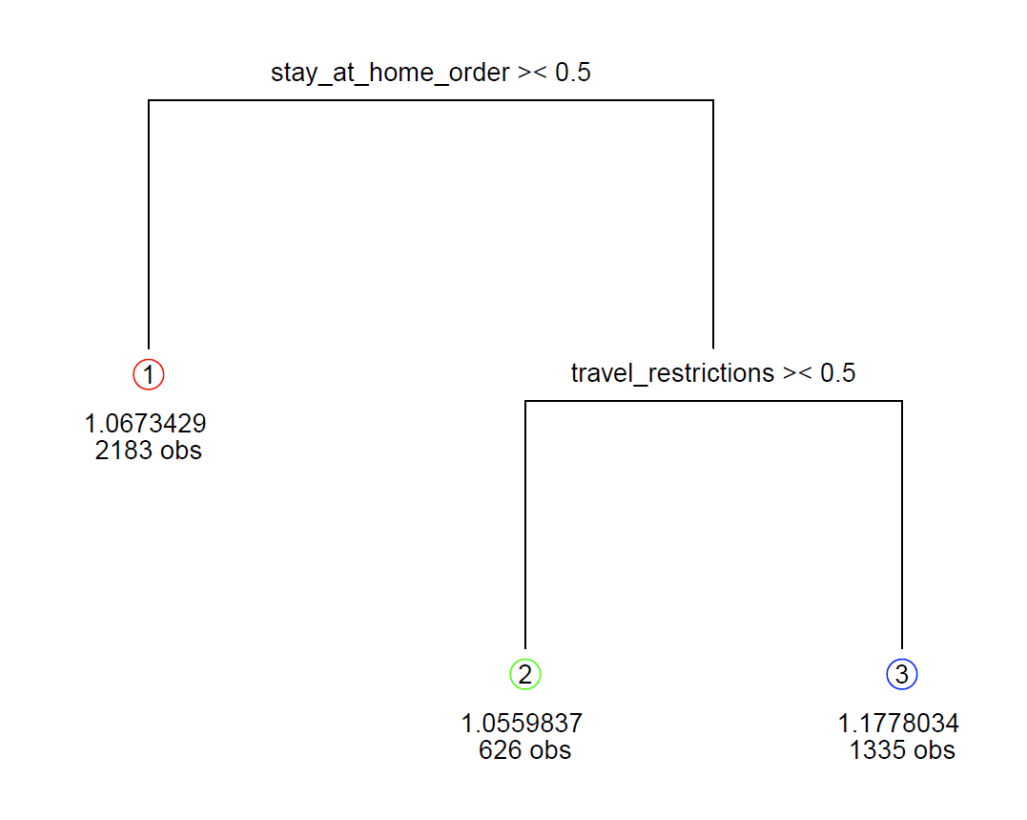
```
        Lagrange Multiplier Test - (Breusch-Pagan) for unbalanced panels

data:  f
chisq = 5160.9, df = 1, p-value < 2.2e-16
alternative hypothesis: significant effects
```

We use panel data to build 3 different models (pooled OLS/fixed effect /random effect model) to estimate the effects of the parameters. After using the LM test for model selection, the result indicates that the random effect model is the most effective and suitable model for our research topic.

From our analysis, the stay at home order, mask mandates, and travel restrictions significantly reduce the infection rate which consistent with our initial prediction. As implement “stay at home order” policy, each day the changing rate of infections will decrease by 0.0357323 on average. The same interpretation will apply to mask mandates(-0.0352082) and travel restrictions(-0.0314276). However, the prediction also indicates that the restaurant limit policy will increase the infection rate by 0. 0165912 on average each day, which violates our initial thought.

Note: detailed analysis in one-page write up

Decision Tree Learning for Panel Data



We use a decision tree to go from observations about an item to conclusions about the item's target value. We use greedy algorithm to train the tree.

We found that the policies of stay-at-home and travel restrictions have more significant impact in the model. If the government do not take any restrictions, the infection rate will get higher. But if the government forces people to stay at home, then the infection rate will be controlled. An interesting thing is that if government only restricts travel but allows people to go outside, the infection rate is lower than the case where government only prohibits people from going outside but allows travel, indicating that travel restriction is more effective.

Note: detailed analysis in one-page write up

Conclusions

Combining the results of two models, this study indicates that stay-at-home and travel restrictions are two effective policies that reduce the growth of COVID-19 infection in each state. Consisting with our hypothesis, mask mandates also have a significant impact on the model considered each state's random effect. Meanwhile, limitations on restaurants unexpectedly increase the infection rate. Since COVID-19 spread through people's interaction with each other, takeout and delivery of food can only reduce the interaction within the restaurants but indirectly increase the interaction, as well as infection rate of delivery servers, and then spread as deliveries get served.

Thus, we suggest governments use more effective policies such as stay-at-home order and travel restriction and pay more attention to the isolation effect of other existing policies.