

COVID-19 Responses

Charlie Zhang

12/01/2020

Background

The COVID-19 pandemic has severely hit the world. Since January, there have been over 63 million reported cases with over a million deaths globally.¹ In light of pandemics and policy toolkits, countries worldwide have adopted similar and diverse measures. Social distancing measures, ranging from full-scale lockdown to voluntary self-compliance measures, have been frequently implemented for “flattening the curve.”²

The project intends to evaluate what factors affect a country's COVID-19 situations. At this moment, I think there are at least two categories— state capacity and proper responses.

¹World Health Organizations. *Coronavirus disease (COVID-19) pandemic*, 2020. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>.

²Brodeur, Abel, David M Gray, Anik Islam, and Suraiya Bhuiyan. “A Literature Review of the Economics of COVID-19,” 2020.

Background

Supposing:

- ▶ a country with the weak capacity and improper responses, it would be doomed;
- ▶ a country with the strong capacity and proper responses, it would not have much damage.

However, the above are extreme examples. Most countries fall between these two “ideal” types. Then, which category is more important? State capacity? Or proper responses?

Moreover, what factors affect the policy? Like a country relies heavily on tourism could be reluctant to impose travel restrictions (it is not necessarily true).

Data Source

- ▶ *CoronaNet Research Project*
 - ▶ The dataset includes certain features of a policy implemented by the country, like start date, end date, policy level (national, provincial, or municipal), compliance (mandatory vs. voluntary)
- ▶ COVID-19
- ▶ *World Development Indicator* from the World Bank
 - ▶ The dataset includes several variables that might be crucial to understand covid-19, e.g. *population, GDP, health expenditure, hospital bed per 1k people, immunization, life expectancy*, etc.

Methods

- ▶ What is the unit of analysis?
 - ▶ Country-Week
 - ▶ Weekly data seems more stable and creates less outliers than daily data.
- ▶ Autocorrelation could be a problem!
- ▶ For each country c in the week w (for any $w > 1$), we have the rate of change (ROC):
 - ▶ $ROC_{c,w} = \frac{Cases_{c,w} - Cases_{c,w-1}}{Cases_{c,w-1}} * 100$
 - ▶ $ROC_{c,w} = \frac{Deaths_{c,w} - Deaths_{c,w-1}}{Deaths_{c,w-1}} * 100$

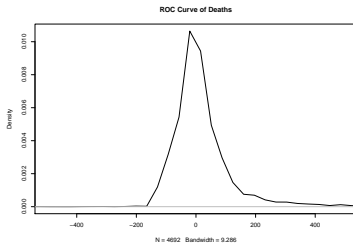
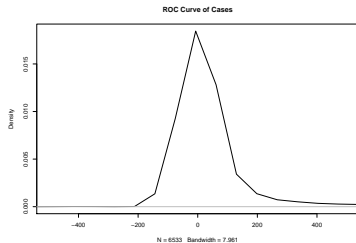
Cont'd

► Explanatory Data Analysis (EDA)

- DVs: case_roc and death_roc
- Key IVs: gdp_per_capita_log, hospital_bed_per1k, lifeexp, active_policy (dummy variable), policy_strength (mandatory vs. voluntary), and median_policy_activity_index
- A row should look like the below (policy-related variables are still under wrangling)

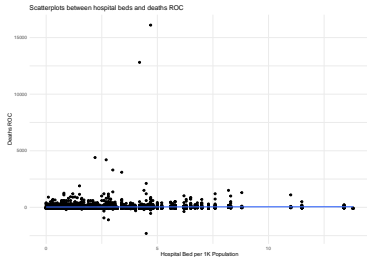
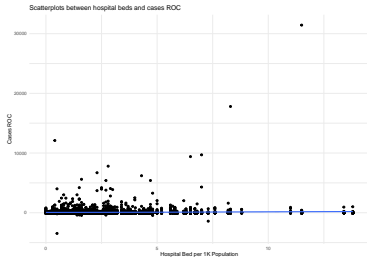
```
##      iso_code      date new_cases new_deaths last_week_
## 4854      MDA 2020-10-27      4502          93
##      cases_roc deaths_roc week_number      pop      gdp
## 4854 -3.865044  -23.77049          44 2657637 9875157730
##      health_per_capita hospital_bed_per_1k
## 4854          191.1858          5.8
```

Preliminary Analysis



- ▶ The density plots above reflect that the sample size is acceptable for machine learning.
 - ▶ When ROC = 400, it represents that cases or deaths of this week are 4 times larger than that of last week.

Cont'd



- It is hard to conclude any relationship from the above plots. Moreover, `hospital_bed_per_1k` and `health_per_capita_log`, both of which reflect the public health infrastructure, seem to produce positive effects on cases- and deaths-ROC. Although these findings are counter-intuitive, it at least tells that state capacity is not the sole category of factors that affect the COVID-19 cases and deaths.

Problems

- ▶ The maximum ROC is 31433.333 (`cases_roc`), which, given the infectivity of COVID-19, is understandable. However, it might create the problem for the learning.
 - ▶ One potential solution is to convert ROC to ordinal levels.
- ▶ Other than state capacity and proper responses, what are the other relevant factors under COVID-19 responses?
 - ▶ Restricted to the above variables would provide no other insights other than the two major categories;
 - ▶ Try to incorporate as many relevant variables as possible other than the above categories into the dataset, like *tourism share of gdp*, *federalism*.