Covid-19 Pandemic Simulation

Shivam Mishra, Habtu Habtu
Institute of Systems Research
University of Maryland
College Park, MD 20740

1 Introduction

The covid-19 virus has infected more than 4 million and claimed more than 300 thousand lives globally[1]. The virus primarily spreads through human-to-human contact. The scale and transmissibility of the virus has forced the governments all across the world to impose severe restrictions on gathering of people. People are not allowed to leave their homes until it's for some essential or necessary thing. This lockdown all across the world has led to a major halt in economic activities. The plots in figure 1 outlines the impact of covid-19 on world economy. The coronavirus pandemic is said to have caused the biggest global economic crisis since the Great Depression[2]. Amid this covid-19 situation more than 30 million people in USA has filed for Initial Unemployment claims[refer figure 2].

The authorities which initially imposed these lockdown measures are now worried about it's impact on the economy. Pushing for greater lock-down periods may push the economy to recession, and lifting the lock-down restrictions too early may lead to a very high number of infected cases in the country leading to a very high number of people losing their lives. The stakeholders who take these decisions has to trade-off between the economic and health impact.This project explores this trade-off analysis.

This is done by forecasting the health impact and the economic impact of the covid-19, using the available trends and then by performing a MAVF trade-off analysis for a surrounding similar to USA.

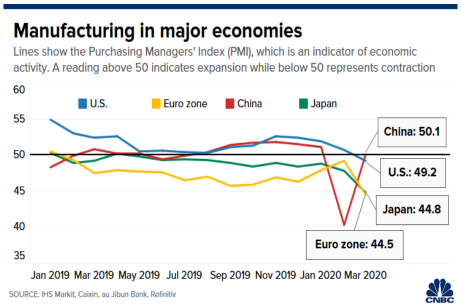
2 System Description

The system replicates the people living in the United States and the system consists is divided into three categories, infected, non-infected and immune. The infected population in the system acts as carrier of the covid-19 virus. The non-infected population are the ones at risk of catching the virus, this is the population that needs to be saved from exposure to the virus. The immune category of population are the people who have caught the coronavirus ones in the past, has recovered and now are immune to its effects. For this simulation, it is assumed that this category of population would not become ill from the virus again or will also not act as carrier of the virus.

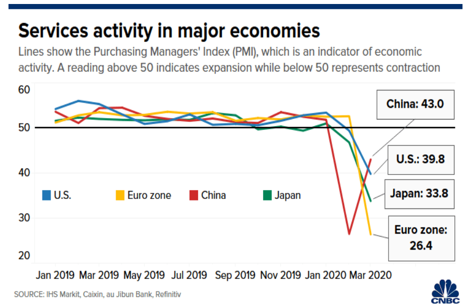
The system also assumes that the population in the United States is homogeneously distributed and a person in the USA can interact with any random subset of the US population. For the modelling of coronavirus spread, it is assumed that the US population consists of a fixed number of infected people on day0. The infected people then gets exposed to a random subset of US population, who on exposure may or may not get infected.The probability of an exposure becoming an infection is given by p (infection probability), and the average number of people an infected person interacts with is given by E (Exposure factor).

On a given day, some new people would get infected by the virus, some old patients may recover from it and some may lose their life due to the virus. To replicate the real life situation, the system is also assumed to be under lockdown, halting the economic activities in the system and resulting in the loss of jobs on the given day. All these factors is used to determine the state of the system. The

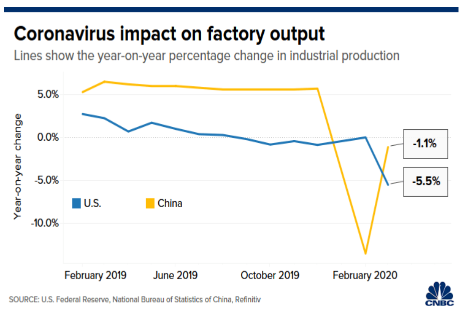
Preprint. Under review.



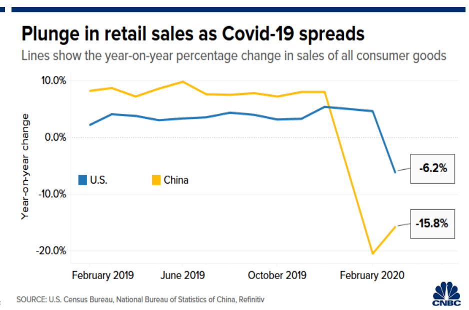
(a) Manufacturing



(b) Services



(c) Industrial Production



(d) Retail sales

Figure 1: Covid-19 impact on economy

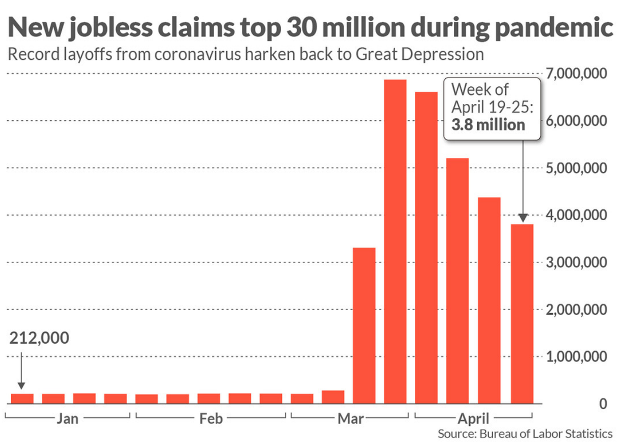


Figure 2: Initial Unemployment Claims

metrics between which we want to trade-off is the number of people losing their jobs on a given day and the number of people losing their lives on a given day. The project specifically wants to explore that what weight the stakeholders may assign to economy against people's lives that can justify the decision of opening up.

The following section defines all the metrics and factors used to define the state of the system.

- **Number of Cumulative Confirmed Cases (N(d))**
  Denotes the total number of confirmed cases of Covid-19 on day d, in United States from the day first case was confirmed (22 January for our case).

- **Number of Cumulative Death Cases**
  Denotes the total number of death cases of Covid-19 on day d, in United States from the day first case was confirmed.
- **Number of Cumulative Recovered Cases**
  Denotes the total number of recovered cases of Covid-19 on day d, in United States from the day first case was confirmed.
- **Growth Factor**
  Factor by which cumulative confirmed cases increases on day d+1, based on the cumulative confirmed cases on day d.

  Mathematically, Growth Factor = N(d+1)/N(d)
- **Infection Probability (p)**
  Probability of an exposure with a Covid-19 positive person to become an infection.
- **Exposure factor (E)**
  Average Number of people a Covid-19 positive person is exposed to each day. Please note that the product pE is an indication of "lockdown" or "mitigation" measures.
- **Initial Unemployment Claims (IUC4)**
  Initial Claims are an employment report that measures the number of new jobless claims filed by individuals seeking to receive unemployment benefits. Higher initial claims correlate with a weakening economy. Initial claims typically rise before the economy enters a recession and decline before the economy starts to recover.

The following section lists down the summarizes key assumptions for the system

1. The Simulation is based on Random Shuffle. It assumes that a covid-19 positive person interacts with a random subset of US population.
2. The product of infection probability (p) and Exposure Factor (E), represented by "pE" is an indication of lockdown or mitigation efforts. The product pE decreases with the mitigation efforts and increases with the lack of it.
3. The mitigation efforts include lockdown orders... tracing , tracking and isolation of covid-19 positive patients.
4. The Growth Factor decreases with the increase in the mitigation effort.
5. The Growth factor follows an exponential probability density.
6. The cumulative number of recovered and death cases linearly vary with the cumulative number of confirmed cases and number of days since the first confirmed case in the United States.

Please note that the last two assumptions are made based on the trends observed in the past data. The past data consists of the numbers that has been reported by the countries prior to April 20, 2020. These assumptions can be observed in the reported data for almost all of the countries struggling with the covid-virus situation, and having cumulative number of confirmed cases greater than 10,000. The past observations trend for some of the countries are plotted in figure 3, 4 and 5.

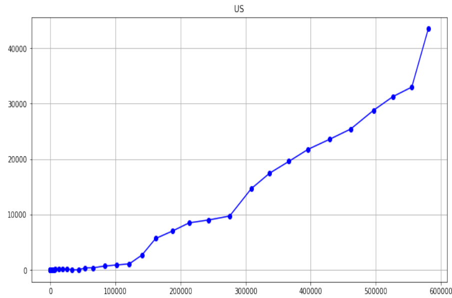
Figure 3 and Figure 4 depict the linear variation of cumulative number of recovered and cases with the cumulative number of confirmed cases. Figure 5 shows that the probability density for the growth factor can be very accurately approximated with the exponential probability density.

3 Analysis Approach

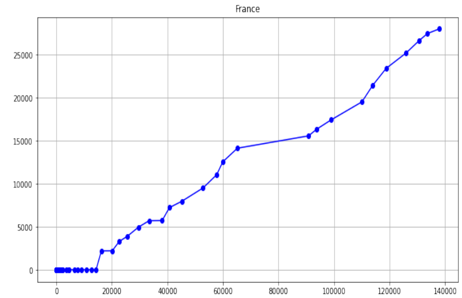
The analysis assumes that the initial number of virus carriers in the system is N0. The number of cumulative confirmed cases on the next day is then calculated by the following equations:

$$N(d+1) = N(d) + \Delta N(d+1) \tag{1}$$

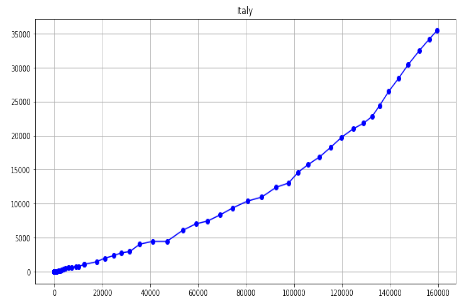
$$\Delta N(d+1) = N(d)pE \tag{2}$$



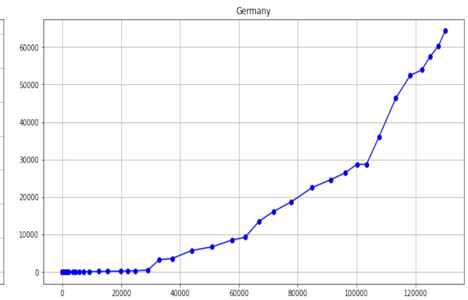
(a) US



(b) France



(c) Italy



(d) Germany

Figure 3: Cumulative recovered cases vs Cumulative Confirmed cases

$$N(d+1) = N(d)(1+pE) = N(d) * growth_factor \tag{3}$$

N(d+1) is the cumulative confirmed cases on day d+1 and N(d) is the cumulative confirmed cases on day d. p is the infection probability and E is the exposure factor. The equation says that the change in cumulative confirmed cases on day d+1 is directly proportional to the cumulative confirmed cases on day d, infection probability and Exposure factor.

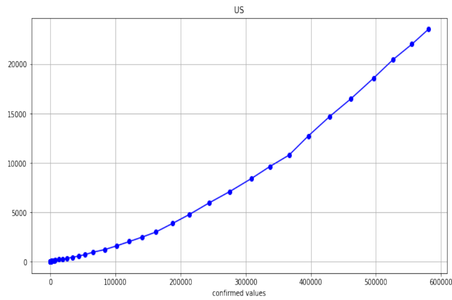
The growth factor on day d+1, is extracted from the exponential probability density. The parameters for the probability density are obtained using the available data for covid-19 pandemic.Now since the exponential probability density is based on the past data, it can only account for the scenario that has happened in the past and that includes a major portion of days when the lockdown restrictions were not in place. In essence the exponential probability distribution based on past data can not be directly utilized to model the present scenario when there are a lot of restrictions to reduce the factors E(exposure factor) and p(infection probability). To take this into account, a factor λ is multiplied to the "pE" term in equation 3. This limits the max value of the growth factor and ensures that the growth factor values amid restrictions would be less than the growth factor without restrictions (to which the data used mainly corresponds to). The modified equation is summarized below:

$$N(d+1) = N(d)(1+\lambda pE) \tag{4}$$

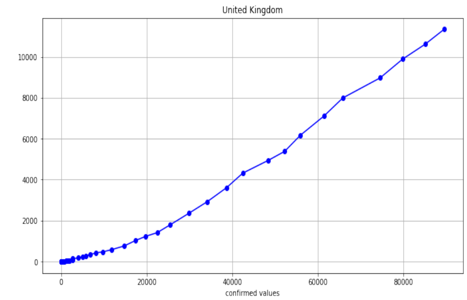
Please note that now the λ factor would be used to replicate the severity of restrictions. The λ value of 1 depicts the time when lockdown restrictions were not in place. The lambda decreases with the increase in lockdown restrictions.

Based on the cumulative number of confirmed cases and number of days since the first case was confirmed in the United States, cumulative number of death and recovered cases is approximated using Linear Regression. The regression parameters are also calculated based upon the past data available for covid-19 pandemic.

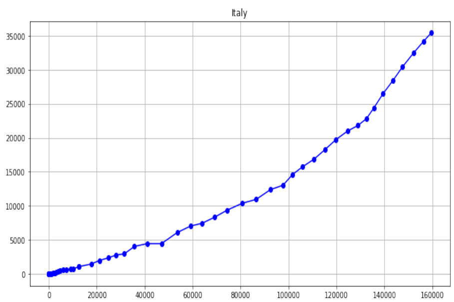
The trade-off analysis is done between the Initial Unemployment claims (4 weeks moving average) and number of death cases calculated by the model.



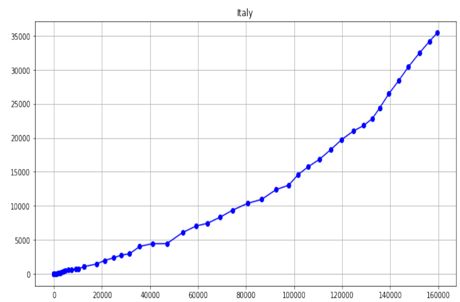
(a) US



(b) UK



(c) Italy



(d) Iran

Figure 4: Cumulative Death cases vs Cumulative Confirmed cases

Table 1: Important Dates

Event	Day Number	Date
First Case	0	22 January 2020
Lockdown Begins*	70	30 March 2020
Present Day	93	23 April 2020
Prediction for	121	21 May 2020

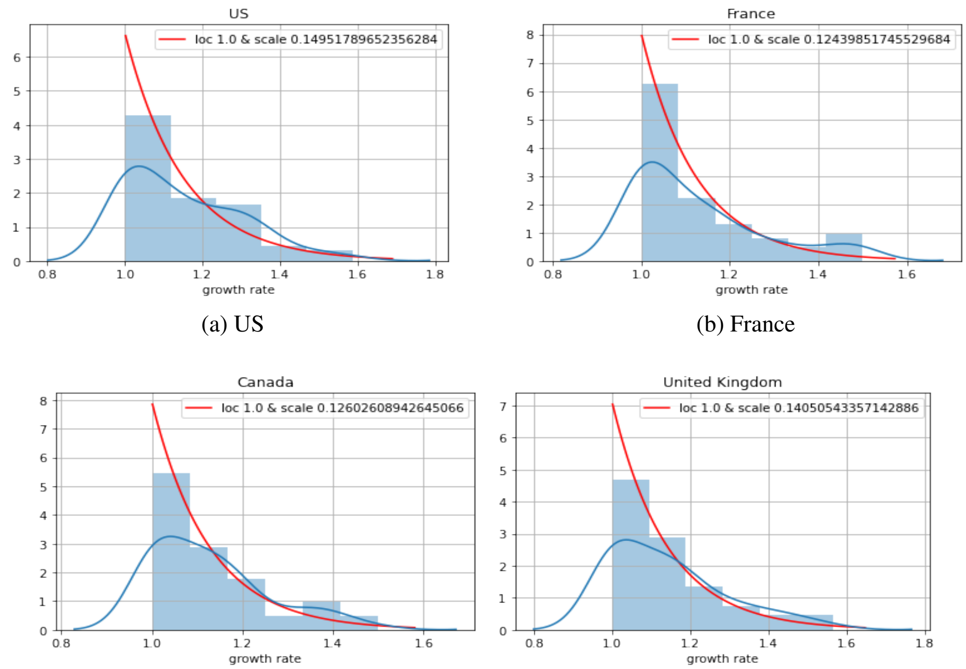
The 4 week moving average of initial unemployment claims is taken from the Bureau of Labour Statistics.The IUC4 number on the last week of April and first week of May represents the full-lockdown scenario when all the restrictions are heavily enforced. The 4 week moving average of initial unemployment claims at the week ending on April 25, 2020 was 5,040,000, and the same number on the first week of May 2019 was 221,000 [3]. Please note that for the purpose of this project the IUC4 values of 5,000,000 and 500,000 are taken as the extremes. The 5 million represents the IUC4 value when the restrictions are heavily imposed and 500,000 in the normal situation.

The IUC4 (initial unemployment claims) value is assumed to vary linearly with the λ. The rationale for the assumption lies in the fact that **lambda** represents the strength of lockdown restrictions because of which the economic activities are on a halt. If restrictions are eased economic activities would start growing. The growth relation here is assumed linear.

The trade- off analysis explores the scenario 28 days after 23 April 2020 (the last available data used for modelling), it models confirmed values, death values and unemployment claims at the 4 week end with different λ parameters.

Table 1 lists down the dates for various covid related events that happened in the United States and that the system tries to emulate. Lockdown begin date varies across the country, it is assumed to be 30 March (lockdowns have started plus/minus one week from the date mentioned).

Please note that the mote-carlo simulation is done in three parts. First, the model generates the values for λ = 1 for all the days before day 70 (March 30, 2020) when the lockdown in the US is assumed to have been started. After lockdown started, that is after the March 30, 2020 the model is tuned by choosing the correct λ value to match the actual data after 30 march, 2020 to the present day 23



(a) US

(b) France

(c) Canada

(d) UK

Figure 5: Growth factor Probability Density

April, 2020. Then, the model predicts the cumulative number of confirmed cases for different values of λ, simulating the different conditions on lockdown restrictions. Please note that the λ value of 0.2 is used to depict the highest lockdown measure. This is done because the λ parameter used to match the actual cum. confirmed cases for the lockdown-begins to present day period is 0.4, and if the lockdown continues further the λ value would definitely be less than 0.4...that is conservatively taken as 0.2. Similarly, for the full ease of lockdown the λ value is assumed to 0.7 instead of 1.

4 Supporting Models and Simulations

4.1 Key Equations

The following section lists down the major equations used in the project:

4.1.1 Cumulative Confirmed Cases Equation

$$N(d+1) = N(d) + \Delta N(d+1) \tag{5}$$

$$\Delta N(d+1) = N(d)pE \tag{6}$$

$$N(d+1) = N(d)(1+pE) = N(d) * growth_factor \tag{7}$$

N(d+1) is the cumulative confirmed cases on day d+1 and N(d) is the cumulative confirmed cases on day d. p is the infection probability and E is the exposure factor.The equation says that the change in cumulative confirmed cases on day d+1 is directly proportional to the cumulative confirmed cases on day d, infection probability and Exposure factor.

4.1.2 Recovered and Death Cases Equation

$$X(d) = A(N(d)) + B(d) + C \tag{8}$$

Table 2: Model Architecture Summary

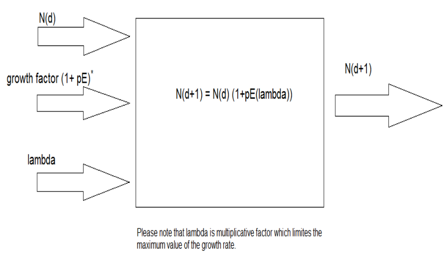
Case	coef confirmed	coef days	intercept
Death	0.054	-127.97	2594.95
Recover	0.0924	-127.97	2594.95

where, A, B and C are constant. X(d) is the number of death cases or recovered cases. This essentially tells that the cumulative number of death cases or recovered cases is a linear combination of cumulative confirmed cases and number of days since the first case was confirmed (22 January for this case). The parameters used for regression are listed in Table2.

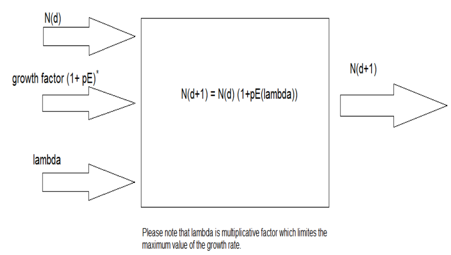
4.1.3 Growth Factor Modification Equation

$$GrowthFactor = N(d+1)/N(d) = 1 + pE\lambda \tag{9}$$

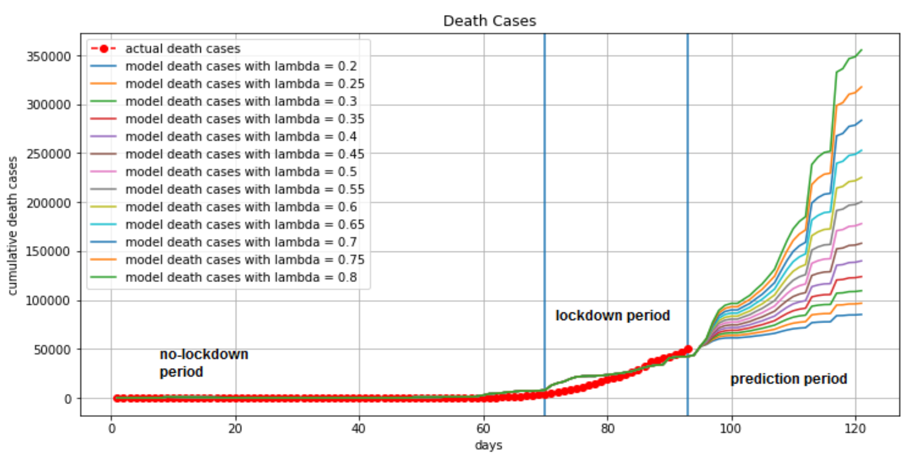
Lambda is a multiplicative factor that limits the max value of pE. The Growth factor obtained is from a probability density that is based on the past data. The lambda is assumed to be equal to 1 for past cases. It is assumed that with the proper mitigation efforts, the max value of growth factor would decrease in the future. . . to take that into account pE is multiplied by **lambda**



(a) Response model1



(b) Response model2



(c) Death Cases

4.2 Simulation Results

The simulation results for cumulative confirmed, death, recovered and Initial Employment Claims (IUC4) are summarized in Table 3.The **lambda** values represents the different scenarios or the intensity of the lockdown restrictions. The results are also summarized in the Figure 6.

Table 3: Simulation Results

λ	cum confirmed cases	cum. death cases	IUC4 values
0.2	1568642	85107	5000000
0.25	1778621	96562	4550000
0.3	2014024	109404	4100000
0.35	2277616	123783	3650000
0.4	2572426	139866	3200000
0.45	2901767	157832	2750000
0.5	3269260	177880	2300000
0.55	3678859	200224	1850000
0.6	4134870	225101	1400000
0.65	4641980	252765	950000
0.7	5205282	283494	500000

5 Possible Scenarios or Design Options

The possible scenarios after 28 days serve as design options for this project. The different values of the **lambda** corresponds to the amount of restrictions authorities are willing to lift. The **lambda** can be associated with the various restrictions the authorities has imposed. The extremes are intuitive but even the mid-values can be carefully modeled to the various imposed restrictions. This is not done as the part of this project. This is left as the future work. The scenarios listed in the Table 3 serves as the design options for the MAVF analysis.

6 Analysis and Results

The results from MAVF analysis are reported in Figure 6. Vd is the individual value function for cumulative number of deaths, Vi is the individual value function for economy. V is the multi-attribute value function. The presented MAVF analysis is done by assigning a rank of 100 to health and a rank of 60 to economy. The optimum **lambda** that should be chosen based on 0.3 with multi attribute value of 0.623455.

A sensitivity analysis is also done on the Ranks assigned to the Economy and the health. Those results are summarized in the Figure 8. The Ri is the rank associated with the economy and Rd is the rank associated with cumulative number of deaths. The range of rank that can be assigned to Death is in the range [100, 70] and rank range for the economy is [40,70]. The sensitivity analysis shows that the value function is very sensitive to the rank assigned to economy.

Another analysis is done to see what relative importance one would give to the economy to allow the full restrictions lift-up. The restrictions lift-up are associated with λ = 0.7 in this project. For this purpose, the rank of health is fixed at 100 and the rank for economy is varied. The results are summarized in the Figure 9. On the x-axis it has the rank for the economy and on the y-axes it has the corresponding **lambda** value. The analysis reveals that the economy must be at least 1.6 times more important to the stakeholders than health for them to take a decision to completely lift up lockdown among the pandemic.

7 Recommended Courses of Action

The recommended course of action based on the simulation done and the ranks assumed is not to rush into opening up places at ones. For most of the sensible values of relative importance between the economy and the health, it does not make sense to lift up all the restrictions so quickly. The simulations projects that even with the 0.2 lambda value one can reach the cum death numbers of 85k by May 23. At the time of writing this report, that is on 18 May, the number has already crossed the 90k mark.

The other rank analysis points out that the relative importance for the economy needs to 1.6 times the health for authorities to completely lift-up restrictions. This relative importance may be justified for some of the poor nations where prolonged economic activity closure may itself result into the multiple number of deaths, but can't be justified in a developed country like USA.

8 Future Works

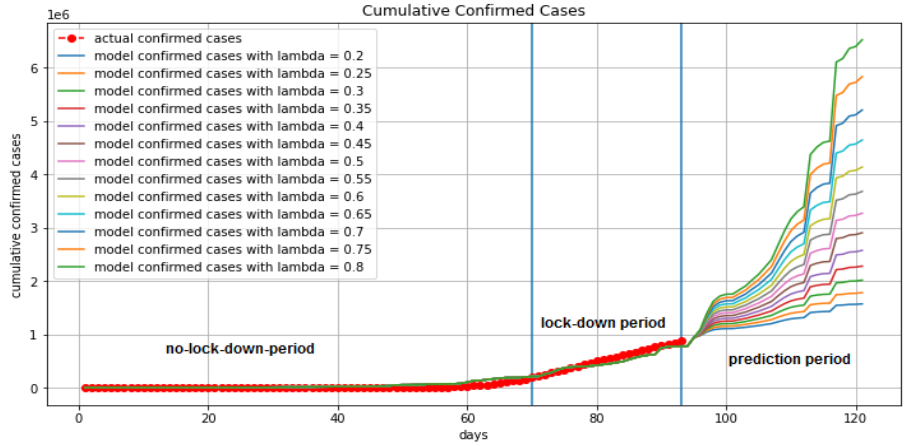
The future work includes associating the **lambda** with the real imposed restrictions by different States in the USA. The other work is to eliminate the random shuffle and even population distribution assumptions.

References

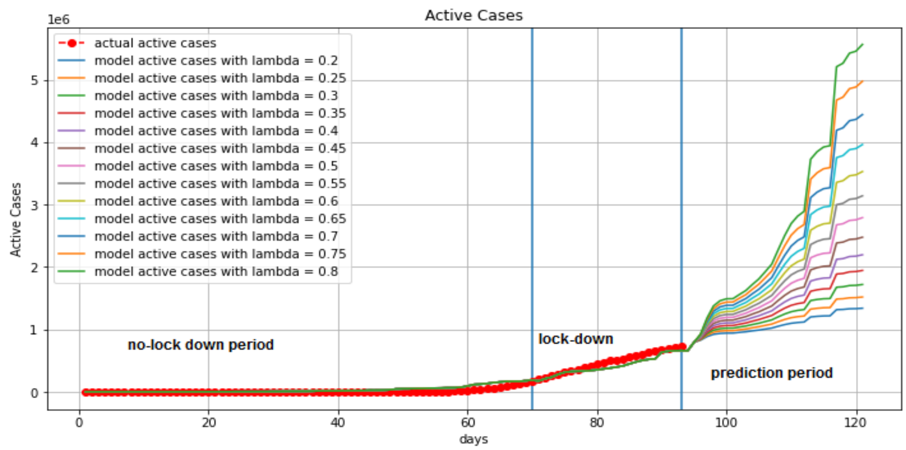
[1] John Hopkins University CSSE. Covid-19 dashboard.

[2] IMF. The great lockdown: Worst economic downturn since the great depression.

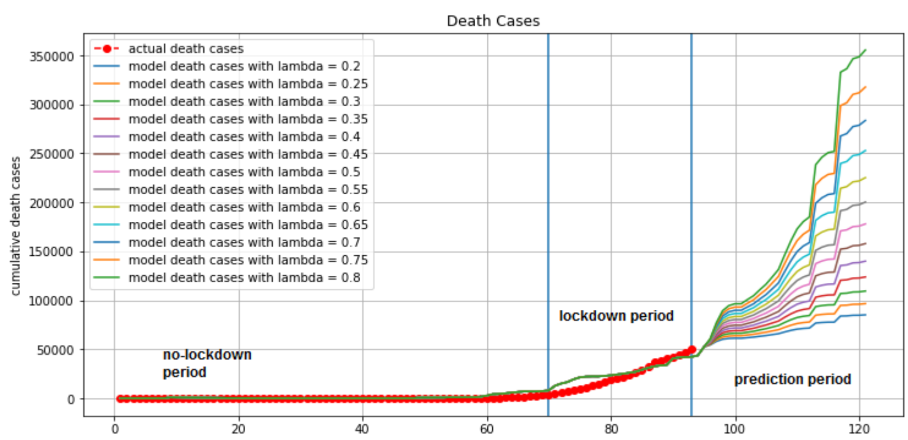
[3] Bureau of Labor Statistics. Unemployment insurance weekly claims, news release.



(a) Confirmed Cases



(b) Active Cases



(c) Death Cases

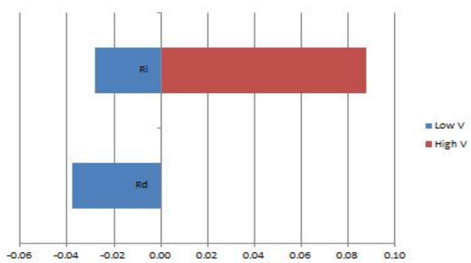
Figure 6: Simulation Results

lambda	death_cases	IUC4	Vd	Vi	V
0.2	85107	5000000	1	0	0.625
0.25	96562	4550000	0.942259	0.1	0.626412
0.3	109404	4100000	0.877527	0.2	0.623455
0.35	123783	3650000	0.805048	0.3	0.615655
0.4	139866	3200000	0.723979	0.4	0.602487
0.45	157832	2750000	0.633419	0.5	0.583387
0.5	177880	2300000	0.532364	0.6	0.557727
0.55	200224	1850000	0.419735	0.7	0.524834
0.6	225101	1400000	0.294339	0.8	0.483962
0.65	252765	950000	0.154894	0.9	0.434309
0.7	283494	500000	0	1	0.375

(a)

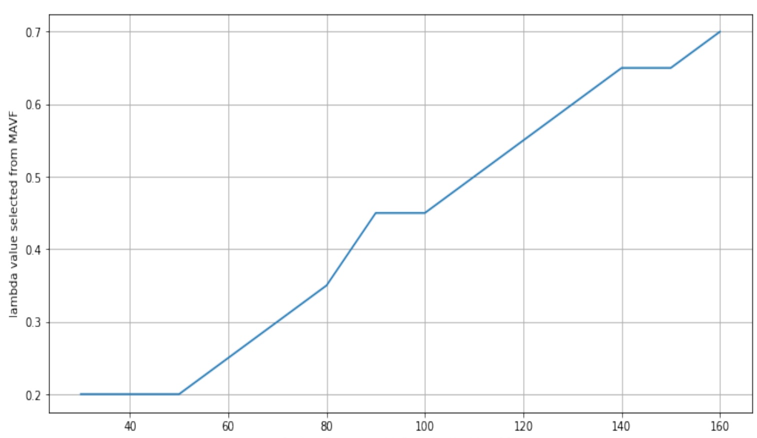
Figure 7: MAVF results

	Base MAVF =	0.63	Rd = 100, Ri = 60	
	Low V	High V	High Rank	Low Rank
Rd	-0.04	0.00	100	70
Ri	-0.03	0.09	70	40



(a)

Figure 8: Sensitivity Analysis Results



(a)

Figure 9: Relative importance of Economy, R(health) 100, relative importance is R(economy)/R(health)