Different strategies against Covid-19

A review of the literature

Siming, Su

June 5, 2020

Abstract

This paper compares two research papers about the prediction of COVID-19. We investigate two different statistical models and how these two models predict and work divergently in different scenarios given by authors. These two models are dynamic panel SIR (DP-SIR) introduced by Xiaohui Chen and Ziyi Qiu and SIRD models implanted by Jesús Fernández-Villaverde and Charles I. Jones. I will discuss how these two models work in the role of prediction of the infectious population and also talk about the difference and the different values of these two models. Then, by investigating these two models' functionalities, the sense of when and how different strategies should be applied in different time stages should be clear. Both models show that government policies, including either applying non-pharmaceutical interventions or enforcing social distance, will have a large-scale impact on the future infectious population.

1.Introduction

A global pandemic, Covid-19, has attacked Wuhan, China first around December 2019. As a transit center in China, this local-level disease suddenly transformed into a global issue. Starting from the first confirmed case being reported in different countries, more and more countries were forcefully taken into this battle with diseases. Numerous countries have executed different policies such as enforcing social distancing, national lockdown, or school closure. We will take a close look at how two different articles address this issue in the lens of statistics.

Article1: By examining at how different non-pharmaceutical interventions' and its effects on stopping the spread of Covid-19, Chen and Qiu figure out and can generate the helpful advice about the most effective non-pharmaceutical interventions (NPIs) combination for different counties by using DP-SIR model [1].

Article2: In the other article, Jones and Villaverde have also successfully used the SIRD model with social distancing to get a result that by assigning different values into the variable, predictions have become different. They examined how the death rate can affect the ever-infectious population because people will change their behavioral patterns, such as staying at home, accordingly. They also discussed how the economy re-opening can transform the direction of the disease spreading and in what percentage can we open our economy to normal on the basis of on outbreaks [2].

Motivations: Both of these two articles shed a light on the correlation of the intervention strategies either in the country-level or person-level with disease control. I will compare some of their works including their statistical models and their research results to see how these articles have yielded different practical suggestions for those countries experiencing serious pandemic. It is important because many countries are being influenced by this pandemic. Based on different situations, policymakers should adopt different strategies against COVID-19.

2. Models Comparison

2.1.1 Dynamic Panel SIR model. To Summarize this model, three equations play an important role in this prediction. On the basis of SIR model with time varying parameters, he made each term S(t), R(t), and I(t) in SIR model as $S_j(t)$, $R_j(t)$, and $I_j(t)$, where each represents the Susceptible, Recovered, and Infectious population in country j at time t. The following equation comes from the original SIR model.

$$R_t = \frac{\beta(t)}{\gamma(t)}, \quad (1)$$

Where $\beta(t)$ is the transmission rate of the disease and $\gamma(t)$ is the recovered rate at time t.

Chen and Qiu also construct the coefficient $\beta_j(t)$ as following:

$$\beta_j(t) = exp(\alpha_j + \sum_{k=1}^{K} \beta_k NPI_{tjk}), \quad (2)$$

Where NPI is the country interventions, t represents time, j represents the country, and k represents which intervention they use. Summing them up, we can get the total effect of those NPI by looking at $\beta_j(t)$, which is the transmission rate.

And they define NPI_{tjk} as the flowing:

$$NPI_{tjk} = \begin{cases} 1 & if\ 0 \leq t_{jk}^* \\ exp\left(-\frac{t-t_{jk}^*}{\tau}\right) & if\ t_{jk}^* \leq t \leq T \end{cases}, \quad (3)$$

Where τ controls the time that NPIs start to have an effect. By combining (1) and (2) equations, the author uses OLS to estimate $\theta = (\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_K, \gamma_1, \ldots, \gamma_p)$. By using each OLS parameter and plugging back into formula 2, we can get approximately how NPIs affect the diseases spreading.

2.1.2 interpretation: I think what is delicate about this model is that it added the term NPI so that $\beta_j(t)$ can reflect the effect of NPIs. Then, we can plug $\beta_j(t)$ back into R_t to see if there is an outbreak or not. The reason why R_t can reflect the outbreak is that this is the proportion of transmission rate $\beta(t)$ and recovery rate $\gamma(t)$ at time t. If $\beta(t)$ is far greater than $\gamma(t)$, then the outbreak of Covid-19 would happen. Thus, by adjusting the combination of NPIs, we can control the $\beta(t)$ to prevent the spreading.

2.2.2 SIRD model. To summarize how this model works in general, Jones and Villaverde have introduced two new terms which are "Resolving" and "Dead", compared with the conventional

SIR model. The computations are complicated, and I am just going to show the result and interpret it. The final two results the author has derived are

$$\beta_t = \frac{N}{S_t}\left(\gamma + \frac{\frac{1}{\theta}\Delta\Delta d_{t+3} + \Delta d_{t+2}}{\frac{1}{\theta}\Delta d_{t+2} + d_{t+1}}\right) \qquad (4)$$

and,

$$S_{t+1} = S_t\left(1 - \beta_t \frac{1}{\delta\gamma N}\left(\frac{1}{\theta}\Delta d_{t+2} + d_{t+1}\right)\right) \quad (5)$$

2.2.3 interpretation: the most interesting formula to notice is that for β_t is determined by $\Delta\Delta d_{t+3}, \Delta d_{t+2}, and\ \Delta d_{t+1}, and\ d_t$ is the dead population per million at day t. This shows that we need future death statistics to tell us the transmission rate today. Same as susceptible populations, we will need future death numbers to get the estimate of susceptible populations. For the other parameters, which are γ, θ, α, and δ, what the researchers do is to fix these parameters into a certain number that makes sense biologically. The parameter α is what the researchers use to estimate the future outcomes adjusted for daily death per million people. The formula is $R_{0t} = constant * d_t^{-\alpha}$ which shows that by changing the daily death population, R_{0t} will also change accordingly. For example, if the daily death population has increased, then people's reactions will be to stay home and reduce contact with others. Then, R_{0t} will decrease. The researchers have applied this behavioral pattern to adjust $\beta(t)$. Therefore, we can compare the R_{0t} by plugging in $\beta(t)$ with the initial R_0 to see how we can adjust our social distancing. This will be further discussed in the fourth section.

2.3 comparison. For these two models, they both developed based on the conventional SIR model with varying time. However, what makes them differ is for the first model (Dynamic Panel SIR model), Chen and Qiu heavily draw β_t on NPIs effect in order to see how different government interventions have caused different levels of transmission rate. However, for the SIRD model, as its name suggests, it added the new term dead and also depends on the important factor β_t on the dead populations heavily and also its mortality rate, δ. We will take a close look at how these two models work in numerous states and countries below.

3. Data Description

After the brief introduction of both models, let us see how these models work in the real data. For both articles, the authors use similar data to implant their statistical models. The only difference is the publication date difference. Jones and Villaverde's article was published a little bit later than Chen and Qiu's article, so the data Jones and Villaverde used would be more current than Chen and Qiu's. They both used the data from Johns Hopkins university CSSE (Center for Systems Science and Engineering).

Different approaches:

Chen and Qiu: besides using the data from John Hopkins, they have collected other information from the government website or newspaper to estimate their variable, NPI. Because they should know what NPIs different countries took. Six non-pharmaceutical interventions (NPI) are summarized, they are respectively: travel restriction (TR), mask wearing (MW), lockdown (LD), social distancing (SD), school closure (SC), and centralized quarantine (CQ).

Jones and Villaverde: they instead manipulated some of the death populations. They have increased the death population a little to let it be more realistic. And they also remove so-called "weekend effect", where 0 death is reported at the weekend.

4. Estimation

4.1 Estimation of Panel Dynamic SIR model

4.1.1 Estimation for NPI variable (Chen and Qiu's article)

According to formula (2), by applying one statistical concept: confidence level, there are two ineffective NPIs coefficients which are SD and TR according to the following confidence interval table.

NPI	Estimated coefficient	95% CI
Travel restriction (TR)	-0.343	[-0.786, 0.100]
Mask Wearing (MW)	0.651	[0.009, 1.294]
Lockdown (LD)	1.063	[0.427, 1.699]
Social Distancing (SD)	-0.279	[-0.986, 0.427]
School Closure (SC)	0.972	[0.339, 1.604]
Centralized Quarantine (CQ)	2.042	[1.493, 2.592]

Table (1)

Notice that the 95% confidence interval of TR and SD is cross 0, which means that it has little effect on estimating β_t. For the travel restriction, it makes sense because what numerous countries did was to restrict the entrance of Chinese people. However, this is a global pandemic. Other countries' residence also has a chance to enter the border and spread the disease. My interpretation of why social distancing will be excluded in the model is that it has an overlapped section with LD, SC, and CQ regarding the prevention of diseases spreading. This means that by doing national lockdown or school closure, it also promotes social distancing. That is why the confidence interval for Social Distancing is cross 0.

4.1.2 NPI choices

The NPI selection will play a significant role in controlling Covid-19. Two combinations of NPIs are selected by the authors, one is MW+LD+SC+CQ and the other is MW+SC+CQ. It is practical to set LD as our control variable in the real world. Since implanting LD, national lockdown, is highly expensive regarding a country's economy, so we take this factor out to see if we can get a solution that is not hurting the economy too much and meanwhile controlling Covid-19 virus.

4.1.3 different NPI graphs and interpretation

By collecting 9 countries' data and fit them into the Dynamic Panel SIR model, we can get a sense of how NPI variables impacted the COVID-19. Figure (1) and figure (2) has shown us that for the countries which have passed its peak, or near its peak, then LD no longer makes differences. The graph looks similar to these 9 countries. Then, this may suggest that the effect of LD for those countries which are near its peak or have passed its peak is not strong in controlling the diseases. What is more, LD is a policy that hurts one country's economy to a great extent. To remove this factor, other businesses can actually operate normally, and the diseases can get controlled by the other three NPIs. Someone may think that this is just predictive, and in the real world, it is unstable

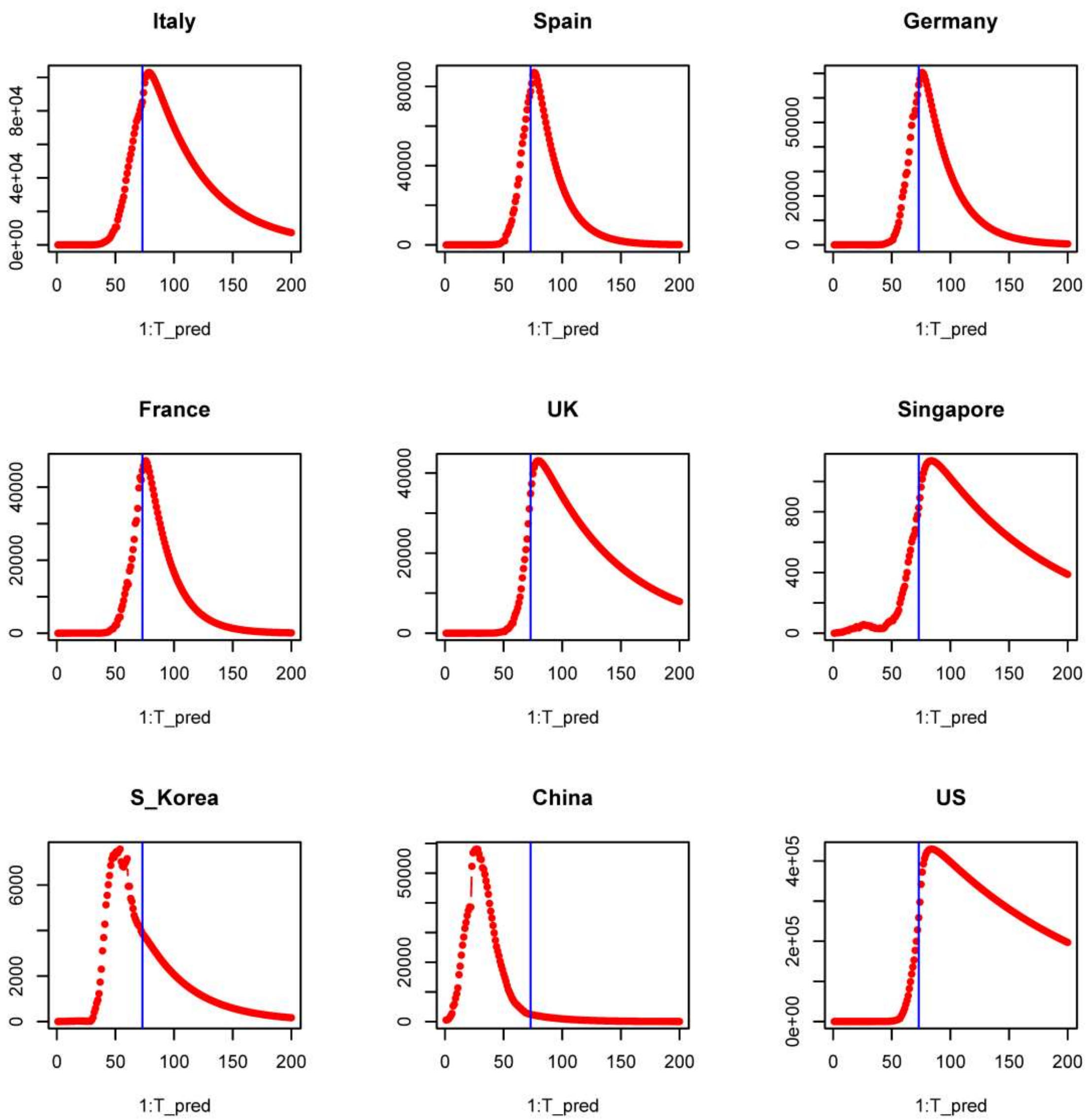

Figure (1)

Predicted confirmed cases using NPI: MW+LD+SC+CQ

the verticla blue line is the date 04/03/2020, and the x-axis is ranging from 01\22\2020 to 08\09\2020

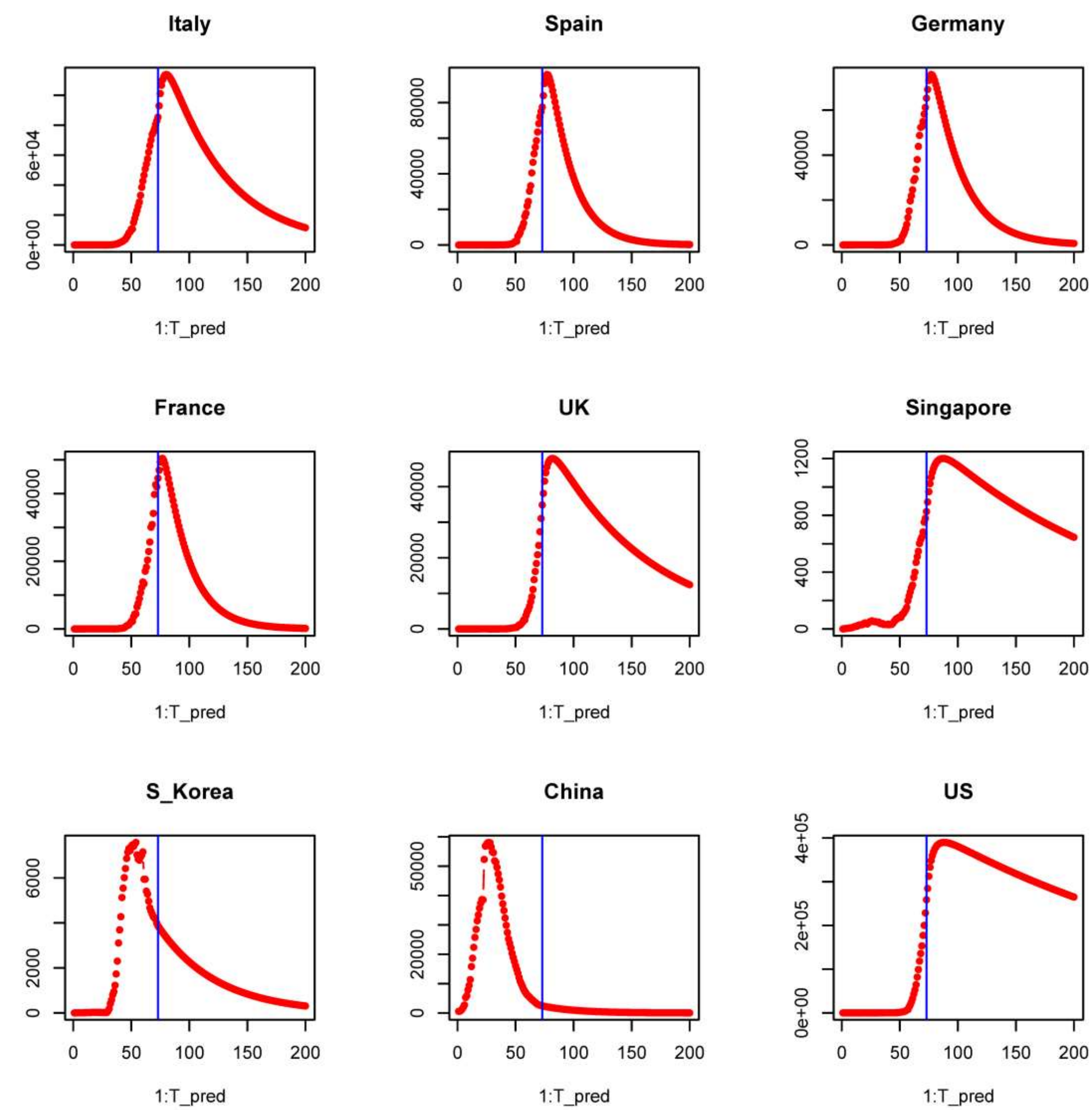

Figure (2)

Predicted confirmed cases using NPIs: MW+ SC + CQ

the verticla blue line is the date 04/03/2020, and the x-axis is ranging from 01\22\2020 to 08\09\2020 and unpredictable regarding disease control. However, we can take a look at the real-world cases and we can know that this prediction is pretty close to reality. South Korea is an example who implanting NPIs including MW+SC + CQ without LD before reaching its peak. Compared to China, which has utilized all four NPIs strictly to a nation-level, South Korea reached its peak less than a month late. Considering the large economical loss South Korea has avoided, MW+SC+CQ probably is also economically effective in fighting COVID-19.

4.2 Estimation of SIRD model (Fernandez-Villaverde, and Jones)

4.2.1 the accuracy of SIRD model

At the begging, Fernandez-Villaverde and Jones briefly used the example of New York City to talk about the quality of the SIRD model in terms of predictions, and it turns out that this model corresponds with current situations to a large degree. For example, the baseline estimation has been done by Jones and Fernandez-Villaverde, (the baseline estimation is when we assign 0.01, 0.05, and 0.1into our parameters $\delta, \alpha, and\ \theta$). Then, they use the formula (4) and formula (1) to calculate the R_{0t} which is around 2.7 in New York. This means at the beginning, every infected person would spread the diseases to approximately 2.7 people. This result generated by baseline estimation is close to what has been observed in New York. They also estimated the infectious population, the result also agrees with the antibody test which has been done in New York City.

4.2.2 the estimation on New York and California

In this section, Fernandez-Villaverde and Jones started to apply this model to predict future statistics. There are some properties worth mentioning. if the city or state has passed the peak value, then the prediction generated by SIRD model will converge, otherwise, the predictions will be pretty "noisy" when some values are changed. Next, I will give two figures that show this property. The examples are New York City, which has passed its peak value, and California, which has not experienced its peak in May. Figure(3) and Figure(4) illustrates how the new data of deaths come in and change how the model behaves. Fort the Figure(3), we can see those prediction lines all converge in one direction. However, for the state California, notice that with new deaths coming into the data, those lines behave almost randomly. This result appears due to the effect of α we added in SIRD model. Recall the formula we have talked about in section 2.2.3, $R_{0t} = constant * d_t^{-\alpha}$. As d_t comes in high value, then it will result in less R_{0t}, which basically means that people will react to the high values of deaths and reduce their contacts with people. This is why R_{0t} will reduce. Then, the chance of getting diseases will lower. This relationship is what we are going to use in the following topic about economy reopening and herd immunity.

4.2.3 economy reopening and herd immunity possibility

Let us bring in some numbers and analyze what this model can inform us. The main point for this topic is that the infectious population is closely associated with death rate δ, as I explained

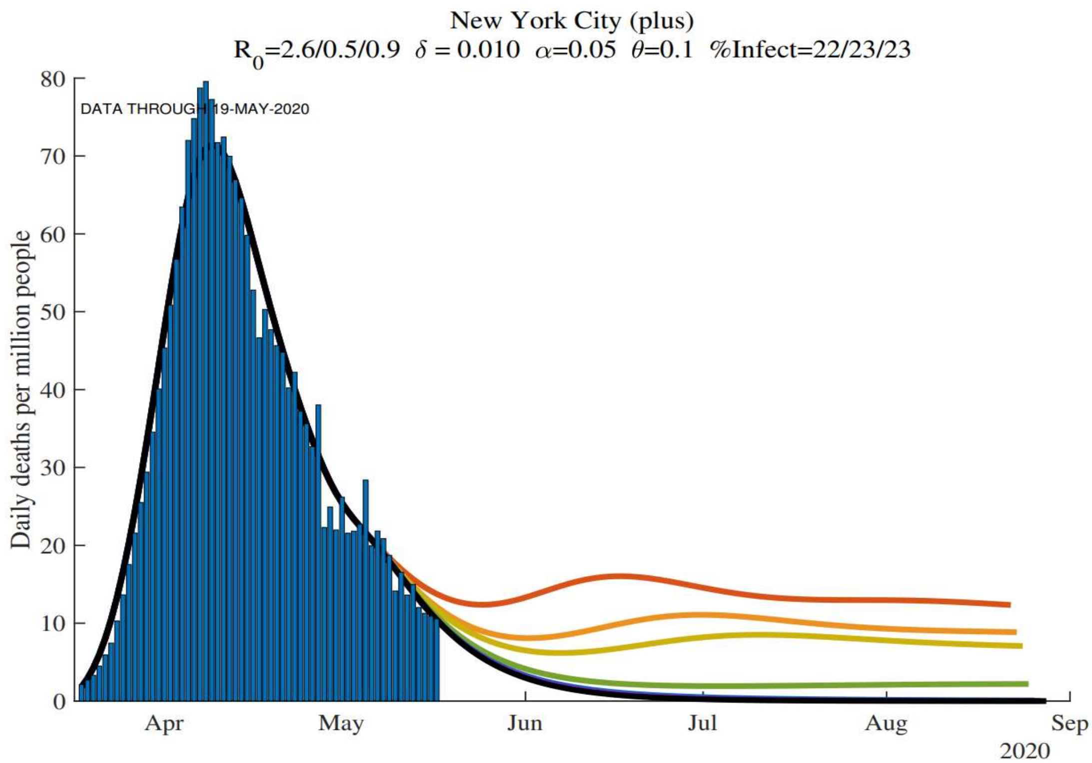

Figure (3)

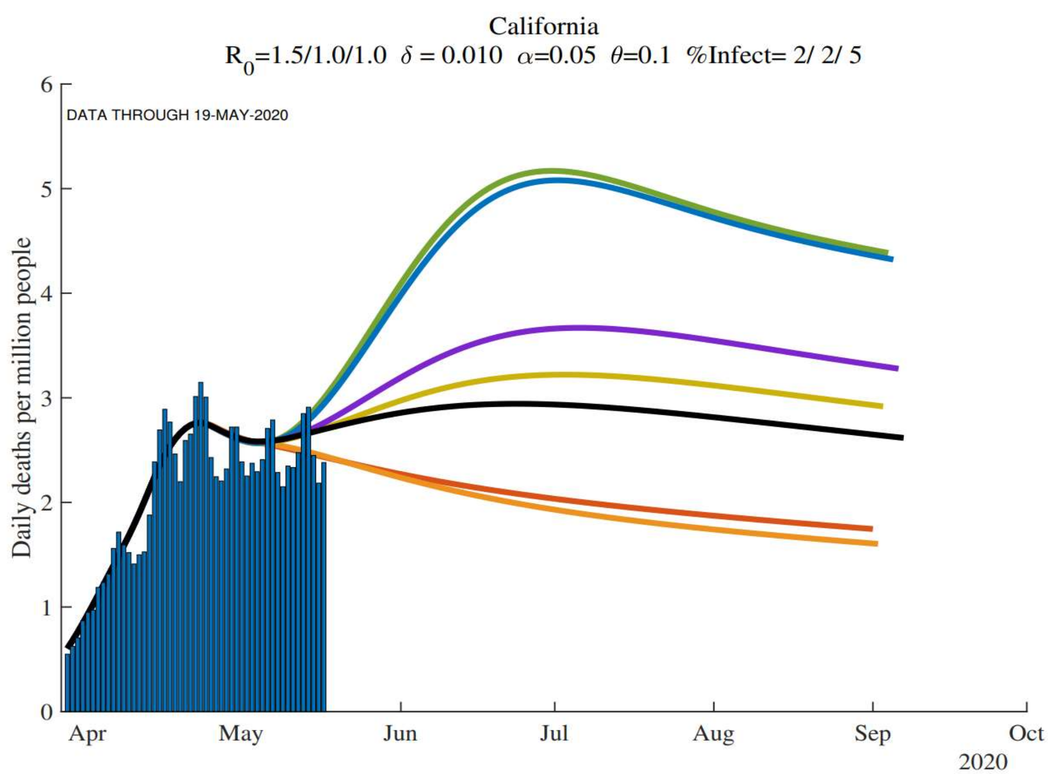

Figure (4)

Above. However, how close are they related to each other? In fact, they are almost bond with each other. The following table which is generated by SIRD model shows the relationship.

	-percent ever infected today		
	$\delta = 0.5\%$	$\delta = 1.0\%$	$\delta = 1.2\%$
New York City	51	26	22
Lombardy, Italy	43	22	19
Madrid, Spain	36	18	15
Detroit	36	18	15

(table 2)

I only selected some of the cities from the articles, but we can obtain in the table that if you cut the death rate into half, then the percent of infectious population will double. In the case of New York City, when the death rate decreased from 1.0% to 0.5%, the ever-infectious population increases from 26 to 51. Probably it sounds not good in the first place, but if you think in terms of herd immunity, then it means that herd immunity is almost a possibility. Not only those recovered people can protect the rest of the susceptible population, but there is also not possible to have a second outbreak since half of the population was immumed by COVID-19. This actually again shares some strategies on how to relax social distancing. For those countries having a lower percentage of susceptible populations, they can relax social distancing such that more people get infectious to create herd immunity.

Relaxing social-distancing:

Again, we take one example to see how this works. According to the formula (1), we can generate the other formula commonly used in SIR model. The formula will be the following:

$$\boldsymbol{R_0(t) * \frac{S(t)}{N} < 1} \quad (6)$$

We may want to control $R_0(t)$ to ensure that the whole term is less than 1 so that the diseases will eventually die out. Or otherwise, there will be an outbreak according to the traditional SIR model. For example, for the city, Lombardy, Italy, after using the model prediction, the susceptible population 77.5 in a month. Then, we are going to control the $R_0(t)$ to be $\frac{N}{S(t)}$ which in this case is 1.3. What we are going to do is to compare with current $R_0(t)$ which is 0.9 to calculate what proportion can we relax our social distancing. The initial R_0 for this city is 2.5. In this case, we can relax our social distancing by about $\frac{1.3-0.9}{2.5-0.9} = 0.25$. Thus, for the city Lombardy, we can relax our social relaxing policy about 25% on the prerequisite that the outbreak will not occur. However, for those cities which have a larger amount of susceptible population, it is probably not a good idea to loosen the social-distancing policy because $\frac{N}{S(t)}$ will be close to 1, which means restricting social distancing is still necessary. Depending on each situation that of different cities, policymakers would need to adjust their policies.

Economy re-opening:

The other concept regarding controlling the diseases is economy re-opening. According to the previous statement about social-distancing, if herd immunity is so useful, why do not we just reopen our economy so that we can get the infectious population as quickly as possible to form herd immunity? However, this is not the case. This is because when we control our infectious population to reach herd immunity, we do not want extra people to be infected and cause the deaths to increase. For example, suppose we only need 60% of infectious individuals to form herd immunity. If we reach our S(t) 40% too fast, the consequences will be that 90% of the population is getting infected. That is being said, 30 percent more people get infected and more people would die. The following graph about Italy will be sufficient to support this argument. In this graph, if Italy reopened its economy in April, then the daily death surges to almost 90 per million people, compared to 20 daily death before. This graph suggests that we need to reinitiate our economy gradually but not open up everything suddenly. If so, it will be a disaster of pandemics.

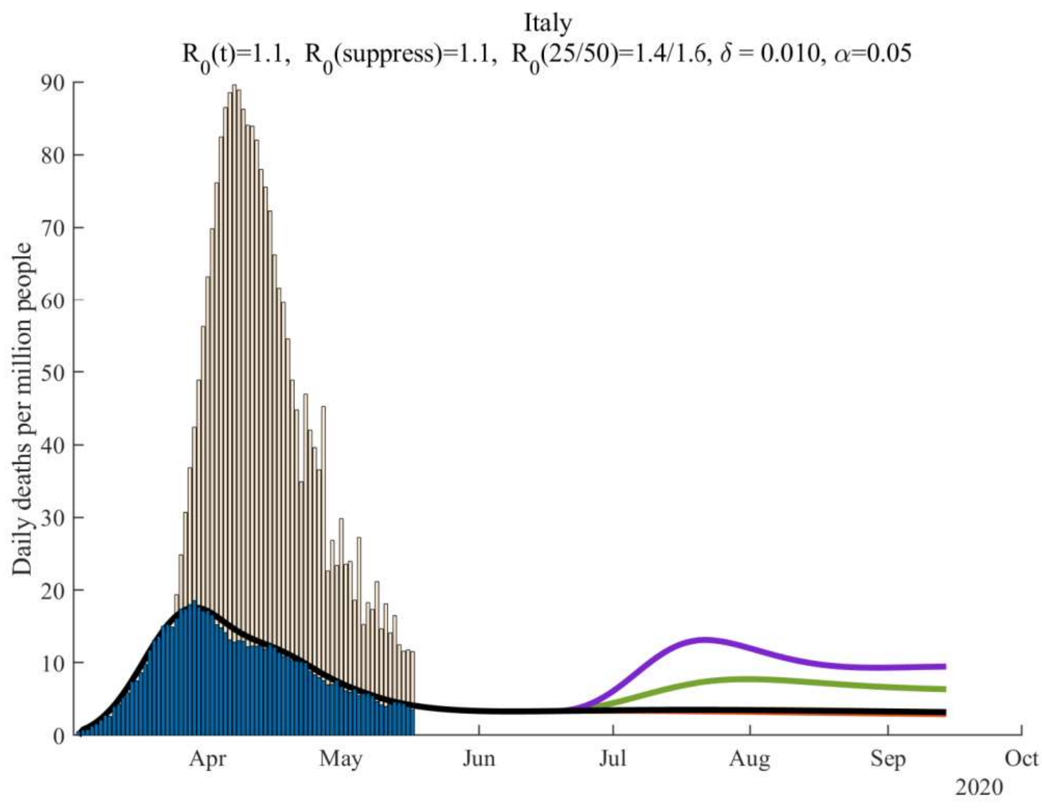

Figure (5)

5. Conclusion:

These two papers all generate good pieces of advice for policymakers and government about how to control the diseases more efficiently in an economical way or in terms of herd immunity. For Chen and Qiu's article, they consider numerous NPIs. By comparing the effects of different combinations of NPIs, they ascertained the most economical and effective way to prevent the disease from spreading, which is MW + SC + CQ. They have also considered the special case, the US, which was experiencing huge growth at that time. Their suggestions are to make the policy more strict so that the new infectious cases will reduce. This paper, I think, is quite insightful and practical. By examing each country's situation, they are able to offer different combinations to NPIs to mitigate the COVID-19 spreading. Compared to Chen and Qiu’s article, Fernandez-Villaverde, and Jones also offered good suggestions for policymakers. In his article, they first discuss how this SIRD model can be fitted into and how its result compares with reality. Then, by doing more stimulations, They can tell how much the government can loosen social distancing by calculating the ratio of $R_0(t)$ meanwhile making sure there is no outbreak of the virus. They also shed a light on addressing the problems of economy re-opening and states that economy re-opening probably is not the right approach for policymakers to take. However, these two articles address questions by the different lens, the baseline of Chen and Qiu’s article is to make infectious population as low as possible to prevent the spreading of COVID-19. However, for Fernandez-Villaverde, and Jones' article, they attach more importance to controlling the disease in terms of herd immunity. Their baseline is as long as there is no outbreak and daily deaths per million people are decreasing, we can re-initiate our society by a certain percentage accordingly. In general, they both expressed some insightful suggestions and strategies for how are we going to behave next regarding controlling the diseases. Under current situations, every process of decision-making is significant and should be taken seriously. More insightful works should be produced and done in the light of decision making.

Reference

[1] Xiaohui Chen and Ziyi Qiu. Scenario analysis of non-pharmaceutical interventions on global covid-19 transmissions. arXiv preprint arXiv:2004.04529, 2020.

[2] Jes´us Fern´andez-Villaverde and Charles I Jones. Estimating and simulating a sird model of covid-19 for many countries, states, and cities. Technical report, National Bureau of Economic Research, 2020.