

CIRD-F: Spread and Influence of COVID-19 in China

ZHOU Lingyun^{a†} (周凌云), WU Kaiwei^{b†} (吴凯伟), LIU Hanzhi^{c†} (刘涵之),
GAO Yuanning^c (高远宁), GAO Xiaofeng^{c*} (高晓枫)

(a. Antai College of Economics and Management; b. SJTU-ParisTech Elite Institute of Technology; c. School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

© Shanghai Jiao Tong University and Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract: The outbreak of coronavirus disease 2019 (COVID-19) has been spreading rapidly in China and the Chinese government took a series of policies to control the epidemic. Therefore, it will be helpful to predict the tendency of the epidemic and analyze the influence of official policies. Existing models for prediction, such as cabin models and individual-based models, are either oversimplified or too meticulous, and the influence of the epidemic was studied much more than that of official policies. To predict the epidemic tendency, we consider four groups of people, and establish a propagation dynamics model. We also create a negative feedback to quantify the public vigilance to the epidemic. We evaluate the tendency of epidemic in Hubei and China except Hubei separately to predict the situation of the whole country. Experiments show that the epidemic will terminate around 17 March 2020 and the final number of cumulative infections will be about 78 191 (prediction interval, 74 872 to 82 474). By changing the parameters of the model accordingly, we demonstrate the control effect of the policies of the government on the epidemic situation, which can reduce about 68% possible infections. At the same time, we use the capital asset pricing model with dummy variable to evaluate the effects of the epidemic and official policies on the revenue of multiple industries.

Key words: coronavirus disease 2019 (COVID-19), epidemic prediction model, negative feedback, capital asset pricing model, dummy variable

CLC number: O 193 **Document code:** A

0 Introduction

While coronavirus disease 2019 (COVID-19) has been spreading rapidly in China and other regions, this virus outbreak has received substantial attention from the global public and was listed by the World Health Organization (WHO) as public health emergency of international concern (PHEIC). Given this situation, there is an urgent need to predict the development of the epidemic. To control the epidemic, the Chinese government took actions rapidly^[1]. They restricted the traffic, extended the Spring Festival holiday, and even closed Wuhan and other cities in Hubei. These measures are actually double-edged swords. If the measures are not strong enough, the virus transmission will become more

difficult to control. Under this circumstance, measures such as traffic control and regional closures are believed to have positive impacts on controlling the epidemic. However, the national economy may be impacted negatively. Thus, some industries may be negatively impacted. Therefore, studying and quantifying the effect of official policies can help us better understand these policies and prepare for the future. In general, we focus our research on the epidemic prediction and influence of official policies.

Traditionally, several existing models, such as cabin models based on propagation dynamics, individual-based models, and machine learning, can be used to predict the tendency of the epidemic. For example, the model of the susceptible, the exposed, the infected and the recovered (i.e., SEIR model) is a cabin model. It divides the population into the susceptible, the exposed, the infected and the recovered, using differential equations to explain the transmission of the virus between these groups. Cellular automation is a commonly-used individual-based model^[2]. It sets detailed rules for population mobility, which accurately simulates the situation in a specific group of people. However, the former type for the predicting models always oversimplifies the

Received date: 2020-02-28

Foundation item: the National Key Research and Development Program of China (No. 2018YFB1004700), the National Natural Science Foundation of China (Nos. 61872238 and 61972254), the Shanghai Science and Technology Fund (No. 17510740200), and the CCF-Huawei Database System Innovation Research Plan (No. CCF-Huawei DBIR2019002A)

***E-mail:** gao-xf@cs.sjtu.edu.cn

[†]These authors contributed equally to this work.

epidemic, leading to inaccurate prediction, while the latter type is too meticulous to simulate the large-scale situation such as the nationwide prediction. Furthermore, there are few related researches on the influence of official policies. Most existing analyses focus on the impacts of the epidemic itself^[3-5]. Also, some researches confuse the influence of official policies and the influence of the epidemic after the policies were published. It may give rise to some misunderstandings of the influence of official policies.

In this paper, we give a reality-based analysis of the transmission process of COVID-19 at first. The analysis considers four groups of people: the incubated, the infected, the recovered, and the dead. It also takes public vigilance to the number of infections and the capacity of virus transmission into consideration, which makes the analysis more concise than SEIR model. Then we build four differential equations and one functional equation, which is called the model of the incubated, the infected, the recovered, the death and an infection function (i.e., CIRD-F model), to describe the transmission of the epidemic. Using the data of cumulative infections, we can conduct regression by the least square method to quantify the public vigilance and the capacity of virus transmission. After that, we can draw curves to predict the number of future infections, incubations, recoveries and deaths. We also use a capital asset pricing model with dummy variable^[6-7], which is called CAPM-DV model, to quantify the influence of official policies on different industries. We first set the value of the dummy variable according to the start time of the epidemic to quantify the influence of the epidemic^[8-9]. Then, we set the value according to the day when the policies were published to quantify the influence of the epidemic after the policies. Therefore, by subtracting the former influence from the latter, we can figure out how the policies influence different industries.

In our experiments, we predict the tendency of the epidemic at first. We fit CIRD-F model in Hubei and China except Hubei separately. The results show that the situation in Hubei is much more severe than the situation out of Hubei. By combining the two situations, we draw curves to predict the tendency of the epidemic. Then we analyze influence of official policies in two aspects. For one thing, we simulate the tendency of the epidemic without the closure policy. If Wuhan and other cities in Hubei were not closed, the situation of the whole country will be as severe as that in Hubei. Therefore, we use CIRD-F model for Hubei to predict the tendency of the epidemic in China, which shows that the policies help reduce about 68% possible infections. For another, we use CAPM-DV model to quantify the economic influence of the policies on the healthcare, transportation, construction, entertainment and tourism industries.

1 Related Work

Existing related work can be divided into three groups: ① epidemiological and clinical features of COVID-19; ② mathematical methods of predicting the epidemic; ③ influence of the official policies.

Epidemiological and Clinical Features Existing researches are mostly based on cases^[10-11] and calculate medical parameters such as the cure rate, the mortality, the incubation period and the basic reproduction number. These data are an important basis for building mathematical prediction models.

Mathematic Models^[2,8,12] The methods of modeling are used to predict the spread of the epidemic. These methods are also used during SARS^[8], Ebola^[12] and so on^[2]. A commonly-used model is the cabin model based on propagation dynamics. One of the most basic models is SEIR model and its modification. This type of models divides people into different categories and describes the transmission of virus using differential equations. The accuracy of parameters and the description of the transmission determine the effectiveness of the model. Another method is to use individual-based models like cellular automaton to simulate the transmission of virus^[2], which is direct and vivid. However, it is difficult to simulate a domestic situation by models of this type. Furthermore, machine learning is also used to conduct epidemic prediction^[12]. This method is based on big data. Besides, a more mature model considers the public vigilance to the epidemic^[8], using negative feedback. However, the model omits the people in incubation period, which oversimplifies transformation process of the infected population in different stages.

Influence of Official Policies For one thing, WHO and the global public fully affirm and praise the performance of the government in the epidemic prevention and control. For another, while some literature points out the negative economic effect of the epidemic^[3], most Chinese newspapers believe that the economy will not be greatly affected. More studies are focused on the influence of the epidemic rather than the effect of official policies. For example, Ref. [8] uses a CAPM-DV model to analyze the impact of the epidemic on economy during SARS.

In this paper, we use a cabin model based on propagation dynamics, which is modified from the model with negative feedback to predict the development of COVID-19^[8]. We also adopt the CAPM-DV model to analyze official policies in an innovative way.

2 CIRD-F Model

To solve the shortcomings of previous models, we build CIRD-F model and obtain more accurate prediction. CIRD-F model is actually a series of differential

equations explaining the transformation process of the infected population in different stages. Four stages of the infected population involve the incubation period, the infection period, the state of recovery from COVID-19, and the state of death due to COVID-19. CIRD-F represents incubation (C), infection (I), recovery (R), death (D) and a function (F) involving a feedback process to explain the public vigilance to the epidemic, as shown in Fig. 1.

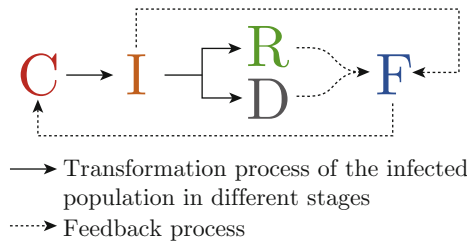


Fig. 1 Transformation process of the infected population in different stages with the feedback process

To simplify and formulate the transformation process of the infected population in different stages, we make some basic assumptions. First of all, since the Chinese government rearranges the medical resources timely and has built two makeshift hospitals, there would be enough beds in the hospital and each patient would be quarantined immediately. Therefore, we assume that the diagnosed patients will be immediately quarantined into the hospital, and then they will not be able to infect others. Secondly, since there is no reported reinfection right now, we assume that the cured patients will not be reinfected and will no longer have the ability to spread the disease. Finally, since there is no specific drug or available vaccine, we assume that the cure rate and the mortality rate remain unchanged in this paper. Namely, we do not consider the occurrence of specific drugs for the time being.

On the basis of the assumptions, we define the four stages of infected population more precisely. The incubation period is the first stage which a person experiences after he is infected. People in the incubation period show no symptoms and act like the healthy people. Therefore, they may infect the healthy people and turn the healthy into those in the incubation period. According to the statistics recently collected^[6], the length of incubation period is three days. The people in the infection period are those who have undergone the incubation period. People in the infection period show obvious symptoms, so they will be quarantined in the hospitals immediately and be not able to infect the healthy. Since different patients get different therapeutic effects in the process of hospital treatment, some people in the infection period will die and others will survive, which corresponds to the state of recovery from COVID-19 and the state of death due to COVID-19. People in the

states of both recovery and death will not infect the healthy.

We reckon that the cumulative number of confirmed cases (the total number of people in the infection period, the state of recovery and the state of death) will affect public vigilance to the disease. With the increase of the number of confirmed cases, public vigilance is constantly improved. Because of the increased public vigilance, the same number of cases of persons in the incubation period will lead to fewer healthy people being infected than the previous situation. In other words, the growth of the confirmed cases decreases the growth rate of people in the incubation period, which is called a feedback process. The qualification of the feedback process will be explained later.

Figure 1 shows the transformation process of the infected population in different stages with the feedback process, which is the basic idea of CIRD-F model. In this figure, we can see two types of lines, the dashed lines and the solid lines. The solid lines connect different stages of the patients and show the transformation process of the infected population in different stages. The growth rates of the people in the incubation period and the states of recovery and death remain unchanged, which will be later explained in detail. The dashed lines denote the feedback process describing the infection function. The infection function is the growth rate of people in the incubation period which is affected by the confirmed cases and affects the rate of incubated people infecting the healthy.

Based on the above analysis, CIRD-F model is shown as follows:

$$F = a(n_I + n_D + n_R) + b, \tag{1}$$

$$\frac{dn_C}{dt} = -\alpha n_C + F n_C, \tag{2}$$

$$\frac{dn_I}{dt} = \alpha n_C - \beta n_I - \gamma n_I, \tag{3}$$

$$\frac{dn_R}{dt} = \gamma n_I, \tag{4}$$

$$\frac{dn_D}{dt} = \beta n_I, \tag{5}$$

where

$$\alpha = \frac{1}{C_t}, \quad \beta = \frac{1}{T} \sum_{i=2}^T \frac{n_{D,i} - D_{D,i-1}}{I_i},$$

$$\gamma = \frac{1}{T} \sum_{i=2}^T \frac{n_{R,i} - D_{R,i-1}}{I_i}.$$

Here, F represents the infection rate, n_C represents the number of people in the incubation period, n_I represents the number of real-time infections, n_R represents the number of cumulative cures, n_D represents the number of deaths, and t represents the outbreak duration;

a represents the public vigilance to the number of infections, and b represents the capacity of virus transmission; α , β and γ represent the proportion of people becoming infected after the incubation period every day, the death rate and the cure rate, respectively; C_t is the length of the incubation period. The subscript i represents the i th day after 23 January 2020, and T is the maximum value of i . What needs to be emphasized here is that F is a feedback function.

The meanings of the five equations are explained as follows. As to the first equation, F is impacted by n_R , n_D and n_I together. Here, we set $n_I + n_D + n_R = n_T$, where n_T is the number of the total infections (cumulative infections). With the increase of the total number of infected people, we believe that the government will take more stringent measures to control the epidemic, and the people will have a deeper understanding of the epidemic. Therefore, with the growth of n_T , F will decline, which represents that the infectious ability of disease in human society is gradually weakened. The second formula shows the law of population growth in the incubation period in unit time. Some people in the incubation period become confirmed patients (infections), and the ratio of this process is α . Some healthy people are still in the incubation period, which is directly affected by the infection index F . The third formula shows the increase in the number of diagnosed infections per unit of time. The proportion of people with a ratio of α changes from incubation period to confirmed patients. There are also a group of people cured and another died as a result of failed rescue efforts. The fourth and fifth equations are very simple, indicating that a certain number of patients will turn into healthy people and death cases under a certain cure rate and mortality in a unit time.

We use the data from 23 January 2020 to conduct the calculation. The initial values of n_C , n_I , n_R and n_D are set as follows: $n_{I,0}$, $n_{R,0}$ and $n_{D,0}$ are the numbers of the ones on 23 January 2020. Since the incubation period is three days, $n_{C,0}$ is the difference of the cumulative number of infections on 25 January 2020 and 26 January 2020. In our fitting process, the sum of squares of residuals is required to be minimized to achieve the best result, so we use the least square method to fit.

Here are the strengths and weaknesses of CIRD-F model. In terms of strengths, our model takes the incubation period into account. Therefore, the model is more accurate and closer to the reality. We set up a negative feedback function, in which a and b can well reflect the attitude of the government and people to the epidemic, making the later policy analysis more convenient. Our model is not a traditional warehouse model, which requires less data and has higher accuracy. There are also some weaknesses of our model. In the modeling process, it is assumed that people would fully follow the instructions of the government. Such a prediction may

be more ideal than the real situation. In our calculation process, the calculation of cure rate and death rate is not very rigorous. In reality, both the cure rate and the death rate are functions of time, which is not reflected in the calculation process.

3 Mechanism of CAPM-DV Model

Through CIRD-F model, we can figure out the development of the epidemic and the effectiveness of official policies, but it is still difficult to quantify the economic impact of official policies. Therefore, we need a capital asset pricing model with a dummy variable to solve this problem. Traditional capital asset pricing model (CAPM) mainly studies the relationship between expected return of assets in the securities market and systematic risk^[4]. Therefore, it shows how a certain capital asset is influenced in the background of a certain market. The dummy variable is a qualified independent variable^[3], which is used to analyze the effect of abnormal factor. The model is

$$(r_E - r_f)_i = \omega(r_M - r_f)_i + \lambda n_{D,i} + \epsilon, \quad (6)$$

where r_E means the daily expected return of certain capital asset, r_f means the daily risk-free interest rate, r_M means the daily expected return of the certain market, ω is system risk which is defined in the traditional CAPM, ϵ is the disturbance, and λ is the influential factor. We need to fit ω and λ . The focus of this analysis is λ . If λ is significantly greater than 0, it indicates that the capital asset is influenced by the abnormal factor positively; if λ is significantly less than 0, it indicates that the capital asset is influenced by the abnormal factor negatively.

Recent study with auto-regressive distributed lag model shows that the increase and decrease in economic factors impact the stock returns, which are very sensitive, in China^[13]. Furthermore, a study which focuses on several Central and Eastern European countries shows that there is a high level correlation between each country's gross domestic product (GDP) and its national stock exchange index^[14]. Since the GDP or GDP per capita of most of these countries is similar to that of China, we infer that the result also applies to China. On the basis of these studies, we assume that there is a strong relationship between the stock market and the economy in China. Therefore, we focus on the Shanghai Stock Exchange and reckon that the stock indices in certain industry represent the economy situation of certain industry.

In this way, "certain capital asset" of CAPM-DV model means "certain industry". "Certain market" is referred to in particular as Shanghai Stock Exchange in this paper. Furthermore, we use the five-year government bonds to represent the risk-free interest. Namely, $r_{E,i}$ means the daily income rate of the stock in certain

industry on the i th day; $r_{f,i}$ means the daily income rate of five-year government bonds, which remains unchanged; $r_{M,i}$ means the daily income rate of Shanghai Composite Index on the i th day. We use equal weight method to calculate the indices $r_{E,i}$ and $r_{M,i}$ of each industry separately. The complete calculations are

$$P_{E,i} = \sum_{j=1}^m P_{i,j}, \tag{7}$$

$$r_{E,i} = \ln \frac{P_{E,i}}{P_{E,i-1}}, \tag{8}$$

$$r_{M,i} = \ln \frac{P_{M,i}}{P_{M,i-1}}, \tag{9}$$

where $P_{i,j}$ means the closing price of the j th stock in certain industry on the i th day, $P_{E,i}$ means the daily income of certain industry on the i th day, m means the number of stocks in certain industry, and $P_{M,i}$ means the closing price of Shanghai Composite Index on the i th day.

After we set the range of the subscript i and the dummy variable $n_{D,i}$, which will be explained in Subsection 4.2.2, we use linear regression to fit ω and λ . If $n_{D,i}$ equals to 0 and 1 before and after the epidemic broke out respectively, the absolute value of λ means the level of influence of the epidemic on certain industry and the sign of λ reflects whether the epidemic influences the industry positively or negatively. Similarly, if $n_{D,i}$ equals to 0 and 1 before and after the government published official policies respectively, the absolute value of λ means the level of influence of the epidemic after official policies were published on the industry and the sign of λ reflects whether the influence on the industry is positive or negative. It should be noted that ω means the level of risk of the stocks in certain industry and it is not the focus of our research.

4 Experiments and Analyses

We conduct several experiments to predict the tendency of the epidemic and evaluate the influence of official policies. In Subsection 4.1, we collect the cumulative number of infections from 23 January 2020 to 1 March 2020 from the National Health Commission of the People’s Republic of China. On the basis of these data, we use CIR-D-F model to predict the tendency of the epidemic. In Subsection 4.2.1, we use the result in Subsection 4.1 to evaluate how many possible infections will be reduced by official policies. In Subsection 4.2.2, we randomly collect closing prices of 42

listed company stocks in Shanghai Stock Exchange and Shanghai Composite Index from 2 January 2019 to 20 February 2020 from China International Capital Corporation (CICC) wealth. According to these data, we evaluate the economic influence of official policies on five industries separately.

4.1 Epidemic Prediction

In this section, we separate the procedure into two different parts: the data prediction of Hubei and the data prediction of the whole country except Hubei. The reason why we separate the procedure is that most cities in Hubei have taken strict traffic control. Also, Hubei is a major epidemic area and the medical resources are saturated. Therefore, the cure rate γ , the death rate β , the public vigilance a , and the virus transmission capacity b in Hubei are evidently different from the ones in the other areas. After making the predictions separately, we figure out the trend of the epidemic of the whole country.

For the data prediction of China except Hubei, we use the total number of infections (i.e., n_T) from 23 January 2020 to 1 March 2020 to fit the parameters a and b . For Hubei, since the government revised the diagnosis standard on 13 February 2020 and the total number of infections that day increased by over ten thousand, we modify the range of data used for fitting accordingly. We assume that the reason why the government revised the standard was that as the epidemic developed, the old diagnosis standard lost its effectiveness. Therefore, we reckon that the new standard after 13 February 2020 is effective and the old standard stayed valid until 6 February 2020. On the basis of the assumption, we use the total number of infections from 23 January 2020 to 6 February 2020 and from 13 February 2020 to 1 March 2020 to fit the parameters a and b .

Using the corresponding data in Hubei and China except Hubei, we acquire the values of a and b separately. Since we set the steps of a and b as 5×10^{-7} and 5×10^{-3} respectively, we also set the prediction intervals of a and b as $\pm(5 \times 10^{-7})$ and $\pm(5 \times 10^{-3})$ respectively. The values and the prediction intervals of the used parameters are shown in Table 1. Figure 2 shows the tendency and changes of different populations (recovered people, death toll and so on) in Hubei. Figure 3 shows the tendency in China except Hubei. In the two figures, the solid line means the estimated number of infections in different stages, the shadow means the prediction intervals and the scatter means the data used to fit the model.

Table 1 The values of the used parameters in CIR-D-F model

Region	α	$\beta \times 10^3$	$\gamma \times 10^3$	$n_{C,0}$	$n_{I,0}$	$n_{R,0}$	$n_{D,0}$	$a \times 10^6$	$b \times 10^2$
Hubei	1/3	4.64	24.03	371	494	31	24	$-(9 \pm 0.5)$	63 ± 0.5
China except Hubei	1/3	0.56	53.14	398	277	3	1	$-(49 \pm 0.5)$	64 ± 0.5

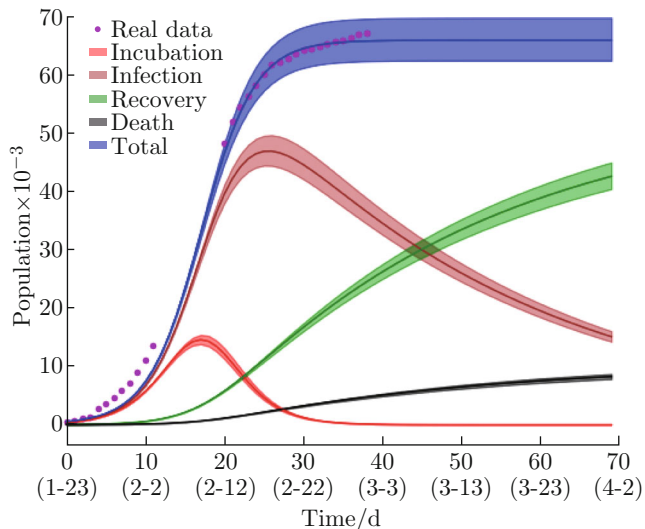


Fig. 2 Data prediction of Hubei from 23 January 2020 to 2 April 2020

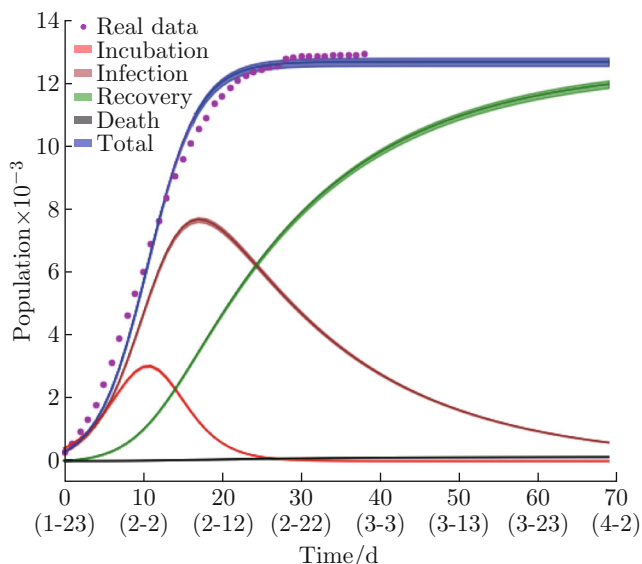


Fig. 3 Data prediction of China except Hubei from 23 January 2020 to 2 April 2020

Through comparing the values of a and b in Hubei and China except Hubei, we find that the absolute value of a in China without Hubei is significantly greater than that in Hubei and the values of b in both regions are basically the same. The reason to the first fact is that the absolute value of a represents the public vigilance to the number of cumulative infections, and since the epidemic is more severe in Hubei, the vigilance in Hubei is weaker than that in China without Hubei. The reason to the second fact is that the value of b represents the capacity of virus transmission which only depends on the virus itself and remains unchanged.

We assume that the infected people are quarantined in hospitals. Therefore, we define the end of the epi-

demical as the day when the number of people in the incubation period drops to zero. According to our prediction, the number of people in the incubation period in Hubei will drop to zero in the middle of March, about 17 March 2020. Furthermore, if the situation in Hubei is not taken into account, the national epidemic will end around 2 March 2020. In other words, the last patient in the incubation period in the country will be diagnosed and quarantined on 2 March 2020.

The actual data of the number of infections can be used to prove the accuracy of our prediction. Comparing the estimation interval of the total number of infections (the blue shadow) with the actual data (the pink scatters) in Figs. 2 and 3, we find that all the data predictions are in the estimation interval.

We assume that the circulation of personnel in Hubei and other areas of the country is seriously blocked, so we simply add the results of the previous two predictions and get the prediction of the whole country. The tendency of the epidemic is shown in Fig. 4. Figure 5 shows the timeline of the development of the epidemic. From the figure, several important time nodes and corresponding data of the national epidemic can be obtained. As the reality, the peak of the national epidemic came around 17 February 2020, when the total number of infected people was about 72 265 (prediction interval, 70 453 to 73 684). The national epidemic will end on 17 March 2020, which is the same as the end date of the epidemic in Hubei. By that time, about 78 191 (prediction interval, 74 872 to 82 474) people will be infected.

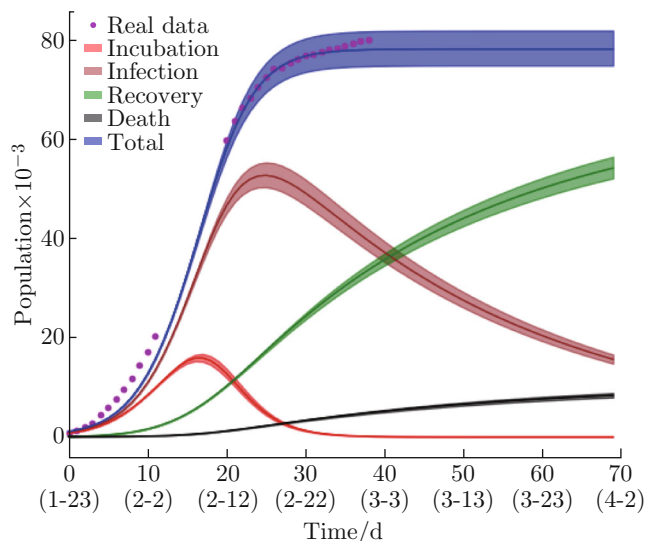


Fig. 4 Data prediction of the whole country from 23 January 2020 to 2 April 2020

We also calculate the goodness-of-fit of R^2 for the models of Hubei and China except Hubei. We arrange the actual data of daily infection total (i.e., $n_{I,i} + n_{D,i} + n_{R,i}$, recorded as $n_{T,i}$), which is defined in Eq. (2) and

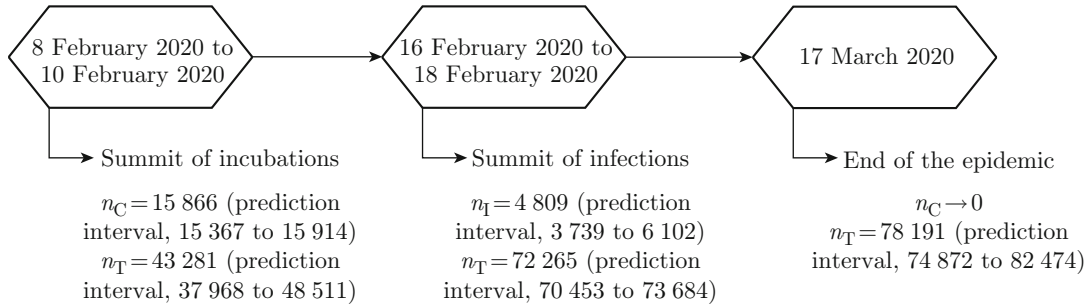


Fig. 5 Timeline of the trend of the epidemic

used for fitting, in chronological order as $n_{T,1}, n_{T,2}, \dots, n_{T,K}$, where K means the total number of the samples. The coefficient of determination is

$$R^2 = 1 - \frac{\sum_{k=1}^K (n_{T,k} - \hat{n}_{T,k})^2}{\sum_{k=1}^K (n_{T,k} - \bar{n}_{T,k})^2}, \quad (10)$$

where $n_{T,k}$ means the number of daily infection total of the k th sample, $\bar{n}_{T,k}$ means the average of all the samples, and $\hat{n}_{T,k}$ means the estimated number of daily infection total corresponding to the date of the k th sample. The values of R^2 for the models of Hubei and China except Hubei are 0.99599 and 0.99078 respectively, which are both close to 1 and show the high goodness-of-fit. However, while the actual number of the daily increased infections ($n_{T,i} - n_{T,i-1}$) in China except Hubei remains between 6 and 10 for about ten days from 23 February 2020, the corresponding estimated numbers of daily increased infections decrease rapidly to zero. It shows that the model cannot describe the ending period of the epidemic very accurately.

4.2 Influence of Official Policies

In this section, we consider the policy impact of the government from two aspects. These two aspects are the control of the epidemic and the impact on the economy. To analyze the impact of epidemic control, we use the previous model. As to the economic impacts, we will use a CAPM with dummy variable to assess the impact on different industries. Since the Chinese government mainly published its policies on 23 January 2020, we assume that the influence of official policies started from 23 January 2020.

4.2.1 Reducing Possible Infections

In the previous work, we have predicted the epidemic situations in Hubei and China except Hubei, and obtained some crucial parameters. Since the epidemic in Hubei is much more severe than that out of Hubei, the government published a series of policies on 23 January 2020. If these policies were not published, the situation of the whole country would become as severe as that of

Hubei. To simulate the above situation, we use the values of the parameters α, β, γ, a and b in Hubei and the initial numbers $n_{C,0}, n_{I,0}, n_{R,0}$ and $n_{D,0}$ in the whole country and solve the differential equations. The values of the parameters mentioned above are shown in Table 1. Figure 6 shows the development of the epidemic if the policies were not published. As shown in Fig. 6, the predicted number of cumulative infections greatly exceeds the actual number of cumulative infections, which means the control of the government is effective. If it is not controlled, there will be 131 594 (prediction interval, 124 655 to 139 350) cases of infection at the end of the epidemic. The comparison between the simulation and the current situation shows that official policies decrease the number of possible infections by about 68%.

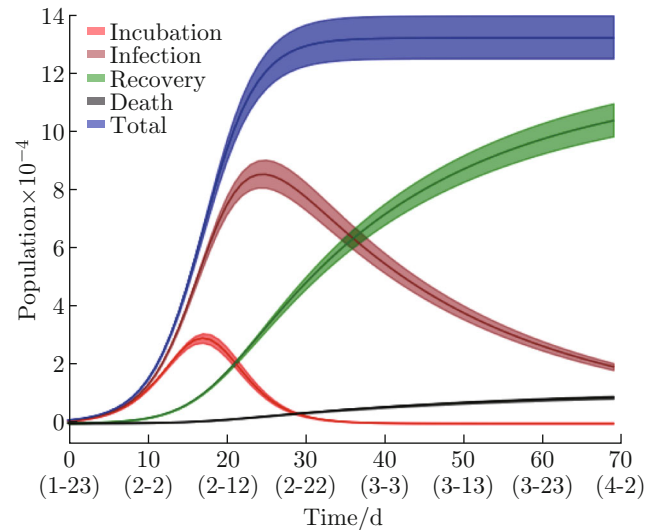


Fig. 6 Situation of the whole country if the policies were not published from 23 January 2020 to 2 April 2020

Moreover, we use the values of the parameters α, β, γ, a and b in China without Hubei and the initial numbers $n_{C,0}, n_{I,0}, n_{R,0}$ and $n_{D,0}$ in the whole country and solve the differential equations. The values of the parameters mentioned above are shown in Table 1. The above simulation reflects the situation that Hubei acts like other regions in China. Figure 7 shows the

situation that Hubei implements strict control at the early stage of the outbreak. It can be seen that the total cumulative infections will be far lower than the current cumulative infections. If Hubei takes actions like other regions, there will be only 24787 (prediction interval, 24534 to 25044) cases of infections at the end of the epidemic. The comparison between the simulation and the current situation shows that if the Hubei takes action as quickly as the other regions, the number of possible infections will decrease by about 68%.

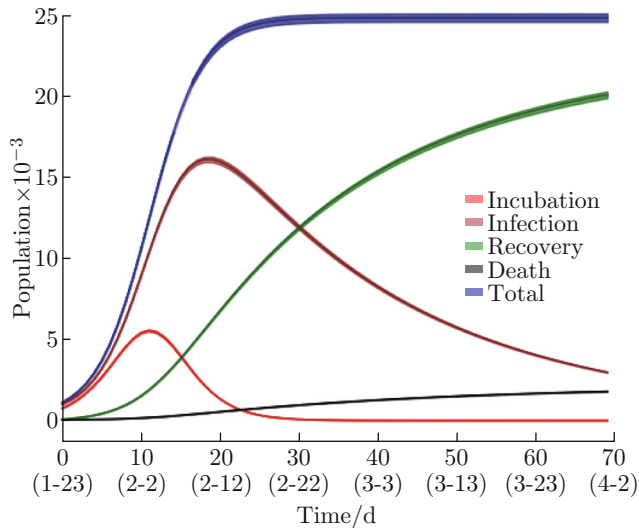


Fig. 7 Situation of the whole country if Hubei takes actions more timely from 23 January 2020 to 2 April 2020

4.2.2 Economic Impacts

In this section, we use CAPM-DV model to calculate the economic impacts of the epidemic and official poli-

cies on different industries. The industries are health-care, transportation, construction, tourism and entertainment. Since the stock index can represent the performance of certain industry in a large economic environment, we randomly collect closing prices of 42 listed company stocks in the five industries from 2 January 2019 to 20 February 2020. The companies are all listed in Shanghai Stock Exchange and located in all over the world. We list the stock code and its corresponding industry in Table 2.

After collecting the stock indices of the above stock codes from CICC wealth, we gain the values of $r_{f,i}$, $P_{M,i}$ and $P_{i,j}$. The usual closing price $P_{M,i}$ of Shanghai Composite Index ranges from 2000 to 3000, the usual closing price $P_{i,j}$ of each stock ranges from 0 to 15 yuan, and $r_{f,i}$ equals to 0.117%. The range of the subscript i will be explained later. Assume that we have determined the range of i . On the basis of the above data, we use the equal weight method, which is described in Section 3, to calculate the parameters $r_{E,i}$ and $r_{M,i}$ in CAPM-DV model.

Economic Influence of the Epidemic First, we analyze the impact of the epidemic on each industry by comparing the performances of the stock market before and after the epidemic broke out. Therefore, the range of subscript i is from 2 January 2019 to 20 February 2020 and the dummy variable $n_{D,i}$ is set as: when i is later than 31 December 2019, $n_{D,i}$ equals to 1; when i is between 2 January 2019 and 30 December 2019, $n_{D,i}$ equals to 0. December 31, 2019 is the day when Wuhan reported the epidemic. Therefore, we can use linear regression to calculate the influential index λ for each industry separately. The results are shown in Table 3.

Table 2 Data source of CAPM-DV model

Industry	Stock code
Healthcare	600222 600223 600380 600466 600789 600796 600812 600866
Transportation	600029 600386 600611 600708 600834 601111 601766
Construction	600039 600170 600463 600502 600820 600970 600986 601226 601789
Tourism	600054 600258 600555 600593 600754 601007 603043 603099 603869
Entertainment	600037 600088 600246 600386 600633 600716 600831 601801

Table 3 Influential index of the epidemic

Industry	$\lambda \times 10^3$	Description
Healthcare	3.73	The epidemic brings 0.373% additional income to the healthcare industry
Transportation	-3.48	The epidemic causes 0.348% additional loss to the transportation industry
Construction	-2.45	The epidemic causes 0.245% additional loss to the construction industry
Tourism	-0.58	The epidemic causes 0.058% additional loss to the tourism industry
Entertainment	—	The epidemic does not influence the entertainment industry directly

Economic Influence of Official Policies Then, we analyze the impact of the epidemic on each industry when official policies were conducted by comparing the performances of the stock market before and after the policies were conducted. Therefore, the range of subscript i is from 31 December 2019 to 20 February 2020 and the dummy variable $n_{D,i}$ is set

as follows: when i is later than 23 January 2020, $n_{D,i}$ equals to 1; when i is between 31 December 2019 and 22 January 2020, $n_{D,i}$ equals to 0. January 23, 2020 is the day when the Chinese government started publishing policies. Similarly, we conduct the calculation for each industry and show the results in Table 4.

Table 4 Influential index of the epidemic after the policies were published

Industry	$\lambda \times 10^3$	Description
Healthcare	-0.64	The epidemic causes 0.064% additional loss to the healthcare industry
Transportation	-2.60	The epidemic causes 0.260% additional loss to the transportation industry
Construction	-3.45	The epidemic causes 0.345% additional loss to the construction industry
Tourism	-2.20	The epidemic causes 0.220% additional loss to the tourism industry
Entertainment	—	The epidemic does not influence the entertainment industry directly

If we subtract the values of λ in Table 3 from the corresponding ones in Table 4, we can calculate the economic impact of official policies. On the positive side, the policies bring 0.088% additional income to the transportation industry. On the negative side, the policies cause 0.437%, 0.1% and 0.162% additional losses to the healthcare, construction and tourism industries, respectively.

There are some interesting points. ① Although the epidemic stimulates the healthcare industry, official policies negatively influence this industry. Possibly, the reason is that when the epidemic continues, public emotions have stabilized. Thus, most people stay at home, no longer buying drugs blindly. ② Although the epidemic negatively impacts the transportation industry, the policies show positive to this industry. Possibly, the reason is that the epidemic has reduced the tourism, but the deployment of materials and personnel after the policies were conducted relatively stimulates the transportation industry.

The model provides a universal method to calculate the impact on industries. As long as we get the stock indices of several companies in a certain industry, we can calculate the impact of the government on this industry quickly. More commonly, the dummy variable can be used to calculate the impact of a certain period on a certain industry. We can set different interval of i to determine when the dummy variable $n_{D,i}$ equals to 0 and when it equals to 1. However, in this paper, we do not collect the data of all listed companies in certain industry, so errors probably exist.

5 Conclusion

We provide CIRD-F model to analyze the spread and influence of COVID-19. We use different parameters to predict the trend of the epidemic in and out of

Hubei. The influence of official policies is divided into two parts: virus transmission, and economy.

When predicting the tendency of the epidemic, we adopt the negative feedback to quantify public vigilance and consider the difference between people in the incubation and infection period. We also take different situations in and out of Hubei into consideration. The results can be used to evaluate the possible infections which are reduced by official policies. When quantifying the economic influence of the policies on different industries, we calculate the influence of the epidemic and the influence of the epidemic after the policies were published. Using these results, we can evaluate how the policies influence different industries.

According to our calculation, the termination date of the epidemic in Hubei are different from that of other parts of China. The former is around 17 March 2020, while the latter is around 2 March 2020. Therefore, the epidemic in China will terminate on 17 March 2020. It means that the outbreak will not last long. According to the different situations in and out of Hubei, we find that official policies are very helpful in reducing the number of infections. From the perspective of China, if the control is not timely, there will be about 50 000 more infected people. If Hubei acts like other regions in China, the number of infected people will be greatly reduced, which further confirms the importance of early prevention and control of the epidemic. Furthermore, we use CAPM-DV model to calculate the economic impacts of the epidemic and official policies on different industries. The epidemic influences the healthcare industry positively and impacts the transportation, construction and tourism industries negatively. By contrast, the policies relatively influence the construction industry positively and impact the transportation, healthcare and tourism industries negatively.

However, when using the CIRD-F model, we assume that the cure rate and the death rate remain unchanged to simplify the calculation. In the future work, we can treat the cure rate and the death rate as a function of time to give a more accurate prediction.

References

- [1] KHAN S, ALI A, SIDDIQUE R, et al. Novel coronavirus is putting the whole world on alert [J]. *Journal of Hospital Infection*, 2020, **104**: 252-253.
- [2] SIRAKOULIS G C, KARAFYLLIDIS I, THANAILAKIS A. A cellular automaton model for the effects of population movement and vaccination on epidemic propagation [J]. *Ecological Modelling*, 2000, **133**: 209-223.
- [3] LUM L H W, TAMBYAH P A. Outbreak of COVID-19: An urgent need for good science to silence our fears? [J]. *Singapore Medical Journal*, 2020, **61**(2): 55-57.
- [4] DE LA SEN M. On the design of hyperstable feedback controllers for a class of parameterized nonlinearities: Two application examples for controlling epidemic models [J]. *International journal of Environmental Research and Public Health*, 2019, **16**(15): 2689.
- [5] GUAN W, NI Z, HU Y, et al. Clinical characteristics of coronavirus disease 2019 in China [J]. *The New England Journal of Medicine*, 2020. <https://doi.org/10.1056/NEJMoa2002032> (published online).
- [6] JOHNSON Q E, ESSI I D. Dummy variable regression model and two-stage nested design of agricultural variables [J]. *International Journal of Applied Science and Mathematical Theory*, 2018, **4**(1): 15-41.
- [7] KUEHN L A, SIMUTIN M, WANG J J. A labor capital asset pricing model [J]. *The Journal of Finance*, 2017, **72**(5): 2131-2178.
- [8] ZOU Y T, ZHEN X L, LIAO X H, et al. The mathematical principle of the spread of SARS and its application on forecasting and controlling SARS epidemic [J]. *Journal of Engineering Mathematics*, 2003, **20**(7): 29-34 (in Chinese).
- [9] SAFARI S, ERFANI A R. A new method for fuzzification of nested dummy variables by fuzzy clustering membership functions and its application in financial economy [J]. *Iranian Journal of Fuzzy Systems*, 2020, **17**(1): 13-27.
- [10] ZHANG Y P. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China [J]. *Chinese Journal of Epidemiology*, 2020, **41**(2): 145-151 (in Chinese).
- [11] CHEN N S, ZHOU M, DONG X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: A descriptive study [J]. *The Lancet*, 2020, **395**: 507-513.
- [12] COLUBRI A, SILVER T, FRADET T, et al. Transforming clinical data into actionable prognosis models: Machine-learning framework and field-deployable App to predict outcome of Ebola patients [J]. *PLoS Neglected Tropical Diseases*, 2016, **10**(3): e0004549.
- [13] KHAN M K, TENG J Z, PARVIAZ J, et al. Nexuses between economic factors and stock returns in China [J]. *International Journal of Economics and Finance*, 2017, **9**(9): 182-191.
- [14] PRZEKOTA G, REMBEZA J, MENDEL G, et al. The relationship between the stock market and the economy: Evidence from Central and Eastern European countries [J]. *Transformations in Business & Economics*, 2019, **18**(2): 397-415.