Economic Data Analysis of 50 States in USA by Considering the Influence of COVID-19

Yuechen Mu
muyueche@msu.edu

Department of Mathematics, Natural Science College, Michigan State University, East Lansing, Mi, USA

Abstract----The world was met by a challenge from Covid-19 since 2020. The biggest loss of the pandemic is often considered to be social-economic. This paper mainly focus on the impact of Covid-19 on the economic growth of the United States with a consideration of several key parameters such as GDP, personal income and industry earnings. In order to find a more accurate time reference. All data are dcollected from the national economic report from the first quarter of 2018 till the last of 2020. After that we
At first, the paper compared the variation of key economic parameters of 50 states before and after the burst of corona virus. After that, the economic status of each state was categorized into three groups that reflect the overall financial situation in the light of K-means analysis. This paper further investigated the dynamic changes of these groups and predicted future trend.
As a result, the GDP by industry witnessed a huge decline due to the propagation of COVID-19. It reached the lowest value in the 2020 second quarter and increases after that time. In addition, the total personal income began to decrease in the second quarter of 2020. When the GDP plummeted to the all-time low, the personal income also decreases. The Finance and Insurance Earnings may not be impacted by the COVID-19, or the COVID-19 stimulate its growth. The reason may be caused by the online style of these businesses.

Keywords: Clustering; prediction; Mathematical analysis

1 INTRODUCTION

2020 is a year of great challenges facing the world. In order to specifically assess the specific impact of the COVID-19 on the social economy, this article aims to process and analyze existing economic data, and then obtain relatively intuitive results through mathematical modeling, and explain the economic impact of the covid-19 more rationally and academically. In order to prepare for my research, I first studied the knowledge of data analysis and understood the composition and processing methods of mathematical data. I believe this can help me think better about the problem I want to study.

With the continuous spread of the COVID-19, the growth of the sick population has greatly hindered economic development. For Timothy W Russell, Joseph T Wu, Sam Clifford, W John Edmunds, Adam J Kucharski, and Mark Jit's mathematical group (2020) [1], they defined the internationally imported by trade or other economic activates of COVID-19 relate to the modeling. Ashleigh R. Tuite, David N. Fisman, and Amy L. Greer (2020) [2] set up the modeling of the COVID-19 in the fixed place to explain the transmission and mitigation strategies of the economy. In addition, many scholars have conducted academic discussions on

ICEMME 2022, November 18-20, Nanjing, People's Republic of China

DOI 10.4108/eai.18-11-2022.2327199

the relationship between the COVID-19 and the social economy. Many of them put forward their own different views on the same topic. Richard Baldwin and Beatrice Weder di Mauro (2020) [3] defined the relationship between covid-19 and the aggregate demand and an aggregate supply shock of China through the data analysis. Ang Abel Brodwur's group (2020) [4] uses the data of COVID-19 deaths to defend the economic consequences of COVID-19 also influence the government response. And RakeshPadhana and K. P. Prabheeshb (2021) [5] focused on the same topic about the government policies, they also observe the adverse economic effects during the COVID-19 and survey the business the both national and international. In addition, in Daniel Susskind and David Vines' article (2020) [6], they further explain the economic interventions.

2 Materials and Methods

2.1 GDP by Industry

GDP is generally regarded as an important indicator to measure the economy. Through data collection, in Figure 1, it shows some examples of the GDP by Industry from 2018Q1 to 2020Q4GDP by Industry in 50 states. All the data are real and valid, they come from the authoritative national economic data website.

	2018Q1	2018Q4	2019Q3	2019Q4	2020Q1	2020Q2	2020Q3	2020Q4
Alabama	$217,867.90	$224,372.70	$229,328.20	$230,750.10	$229,831.40	$209,852.30	$228,062.30	$231,736.30
Alaska	$53,536.90	$54,765.10	$54,449.90	$54,674.70	$52,864.40	$45,644.40	$50,412.90	$52,065.00
Arizona	$344,512.80	$357,076.80	$373,173.00	$379,018.80	$377,475.90	$350,140.90	$378,296.50	$383,930.50
Arkansas	$125,956.30	$129,333.40	$131,556.20	$132,596.40	$131,818.00	$120,812.40	$130,709.50	$132,955.60
California	$2,918,947.60	$3,023,522.60	$3,143,837.30	$3,205,000.10	$3,189,702.60	$2,893,053.90	$3,120,386.10	$3,164,343.40
Texas	$1,755,051.10	$1,824,833.20	$1,850,626.40	$1,861,581.90	$1,818,394.50	$1,628,185.00	$1,772,131.60	$1,820,226.50
Utah	$176,877.80	$184,972.20	$193,976.00	$196,639.40	$195,594.40	$182,894.60	$198,630.20	$202,823.90
Vermont	$32,554.50	$33,320.90	$34,117.20	$34,320.20	$34,016.60	$30,174.80	$33,277.80	$33,717.50
Virginia	$75,294.10	$79,484.80	$79,078.50	$78,480.50	$76,366.30	$68,000.80	$74,510.90	$75,958.70
Washingtor	$560,333.30	$585,621.80	$618,623.00	$624,861.40	$625,120.60	$579,694.90	$632,013.50	$637,990.70
West Virgin	$75,294.10	$79,484.80	$79,078.50	$78,480.50	$76,366.30	$68,000.80	$74,510.90	$75,958.70
Wisconsin	$333,080.90	$342,197.30	$350,017.40	$353,935.50	$348,021.00	$314,026.50	$344,500.00	$348,165.90
Wyoming	$38,954.80	$40,227.20	$40,500.80	$40,764.30	$38,680.20	$33,233.10	$35,999.50	$37,053.30

Figure1 Examples of GDP by Industry from 2018Q1 to 2020Q4

Each state presents a different situation. However, most states showed their maximum value in the fourth quarter of 2019 with blue high light, which may be the peak of the industrial economic development in the United States. It then began to decline at the end of 2019 and reached a record low in the second quarter of 2020. This may mean that the impact of theCOVID-19 is gradually expanding.

This observation proves the impact of the COVID-19 on the GDP by the industry. The situation has improved after the minimum point of third quarter of 2020, which may mean that people have learned from the unprepared plague.

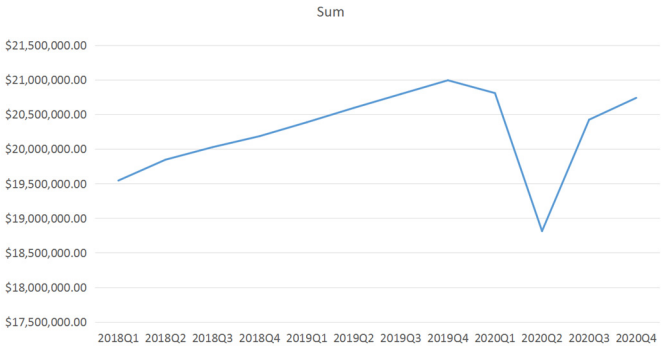

Figure 2 Sum of GDP by Industry of 50 states

In order to make the data clearer, I made a statistical graph of the sum of all states and time period which show in Figure 2. We can clearly see the economic trend of the U.S. nationwide based on this line chart. 2020Q2 is the worst time to be hit by the COVID-19. And when we go back to the Figure 1, those states whose maximums (red high light) appear in the last quarter of 2020 mean that they are recovering quickly from the COVID-19. They are Alabama, Arizona, Arkansas, Colorado, Georgia, Idaho, Indiana, Iowa, Mississippi, North Carolina, South Dakota, Utah, Washington. The best example is Arkansas which shows in Figure 3 below.

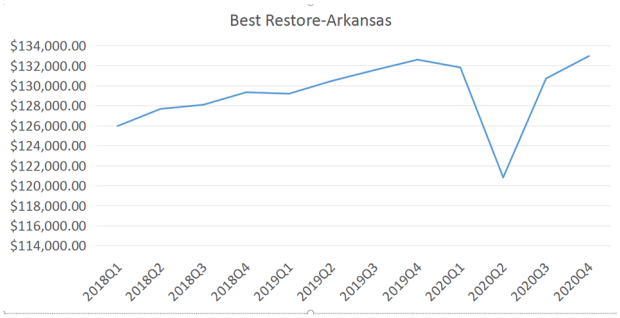

Figure 3 Sum of GDP by Industry in Arkansas

When we analyze the growth rate after the second quarter of 2020 and the decline rate in the fourth quarter of 2019, we can get the impact of the emergence of the COVID-19 on the overall industrial GDP. Figure 4 shows the increase and decrease rates of GDP by Industry in different states. These uneven lines show the different economic growth or decline of the fifty states in the United States.

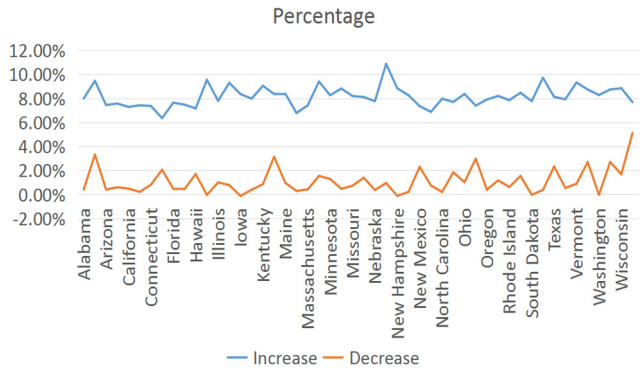

Figure 4 The increase and decrease rates of GDP by Industry

2.2 Total Personal Income

Apart from the research on macro GDP, personal income is also indispensable in economic life. I hope to make an effective mathematical analysis through the difference in personal income in each state. When I collected and sorted out the set of data of total personal income, all of the state's distribution was regular. The maximum (red high light) total income of all units is concentrated in the second quarter of 2020 and begins to fall after this node. The previous situation may be that as society progresses, people's income has been steadily increased. The subsequent decline may be due to the impact of the epidemic on the economy causing more people to be unemployed and forced to suspend work, so incomes will decline. The data collection examples show in Figure 5.

	2018Q1	2018Q2	2019Q3	2019Q4	2020Q1	2020Q2	2020Q3	2020Q4
Alabama	$206,139.70	$207,809.10	$217,250.70	$218,521.50	$221,671.80	$241,592.40	$231,273.30	$228,905.10
Alaska	$43,698.90	$44,029.60	$46,045.10	$46,307.90	$46,271.00	$49,548.90	$47,301.80	$46,336.90
Arizona	$313,429.00	$316,844.00	$336,776.00	$340,736.30	$345,326.70	$378,050.30	$368,116.60	$361,603.10
Arkansas	$128,233.00	$129,662.60	$135,222.70	$136,000.30	$137,387.90	$151,498.70	$141,230.10	$140,942.40
California	$2,463,176.70	$2,500,871.90	$2,633,945.50	$2,673,409.70	$2,703,254.40	$2,877,796.00	$2,861,899.30	$2,813,093.50
Texas	$1,434,778.60	$1,454,607.40	$1,538,237.00	$1,550,113.60	$1,560,918.10	$1,664,525.60	$1,615,459.60	$1,599,823.00
Utah	$144,934.90	$147,383.00	$157,503.20	$159,585.70	$161,545.60	$178,825.70	$169,470.80	$169,397.00
Vermont	$32,958.60	$33,324.20	$34,542.80	$34,689.70	$35,157.80	$38,942.20	$36,502.20	$35,635.80
Virginia	$485,360.30	$490,143.60	$510,425.20	$515,184.40	$521,913.40	$549,707.50	$537,891.30	$533,397.50
Washington	$458,398.10	$463,894.90	$495,061.90	$500,377.50	$506,274.70	$543,589.70	$530,579.90	$522,129.00
West Virginia	$72,796.60	$73,725.50	$75,866.10	$75,656.30	$76,478.50	$87,393.60	$79,846.00	$78,322.70
Wisconsin	$295,958.70	$297,589.70	$310,060.30	$312,743.40	$314,128.80	$338,487.10	$320,270.30	$321,652.00
Wyoming	$34,456.90	$34,865.10	$36,107.50	$36,188.60	$36,326.60	$38,305.40	$36,396.90	$36,331.10

Figure 5 Examples of total Personal Income from 2018Q1 to 2020Q4

To make the line chat between the sum values of all states and time period which shows in Figure 6, it is easy to prove that it increases before the 2020 second quarter, and has an extreme increase from the 2020 first quarter to the second one. And have a huge decrease from the second to the fourth. This means that the COVID-19 has had a huge impact on people's income, and the main factor for the sharp decline may be caused by break out the COVID-19.

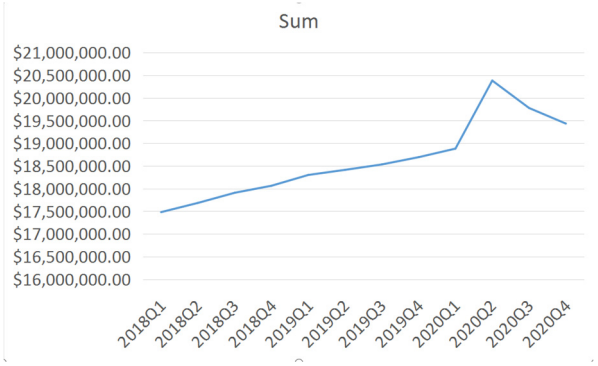

Figure 6 Sum of Total Personal Income of 50 states

When I analyze the increase and decrease before and after the maximum value, the increase of all the states is a unified trend, rising regularly as the times progress. The Figure shows in Figure 7. We can find that in this set of economic data, the differences are relatively large. This proves that compare and contrast personal income compared and GDP, there are obvious differences between states.

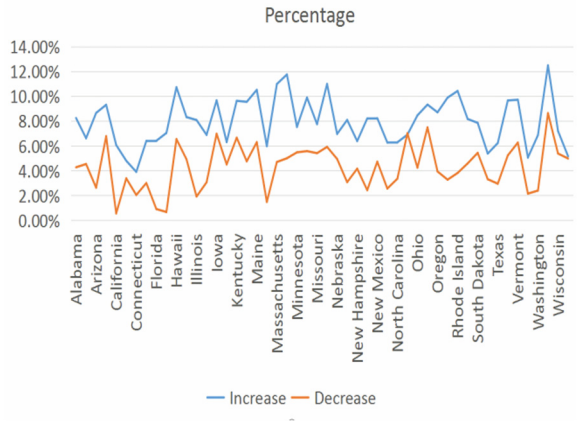

Figure 7 The increase and decrease rates of Total Personal Income

However, all states have suffered heavy losses after the second quarter of 2020, no one could be the exception. To evaluate all the losses by the constant criteria, the least affected is South Dakota which has the smallest drop.(Figure 8)

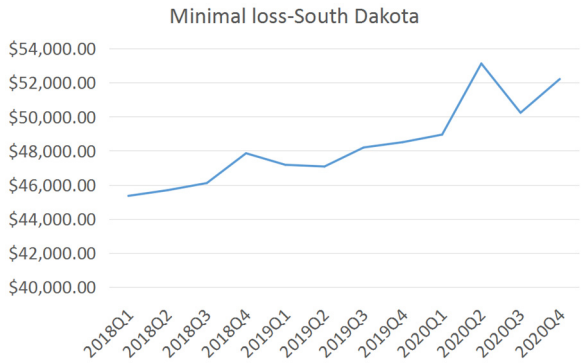

Figure 8 The sum of Total Personal Income in South Dakota

2.3 Finance and Insurance Earnings

Finance and Insurance Earnings have been the last parameters. There is all the data of Finance and Insurance Earnings of fifty States in U.S from the first quarter of 2018 to the last quarter of 2020. The minimum in most states was in the first quarter of 2018, while the maximum is concentrated in the last quarter of 2020. The data collection examples show in Figure 9.

However, for those states where the maximum does not appear in the last, they are affected by the COVID-19. They are: Alaska, Indiana, Kentucky, Louisiana, Massachusetts, Mississippi, Montana, New Hampshire, North Dakota, Rhode Island, Wisconsin and Wyoming.

	2018Q1	2018Q2	2018Q3	2018Q4	2019Q3	2020Q2	2020Q3	2020Q4
Alabama	$7,212,548.00	$7,302,092.00	$7,284,044.00	$7,169,716.00	$7,319,560.00	$7,680,482.00	$7,914,982.00	$8,131,303.00
Alaska	$675,192.00	$697,164.00	$681,184.00	$667,176.00	$689,772.00	$724,540.00	$683,674.00	$697,886.00
Arizona	$16,913,568.00	$17,310,108.00	$17,320,772.00	$17,140,660.00	$18,693,844.00	$20,577,235.00	$20,122,609.00	$20,698,606.00
Arkansas	$3,034,764.00	$3,051,804.00	$3,073,088.00	$2,978,892.00	$3,167,096.00	$3,399,249.00	$3,519,565.00	$3,579,279.00
California	$93,012,100.00	$94,321,192.00	$94,670,784.00	$93,362,712.00	$96,445,736.00	$101,574,168.00	$103,163,552.00	$105,978,791.00
Texas	$69,290,928.00	$69,189,408.00	$69,955,832.00	$69,281,712.00	$73,400,072.00	$77,491,827.00	$77,115,814.00	$80,017,290.00
Utah	$6,567,536.00	$6,648,680.00	$6,626,788.00	$6,390,776.00	$6,912,720.00	$8,113,135.00	$7,916,646.00	$8,276,746.00
Vermont	$955,796.00	$959,116.00	$984,612.00	$961,000.00	$984,396.00	$1,079,791.00	$1,019,184.00	$1,051,998.00
Virginia	$18,249,932.00	$18,060,172.00	$18,789,180.00	$18,195,292.00	$18,698,108.00	$20,176,822.00	$20,230,323.00	$20,801,179.00
Washington	$12,277,824.00	$12,414,236.00	$12,506,164.00	$12,228,544.00	$13,120,188.00	$14,356,772.00	$14,268,596.00	$14,662,054.00
West Virginia	$1,395,520.00	$1,477,196.00	$1,455,200.00	$1,445,996.00	$1,456,004.00	$1,503,995.00	$1,505,145.00	$1,542,521.00
Wisconsin	$12,897,000.00	$13,010,964.00	$13,187,360.00	$13,003,532.00	$13,362,332.00	$14,118,565.00	$13,743,392.00	$14,049,130.00
Wyoming	$551,524.00	$553,392.00	$548,804.00	$535,820.00	$618,676.00	$601,956.00	$598,774.00	$607,367.00

Figure 9 Examples of finance and Insurance Earnings from 2018Q1 to 2020Q4

When we integrate the total value of the fifty states and make a line chart with time (Figure 10), we can clearly see that the overall trend is rising, which means that the Finance and Insurance Earnings has not had a negative impact due to the COVID-19, it has still shown a steady upward trend.

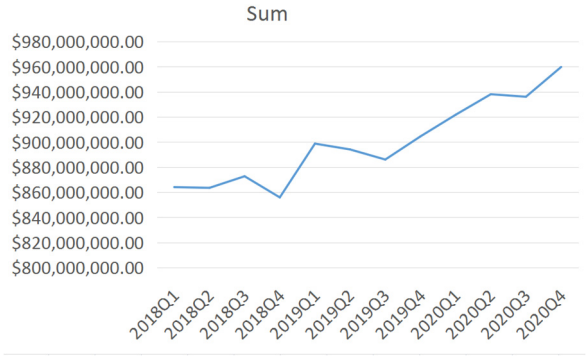

Figure 10 Sum of Finance and Insurance Earnings of 50 states

When we analyze the growth rate (Figure 11) when approaching the maximum, we can get the following figure. Among them, Louisiana (Figure 12) has the largest negative growth rate, which means that this COVID-19 may have caused an unprecedented to the Finance and Insurance Earnings in this state.

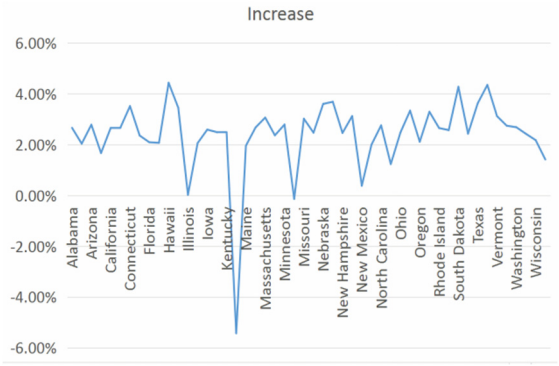

Figure 11 The increase rate of Finance and Insurance Earnings

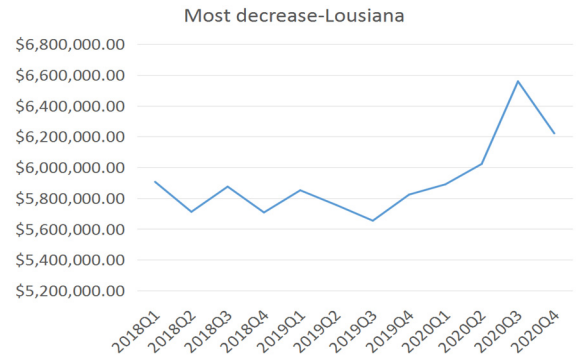

Figure 12 The sum of Finance and Insurance Earnings in Louisiana

In order to better think about the different economic impacts of the COVID-19 on 50 states, I decided to integrate three parameters to cluster 50 samples, making it more convenient to make complex data simpler and more intuitive.

3 RESULTS & DISCUSSION

3.1 Cluster

K-mean is the best way to analysis this model. In order to eliminate the difference of different parameters, I chose the rate of increase as a reference standard. The data collection shows in Figure 13. So the first step is to set up the state and the increase rate of GDP of industry, total income, and finance and Insurance Earnings. Then according to the three parameters, the fifty states are divided into three groups.

State	Sort1	Sort2
Florida	1	weak
Kansas	1	weak
Kentucky	3	mid
Louisiana	2	high
West Virginia	3	mid
Wyoming	1	weak

Figure 13 The example data of increase rate

Then we use these three parameters to clustering the state in the three groups, there are high, middle, and weak. The calculation of the clustering center told us that number 2 means high economic level and number 3 means middle economic level and number 1 means the weak economic level for the GDP of industry and total income. However, the finance and insurance earning is the opposite situation.

At the end of the final report, there is a total of 24 weak states and 25 middle states and only one high state. Only Louisiana has enough economic base to distinguish it from other states, so there is only one item in the high group. And the high economy is hard to achieve. That means the high economic lower depend on the finance and insurance earnings. The result shows in Figure 14.

Through classification analysis, we can find that weak states have obvious performance in both the decline in GDP and the decline in personal income. When making adjustments to the economies of their group of states, they can learn from each other, because combining the three data indicators, their situation is similar.

clustering	1	24.000
	2	1.000
	3	25.000
valid		50.000
Missing		1.000

Figure 14 The result of number of clustering

After analyzing the known data of 50 states, reasonable predictions for the future can help us better observe the impact of the COVID-19 on the economy, so the following part will be based on the current data and the first quarter of 2021, which has no accurate basis. The situation is predicted.

3.2 Prediction

Based on the currently known data, I tried to use mathematical methods to make a reasonable forecast for the first quarter of 2021. Prediction can help us to get a more intuitive view of the future situation, in order to adjust our policies in life. In order to get the macro results more intuitively, we use the sum of all states in each quarter as a reference. There are the following steps:

(1) Calculate the sum of the data for each time period to obtain 12 known values.

(2) Select one smooth factor for the pretest.

(3) In this case, the predicted value of 2018Q1 is the same as the actual value. The forecast formula for 2018Q2 is: smooth factor× ∑ 2018Q1+(1- smooth factor)=predicted value

(4) Calculate the mean square error(MSE) the formula is: MSE= (average actual value- average pretest value) ^{2}and use the second formula to test at the same time, their data results are the same.

$$MSE=\frac{\sum(\text{actual value}-\text{pretest value})^2}{number\ of\ month}$$

(5) Calculate the MSE change of the smooth factor from 0.1 to 0.6, and compare it to get the minimum MSE value.

(6) Obtain the optimal smoothing factor through the smallest MSE value.

(7) Perform accurate prediction again, using the optimal smooth factor ×∑ 2018Q1+(1- optimal smooth factor) = optimal prediction value.

(8) The results show in Figure 15, 16 and 17. Only the personal income goes increase in 2021Q1, the other two parameters are all decrease. The reason may be that government subsidies have increased people's income, but the economic situation in the general environment has not yet eased.

Month	Sum of GDP of industry	Test prediction	Optimal prediction			
2018Q1	$19,544,133.00	$19,544,133.00	$19,544,133.00			
2018Q2	$19,845,433.40	$19,544,133.00	$19,544,133.00		smooth factor	0.2
2018Q3	$20,026,604.00	$19,604,393.08	$19,544,133.00		MSE	4.85144E+11
2018Q4	$20,185,628.00	$19,688,835.26	$19,568,237.03		MSE min	4.59779E+11
2019Q1	$20,385,849.70	$19,788,193.81	$19,616,476.32		Optimal smooth factor	0.4
2019Q2	$20,594,220.60	$19,907,724.99	$19,685,163.32		2021Q1 predicted value	$20,196,250.26
2019Q3	$20,794,572.70	$20,045,024.11	$19,774,187.99			
2019Q4	$20,993,762.10	$20,194,933.83	$19,882,522.44		MSE test	
2020Q1	$20,808,433.50	$20,354,699.48	$20,007,486.99		smooth factor	4.8514E+11
2020Q2	$18,812,711.70	$20,445,446.29	$20,146,371.99		0.1	5.7418E+11
2020Q3	$20,423,825.00	$20,118,899.37	$20,266,001.71		0.2	4.8514E+11
2020Q4	$20,739,162.40	$20,179,884.50	$20,207,160.77		0.3	4.6094E+11
2021Q1		$20,291,740.08	$20,196,250.26		0.4	4.5978E+11
					0.5	4.6913E+11
					0.6	4.8442E+11

Figure 15 Prediction of GDP of industry for 2021Q1

Month	Sum of Total personal income	Test prediction	Optimal prediction			
2018Q1	$17,479,532.30	$17,479,532.30	$17,479,532.30			
2018Q2	$17,682,304.00	$17,479,532.30	$17,479,532.30		smooth factor	0.2
2018Q3	$17,906,077.70	$17,520,086.64	$17,601,195.32		MSE	6.64799E+11
2018Q4	$18,061,122.60	$17,597,284.85	$17,784,124.75		MSE min	2.75856E+11
2019Q1	$18,299,057.90	$17,690,052.40	$17,950,323.46		Optimal smooth factor	0.6
2019Q2	$18,411,875.70	$17,811,853.50	$18,159,564.12		2021Q1 predicted value	$19,563,520.52
2019Q3	$18,531,211.10	$17,931,857.94	$18,310,951.07			
2019Q4	$18,691,447.20	$18,051,728.57	$18,443,107.09		MSE test	
2020Q1	$18,880,185.10	$18,179,672.30	$18,592,111.16		smooth factor	6.64799E+11
2020Q2	$20,381,187.90	$18,319,774.86	$18,764,955.52		0.1	1.08171E+12
2020Q3	$19,776,601.30	$18,732,057.47	$19,734,694.95		0.2	6.64799E+11
2020Q4	$19,432,641.70	$18,940,966.23	$19,759,838.76		0.3	4.64117E+11
2021Q1		$19,039,301.33	$19,563,520.52		0.4	3.61296E+11
					0.5	3.06148E+11
					0.6	2.75856E+11

Figure 16 Prediction of sum of total personal income for 2021Q1

Month	sum of Finance and Insurance Earnings	Test prediction	Optimal prediction			
2018Q1	$864,154,684.00	$864,154,684.00	$864,154,684.00			
2018Q2	$863,566,144.00	$864,154,684.00	$864,154,684.00		smooth factor	0.2
2018Q3	$872,794,872.00	$864,036,976.00	$863,801,560.00		MSE	8.79896E+14
2018Q4	$855,961,852.00	$865,788,555.20	$869,197,547.20		MSE min	3.43446E+14
2019Q1	$898,786,316.00	$863,823,214.56	$861,256,130.08		Optimal smooth factor	0.6
2019Q2	$894,143,472.00	$870,815,834.85	$883,774,241.63		2021Q1 predicted value	$948,983,678.98
2019Q3	$886,101,440.00	$875,481,362.28	$889,995,779.85			
2019Q4	$904,538,140.00	$877,605,377.82	$887,659,175.94		MSE test	
2020Q1	$921,750,190.00	$882,991,930.26	$897,786,554.38		smooth factor	8.79896E+14
2020Q2	$938,110,321.00	$890,743,582.21	$912,164,735.75		0.1	1.36028E+15
2020Q3	$936,110,422.00	$900,216,929.97	$927,732,086.90		0.2	8.79896E+14
2020Q4	$959,800,073.00	$907,395,628.37	$932,759,087.96		0.3	6.20412E+14
2021Q1		$917,876,517.30	$948,983,678.98		0.4	4.74861E+14
					0.5	3.91272E+14
					0.6	3.43446E+14

Figure 17 Prediction of sum of Finance and Insurance Earnings for 2021Q1

4 CONCLUSION

After sorting out the three kinds of data of 50 states in the United States, we can observe the similarities and differences in all of them. The analysis of these special points gives us the capability to detect the impact of the COVID-19 on the social economy and to verify some of my hypotheses. My hypothesis is basically correct, and the second quarter of 2020 is the most critical point.

The GDP by industry decreases during the COVID-19 time. It reaches the lowest in the 2020 second quarter and increases after it.

The total personal income decreases during the COVID-19 time and begins to decrease in the second quarter of 2020. When the GDP reaches the lowest, then the personal income decrease.

The Finance and Insurance Earnings may not be impacted by the COVID-19, or the COVID-19 encourages the increase of it. The reason may be caused by the online style of these businesses.

Acknowledgment. Yuechen Mu thanks all the data from Federal Reserve Economic Data.

REFERENCE

[1] Timothy W Russell, Joseph T Wu, Sam Clifford, W John Edmunds, Adam J Kucharski, Mark Jit, "Effect of internationally imported cases on internal spread of COVID-19: a mathematical modelling study" *in the Lancet public health*, Volume 6, Issue 1, January 2021, Pages e12-e20

[2] Ashleigh R. Tuite, David N. Fisman, Amy L. Greer, "Mathematical modelling of COVID-19 transmission and mitigation strategies in the population of Ontario, Canada," *in CMAJ,* May 11, 2020

[3] Richard Baldwin, Beatrice Weder di Mauro, "Economics in the time of COVID-19: A neweBook," *in VOX, CEPR Policy Portal,* 06 March 2020

[4] Abel Brodeur, David M. Gray, Anik Islam, Suraiya Bhuiyan, "A Literature Review of the Economics of Covid-19," *in IZA Discussion Paper* No. 13411, 29 Jun 2020

[5] RakeshPadhana, K. P. Prabheeshb, "The economics of COVID-19 pandemic: A survey," *in Economic Analysis and Policy* Volume 70, June 2021, Pages 220-237

[6] Daniel Susskind, David Vines, "The economics of the COVID-19 pandemic: an assessment," *in Oxford Review of Economic Policy*, Volume 36, Issue Supplement_1, 2020, 29 August 2020, Pages S1–S13