A Brief Study of China’s Investor Sentiment Index

Sophia Gan

Tandon School of Engineering

New York University

[szg218@nyu.edu](mailto:szg218@nyu.edu)

September 29, 2020

**Abstract**

Investor's sentiment index is a set of numbers that directly shows how people view the current economic situation, whether they are interested in recent investments or not. By using the data of China's Initial Public Offering numbers (IPON), Closed-end fund (DCEF), new initial account number (NIA), turnover rate (TURN), and consumer confidence index (CCI) from 2003 February to 2020 January, this paper estimates a model representing the trend of consumers' sentiment index. Using principal components analysis (PCA) through python reduces the factors that will influence the final result, leading to a more pure development that shows the direct relationship between the time and consumer sentiment index. The findings through this paper will show how China's consumer sentiment index changed along as China develops.

**Introduction**

In recent years, China's economy has improved a lot after the recession that happened in 2007. Economy condition changed a lot throughout the years, and with the improvements in the economy in total, consumers as individuals also gain more confidence as the economy improves. To look at how people reacted to the economic environment around them, we have to look at both the economy before the recession and afterward. By comparing the different index numbers with their corresponding periods, we can get the corresponding CCI values, which can conclude the trend in the past few years.

There is a lot of research on how other countries' economy and their consumer sentiment index relation changed as time passed. However, little investigations have been done on China's economic situation. Information isn't that transparent in China, making individuals have more difficult access to the data related to the country's economy, mainly when focusing on the consumer's sentiment index. Before doing this research, I didn't know the different types of index numbers in stocks and economies, and even the computer-based learning in python wasn't that easy for me to apply. It was hard when I first started to contact these pieces of knowledge—the basic concepts of each index number and the different usage of the various packages in python.

The first part of this research is by gathering data from different databases, copy them into an Excel worksheet, and generalize them into workable data. The second part will be implementing the data by using python through the principal component analysis and illustrating the trend graph of China's economy throughout 2003-2020.

**Methodology**

By including Initial Public Offering numbers (IPON), Initial Public Offering rate (IPOR), Closed-end fund (DCEF), new initial account number (NA), turnover rate (TURN), and consumer confidence index (CCI) from 2003 February to 2020 January, one way of calculating the sentiment index is by calculating the ISI number by using the equation:

Another way of calculating the sentiment index is by calculating the CICSI number by using the following equation:

Before calculating all the data and the different periods, we shift the data so that each index number matches each other. For example, if the IPOR number is for last month, we need to change the data one period after so that it will be the same period as the other data. Same for the opposite scenario. If the index number is for this month, and the others are all index numbers from last month, we shift the data members one period early. This process can be done in PyCharm by the shift method with a period of 1 or -1 according to the different scenarios using the code shift(period=1 or -1).

By comparing the two data sets from ISI and CICSI, looking at the explained variance, the group with all positive explained variance will be the data set that best fits this analysis. The principal component analysis is a method that helps summarize the information throughout a massive set of data, especially when looking at how the data trend looks like throughout a time, which in this case, is what we are observing in this paper.

By inputting the data in the PCA package in PyCharm and using the matplotlib package to generate the graph of the trend, we can directly observe how the economy changed as time passed by looking at the graph python generated.

**Discussion and Result**

To provide a basic understanding of how the data worked in this research, table 1 and 2 are the data set after shifting the data to its corresponding periods. Each table has the data set that corresponds to last month's data or the current month's data set. There's no combined period in the two sets of data.

Table 1. Data from last month

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | mean | std | 25% | 50% | 75% |
| DCEF\_t0 | -0.008577 | 1.000403 | -0.8148 | 0.301153 | 0.845851 |
| RIPO\_t0 | -0.003933 | 1.004067 | -0.619276 | -0.188707 | 0.177981 |
| IPON\_t0 | 0.007208 | 1.002203 | -0.815894 | -0.251345 | 0.655964 |
| NA\_t0 | 0.007699 | 1.001856 | -0.593711 | -0.177025 | 0.138873 |
| TURN\_t0 | 0.004939 | 1.001834 | -0.667455 | -0.319058 | 0.400773 |
| CCI\_t0 | -0.00868 | 0.990398 | -0.818699 | -0.057954 | 0.395956 |

Table 2. Data from the current month.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | mean | std | 25% | 50% | 75% |
| DCEF\_t1 | -0.011045 | 0.998528 | -0.8148 | 0.301153 | 0.778351 |
| RIPO\_t1 | -0.009841 | 0.998673 | -0.619276 | -0.188707 | 0.177981 |
| IPON\_t1 | -0.000303 | 1.002603 | -0.815894 | -0.251345 | 0.635802 |
| NA\_t1 | 0.004762 | 1.003341 | -0.59593 | -0.177025 | 0.138873 |
| TURN\_t1 | 0.001388 | 1.004516 | -0.671043 | -0.319058 | 0.400773 |
| CCI\_t1 | -0.023149 | 0.976925 | -0.818699 | -0.078241 | 0.370598 |

Table 3 is the explained variance ratio that we got from the data above through PCA, where none of the values were below zero.

Table 3. Explained Variance Ratio

|  |  |  |  |
| --- | --- | --- | --- |
| 99.01598871 | 0.44152431 | 0.13228094 | 0.11693845 |
| 99.01598871 | 99.45751302 | 99.58979396 | 99.70673241 |

Graph 1 reveals the trend of how China's economy was related to consumers' sentiment index. It's clearly shown that before 2008, the economy in China was developing at a rapid speed. However, there was a massive drop in the consumer sentiment index throughout 2008, when the recession hit China. It took some time for China to recover from that colossal breakdown. By looking at the graph, the economy nowadays has been improving in a more fluent tendency. Chart, histogram

Description automatically generated

**Conclusion**

This study reflects how China's economy was shown through the consumers' sentiment index. As consumers have higher confidence in purchasing power, China's economy was more powerful and more developed along with high sentiment index. The data are from 2003 to 2020, which shows a general trend of how China's economy recover after the recession and even before the recession. The empirical results demonstrate the strong influence the recession had on consumers' sentiment index in China since there's a considerable drop in Graph 1.

This research also has limitations that could be improved through further studies later on. If possible, having access to earlier monthly data on China's economy can help us have an even more detailed view of the trend before 2003. Second, if the data can be from one organization or one website with the same method of calculating it, the error between each data will be smaller.

**References**

魏星集, 夏维力, 孙彤彤. 基于BW模型的A股市场投资者情绪测度研究[J]. 管理观察, 2014(33):71-73.

易志高, 茅宁. 中国股市投资者情绪测量研究:CICSI的构建[J]. 金融研究, 2009(11):174-184.