

Research on the Relationship between Investor Sentiment and Stock Price

Yi-ming DU^{1,2,3,a}, Rui DING^{1,2,3,b,*}, Yi-lin ZHANG^{1,2,3,c}
Ting ZHANG^{1,2,3,d} and Tao ZHOU^{1,2,3,e}

¹College of Big Data Application and Economics (Guiyang College of Big Data Finance),
Guizhou University of Finance and Economics, Guiyang, China

²Key Laboratory of Green Fintech, Guizhou University of Finance and Economics, Guiyang, China

³Guizhou Key Laboratory of Big Data Statistical Analysis, Guizhou University of Finance and Economics, Guiyang, China

^a1558851333@qq.com, ^b396124877@qq.com, ^cevalinzyl@163.com,
^d269493818@qq.com, ^e1013373971@qq.com

*Corresponding author

Keywords: Investor Sentiment, Principal Component Analysis, VAR Model.

Abstract. As one of the main contents of behavioral finance, investor sentiment has become a research hotspot in recent years. This paper takes the CSI300 index of China as the observation object, selects five emotional monthly time series data including lag one period from 2016 to 2020. The method of principal component analysis will be used to reduce the dimension of 10 groups of data. After eliminating the macroeconomic factors, the dimension reduction results are analyzed by the second principal component analysis to obtain the comprehensive index of emotion. Furthermore, a Vector Auto Regressive model (VAR) is established to investigate the relationship between ISIO and CSI300 of the stock market. The results show that investor sentiment and stock price interact with each other, but only in the short term. With more and more sufficient market information known, the effect is becoming insignificant.

1. Introduction

Traditional finance assumes that investors in the market are “rational people”, who can always keep rational and aim at maximizing utility (Adam Smith). However, due to the limitation of resource supply, the asymmetry of market information, the lack of human knowledge, and the insufficiency of market mechanism, the hypothesis of “rational man” does not exist in reality. The emergence of financial market anomalies has also questioned the traditional financial theory. Behavioral finance explains, studies and forecasts the development of financial market from the micro individual behavior and psychology. Compared with traditional finance, behavioral finance is closer to the real financial market. The role of investor sentiment in asset pricing and investment decision-making in behavioral finance has become a hot research topic in recent years.

Although the definition of investor sentiment are diversities due to the short development of behavioral finance, a bulk of researchers conceive that investor sentiment has noteworthy influence on capital market. Ma Yong, Yang Wenwei and Jiang Yiqing (2020)[1] bear out that the factors of investment sentiment are obviously driving the fluctuation of the price of China’s stock market by the Asset Pricing Model. Li Yan and Jin Dehuan (2018)[2] exert the method of Fama-Mac-Beth to testify that there is a positive correlation between investor sentiment and stock returns. However, there is existing a contrary voice. For example, Baker, Stein and Wurgler (2003)[3] utilize the Q ratio to discover a negative correlation between investor sentiment and share price moving.

Effective measurement of investor sentiment is the first step to study investor sentiment. There are numerous methods to measure investor sentiment, but they can be summarized into three categories. The first category is objective indicators. Closed-end fund discount (Han Liyan and Wu Yanran, 2007[4]), The volume of IPO and Initial returns (Ljungqvist and William, 2003[5]),

Turnover (Baker and Stein, 2004[6]), etc. The second category is subjective indicators. Consumer confidence index (Salhin et al.,2016[7]), Haodan index (Cheng Kun and Liu Renhe, 2005[8]) and CCTV Kanpan index (Wang Meijin and sun Jianjun, 2004[9]). And the third category is composite index. Principal component analysis (Baker and Wurgler, 2006[10]) was used to combine objective indicators with subjective indicators.

The factors of stock market trading system and market investor's structure will boost the sharp fluctuation of stock price (Wang Chaoyang and Wang zhenxia, 2017[11]). However, what makes investors too high expectations and leads to huge bubbles in stock prices in a short time? In order to study the relationship between investor sentiment and stock market, this paper constructs a comprehensive index of sentiment by using the method of principal component analysis. Then, the VAR model is applied to study the specific relationship between emotion agent variable and stock price.

2. Empirical Research

2.1. Data Sources

Based on the basic situation of China's stock market and the availability of research data, this paper selects five indicators as the basis of constructing investor sentiment index. Then the method of principal component analysis is used to reduce the dimensions of these indicators, so as to construct the proxy index of investor sentiment. Next, we apply the VAR model to further study the relationship between proxy index of investor sentiment and the closing price of CSI300. The data frequency is monthly, the time span is from January 2016 to December 2020, the data source is from Dongfang fortune net and Resset database, and the data processing software is SPSS and Eviews.

2.2. Selection of Indicators

In order to construct the composite emotion index, this paper will select five indicators. The first is market turnover rate (TURN), which refers to the frequency of investors trading stocks in a certain period of time. The higher the turnover rate is, the more active the investment market is. The second is volume (VOL) that means the activity of stock trading in a unit time, which can directly reflect the activity of investors in trading. The third is the number of newly opened accounts (NA) representing the number of new stock accounts opened in the current month. The more accounts opened, the more popular the stock market will be. The fourth is the consumer confidence index (CCI) being released by the National Bureau of statistics. The larger the value is, the consumers are more confident about economic growth. The fifth is price-earnings ratio (MPE) is one of the important indexes to measure whether the stock price level is reasonable.

However, this paper takes the irrational part of investor sentiment as the research object, thus it should eliminate the influence of macroeconomic factors which affect the investor sentiment. This paper selects the Consumer price index (CPI), Producer price index (PPI) and Macroeconomic boom index (MBCI) as the proxy indicators of macroeconomic factors. The level of CPI enacts an indispensable role in the formulation of national economic policies and the regulation of prices, which will bear a significant impact on the investors' expectations of the market and the changes of the capital market. PPI reflects the changes in the factory prices of industrial products. The public could infer the future policies according to the PPI index, which give rise to investor sentiment fluctuation. MBCI refer to the attitude of entrepreneurs towards the future economic trend. Generally, 100 is taken as the benchmark. When the value is greater than 100, entrepreneurs are generally optimistic about the future trend. While the value is less than 100, entrepreneurs are pessimistic.

2.3. Construction of Investor Sentiment Index

2.3.1. Unit Standardization

Considering the influence of time factor on data variables, thus the data lagging one period is selected to be standardized together with the original data. The processing formula is as follows:

$$X_i = \frac{X - E(x)}{\sqrt{D(X)}} \quad (1)$$

Here, X is the original variable, E (x) is the mean value of the variable, and D (x) is the variance of the variable.

2.3.2. The Construction of ISIO

The first step is to use KMO and Bartlett spherical test to judge the validity of factor analysis.

The second step, the standardized ten variables were analyzed by the first principal component analysis to get the emotion comprehensive index ISIO1. Then the correlation between ISIO1 and these ten variables carries on Pearson test, and the ones have higher correlation with ISIO1 were selected, and finally five variables were selected.

The third step is to establish a linear regression model based on selected five variables and the proxy indicators of macroeconomic factors.

The fourth step, the residual term in the regression model is the price earnings ratio after excluding macroeconomic factors. The five residuals are standardized into new variables, and then the five new variables were used for the second principal component analysis to construct the final emotional composite index ISIO.

The fifth step is calculated the final emotion comprehensive index ISIO. The analysis of the explained total variance results, the first and second eigenvalues are greater than 1, and the cumulative interpretability of the first and second principal components is 77.473%, which already contains most of the original information. Therefore, the first and second components are extracted. According to the results of main components, the comprehensive index of investor sentiment can be calculated by weighting, including volume of trading, number of new accounts opened, consumer confidence index, price earnings ratio and turnover rate. The two main components are Z1 and Z2, and the results are as follows:

$$Z1=0.31*MPEa+0.20*NIAa+0.57*TURN1a+0.57*VOL1a-0.46*CCI1a \quad (2)$$

$$Z2=0.54*MPEa+0.76*NIAa -0.17*TURN1a-0.17*VOL1a+0.27*CCI1a \quad (3)$$

Then, we use the variance in the rotation square loading as the weight, W1 = 0.54274, W2 = 0.77473, and get the composite emotion index ISIO:

$$ISIO=0.54274*Z1+0.77473*Z2 \quad (4)$$

Table 1. Total Variance Explained.

Component	Initial Eigenvalues			Extraction Sums of Squared Loading		
	Total	% of Variance	Cumulative%	Total	% of Variance	Cumulative%
1	2.714	54.274	54.274	2.714	54.274	54.274
2	1.160	23.199	77.473	1.160	23.199	77.473
3	.701	14.021	91.495			
4	.425	8.505	100.000			
5	1.579E-5	.000	100.000			

Extraction Method: Principal Component Analysis.

2.4. The Relationship between Investor Sentiment and CSI 300 Index

2.4.1. The Drawing of CSI300 Closing Price and Sentiment Composite Index

The standardized CSI 300 composite index and investor sentiment composite index ISIO are made into a bar chart to intuitively feel the trend of the two indexes and judge their effectiveness. From chart 1, it can be seen that from January 2018 to December 2018, the market has experienced a round of bear market due to the continuous upgrading of Sino-US trade war, the continuous

devaluation of RMB and the shortage of capital in the stock market. From March 2020 to December 2020, under the crisis of the COVID-19, rapidly recovery of China’s economy and implemented policy of monetary easing contribute to abundant funds flowing into the stock market, causing that stock price rise greatly. Chinese investors are mainly retail investors, thus they can hardly be rational when facing the market fluctuation. As a result, the fluctuation of the investor sentiment composite index is stronger than that of the CSI300 index. But in terms of its entirety and basic tendency, the images tend to be consistent.

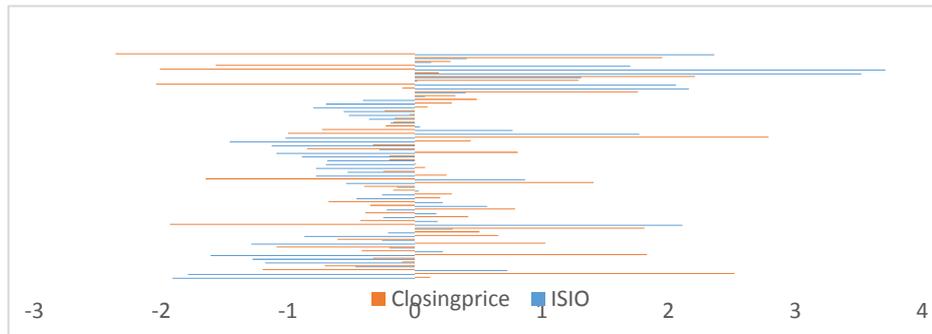


Figure 1. Bar Chart.

2.4.2. ADF Test of Time Series

In order to avoid the heteroscedasticity phenomenon of time series, the closing price of CSI300 has to be processed with logarithm and is recorded as R. The results from the table show that both investor sentiment (ISIO) and R are stable sequences so that can be used for regression.

Table 2. ADF Test.

Variable	Auqmented Dickey-Fuller test statistic	1% level	5% level	10% level
R	-7.803318	-2.605442	-1.946549	-1.613181
ISIO	-3.793972	-2.604746	-1.946447	-1.613238

2.4.3. The Lag Order of VAR Model

In order to get the optimal lag order, SIC and AIC are used to judge the results, and the results are shown in Table 3. According to the “majority principle”, the model should select the third order as the optimal lag order.

Table 3. Results of optimal lag order of VAR model.

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-33.07793	NA	0.012275	1.275561	1.348555	1.303788
1	24.38449	108.6562	0.001757	-0.668527	-0.449545*	-0.583845*
2	28.07213	6.704797	0.001779	-0.657168	-0.292199	-0.516031
3	33.90288	10.17733*	0.001667*	-0.723741*	-0.212784	-0.526150
4	36.12613	3.718879	0.001784	-0.659132	-0.002187	-0.405086
5	36.46018	0.534481	0.002048	-0.525825	0.277109	-0.215324

2.4.4. Model Stability Test

Impulse response diagram can be used to better explain the relationship of ISIO and R. However, there is an assumption that VAR model is stable before using impulse response diagram, it is supposed to test the stability of the model. According to the AR test chart, it is observed that the six unit roots of the model are in the unit circle, so the VAR (3) is stable and can be used for impulse response analysis.

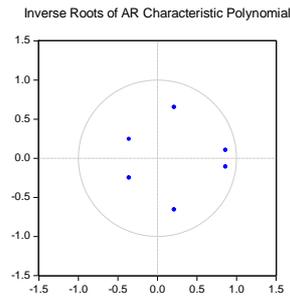


Figure 2. AR Test.

2.4.5. Granger Causality

It can be summarized from the results in Table 4 that at the 5% confidence level, the original hypothesis that R is not the Granger cause of ISIO is rejected, and R is the Granger cause of ISIO. The original hypothesis that ISIO is not the Granger cause of R is rejected, and ISIO is considered to be the Granger cause of R. It indicates that investor sentiment and R are inter-Granger causality and influence each other.

Table 4. Granger causality test.

Null Hypothesis	Obs	F-Statistic	Prob
R does not Granger Cause ISIO	57	3.75611	0.0165
ISIO dose not Granger Cause R		3.87235	0.0144

2.4.6. Impulse Response Analysis Based on VAR Model

After a positive shock on R, ISIO began to rise sharply in the first phase, it began to decrease to zero in the thirty-four phase when reaching the peak in the third period. It demonstrates that the investor sentiment index will rise when the market is positive. Nevertheless, the continuity of this high sentiment is not long. The reason is that individual stock investors account for a large proportion in China's stock market, thus there exists a train of blind chasing investment behaviors. The behavior just only last for a short time , then investor's emotion will present stable and even reverse as they master more and more market information. Giving ISIO a unit of standard deviation shock, the impact of R manifests negative in the first three periods and then it begins to ascend the highest point in the fifth period. Finally, it starts to descend to zero in the 26th period. In the early stage, the negative impact may be that the market is affected by some factors, for example, COVID-19, resulting in negative overall market returns. However, investors still hold optimistic expectations for the development of the market. With the continuous injection of funds, the stock price will rise. After a period of rising, the high mood will gradually become weak until it drops to zero.

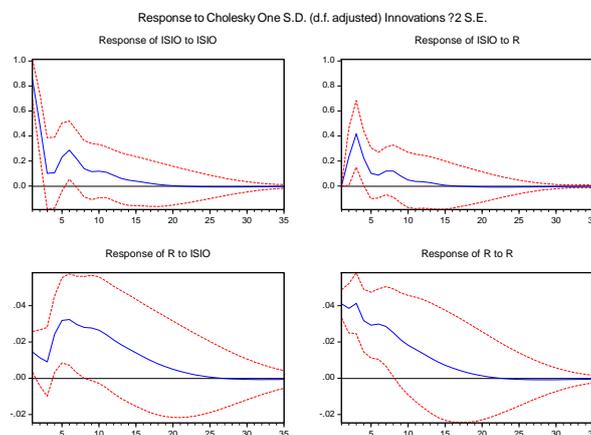


Figure 3. Impulse response diagram.

3. Conclusion

Considering the availability of data and the stock market situation in China, this paper selects 10 groups of data including one lag period based on the research methods of previous scholars. The final investors sentiment composite index will be constructed after eliminating the influence of macro factors. Then the relationship between investor sentiment and CSI300 stock market could be studied by establishing VAR model. As a result, there exist complex interactions between investor sentiment and stock price. High investor sentiment drives the rise of stock price, similarly, the uptrend of stock price will attract more investors. Compared with foreign stock markets, retail investors who are with lower level of education and lack the understanding of financial and economic knowledge are in the majority of the China's stock market. Therefore, some irrational investment behaviors prevail, such as follow the trend blindly and chase sell. However, with more and more sufficient information in the market, the interaction between investor sentiment and stock price will not significant in the long time.

Acknowledgement

This study is funded by National Natural Science Foundation of China (No. 72001053) and Science and Technology Planning Project of Guizhou Province of China (No. Qian ke he ji chu [2020]1Y283).

References

- [1] Ma Yong, Yang Wenwei and Jiang Yiqing, "How Does Investor Sentiment Affect Company Stock Price", *Financial Forum*. 2020: 57–67.
- [2] Li Yan and Jin Dehuan, "An empirical test of the relationship between investor sentiment and stock returns", *Statistics & Decision*. 2018: 166–169.
- [3] Baker M., Stein J.C. and Wurgler J., "When does the market matter? Stock prices and the investment of equity-dependent firms", *Quarterly journal of economics*. 2003: 969–1005.
- [4] Han Liyan and Wu Yanran, "Imperfect Rationality, Sentiment and Closed end fund Puzzle", *Economic Research Journal*. 2007: 117–129.
- [5] Ljungquist A. and Wilhelm W.J., "IPO pricing in the dot-com bubble", *Journal of Finance*. 2003: 723–752.
- [6] Baker M. and Stein J.C., "Market liquidity as a sentiment indicator", *Journal of Financial Markets*. 2004: 271–299.
- [7] Salhin A., Sherif M., and Jone E., "Managerial Sentiment, Consumer Confidence and Sector Returns", *International Review of Financial Analysis*. 2016: 24–38.
- [8] Chen kun and Liu Renhe, "Review on Research Documents about Fluctuation of Stock Market", *Social Sciences Journal of Universities in Shanxi*. 2005: 29–30+41.
- [9] Wang Meijin and Sun Jianjun, "Stock Market Returns, Volatility and the Role of Investor Sentiment in China", *Economic Research Journal*. 2004: 75–83.
- [10] Baker M. and Wurgler J., "Investor sentiment and the cross-section of stock returns", *Journal of Finance*. 2006: 1645–1680.
- [11] Wang Chaoyang and Wang Zhenxia, "Pice Limit, Magin Trading, and Stock Price Volatility: A Comparative Study between A-shares and H-shares", *Economic Research Journal*. 2017: 151–165.