

# Research on the Influence and Risk of Online Investor Sentiment on GEM Stock Price under the Background of Big Data

Yuming Zhang\*, Ziyang Wang

School of Big Data Application and Economics, Guizhou University of Finance and Economics, Huaxi University, GuiYang, China

**Keywords:** Investor sentiment; Fama-French five-factor model; Stock price change; GEM

**Abstract:** This article uses the web crawler technology to crawl the GEM 50 index constituent stock reviews in East Money Stocks bar, and evaluated the stock comments by the software of sentiment analysis to construct the emotional factors. We added the emotion factor into the Fama-French five factor model to form a six factor model in order to explore the impact of investor sentiment on the stock price of the GEM. Through this study, we find that the current investor sentiment will impact the stock price of GEM. When the investor sentiment is high, the next day stock price usually performs well, while when the sentiment is low, the next day stock price does not perform well. As GEM investors tend to pursue high-risk stocks to seek high returns, market value effect and book to market ratio effect are common in GEM market. In view of this kind of risk, this paper puts forward corresponding suggestions.

## 1. Introduction

Since the reform and opening up, China's economy has been developing rapidly and people's incomes have been rising. As the lubricant of the national economy, China's financial industry has been playing an important role in the national economy. China's stock market has been going for nearly 30 years since the Shanghai and Shenzhen stock exchanges were established in 1990. During these 30 years, China's capital market boomed. Gem, as a supplement to the main board market, has more growth stocks and plays an important role in the Chinese stock market. However, compared with foreign countries, the development of China's capital market is still relatively immature. The stock market investment environment, the market investment system, and effective supervision are still insufficient. The phenomenon of "policy market" and "information market" still exists. Gem market because of its own characteristics, the board of the listed company's stock price changes more than the main board, higher investment risk. The phenomenon of herd behavior and buying stocks when they rise and selling them when they are still exist.

Under the traditional financial theory system, the model is built on many assumptions, such as the market is efficient, investors are rational and so on. Investors will pursue higher returns on the premise of certain risks, or try to reduce risks on the premise of certain returns. When asset prices are in an unreasonable state, investors will obtain excess returns through risk-free arbitrage to maintain asset prices at a reasonable level. This resulted in the traditional landmark theories in the field of finance, such as Fama's efficient market hypothesis, Markowitz's Portfolio Theory, Sharp's Capital Asset Pricing Model, Ross's APT model, Black and Scholes' B-S Option Pricing Model and so on. However, with the continuous development of market and theory, many anomalies that cannot be explained by traditional financial theories begin to appear. Behavioral finance has gradually entered people's field of vision and tries to explain phenomena that cannot be explained by traditional finance. Behavioral finance combines finance, sociology, psychology, behavioral science and many other disciplines to theoretically analyze the behavior of financial market traders. Behavioral finance believes that the price of securities mainly depends on the behavior of trading participants, that is, the psychological activities of traders can have a significant impact on the price changes of securities.

From the perspective of exploring investor sentiment, this paper attempts to explore the influence of investor public opinion on stock price by combining the two theories of traditional

finance and behavioral finance. At the same time, according to the characteristics of China's stock market, especially small and medium-sized stocks are susceptible to the impact of public opinion, choose the GEM 50 index stocks as the research object. Compared with the main board stocks, the market value of GEM stocks is relatively small and there are more individual investors, so the phenomenon of "emotional trading" is more obvious. Therefore, taking GEM stocks as the research object can better highlight the impact of investors' public opinions on the stock price fluctuations in China's capital market.

## 2. Literature Review

The earliest research on Investor sentiment can be traced back to 1963, when Investor Wisdom published an article on Investor sentiment that compiled a bear-market sentiment index based on stock market data at the time. It can be seen that researches on public opinions have been conducted for quite a long time. Next, this paper will conduct literature review from three aspects: the concept of public opinions, the fitting of public opinion indicators, and the relationship between public opinions and stock prices.

The concept of investor sentiment was first put forward by foreign scholars, but no unified definition has been formed yet. De Long et al. (1990) explained investor sentiment by using the concept of behavioral finance for the first time, believing that investor sentiment is a psychological deviation caused by the current objective and realistic information failing to meet one's own subjective needs [1]. Brown & Chff (2005) believed that the stock value was derived from the quality and performance of the company. The higher the quality and performance, the higher the stock price should be, and investor sentiment was the principal value tendency of the company [2]. Baker & Wurgler (2006) defined investor sentiment from the perspective of stock yield. They believed that investor sentiment was the subjective belief of risk preference and expected return. Domestic scholars have also made relevant definitions of investor sentiment [3]. Sun Jianjun and Wang Meijin (2004) believed that investors would have perceptual cognition of stock price when conducting stock trading, and such irrational cognition would become a systematic factor affecting stock equilibrium price [4]. Huang Delong (2009) found in his research that investor sentiment is the comprehensive performance of investors' expectation target and expectation probability [5]. Gao Daliang (2015) believes that there are many irrational emotional investors in China's stock market, who have a more serious risk seeking tendency [6].

The characterization of emotional indicators is the basis of public opinion research, which has been studied by many scholars at home and abroad. There are already official investor sentiment indexes published by financial institutions abroad, such as the American association of individual investors index (AAII), which reflects investor sentiment and predicts the stock market in the next six months. The investor intelligence index (II index) reflects the difference in the ratio of bulls to bears in the stock market. There are many index in domestic such as CCTV watch index, consumer confidence index and so on. In addition to direct indicators, many scholars also explore the characterization of investor sentiment indicators by constructing introduction indicators and composite indicators. Through empirical analysis, foreign scholars Lee & Jiang (2002) found that discount rate of closed funds can effectively reflect investor sentiment [7]. Tetlock (2007) expanded the building factors of sentiment indicators, and explored the influence of news media reports on traders' behaviors through emotional analysis of the vocabulary of Wall Street stock columns [8]. Siganos (2017) measured the impact of positive and negative sentiment divergence and stock price volatility in 20 countries and regions by analyzing the status data of Facebook trading investors [9]. Yang Xiao (2016), a domestic scholar, found that the discount rate of closed funds, stock turnover rate, trading volume, falling index and other simple indicators can effectively measure the investor sentiment in China [10]. Lv Yongxia (2018) analyzed the information of stock columns through text mining technology, and constructed investor sentiment indicators [11].

In the end, this paper reviews the literature on the influence of investor sentiment on stock price volatility, and explores the feasibility and rationality of this paper. Sayi & Raham (2015), two foreign scholars, found that investor sentiment can effectively influence the rise and fall of stocks

by analyzing the Turkish stock market [12]. Lin (2018) explored the impact of investor sentiment on the futures market and found that the arbitrage risk and transaction cost of the futures market would increase while the price discovery function would be weakened [13]. Domestic scholars Li Helong, Feng Chune (2014) explored the relationship between index rate of return and investor sentiment. They divided the period into three parts: short-term, medium-term and long-term. The study found that the short-term return rate of the Shanghai composite index was positively correlated with investor sentiment, the medium-term investor sentiment was ahead of the change in the return rate of the Shanghai composite index, and the long-term investor sentiment was behind the change in the Shanghai composite index. However, the shenzhen stock exchange index changes in the same period of investor sentiment [14].

Through sorting out relevant literatures on investor sentiment, we find that investor sentiment has a certain influence on the prices of related products in the securities market. Especially in the immature Chinese capital market, investor sentiment is expressed through public opinion. Nowadays, with the rapid development of Internet technology, it has become an inseparable part of people's daily living standards and a platform for investors to express their views. Therefore, the study of investor sentiment on Internet platforms is still a part of adapting to today's social development. In this study, the Internet platform is used to select the best portal websites that can express public opinions for analysis, and the model more in line with the Chinese market is used to analyze the impact of investor sentiment on stock prices.

### **3. Concept Definition and Framework Research**

#### **3.1 Concept definition**

Since the word "investor sentiment" came into being, almost every relevant research paper has made a clear definition of it. The concept of investor sentiment is very broad, and general generalization can cover many areas. For example, investors' trading behavior in the face of stock price, or verbal expression of emotion, or the thoughts generated by seeing the trading price can be defined as investor sentiment. In the second part of literature review, we found that most of the definitions of investor sentiment in literature were based on existing indicators, such as stock turnover rate, stock trading volume, discount rate of closed funds and so on. These indicators tend to target the sentiment of the investors who are making the trades. In other words, the sentiment expressed by the traders who are actually buying and selling. Such sentiment indicators tend to represent investor behavior in completed trades, but ignore the effect of investor sentiment in non-traded investors on stock prices. With the rapid development of Internet technology, more and more people begin to express their opinions on the Internet platform, which of course includes their personal opinions on stocks. Stock bar, forum, microblog, WeChat public account and other websites and media provide investors with a good platform for the exchange of views. Everyone is not only the producer of information, but also the transmitter of information. Compared with the traditional view of official media, in the era of big data, information spreads faster and the amount of information increases geometrically, so the influence of information also expands dramatically. After receiving a large amount of information about the stock price, investors will have a new view on the current stock price, so they will have new trading behaviors, which will affect the stock price changes. The influence of investors on the stock price does not come from the pure trading behavior of investors themselves. They often express their personal views on the stock price on the Internet platform. This behavior has an impact on the trading behavior of other traders, which has an impact on stock prices. This is why "herd behavior" occurs frequently in the stock market. Therefore, personal views on stocks also belong to the concept of investor sentiment.

Therefore, the investor sentiment in this paper is defined as the investor opinion expressed by text emotion. In the era of big data Internet, investors are receiving massive information from various Internet portals all the time, and constantly influencing the original value judgment of investors. This information will cause certain fluctuations in the stock price. In this paper, investor sentiment is defined as investors' personal cognition of current stock price and expressed through

Internet platform. The significance of emphasizing the textual view is that such emotions must be expressed by investors in the form of text data on the Internet. However, what investors think in their heart or verbally negotiate with others are all forms of data that cannot be measured, and data classification cannot be achieved in today's technology, so it is not considered in this paper.

### 3.2 framework research

The research framework of this paper is shown in the following Figure 1.

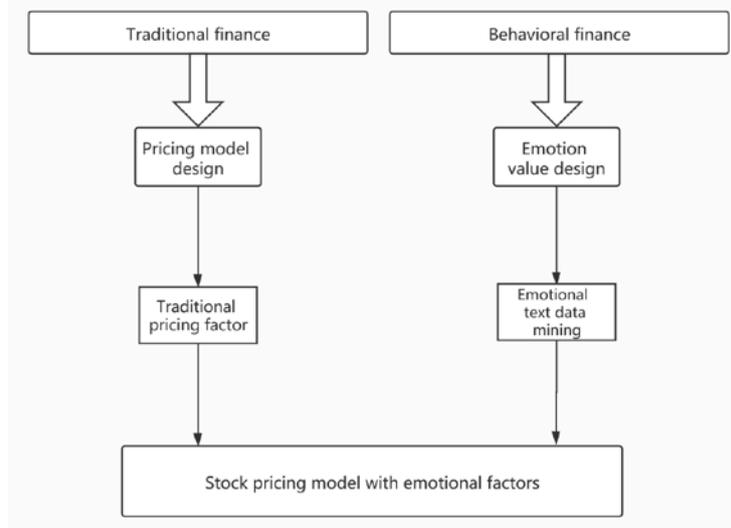


Figure 1 Article structure diagram

The research idea of this paper combines traditional finance and behavioral finance to explore the influencing factors of stock price. In traditional finance, this paper screens the factors that affect stock price and selects the optimal model to predict stock price. The most classical theory of stock price pricing in finance belongs to Capital Asset Pricing Model. This model introduces the value of unit security risk to explore its impact on the rate of return. Then Fama and French proposed a three-factor model as a supplement to CAPM model. This model believes that the influence on stock return mainly includes market excess return, market value factor and book-to-market ratio factor, and proves that this model is effective in predicting stock return through a large number of empirical studies. Then Fama proposed a five-factor model as a supplement to the three-factor model, which added profit factor and investment style factor to the original model. Compared with the three-factor model, the five-factor model has stronger predictive power for stock prices, and gradually replaces the three-factor model to become the mainstream research model for studying the influencing factors of stock prices. More and more scholars begin to use the five-factor model in research. For example, Ma Yueyi (2019) used the five-factor model to conduct empirical research on China's GEM stocks, and found that the scale effect, book-to-market ratio effect, profit effect and investment effect in the model were significant. Compared with the three-factor model, the five-factor model had better predictive ability for stock returns. From the perspective of behavioral finance, this paper collects internet-related stock comments through text mining and analyzes the emotional tone of the comments, showing the investor sentiment values by means of positive and negative and number size. In this way, unquantifiable emotions can be converted into specific values, and emotional factors can be fitted to the traditional pricing model, so as to add investor sentiment value into the pricing model and explore its impact on stock prices.

## 4. Model Setting, Variable Construction and Data Description

### 4.1 Model setting

After reviewing the previous literatures, this paper decided to adopt the Fama-French five-factor model. On the basis of the original model, this paper improves it by adding emotion value factor as the index to measure investor sentiment, and carries out regression analysis to explore its influence

on stock price. The traditional Fama-French five-factor model formula is as follows:

$$IR = \beta RP + aSML + bHML + cRMW + dCMA + \varepsilon \quad (1)$$

IR represents the excess rate of return in period t. RP is the market factor, reflecting the market risk premium. SML: market value factor, the difference of average return between small-market stocks and large-market stocks. HML: the difference between high book-to-market ratio and low book-to-market ratio average rate of return. In the book-to-market ratio B/M, B is the book value of the previous year, that is, the total owners' equity, and M is the total market value at the end of the previous year. RMW: profitability factor. The difference between the average return on a stock portfolio with a high operating margin and a low operating margin. The profitability indicator is measured by the total annual operating profit/equity of t-1; CMA: investment style factor. The difference between the return on a portfolio of stocks with a conservative investment style and an aggressive investment style. The investment index is expressed by the change rate of total assets, that is, the total assets in year t minus the total assets in year t-1 divided by the total assets in year t-1.  $\varepsilon$  is residuals.

After adding emotion value factors that can represent investor sentiment, the modified six-factor model is as follows:

$$IR = \beta RP + aSML + bHML + cRMW + dCMA + eET + \varepsilon \quad (2)$$

Et represents the sum of current investor sentiment values.

## 4.2 Variable construction

First, for the six-factor model after the optimization of the fama-french five-factor model, five factors in the original model of variables need to be constructed. They are market factor RP, market value factor SML, book-to-market ratio factor HML, profitability factor RMW, investment style factor CMA.

$$IR = R_t - R_f \quad (3)$$

$R_t$ : the return rate of the stock in period t;  $R_f$ : risk-free interest rate. In this paper, the three-month lump-sum deposit and withdrawal rate is adopted. This paper adopts the combination construction method in Fama (2015) to construct other factors. In this paper, stocks are divided into two groups according to market value: large market value group and small market value group. Then, in the two groups, stocks are sorted according to the book-to-market ratio. Portfolios are constructed according to 30% and 70% as sub-points. Then, in the same way, they were sorted according to profitability P and investment style IS, with R, N and W respectively representing the combination of aggressive, medium and weak profitability. C, N and A represent conservative, neutral and aggressive portfolios respectively. The factor calculation formula is as follows:

$$SMB = (SH + SN + SL)/3 - (BH + BN + BL)/3 \quad (4)$$

$$SMBP = (SR + SN + SW)/3 - (BR + BN + BW)/3 \quad (5)$$

$$SMBIS = (SC + SN + SA)/3 - (BC + BN + BA)/3 \quad (6)$$

$$SMB = (SMBB/M + SMBP + SMBIS)/3 \quad (7)$$

$$HML = (SH + BH)/2 - (SL + BL)/2 \quad (8)$$

$$RMW = (SR + BR)/2 - (SW + BW)/2 \quad (9)$$

$$CMA = (SC + BC)/2 - (SA + BA)/2 \quad (10)$$

For the construction of emotion factor Et, this paper adopts the method of text mining to crawl the relevant comments of the target stock on that day, and converts each comment about the stock into a value that can represent the opinion of the comment by means of emotion segmentation and emotion assignment. Bullish is positive, bearish is negative. The more intense the emotion, the

greater the absolute value. The sentiment values of all the comments of the day were added together to obtain the total Et of the day's opinions on the stock.

### 4.3 Data description

This study selected GEM stocks as the research object. The research period was from 2016.03.01 to 2019.02.28. The reason for the selection was to avoid the stock market crash in 2015 and the circuit breaker mechanism in early 2016, which caused abnormal stock price fluctuations. At that time, investor sentiment was generally depressed and there were many extreme remarks. Therefore, investor sentiment was not representative of GEM stock price volatility. After March 2016, GEM stock price gradually tends to normal volatility. It is more accurate to study the fluctuation of investor sentiment to stock price. The objects participating in the construction of the five factors are all GEM stocks, and the data source is CSMAR database. For the construction of sentiment value, stock comments are selected from East Money Stock bar. The East Money Stock bar is an old professional website for stock investors to exchange opinions. According to the data of Alex traffic statistics website, its monthly average visits can reach 6912,000, ranking among the best in China's stock bar. With the advantages of authoritative, comprehensive, professional and timely, East Money net has become the most visited and influential financial portal website in China and even the world, ranking first in a number of authoritative surveys and statistics. In terms of key indicators such as effective browsing time, core traffic value and average daily coverage number, East Money net is far ahead and has a very obvious industry advantage.

For the collection of online stock comments, the data volume is very large, and the comments of each stock can reach tens of thousands of pieces. It is not realistic to collect them by copying and pasting one by one. In this paper, the big data collection and processing method -- web crawler technology was used to crawl the stock comments of GEM 50 index stocks from March 2016 to March 2019, and the obtained data was cleaned, which not only greatly shortened the time needed to collect data by traditional methods, but also ensured the accuracy of data.

At present, there are a variety of emotion analysis software. Chinese document emotion analysis software mainly CNKI network emotion analysis tools, ROST emotion analysis software and so on. Although the emotional analysis tool of CNKI network has comprehensive functions, it lacks operability. So, this paper adopts ROST emotional analysis software for emotional analysis. The processing principle of ROST emotion analysis software is based on the weight dictionary of emotion words, negative words, degree words and stop words, etc. According to the Chinese grammar rules, the emotional tendency (or emotional value) of a single sentence can be judged by the position of emotion words and degree words. The target stocks that need emotional analysis should be all the stocks in chinext, but there are 796 stocks in chinext, some of which are illiquid, and the attention of stock bar is very low, with few or no page views or comments. Public opinion on such stocks has little effect on the stock price. GEM 50 index is based on the stock liquidity of gem, the proportion of the average total market value of the stock, the proportion of the average free market value and the proportion of the average transaction amount. By examining the average daily transaction amount of the candidate stock in the last six months, the top 50 stocks are selected to form the sample stock index. Therefore, in order to make the research representative, the members of the GEM 50 index were selected as the research objects, and the stock evaluation and emotion valuation were obtained for the past three years to fit the emotional factors of the GEM stocks. The data acquisition process for this article is shown in Figure 2.

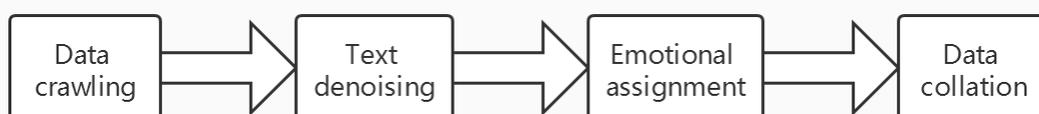


Figure 2 Emotional factors fitting steps

## 5. Empirical Study on the Influence of Network Investor Sentiment on GEM Stock Price

This paper USES the empirical method to study the impact of investor sentiment value on the time effect of gem stock prices. The model adopts the vector autoregressive (VAR) model to conduct empirical analysis through Stata software. The VAR model estimates the dynamic relationship of the joint endogenous variables by regressing the lagged variables of all variables in the current period. Therefore, this model can be used to effectively analyze the impact of investor sentiment on GEM stock price, so as to effectively identify the risk of stock price volatility.

### 5.1 Stability test

In order to determine the stationarity of the time series and ensure the validity and reliability of the following analysis results, the time sequence diagram between each variable is drawn at first, and the stationarity of the time series can be roughly judged by drawing. The time sequence diagram of each variable drawn is shown in figure 3. It can be seen that the variables in the figure fluctuate around the line of 0 without obvious time trend, and the time series is judged to be stable preliminarily.

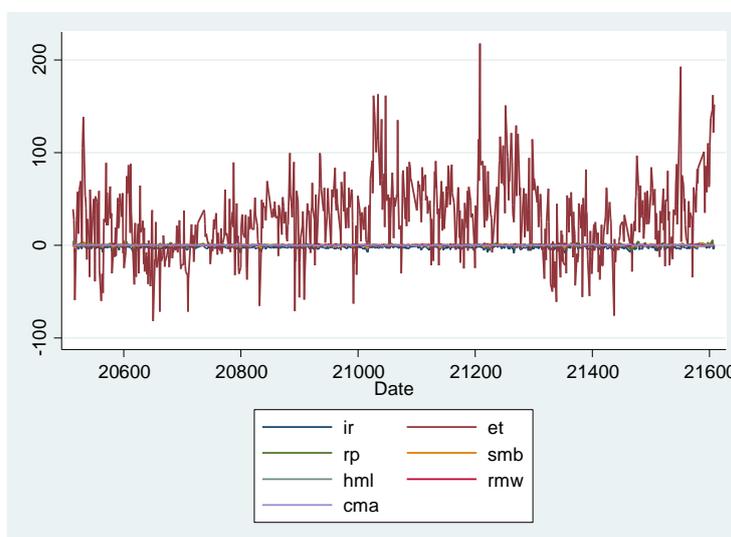


Figure 3 Sequence diagram of each variable

Next, a relatively accurate stationarity test was conducted for this sequence, and the test method was ADF unit root test. The results are shown in table 1. It can be seen from table 1 that the ADF statistics of Ir,Rp, Smb, Rmw, Cma and Et are all less than the corresponding critical value of 5% significance level, indicating that Ir,Rp, Smb, Rmw, Cma and Et are stationary time series. Therefore, VAR model can be directly substituted for regression.

Table 1 ADF test results

Variable	Lag order	ADF Statistics	P Valeue	Conclusion
Ir	1	-7.820	0.0000	stable
Rp	1	-8.126	0.0000	stable
Smb	1	-7.925	0.0000	stable
Hml	1	-7.806	0.0000	stable
Rmw	1	-8.680	0.0000	stable
Cma	1	-8.172	0.0000	stable
Et	1	-3.645	0.0049	stable

### 5.2 Model lag period determination

In the VAR model, it is crucial to determine the appropriate lag order, which usually indicates the time span when the current result is affected by the previous factors. It is generally believed that the larger the lag order, the larger the time span of the influence of the previous factors on the current results. But the larger the lag order, the smaller the number of free variables. In order to

determine the optimal lag order, LR, FPE, AIC, SC, AIC, HQIC and SBIC were used to select the lag order comprehensively in this paper. The results are shown in table 2. It can be seen from table 2 that the best lag order determined by the five judgment indicators is order 1, which indicates that the return rate of the previous GEM 50 index will have an impact on the current index return rate.

Table 2 Determines the model lag results

Lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-1197.9				0.1744	18.1187	18.1805	18.2709*
1	-1106.33	183.13	49	0.000	0.0920*	17.4786*	17.9732*	18.6956
2	-1075.85	60.963	49	0.117	0.1223	17.7571	18.6844	20.0389
3	-1042.53	66.636	49	0.048	0.1571	17.9929	19.3529	21.3369
4	-1007.82	69.411*	49	0.029	0.2004	18.2079	20.0006	22.6195

### 5.3 Model regression analysis

Then, the variables that determine the lag order are put into the VAR model. The regression analysis results are shown in the figure below.

Table 3 Regression results of variables

	Coef.	Std.Err.	Z	P> z	[95% Conf. Interval]	
IrL1.	0.0927	2.4738	0.04	0.970	-4.7558	4.9413
RpL1.	-0.5085	2.8130	-0.18	0.857	-6.0218	5.0048
SmbL1.	-9.2213	3.9665	-2.32	0.020	-16.9953	-1.4471
HmlL1.	4.0421	4.6810	0.86	0.388	-5.1324	13.2166
RmwL1.	-9.7461	5.0973	-1.91	0.056	-19.7366	0.2
CmaL1.	-3.3271	5.9285	-0.56	0.575	-14.9468	8.2925
EtL1.	0.5694	0.0403	0.04	0.970	-4.7558	4.9413
Cons	11.8392	6.8807	1.72	0.085	-1.6467	25.3252

It can be seen from table 3 that the P value of GEM 50 index return rate and GEM investor sentiment value, gem market factor, market value factor, book-to-market ratio factor, profitability factor and investment style factor are all less than 5%. The relationship formula is as follows:

$$\begin{aligned}
 Ir = & 0.0927ir(-1) - 0.5085Rp(-1) - 9.2213Smb(-1) + 4.0421Hml(-1) - 9.7461Rmw(-1) \\
 & - 3.3271Cma(-1) + 0.5694Et(-1) + 11.8392
 \end{aligned}
 \tag{11}$$

From the perspective of the regression equation, changes in book-to-market ratio and investor sentiment will cause positive changes in stock returns, while changes in market excess returns, market value factors, profitability factors and investment style factors will cause stock returns to change in the opposite direction. Among them, the book-to-market ratio factor of the previous period, investor sentiment factor increased by one unit, and the GEM 50 index yield increased by 4.0421 and 0.5694 units respectively. For each unit increase in the excess return rate, market value factor, profit factor and investment factor in the previous phase, the return rate of GEM 50 index decreased by 0.5085, 9.2213, 9.7461 and 3.3271, respectively. However, an increase of one units in the return rate of the last GEM 50 index will promote the increase of 0.0927 units in the current index.

### 5.4 Model stability test

In order to check whether the VAR model established in this paper is stable, it is necessary to test the stability of the model. The principle of stationarity test is as follows: if the reciprocal of all AR root modes of the model built is less than 1, that is, it falls within the unit circle, then the VAR model is stable; otherwise, it indicates that the model is unstable. The test results are shown in figure 4. As can be seen from figure 4, all root modules of the VAR model are located inside the unit circle, indicating that the VAR model is stable and the results obtained by analysis are reliable.

Next, the joint significance of coefficients of each order is tested, and the results are shown in table 4. The results show that except the book-to-market ratio factor, the p-value of other single

equations is less than 5% in the case of lag order, while the p-value of the total equation is less than 5%, indicating that the first-order coefficients are highly significant. This shows that investor sentiment, market excess return rate, company market value, book-to-market ratio, profitability and investment ratio all have an effective and significant impact on stock price, and stock price risk will occur when various factors change. By adjusting the six factors, the risk of investment and stock price fluctuation can be effectively reduced.

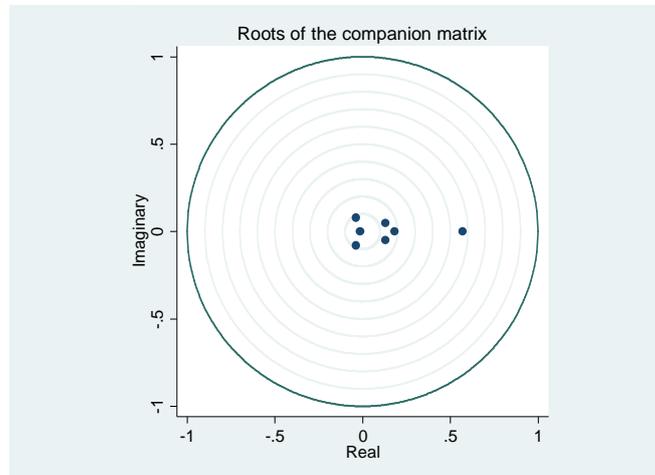


Figure 4 Stability discrimination diagram of VAR system

Table 4 Coefficient significance diagram

Equation	lag	chi2	df	Prob>chi2
ir	1	15.905	7	0.000
rp	1	15.720	7	0.028
smb	1	22.236	7	0.002
hml	1	13.737	7	0.056
rmw	1	28.150	7	0.000
cma	1	27.046	7	0.000
All	1	493.064	49	0.000

Next, the white noise test of residuals is performed. In other words, it is necessary to check whether there is autocorrelation between residuals. The results are shown in table 5. In any case, the p-value is greater than 5%, indicating that the null hypothesis of “no autocorrelation” of residuals can be accepted, that is, the disturbance term is considered as white noise.

Table 5 Residual white noise test

lag	chi2	df	Prob>Chi2
1	64.927	49	0.063
2	50.627	49	0.409

After the above steps, the real model has passed the test, indicating that the model has good stability and certain predictive ability. It indicates that the six factors are effective in predicting stock prices, and the analysis of the six factors can effectively reduce the risk of stock price volatility. So the model construction is successful in theory.

### 5.5 Impulse response function

Impulse response function and impulse response graph are the core empirical steps in VAR model. It means the dynamic impact of a unit standard deviation of a variable on other variables, so it is a relatively short-term dynamic change between variables. The impulse response analysis of the model is performed, and the results are shown in the figure below.

Figure 5 shows that stock yield ir has a positive impact on investor sentiment et, and the impact intensity gradually decreases to the minimum in the 10th phase, indicating the main source of investor sentiment on stocks and changes in stock yield. This emotion will diminish over time.

Figure 6 shows that investor sentiment has a positive impact on stock returns. High positive investor sentiment will lead to an increase in stock yield, but this effect has a short-term effect and gradually weakens after phase 1 and is minimized in phase 5. It indicates that the change of stock price affects the current investor sentiment, and the current investor sentiment will impact the next stock price. When the abnormal stock price occurs, investors are prompted to produce activist sentiment, aggravating the next stock price decline trend, thus exposing the stock to the risk. Therefore, the empirical results show that public opinion risk often leads to stock price risk.

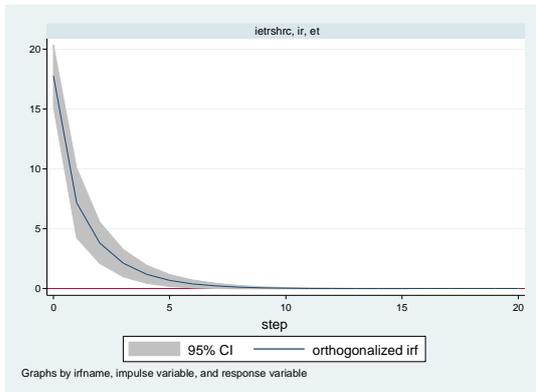


Figure 5 Impulse response diagram of ir to et.

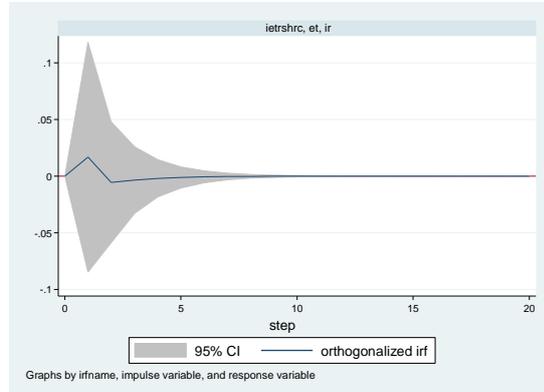


Figure 6 Impulse response diagram of et to ir.

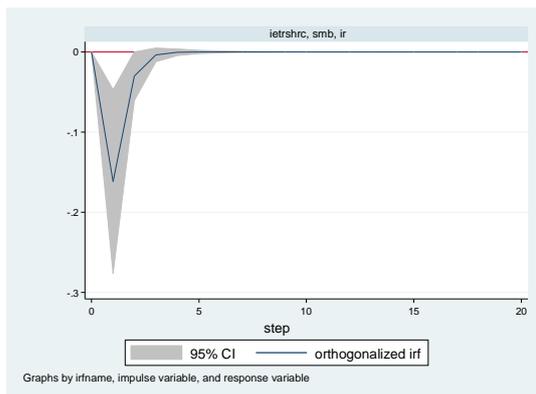


Figure 7 Impulse response diagram of smb to ir.

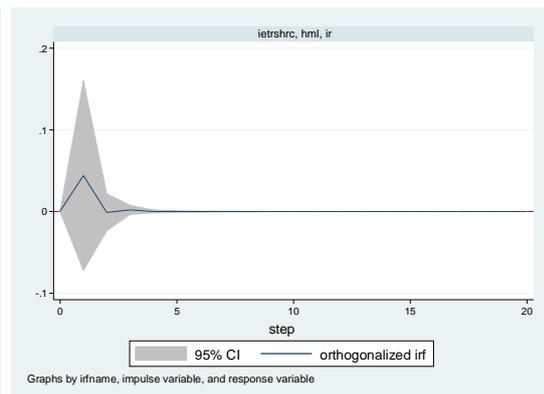


Figure 8 Impulse response diagram of hml to ir.

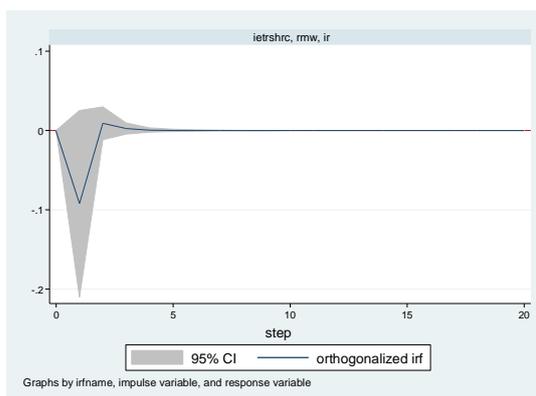


Figure 9 Impulse response diagram of rmw to ir.

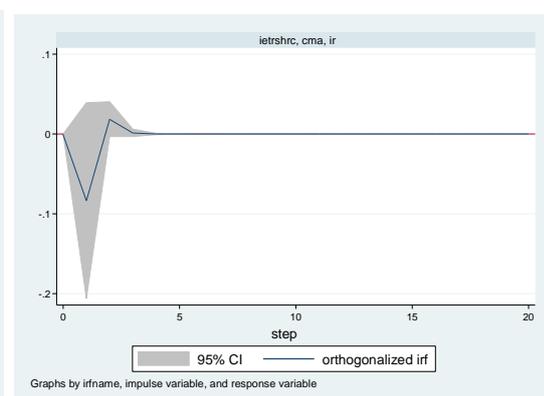


Figure 10 Impulse response diagram of cma to ir.

Figure 7 shows that the market value factor has a negative impact effect on stock returns, that is, the stock returns of the large market value group tend to have a low stock return rate. Such a short-term impact peaks in the first phase and gradually decreases to 0 in the fifth phase. On the contrary, Figure 8 shows that the stocks of companies with a high book-to-market ratio tend to have a high yield, which also confirms the market-to-market effect and book-to-market effect mentioned in Fama's research. Figure 9 and Figure 10 shows that profitability factor and investment style factor have negative impact on stock return. When the profitability factor and investment style

factor were given a unit standard deviation shock, the stock yield of GEM 50 index would fall. The results show that the more aggressive the investment style is, the higher the change rate of profitability is. The stock returns are usually less conservative than those of the investment style with low change rate of profitability. The research shows that investors tend to have a strong risk preference when investing in GEM stocks and bear a higher risk of stock price fluctuation.

## **6. Conclusions and Revelation**

### **6.1 Conclusions**

This study used web crawler technology to obtain the comments of investors of East Money stock bar on the GEM 50 index. After that, this paper uses emotion analysis tools to assign emotion values to stock bar comments and fit investor emotion factors. Emotion factors were added to the Fama-French five-factor model to form the six-factor model, and the model was analyzed in time. Finally, it was found that:

Investor sentiment can have a positive impact on stock returns and keep the same direction, and the lag order is the first order. Through the regression equation, we find that every unit change in investor sentiment in the current period tends to affect the next stock's return rate by 0.5694 units in the same direction. Such shocks tend to have a short-term effect, abating rapidly over time, often falling to zero within a week.

GEM 50 index yields will be affected by the risk of public opinion into a vicious circle. The current yields down cause investors to depression, appear even extreme short remarks. When the consensus is reached, the stock will continue to fall in the next period, which will form a vicious circle in the long run. The risk of public opinion will lead to the increase of stock price, and the stock price will fall to outliers.

The market value effect of stocks in gem still exists, that is, stocks of companies with larger market values tend to have higher yields than stocks of companies with lower market values. The reason is that investors generally believe that companies with smaller market capitalization have higher growth potential and higher development space, so their returns are generally higher than those of large-capitalization companies. And the book-to-market effect also exists in the gem market: stocks of companies with high book-to-market ratios tend to have higher yields than stocks with low book-to-market ratios. Analyze the reasons why companies with high book-to-market ratios tend to have better fundamental performance than companies with low book-to-market ratios. Risk preference shows that investors are more willing to take risks and obtain high returns, while ignoring the stability of company operation, which also indirectly confirms the existence of market value effect. Both the profitability factor and the investment style factor are negative in the model, which also proves that investors tend to choose stocks with small investment proportion and low profitability for investment, in order to expect higher stock returns. Therefore, investors often bear higher stock price risk in GEM stock investment.

### **6.2 Revelation**

In this empirical study, we found that investor sentiment does affect the stock price trend of gem, and the performance is in the same direction. When yields change, investors tend to react accordingly. When expressing opinions on public opinion platforms such as stock bar reveal personal emotions, such emotions are often accompanied by bullish or bearish views, and when other investors browse these views, their own views will be impacted or remain highly consistent. When there is more consensus, it forces investors to trade more quickly, which affects stock prices. This study found that this time interval is relatively short, often only one period. The current gem stock yield affects the current investor sentiment and affects the next gem stock price. Investor sentiment has a short influence on the stock price. When a new price appears, a new investor sentiment emerges and the stock price changes. Fluctuations in stock prices due to the degree of high investor sentiment often force stocks to be exposed to higher risks. When the sentiment of a GEM stock is high, the next stock yield rises and the stock enters an upward cycle. When stock

prices rise to outliers, they run a higher risk of falling.

In this study, we also found that GEM investors' risk appetite is manifested as risk seeking tendency. Investors often prefer to seek out "bottoming out" opportunities in stocks with smaller market capitalisations and temporarily underperforming fundamentals. Such investment idea has bigger risk. It is quite accidental for investors to obtain profits by relying on speculative ideas. The funds raised by speculation, rather than the good technology and strong strength of the company itself, lead to the inflated stock price, which will make the stock bear a higher risk of decline. In the analysis of sentiment, we can find that once the stock yield on GEM declines, investor sentiment will change in the same direction, namely, to put down the current price, which will have a negative impact on the next stock price, prompting more investors to bet against stocks. However, when investors choose to buy stocks, they pay more attention to speculation, while the growth value of stocks is relatively low. Therefore, GEM stocks will enter a period of decline and form a vicious circle. The decline rate of stocks will increase the strength of stocks under the promotion of public opinion risk. Investors will and will be exposed to higher risks and expected losses.

In order to avoid this risk, this paper puts forward the following Suggestions: firstly, investors should sort out the correct investment philosophy; when choosing the target company for investment, they should first analyze the company's financial status and operation status, and judge the company's growth value from the entity perspective. The strategy of replacing speculation with the concept of value investment and giving priority to fundamental analysis over technical index analysis can help investors effectively avoid the high risk appetite of gem investors and force themselves to be exposed to high risks. Secondly, for the risk of public opinion generated by investor sentiment, investors should rationally and objectively evaluate the current status of stocks, and refrain from malicious infusions or rumors to defame irrational remarks. One of the characteristics of the Chinese stock market is that there are many retail investors, and most of them have not learned professional financial knowledge. Therefore, it is easy to stir up emotions when browsing the platform. Bad comments on the stock will often drive other investors to make irrational trading behaviors and have an impact on the stock price. Promoting investors to standardize their own public opinion evaluation can not only help investors to sort out correct stock cognition, but also avoid abnormal stock prices. At the same time, the company shall ensure the openness and transparency of information, and shall not create false good news and conceal negative news for the purpose of driving up the stock price. In the long run, the stock price does not match the real strength of the company, and investors are exposed to potentially higher downside risk. By timely informing investors of the true operating conditions of the company, investors can effectively identify their own risk environment. This can also help investors to sort out the correct investment values and promote the public opinion environment to develop to an optimal level. Through the above methods to reduce the risk of stock prices, avoid the risk of public opinion, promote the growth enterprise market healthy and vigorous development.

## References

- [1] De Long J B, Shleifer A, Summers L H, et al. Noise Trader Risk in Financial Markets. *Journal of Political Economy*, 1990, 98 (4): 703-738.
- [2] Brown G W, Cliff M T. Investor Sentiment and Asset Valuation. *Journal of Business*, 2005, 78(2): 405-440.
- [3] Baker M, Wurgler J. Investor Sentiment and the Cross-Section of Stock Returns [J]. *Journal of Finance*, 2006, 61(4): 1645-1680.
- [4] Wang Meijin, Sun Jianjun. Returns, Volatility and Investor Sentiment in China's Stock Market. *Economic Research*, 2004 (10): 75-83.
- [5] Huang Delong, Wen Fenghua, Yang Xiaoguang. Investor Sentiment Index and Chinese Stock Market. *Systems Science and Mathematics*, 2009, 29(1): 1-13.

- [6] Gao Xiaoliang. Investor Sentiment, Average Correlation and Stock Market Returns. *China Management Science*, 2015, 23 (2): 10-20.
- [7] Lee W, Jiang C, Indro D. Stock Volatility, Excess Returns, and the Role of Investor Sentiment. *Journal of Banking and Finance*, 2002, 26(12): 2277-2299.
- [8] Tetlock P C. Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of finance*, 2007, 62(3): 1139-1168.
- [9] Siganos A, Vagenas-Nanos E, Verwijmeren P. Divergence of Sentiment and Stock Market Trading. *Journal of Banking & Finance*, 2017, 78: 130-141.
- [10] Yang Xiao. Empirical Analysis of the Relationship between Investor Sentiment and market turnover. *Statistics and decision-making*, 2016 (24): 163-166.
- [11] Lv Yongxia. An Empirical Study on the Impact of Network Public Opinion on Stock Market Returns -- based on the perspective of investor sentiment. *Finance and Accounting Newsletter*, 2018, 11 (3): 13-17.
- [12] Sayim M, Rahman H. The Relationship between Individual Investor Sentiment, Stock Return and Volatility. *International Journal of Emerging Markets*, 2015, 10(3): 504-520.
- [13] Lin C B, Chou R K, Wang G H K. Investor Sentiment and Price Discovery: Evidence from the Pricing Dynamics between the Futures and Spot Markets. *Journal of Banking & Finance*, 2018, 90: 1-31.
- [14] Li Helong, Feng Chune. Research on the Relationship between Investor Sentiment and Stock Index Fluctuation Based on EEMD. *System Engineering Theory and Practice*, 2014, 34 (10): 2495-2503.