ISSN: 2710-0723

Investor Sentiment and Stock Market Returns: An Empirical Study based on Market Capitalization Scale Perspective

Di Zou

School of Economics and Management, Southwest Petroleum University, Chengdu, 610500, China

Abstract

With the rise of behavioral finance, more and more scholars have begun to recognize the assumption of "limited rationality" of investors. Based on this, this paper employs empirical analysis methods to explore the mutual influence between investor sentiment and stock market returns across different market capitalization scales. The research results indicate that: (1) investor sentiment has a dual impact on the stock market returns of large, medium, and small-cap stocks, initially positive and subsequently negative, and the impact intensity increases as the market capitalization of the stock increases; (2) The returns of large-cap and mid-cap stocks have a significant positive impact on investor sentiment, while the impact of small-cap stock returns on investor sentiment is minimal.

Keywords

Investor Sentiment; Stock Market Returns; E-Business; Principal Component Analysis; VAR Model.

1. Introduction

China's stock market started late and still faces many problems after more than three decades of development, the most serious of which is the irrational investor structure, with a large number of individual investors dominating the stock market for a long time. On the whole, institutional investors have incomparable advantages over individual investors, such as more professional management, more reasonable asset allocation, and more standardized investment behaviors, while individual investors are highly susceptible to emotional influences and irrational trading behaviors.

Behavioral finance believes that investors have limited rationality. Investors in the stock market often show complex motives, on the one hand, investors will analyze the information obtained rationally, on the other hand, investors will also be affected by their own irrational factors, such as blindly listening to the opinions of other people, preferring a certain industry and so on.

By studying investor sentiment, we can gain a better understanding of investor trading behavior. This article selects direct and indirect indicators as proxy variables, employs multiple linear regression to eliminate the influence of macroeconomic factors, and utilizes a three-stage principal component analysis method to construct a comprehensive investor sentiment index. Subsequently, by constructing a VAR model between the comprehensive index of investor sentiment and stock index returns, this paper empirically investigates the correlation between investor sentiment and stock market returns across different market capitalization scales. Based on the findings, it draws conclusions that provide a new theoretical framework and methodology for the regulatory mechanism of the stock market.

ISSN: 2710-0723

2. Literature Review

2.1. A Review of Investor Sentiment Research

In the stock market, people often tend to exaggerate their actual abilities. When they make a profit, they attribute it more to their personal ability; when they face a loss, they tend to blame it on bad luck. Daniel and Titman [1], in their study of how investors' overconfidence affects their investment decisions, found that overconfidence generates stock return momentum, and that stocks whose valuations require the interpretation of ambiguous information have a stronger momentum effect.

When people face decisions or judgments, they always tend to take a certain "initial value" as a reference, although this "initial value" may not have any relevance to the decision itself, but it is like an anchor at the bottom of the sea that restricts people's thinking.

Psychology tends to believe that people feel more secure when faced with familiar people and things, and that a large part of people's liking for things is influenced by familiarity. Benartzi [2] cites the example of Coca-Cola to explore this, and finds that the employees of Coca-Cola invested no less than 76% of their pensions in the Company's stock.

2.2. A Review of Research on the Relationship between Investor Sentiment and Financial Market Anomalies

Traditional financial theories cannot explain many financial market anomalies, so more and more scholars began to study these market anomalies from the perspective of behavioral finance. Pontiff [3] found that the fund price risk was significantly greater than the risk of the assets invested in the fund by testing the excessive volatility of closed-end fund prices, so it was speculated that there was investor sentiment risk. Lee et al. [4] proposed the theory of investor sentiment, arguing that changes in the discount rate stem from fluctuations in sentiment, which can fully explain the mystery of discount.They studied the problems related to the closed-end fund premium mystery, and got the results showing that when investor sentiment is extremely optimistic, the fund will have a premium situation, however, as time passes, investor sentiment gradually calms down, and in the late stage of the fund's existence, the fund's discount magnitude decreases or even disappears along with the reduction of emotional risk.

In addition, some domestic scholars, such as Wu and Han [5], use investor sentiment to explain the mystery of IPOs in China, respectively, and conclude that investor sentiment is an important influence on asset pricing.

2.3. A Review of Studies on the Relationship between Investor Sentiment and Stock Market Returns

The theory that investor sentiment has an impact on the stock market comes first from Black [6], who suggests that it is the presence of noise and noise traders that allows for increased liquidity in the market and is a prerequisite for the existence of the market, however, noise traders also bring more risk to the market. De Long et al. [7], further proposed a theoretical model of noise trading (DSSW) on top of Black's hypothesis, pointing out that investor sentiment is a systematic factor affecting stock prices, and the results of the study imply that the theoretical system centered on noise traders is initially formed. De Bondt's [8] study shows a significant positive correlation between individual investor sentiment and stock market returns.

ISSN: 2710-0723

3. Research Methods and Data Sources

3.1. Research Methods

3.1.1. Principal Component Analysis

Principal Component Analysis (PCA)'s core idea is to map the original data into a new, lower-dimensional space, which is composed of a set of orthogonal new feature vectors, also known as principal components. These principal components are linear combinations of the original feature vectors and aim to retain the maximum variance of the original data. The working principle of PCA is to sequentially find a set of mutually orthogonal axes from the original space. The first axis is chosen to be the direction with the maximum variance in the original data, and the second axis is the direction in the plane orthogonal to the first axis that maximizes the variance. This process continues iteratively, selecting subsequent axes to maximize variance while remaining orthogonal to previous axes [9]. The general model of principal component analysis can beexpressed by formula (1):

$$\left.\begin{array}{c} F_1 = a_{11}X_1 + a_{12}X_2 + a_{13}X_3 + \cdots + a_{1n}X_n \\ F_2 = a_{21}X_1 + a_{22}X_2 + a_{23}X_3 + \cdots + a_{2n}X_n \\ \cdots\cdots \\ F_m = a_{m1}X_1 + a_{m2}X_2 + a_{m3}X_3 + \cdots + a_{mn}X_n \\ Y = \dfrac{\lambda_1}{\sum_{i=1}^{k}\lambda_i}F_1 + \dfrac{\lambda_2}{\sum_{i=1}^{k}\lambda_i}F_2 + \dfrac{\lambda_3}{\sum_{i=1}^{k}\lambda_i}F_3 + \cdots + \dfrac{\lambda_k}{\sum_{i=1}^{k}\lambda_i}F_k \end{array}\right\} \tag{1}$$

X1,X2,...,Xn are measured variables; F1,F2,...Fm are main components; aij(i=1,2,...,m; j=1,2,...,n) are factor loadings which mean the load of measured variable j on the main component i; Fk is the selection of principal components; λ1,λ2,...,λk are the variance contribution rate of the selected principal component Fk; Y is the final score for the composition of the selected principal components.

In the field of economics, Principal Component Analysis (PCA) is widely used to construct composite indicators due to its ability to identify the main structures within a dataset. These composite indicators include measures such as the Renewable Energy Development Index [10], Household Wealth Index [11], and Socioeconomic Status Index [12].

3.1.2. VAR Model

Vector Autoregressive (VAR) model is a non-structural equation model used to estimate dynamic relationships among multiple variables. The VAR model constructs the model by treating each endogenous variable in the system as a function of lagged values of all endogenous variables in the system. This extension allows for the generalization of univariate autoregressive models to "vector" autoregressive models composed of multivariate time series variables. The VAR model is commonly used for forecasting interrelated time series systems and analyzing the dynamic effects of random disturbances on variable systems. For example, it has been applied to forecast the real price of oil [13], among other studies.

The research methods closely associated with the VAR model include impulse response analysis and variance decomposition analysis. The impulse response function reflects the dynamic impact on other variables in the model when a variable in the VAR model experiences an "external shock." Variance decomposition analyzes the contribution of each structural shock to the variation of endogenous variables, further assessing the importance of different structural shocks.

ISSN: 2710-0723

3.2. Data Sources

3.2.1. The Proxy Variables of the Investor Sentiment Composite Index

In this paper, we draw inspiration from the construction methods of the BW index [14] and the CICSI index [15]. Based on the availability of data and the actual situation of the Chinese stock market, we selected seven proxy variables which including the Shanghai and Shenzhen A-share stock market's new investor accounts, price-earnings ratio, daily average trading volume, daily average trading amount, turnover rate, IPO number, and the investor confidence index. Subsequently, we employed principal component analysis to construct a comprehensive investor sentiment index (ISIrt). The names, symbols, and descriptive statistics of the proxy variables for investor sentiment are presented in Table 1:

Table 1. Descriptive statistics for seven proxy variables

Variable name	Indicator symbol	Minimum	Maximum	Mean	Std. Deviation
new investor accounts	NIA	28.73	497.50	140.95	78.21
price-earnings ratio	PE	20.80	48.71	30.28	6.62
daily average trading volume	CJL	138	1107	531.45	226.72
daily average trading amount	CJY	1310	17458	6690.61	3369.07
turnover rate	TURN	10.7901	65.13	24.40	10.71
IPO number	IPO	0	104	20.08	14.82
Investor confidence index	CICI	41.3	71.2	56.98	6.19

3.2.2. The Proxy Variables of the Macroeconomic Influence

In the stock market, the trading behavior of limited rational investors not only includes the influence of their irrational emotions, but also includes the rational expectation of macroeconomic changes [14]. Therefore, when constructing the comprehensive investor sentiment index, it is necessary to exclude the influence of macroeconomic factors on sentiment. Otherwise, it may disturb the accuracy of sentiment measurement, thereby affecting the objectivity of the conclusions. In this study, the Consumer Price Index (CPI), Producer Price Index (PPI), and Macroeconomic Business Climate Index (MBCI) were chosen as proxy variables for macroeconomic factors, and a multiple linear regression model was used to remove the influence of macroeconomic factors.

3.2.3. Other Variables and the Description of Data Sources

In the subsequent empirical study on investor sentiment and stock market returns, to conduct an in-depth exploration from the perspective of different market capitalization sizes, this paper selects the return rate of CSI300 Index (CSI300), CSI500 Index (CSI500) and CSI1000 Index (CSI1000) to represent the stock returns of large-cap, mid-cap, and small-cap stocks respectively. The CSI 300 Index reflects the stock price performance of the 300 largest and most liquid A-share companies in the Shanghai and Shenzhen stock markets. The CSI 500 Index covers the 500 enterprises with good liquidity and relatively large size outside the CSI 300 Index. Furthermore, the CSI 1000 Index includes an additional 1000 companies with higher market capitalization and good liquidity, excluding those covered by the first two indices.

Based on the availability of data, the selected sample data for this study ranges from January 2014 to July 2023, with all monthly data sourced from Resset Database, China Securities Depository and Clearing Corporation, National Bureau of Statistics of China, and Eastmoney.com.

ISSN: 2710-0723

4. The Construction of Investor Sentiment Composite Index

4.1. Selection of "Advance" and "Lag" Variables

Since the reflection of different indicators on investor sentiment may have the relationship of "advance" and "lag" in time, it is necessary to determine whether the current value or the lagging value of the proxy variable participates in the principal component analysis. First, a principal component analysis (PCA) is conducted on the leading and lagging variables of each indicator (the variables have been standardized to eliminate the impact of unit differences). This allows us to construct an investor sentiment index (isi_t) that comprises 14 variables. To capture as much information as possible, this paper strictly adheres to the statistical criterion of achieving a cumulative variance explanation rate of at least 85% [15]. The results show that the cumulative variance explanation rate of the five principal components in this case reaches 88.33%. The value of isit is calculated by taking the weighted average of the five principal components and referring to the scoring coefficient table of each principal component. Next, a Pearson correlation analysis was performed between isi_t and the 14 variables, and accordingly, seven variables with relatively large correlation coefficients were selected as proxy variables for constructing the comprehensive investor sentiment index in subsequent steps.

According to the results shown, isi_t is highly correlated with $CICI_{t-1}$, NIA_{t-1}, PE_{t-1}, CJL_t, CJY_t, $TURN_t$, and IPO_t. In addition, the investor confidence index, the new investor accounts and the price-earnings ratio all reflect investor sentiment in advance. Next, we select these 7 variables as the ultimate source for constructing the basic investor sentiment index (ISIt).

4.2. The Construction of the Basic Investor Sentiment Index (ISIt)

The results of the second principal component analysis show that the cumulative variance explanation rate of the first to fourth principal components reaches 91.93%, and the basic investor sentiment index ISIt is calculated. Table 3 shows its correlation coefficient with 7 proxy variables:

Table 2. Correlation between ISIt and the 7 variables

	ISIt	$CICI_{t-1}$	NIA_{t-1}	PE_{t-1}	CJL_t	CJY_t	$TURN_t$	IPO_t
ISIt	1							
$CICI_{t-1}$	0.592**	1						
NIA_{t-1}	0.741**	0.305**	1					
PE_{t-1}	0.610**	0.077	0.514**	1				
CJL_t	0.609**	0.476**	0.416**	-0.042	1			
CJY_t	0.726**	0.461**	0.542**	0.123	0.952**	1		
$TURN_t$	0.659**	0.423**	0.606**	0.351**	0.678**	0.750**	1	
IPO_t	0.690**	0.201*	0.275**	0.293**	0.324**	0.405**	0.121	1

Note:*,**respectively represent 5% and 1% significance levels (bilateral); the same is true for tables below.

It can be observed from the table 2 that the correlation between the basic investor sentiment index (ISIt) and each proxy variable is positive, which is consistent with the actual situation.

4.3. The Construction of the Comprehensive Investor Sentiment Index (ISIrt) after Excluding Macroeconomic Impacts

In this section, we utilize a multiple linear regression model to eliminate the impact of macroeconomic factors (with all variables standardized before the regression). We regress the seven sentiment proxy variables from the previous section against three macroeconomic proxy

ISSN: 2710-0723

variables: CPI, PPI, and MBCI. The residual series resulting from these regressions are represented by CICIrt-1, NIArt-1, PErt-1, CJLrt, CJYrt, TURNrt, and IPOrt. Subsequently, we conduct a third principal component analysis on these seven residual series, ultimately obtaining four principal components with a cumulative variance explanation rate of 91.40%. Following the same calculation method as for isit mentioned previously, we obtained the comprehensive investor sentiment index after controlling for macroeconomic impacts, ISIrt. Its expression and the correlation coefficients with various variables are presented in formula (2) and Table 4:

$$\begin{aligned} ISIr_t &= 0.209CICIr_{t-1} + 0.155NIAr_{t-1} + 0.208PEr_{t-1} \\ &+0.025CJLr_t + 0.057CJYr_t + 0.098TURNr_t + 0.178IPOr_t \end{aligned} \tag{2}$$

Table 3. Correlation between $ISIr_t$ and the 7 variables

	$ISIr_t$	$CICI_{t-1}$	NIA_{t-1}	PE_{t-1}	CJL_t	CJY_t	$TURN_t$	IPO_t
$ISIr_t$	1.000							
$CICI_{t-1}$	0.611**	1.000						
NIA_{t-1}	0.751**	0.248**	1.000					
PE_{t-1}	0.631**	0.025	0.493**	1.000				
CJL_t	0.566**	0.496**	0.398**	-0.063	1.000			
CJY_t	0.684**	0.462**	0.523**	0.100	0.948**	1.000		
$TURN_t$	0.733**	0.369**	0.609**	0.364**	0.696**	0.773**	1.000	
IPO_t	0.664**	0.286**	0.309**	0.337**	0.288**	0.384**	0.262**	1.000

From the results in the table above, it can be seen that ISIrt largely retains the characteristics of ISIt, and a test revealed that the correlation between the two indices reaches 92.9% (significant at the 1% level).

5. An Empirical Analysis of the Mutual Influence of Investor Sentiment and Stock Market Returns under Different Market Capitalization Scales

5.1. Stability Tests for Data

A stationary time series can effectively avoid the occurrence of "pseudo-regression" phenomenon, therefore, before constructing the VAR model, the unit root test was first conducted on the investor sentiment index (ISIrt) , CSI 300 Index Return (CSI300), CSI 500 Index Return (CSI500), and CSI 1000 Index Return (CSI1000) by Eviews software(all data have been standardized). The results are presented in Table 5:

Table 4. Augmented Dickey-Fuller test results of each variables

Variable	t-Statistic	Prob.	Conclusion
$ISIr_t$	-4.8405	0.0000	Stationary
CSI300	-8.8818	0.0000	Stationary
CSI500	-8.3825	0.0000	Stationary
CSI1000	-8.5827	0.0000	Stationary

The table 4 shows that the p-values for ISIrt, CSI300, CSI500, and CSI1000 all are 0.0000, which is below 0.05, indicating that all the data is stationary and can proceed to the next step of analysis.

ISSN: 2710-0723

5.2. Granger Causality Test

To observe more clearly the relationship between investor sentiment and stock market returns under different market capitalization scales, we performed a two-lag Granger causality test between ISIrt and CSI300, CSI500, and CSI1000. The test results are presented in Table 5:

Table 5. Granger causality test

Null Hypothesis:	Lag	F-Statistic	Prob.
CSI300 does not Granger Cause ISIrt	1	57.9972	0.0000
	2	29.4487	0.0000
ISIrt does not Granger Cause CSI300	1	9.70705	0.0023
	2	5.97527	0.0035
CSI500 does not Granger Cause ISIrt	1	114.827	0.0000
	2	57.8527	0.0000
ISIrt does not Granger Cause CSI500	1	7.81230	0.0061
	2	3.06500	0.0508
CSI1000 does not Granger Cause ISIrt	1	100.612	0.0000
	2	47.6670	0.0000
ISIrt does not Granger Cause CSI1000	1	6.02912	0.0156
	2	2.10521	0.1268

The data in the table 5 shows that, at a significance level of 5%, ISIrt and CSI300 consistently serve as Granger causes for each other, indicating a significant mutual influence between investor sentiment and the return rate of large-cap stocks. Additionally, ISIrt and CSI500, as well as ISIrt and CSI1000, are Granger causes for each other at the first lag. But at the second lag, only CSI500 and CSI1000 are Granger causes for ISIrt. By observing the p-values and F-values, it is found that the influence of stock market returns on investor sentiment is significantly greater than the influence of investor sentiment on stock market returns, a feature that is particularly evident in the medium and small-cap stock markets.

5.3. Building a VAR Model

Before establishing the VAR model, we need to determine the optimal lag order of the model. The results of the five evaluation criteria, including LR, FPE, AIC, SC, and HQ shown that the optimal lag order is 1. Therefore, this article chooses to construct a VAR model with a lag order of 1.

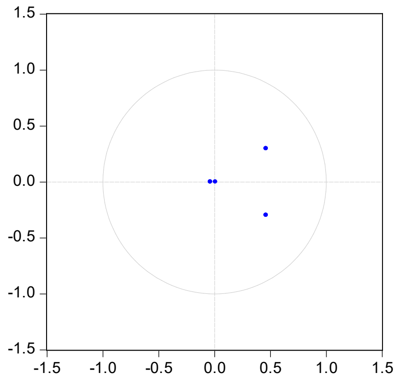

Figure 1. Inverse Roots of AR Characteristic Polynomial

By examining the unit root diagram, we can assess the reliability of the VAR model. As can be seen from the plot, all the roots of the VAR model fall within the unit circle, indicating that the model has passed the stability test and meets the stationarity requirement. This allows us to proceed with subsequent analysis, such as impulse response analysis and variance decomposition analysis.

5.4. Impulse Response Function Analysis

5.4.1. Impulse Response Function Analysis of CSI300\CSI500\CSI1000 to ISIrt

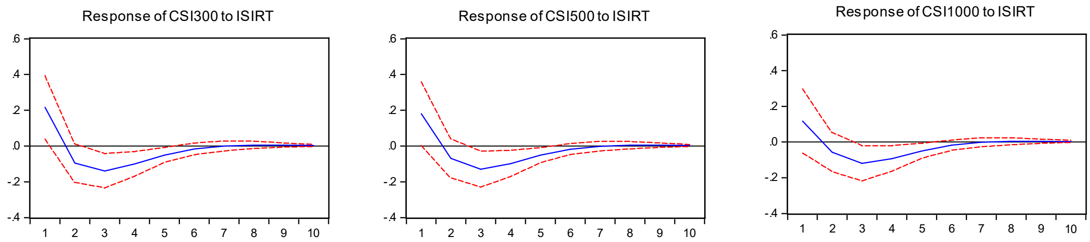

Figure 2. Impulse response function of CSI300\CSI500\CSI1000 to $ISIr_t$

Figure 2 demonstrates that, overall, when stock market returns are subjected to a one-standard-deviation shock in investor sentiment, the impact reaches its maximum in the first period. However, in the subsequent period, it turns negative and persists for a relatively long time, gradually converging to 0 from the sixth period onwards. This indicates that while investor sentiment can positively boost stock market returns in the short term, over time, it starts to have a negative corrective effect on market returns. In the long run, this corrective effect gradually diminishes.

From the perspective of indices with different market capitalization sizes, the trends of CSI300, CSI500, and CSI1000 after being impacted are roughly similar. Among them, CSI300 experiences the greatest impact in the first period, reaching 0.22, and it also has the smallest impulse value of -0.14 in the third period. This indicates that CSI300 is more sensitive to shocks in investor sentiment.

5.4.2. Impulse Response Function Analysis of ISIrt to CSI300\CSI500\CSI1000

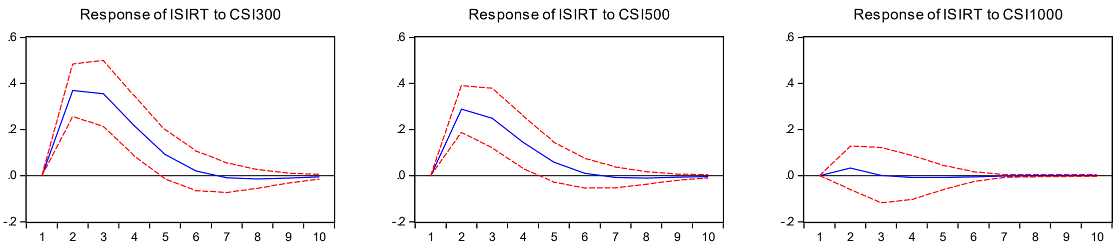

Figure 3. Impulse response function of $ISIr_t$ to CSI300\CSI500\CSI1000

As can be seen from Figure 3, among large-cap and mid-cap stocks, investor sentiment exhibits a similar trend of change. When it is impacted by a one-standard-deviation shock in stock market returns, the impulse value in the first period is 0, reaching its maximum in the second period, which are 0.37 and 0.29 respectively. Then it gradually declines and converges to 0 starting from the sixth period, displaying an inverted U-shaped positive impact that first expands and then shrinks. This indicates that although stock market returns have a certain degree of promoting effect on investor sentiment, such effect has a lag.

From the perspective of the impact of small-cap stock returns, after being impacted by CSI1000, investor sentiment only showed a slight positive reaction in the second period, with a pulse value of 0.03, demonstrating low sensitivity.

ISSN: 2710-0723

5.5. Dynamic Variance Decomposition Analysis

5.5.1. The Variance Contribution Rate of ISIrt to CSI300\ CSI500\ CSI1000

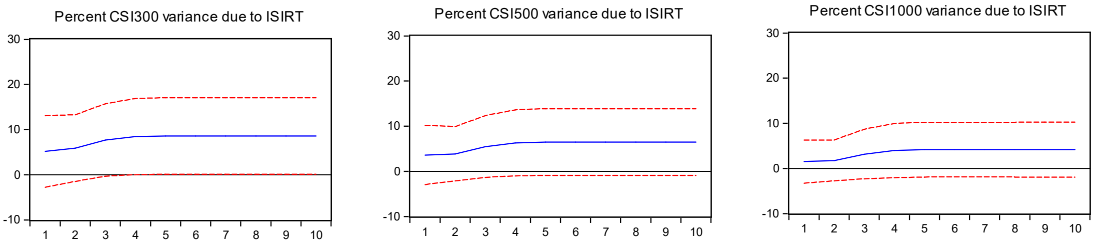

Figure 4. The variance contribution rate of $ISIr_t$ to CSI300\ CSI500\ CSI1000

Figure 4 shows that, overall, although investor sentiment has a certain degree of explanatory power for stock market returns in the first period, the level is relatively low. Although it increases slightly afterwards, the variance contribution rate of investor sentiment does not exceed 9% even by the 10th period. This indicates that although stock market returns are influenced by investor sentiment, the degree of influence is relatively small, and they mainly depend on the increase or decrease of their own intrinsic value. From the perspective of individual indices, the explanatory power of investor sentiment on CSI300, CSI500, and CSI1000 decreases successively, especially for CSI1000, with only 4.15%. This indicates that as the market capitalization of stocks decreases, investors' attention to them also decreases, and thus the impact of investor sentiment on their stock market returns is also smaller. This is consistent with the previous impulse response analysis results.

5.5.2. The Variance Contribution Rate of CSI300\ CSI500\ CSI1000 to ISIrt

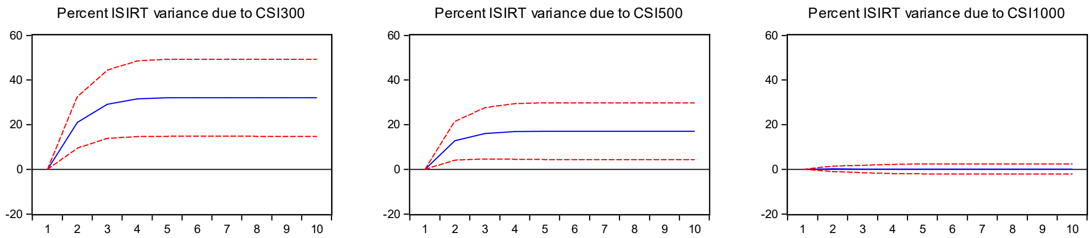

Figure 5. The variance contribution rate of CSI300\ CSI500\ CSI1000 to ISIrt

As can be seen from Figure 5, overall, stocks of different market capitalization sizes have significantly differentiated explanatory power with regard to investor sentiment. In the 10th period, the variance contribution rates of CSI300 and CSI500 are 32.04% and 17.04%, respectively, while the variance contribution rate of CSI1000 is only 0.13%. This suggests that as market capitalization increases, the impact of investor sentiment on stock market returns also becomes greater. This may be because large-cap stocks tend to have a broader investor base and higher liquidity. Additionally, their stock prices are more often determined by internal factors such as the company's own business performance, providing a certain level of assurance. Therefore, they can have a more significant impact on investor sentiment, which is consistent with the previous impulse response analysis results.

6. Conclusion

By utilizing the principal component analysis and multiple linear regression methods, this article constructs a comprehensive investor sentiment index and innovatively establishes a VAR model linking this index with the yield of CSI 300 Index, CSI 500 Index, and CSI 1000 Index. It explores the mutual influence between investor sentiment and stock market returns from the perspective of different market capitalization sizes, and the main conclusions are as follows:

ISSN: 2710-0723

Investor sentiment has a dual impact on the stock market returns of large-cap, mid-cap, and small-cap stocks, initially positive and then negative, and this impact intensifies as the market capitalization of the stocks increases. In the short term, positive investor sentiment will drive an increase in stock market returns, but over time, the stock market returns will gradually be revised in the opposite direction.There are differences in the impact of stock market returns of large-cap, mid-cap, and small-cap stocks on investor sentiment. The stock market returns of large-cap and mid-cap stocks have a significant inverted "U"-shaped positive impact on investor sentiment. In the initial stage of increasing stock market returns, investors need time to interpret positive market news, and then investor sentiment gradually rises, ultimately returning to rationality over time. However, compared to large-cap and mid-cap stocks, small-cap stocks have a minimal impact on investor sentiment due to their relatively low market capitalization.

References

[1] Daniel, K., & Titman, S. (1999). Market efficiency in an irrational world. Financial Analysts Journal, 55(6), 28-40.

[2] Benartzi, S. (2001). Excessive extrapolation and the allocation of 401 (k) accounts to company stock. The Journal of Finance,56(5), 1747-1764.

[3] Pontiff, J. (1997). Excess volatility and closed-end funds. The American Economic Review, 155-169.

[4] Lee, C. M., Shleifer, A., & Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle. The journal of finance, 46(1), 75-109.

[5] Swaminathan, B. (1996). Time-varying expected small firm returns and closed-end fund discounts. The Review of Financial Studies, 9(3), 845-887.

[6] Black, F. (1986). Noise. The journal of finance, 41(3), 528-543.

[7] De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. Journal of political Economy, 98(4), 703-738.

[8] De Bondt, W. P. (1993). Betting on trends: Intuitive forecasts of financial risk and return. International Journal of forecasting, 9(3), 355-371.

[9] Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences, 374(2065), 20150202.

[10] Wang, Y., Zhang, D., Ji, Q., & Shi, X. (2020). Regional renewable energy development in China: A multidimensional assessment. Renewable and Sustainable Energy Reviews, 124, 109797.

[11] Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure data—or tears: an application to educational enrollments in states of India. Demography, 38(1), 115-132.

[12] Vyas, S., & Kumaranayake, L. (2006). Constructing socio-economic status indices: how to use principal components analysis. Health policy and planning, 21(6), 459-468.

[13] Baumeister, C., & Kilian, L. (2012). Real-time forecasts of the real price of oil. Journal of business & economic statistics, 30(2), 326-336.

[14] Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. The journal of Finance, 61(4), 1645-1680.

[15] Yi, Z.G., Mao, N. (2009). The study of investor sentiment measurement in Chinese stock market: the construction of CICSI. Financ Research, 26 (11), 174-184. (in Chinese).