The Influence of Institutional Investors' Clustering Behavior on A-Share Stock

Kai Wang *

Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, 200030, China

* Corresponding Author Email: 1597830016@sjtu.edu.cn

Abstract. Institutional investors' "clustering" behavior can have significant implications for the stock market, especially concerning its stability and the efficiency of the price discovery mechanisms. Recognizing this, this research utilized regression models to rigorously analyze the effects of such clustering practices on the abnormal returns and volatility of stocks that are targeted by Equity-oriented funds. This research also implemented a T-test to differentiate the impacts of various variables in both bull and bear market scenarios. The findings underscore that within China's A-share market, equity-oriented fund tend to favor certain sectors, with a significant inclination towards the manufacturing and finance sectors. Furthermore, the clustering behavior exhibited by these investors has diverse effects on stock abnormal returns and volatility, contingent upon the time scales and prevailing market conditions. This discovery provides insights into institutional investment strategies and the overarching stability of the market. For individual investors, it's crucial to note that stocks with heavy institutional ownership might exhibit increased volatility, particularly in bear market conditions. Therefore, incorporating this understanding could be important in investment decision-making processes.

Keywords: Clustering, Institutional investors, Abnormal returns, Volatility.

1. Introduction

Institutional investors manage substantially larger funds compared to individual investors. They're equipped with specialized research teams and adhere to strict trading protocols. Their investment behaviors are characterized by structured portfolios, professional management, and standardized practices. When institutional investors actively make stock choices based on acquired data, this "clustering" behavior can uncover stock value and discern market price patterns. However, they can also be prone to blindly following trends, overlooking a stock's intrinsic value, and simply mirroring the actions of other institutions [1]. This clustering can induce significant stock price fluctuations, threatening market stability, and even may cause disastrous consequences [2, 3].

Studying the price variations of stocks affected by clustering offers deeper insights into its impact on capital markets. On one hand, if clustering pushes a stock's price significantly higher, it may distort the market's perception of its actual value. On the other, if it leads to a sharp decline, it might trigger market panic. Thus, identifying "herded" stocks provides valuable information for investors, aiding decision-making. Investigating the effects of clustering on stock prices is crucial for maintaining market stability and protecting investors' interests.

2. Literature Review

While scholars have long noticed the "herding" tendency of various institutions to trade in the same direction for the same stock [4], there's a distinct difference between "clustering" and "herding". Generally speaking, when people use the terms "herding" or "herd behavior" in economics and finance, investors refer to the process in which economic agents imitate one another's actions and/or base their judgments on those of others [5]. "Herding" reflects similarities in institutional trading behaviors—whether they trade in the same direction for a particular stock at the same time. In contrast, "clustering" characterizes the phenomenon where multiple institutions simultaneously statically hold the same stocks [6, 7]. Thus, there's a compelling need for further exploration of institutional "clustering".

Regarding the determination of institutional clustering, some scholars have defined it as "two or more funds jointly holding a significant position in the same stock during the same reporting period"[8, 9]. Here, the basis of "significant position" is whether the stock ranks among a fund's top ten holdings for the quarter.

For evaluating the extent of institutional clustering for a particular stock, an indicator called IIGI (Institutional Investors Group Index) has been introduced [9]. Influencing factors of IIGI encompass absolute allocation proportion, relative over-allocation ratio, number of funds holding the stock, number of fund companies, individual stock pricing power, and the average individual stock pricing power.

Existing research indicates that, compared to bull markets, the amplifying effect of fund clustering on stock price volatility is more evident in bear markets [10]. Therefore, in empirical analyses, it's crucial to account for market environment differences, distinguishing scenarios in both bull and bear markets.

3. Data Sources and Processing Procedure

In 2018, the Chinese A-share market experienced a bearish trend, with the Shanghai Composite Index declining by 24.59%. However, by 2019, the market underwent a fundamental shift, with the Shanghai Composite Index surging almost 20% for the year, signaling the advent of a bull market. Given this market backdrop, our study focused on the stock information disclosed by fund companies in the first quarter of 2018 and 2019. The objective was to explore the impact of such information on stock prices over future intervals of 5, 30, 120, and 240 trading days.

Many studies utilize the criteria of stocks co-held by at least two funds to gauge institutional clustering [7, 8]. This approach is justified by:

a) Statistical Reliability: Simultaneous investments by multiple funds in the same stock typically reflect market consensus rather than isolated choices, enhancing research validity.

b) Data Accessibility: A stricter criterion, like stocks co-held by five-plus funds, could reduce the sample size, hindering study feasibility. This two-fund standard ensures ample data.

c) Market Diversity: With varied investment strategies, a high joint-hold criterion might miss stocks overlooked due to strategy differences, ensuring a holistic market view.

d) Trade-off: The number "two" provides a balance. A solo fund's choice may show individual bias, but a too-stringent standard might miss truly clustered stocks. This criterion ensures both reliability and scope.

This research has implemented a series of filtering measures in data processing to enhance the accuracy and consistency of the data. Initially, stocks marked as ST, *ST, or those that have been delisted were excluded. Next, companies missing key financial metrics such as Net Asset Return (ROE), Price-to-Earnings Ratio (PE), Debt-to-Asset Ratio (DTA), the natural logarithm of the company's market capitalization (LMV), and the holding ratio of the top ten shareholders (Holder_T10) were removed. Finally, all stocks were classified based on the industry standards set by the Securities Regulatory Commission.

To investigate whether herding behavior as "collusion" or "consensus" affects stock prices, data was gathered from rating agencies for the years 2018 and 2019 concerning comprehensive stock rating values, number of buy ratings, and the number of research reports issued for the stock. These were then dimensionally reduced into an index PCA_cluster through principal component analysis, exploring the effects of different forms of herding on stock prices.:

In terms of the methodology, the research chose the 15th trading day after each quarter-end as the starting point (T). At T+t periods (t=5,30,120,240), relevant data were collected and calculated, aiming to glean more insights about the herding influence. Then, using each quarter-end's 15th trading day as the starting point (T), this part sought and calculated the following data for T+t periods (t=5, 30, 120, 240).

Table 1. Data item and Symbol

No	Data item	Symbol
1	Average trading volume of the stock for the previous t days	$AV_{i,t}$
2	Turnover rate for the previous t days	$TR_{i,t}$
3	Amplitude for the previous t days	$A_{i,t}$
4	Number of consecutive upper limit days from day T (maximum value)	$SLD_{i,t}$
5	Number of consecutive lower limit days from day T (maximum value)	$DLD_{i,t}$
6	Average number of funds holding the stock as a major position for the four quarters of that year	FN_i
7	Average number of fund companies holding the stock as a major position for the four quarters of that year	FCN_i
8	Proportion of the stock held by institutions	$IS_{i,t}$
9	Return on net assets of the stock company	$ROE_{i,t}$
10	Price-to-Earnings ratio	$PE_{i,t}$
11	Company debt ratio	$DTA_{i,t}$
12	Natural logarithm of the company's circulating market value	LMVi,t
13	Proportion held by the top ten shareholders	$Holder\_T10_{i,t}$
14	Industry the company belongs to, classified according to the China Securities Regulatory Commission	I_i
15	Cumulative abnormal return of the stock compared to the market benchmark during that time period	$CAR_{i,t}$
16	Stock price volatility during that time period	$VP_{i,t}$
17	Average change rate of institutional holdings for that year	F_change_i
18	Comprehensive rating score given by rating agencies for the stock during that year	$Rating_i$
19	Number of agencies issuing a "buy" rating for the stock during that year	RB_i
20	Number of research reports published on the stock during that year	RRn_i
21	Indicator after dimension reduction of Raing, RB, RRn using principal component analysis	PCA_clusteri

The method for calculating the cumulative abnormal return of a stock relative to the market benchmark, CARi, t, is as follows:

First, calculate the cumulative return of the stock, $R_{i,t}$, and the market benchmark, $R_{m,t}$, over a specified interval:

$$R_t = \frac{P_{T+t} - P_T}{P_{T-1}} x100\% \quad (1)$$

Then, based on $R_{i,t}$ and $R_{m,t}$, calculate the cumulative abnormal return of the stock relative to the market benchmark, $CAR_{i,t}$:

$$CAR_{i,t} = R_{i,t} - R_{m,t} \quad (2)$$

The method for calculating the stock price volatility, VPi, t, is:

$$VP_{i,t} = \sqrt{\frac{P_{T+t} - P_T}{t}} \quad (3)$$

Additionally, this research utilized principal component analysis to reduce the three variables: Rating, RB, and RRn, into a single-dimensional variable, PCA_cluster.

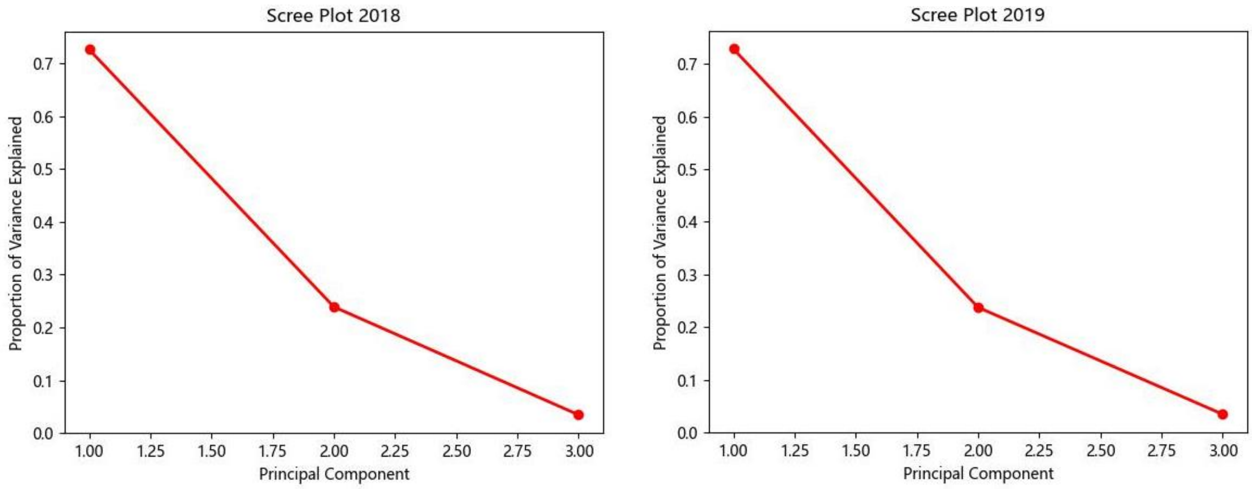

Figure 1. PCA Scree Plot

Photo credit: Original

The two scree plots above show the variance explained by each principal component after performing principal component analysis on Rating, RB, and RRn for the years 2018 and 2019. It's evident that the first principal component can explain around 70% of the variance, suggesting that a single component can effectively represent the impact of the three variables. Furthermore, as previously mentioned, the higher the value of PCA_cluster, the stronger the "consensus" degree of the stock being grouped by institutions.

4. Description of Empirical Design

The objective of the research is to investigate the impact of institutional investors' clustering behavior on stock prices. This paper has selected the cumulative abnormal return relative to the market benchmark (CAR) and the stock price volatility (VP) as dependent variables, reflecting the changes in asset prices. This part focus on examining the influence of three measures of the degree of institutional clustering (FN, FCN, IS), changes in institutional holding proportions (F_change), and the consensus degree indicator for clustered stocks (PCA_cluster) on the cumulative abnormal returns and stock price volatility. When considering the three measures of the clustering degree, this paper believes that due to biases in data collection and limitations in sample size, at least two variables need to be significant either individually or jointly to infer that the degree of institutional clustering has an impact on the cumulative abnormal returns or stock price volatility.

Through the establishment of two regression models, this research aims to separately study the influences on the abnormal returns (CAR) and stock price volatility (VP) of "clustered" stocks to determine which explanatory variables have significant effects.

$$\begin{aligned} CAR_{i,t} = \beta_0 + \beta_1 AV_{i,t} + \beta_2 TR_{i,t} + \beta_3 A_{i,t} + \beta_4 SL_{i,t} + \beta_5 DL_{i,t} + \beta_6 ROE_{i,t} + \beta_7 PE_{i,t} \\ + \beta_8 DTA_{i,t} + \beta_9 LMV_{i,t} + \beta_{10} Holder\_T10_{i,t} + \beta_{11} I_i + \beta_{12} FN_i \\ + \beta_{13} FCN_{i,t} + \beta_{14} IS_{i,t} + \beta_{15} F\_change + \beta_{16} PCA\_cluster + \varepsilon_{i,t} \end{aligned} \tag{4}$$

$$\begin{aligned} VP_{i,t} = \beta_0 + \beta_1 AV_{i,t} + \beta_2 TR_{i,t} + \beta_3 A_{i,t} + \beta_4 SL_{i,t} + \beta_5 DL_{i,t} + \beta_6 ROE_{i,t} + \beta_7 PE_{i,t} \\ + \beta_8 DTA_{i,t} + \beta_9 LMV_{i,t} + \beta_{10} Holder\_T10_{i,t} + \beta_{11} I_i + \beta_{12} FN_i \\ + \beta_{13} FCN_{i,t} + \beta_{14} IS_{i,t} + \beta_{15} F\_change + \beta_{16} PCA\_cluster + \varepsilon_{i,t} \end{aligned} \tag{5}$$

Here, 'i' denotes individual stocks, 't' represents time (t=5,30,120,240),$\beta_0$is the intercept, β_1 to β_{17} are the coefficients of the independent variables and ϵ_{it} is the error term.

5. Model Analysis

Based on the empirical approach presented in section 4, regression equations were constructed to study the impact of the five metrics concerning clustering (FN, FCN, ISt, F_change, PCA_cluster) that measure the clustering degree on the CAR and VP of "clustering-held" stocks.

Below are the regression results for CAR at t=5, 30, 120, and 240 for the years 2018 and 2019:

Table 2. Regression for CAR in 2018

Year: 2018	Regression for CAR Coef (P-Value)			
Variable	t=5	t=30	t=120	t=240
AVt	0 (0.522)	0 (0.176)	0 (0.243)	0 (0.397)
TRt	-0.17 (0.007)	-0.066 (0.01)	0.029 (0.016)	-0.024 (0.005)
At	0.641 (0)	0.809 (0)	-0.721 (0)	0.764 (0)
SLt	2.272 (0.199)	3.905 (0.031)	4.989 (0)	0.938 (0.377)
DLt	-16.04 (0)	-21.935 (0)	-1.115 (0.306)	-6.466 (0)
ROEt	0.244 (0.016)	0.443 (0)	0.563 (0)	0.457 (0)
Pet	0.002 (0.472)	-0.005 (0.337)	0.004 (0.767)	0.021 (0.622)
DTAt	-0.005 (0.762)	-0.073 (0.016)	-0.134 (0.002)	0.011 (0.856)
MVt	-0.505 (0.351)	4.307 (0)	4.874 (0.002)	14.382 (0)
Holder_T10_t	0.045 (0.086)	-0.049 (0.331)	-0.061 (0.457)	-0.438 (0)
I_dummy	-0.014 (0.865)	0.053 (0.732)	0.431 (0.069)	0.047 (0.887)
FN	-0.071 (0.101)	0.096 (0.255)	0.168 (0.194)	0.526 (0.003)
FCN	0.102 (0.368)	-0.458 (0.036)	-0.808 (0.016)	-2.007 (0)
ISt	-0.03 (0.1)	-0.002 (0.945)	0.11 (0.062)	0.069 (0.408)
F_change	-0.025 (0.727)	0.076 (0.583)	0.339 (0.106)	0.668 (0.02)
PCA_cluster	0.136 (0.51)	0.402 (0.328)	1.488 (0.018)	-0.687 (0.435)
Constant	8.643 (0.482)	-107.543 (0)	-86.115 (0.018)	-352.449 (0)

It was found that at t=5, all the metrics concerning clustering were not significant. A joint significance test showed a p-value of 0.1429, indicating joint insignificance. Also, at t=30, only the FCN metric was significant. This might be due to data biases, suggesting that in the short term, the degree of institutional clustering, the change in institutional holdings ratio, and the "consensus" level of clustering-held stocks have no impact on the stock's excess returns. However, at t=120 and 240, the significance of institutional clustering was stronger, indicating an influence on excess returns in the medium to long term.

The coefficient in front of FN was positive, indicating that in bear markets, the more funds cluster holding a stock, the higher the stock's excess return. This might be due to clustering by funds enhancing stock price stability, reducing liquidity, among other reasons.

The coefficient in front of FCN was negative, suggesting that the more fund companies hold a stock, the lower the excess return. This could be because when multiple fund companies hold the same stock, the market consensus for that stock might be too high, potentially leading to stock price bubbles. Moreover, increased participation by more fund companies could lead to increased stock volatility (as later research shows, an increase in FCN has a positive effect on stock volatility), increasing market uncertainty, eroding investor confidence, and subsequently reducing excess returns.

The coefficient for ISt was only significant at t-120, possibly due to multicollinearity.

F_change (change in institutional holdings): The coefficient for F_change was not significant at t=5, 30, and 120. However, at t=240, it was significant and positive, suggesting that in the long run, an increase in institutional holdings might yield higher excess returns. This could be because, during bear markets, fund companies tend to hold onto stocks they believe have potential, anticipating returns when market conditions improve.

PCA_cluster (degree of "consensus" for the stock): The coefficient for PCA_cluster was not significant at t=5, 30, and 240. However, at t=120, it was significant and positive, suggesting that stocks with higher "consensus" levels might yield higher excess returns in the medium term. This could be because, in bear markets, investors might prioritize safety, hence stocks with higher consensus levels could be more appealing. At t=240, the significance of PCA_cluster dwindled, possibly because in the long run, the market might have a full understanding of the value of these "consensus" stocks, thus having a minimal impact on their excess returns.

Table 3. Regression for CAR in 2019

05	Regression for CAR Coef (P-Value)			
Variable	t=5	t=30	t=120	t=240
AVt	0 (0.724)	0 (0.829)	0 (0.837)	0 (0.76)
TRt	-0.04 (0.382)	-0.021 (0.237)	-0.032 (0)	-0.004 (0.607)
At	-0.41 (0)	-0.61 (0)	0.95 (0)	0.863 (0)
SLt	14.848 (0)	9.002 (0)	-1.08 (0.253)	0.615 (0.469)
DLt	-7.096 (0.006)	-3.716 (0.004)	-5.448 (0)	-9.975 (0)
ROEt	0.383 (0.001)	0.316 (0.003)	0.517 (0)	0.58 (0)
Pet	0.013 (0.035)	0.033 (0.002)	0.035 (0.011)	-0.001 (0.91)
DTAt	-0.063 (0)	-0.101 (0)	-0.111 (0.003)	-0.178 (0.004)
MVt	-0.505 (0.297)	0.77 (0.337)	0.651 (0.612)	-6.802 (0.001)
Holder_T10_t	0.046 (0.063)	-0.003 (0.94)	0.016 (0.821)	0.039 (0.73)
I_dummy	0.053 (0.42)	0.133 (0.228)	0.506 (0.004)	0.761 (0.007)
FN	-0.006 (0.842)	-0.003 (0.941)	-0.005 (0.946)	-0.141 (0.028)
FCN	0.073 (0.352)	-0.049 (0.706)	0.165 (0.426)	0.832 (0.011)
ISt	-0.033 (0.066)	0.013 (0.696)	-0.017 (0.726)	0.033 (0.698)
F_change	-0.071 (0.352)	-0.122 (0.334)	0.28 (0.171)	-0.259 (0.441)
PCA_cluster	0.284 (0.19)	0.412 (0.268)	0.516 (0.385)	1.467 (0.12)
Constant	18.794 (0.086)	-1.356 (0.94)	-41.76 (0.143)	126.434 (0.005)

It can be observed that at t=5, only the ISt metric is significant. At t=30 and t=120, none of the five metrics under study are significant. Joint significance tests yielded p-values of 0.7648 and 0.2499 respectively, indicating that in bull markets, in the short to medium term, the degree of institutional clustering, changes in institutional holdings, and the "consensus" level of clustering-held stocks do not have a significant impact on stock excess returns.

At t=240, the coefficients in front of FN and FCN are significant and have opposite signs compared to the bear market. This indicates that FN and FCN have different impacts on stock excess returns in different market conditions. The coefficient of FN is negative, suggesting that in bull markets, the more funds that hold a stock, the lower its excess returns. This might be because 2019 was a bull market, and most stocks selected by funds performed well. Hence, the number of funds holding a stock couldn't significantly differentiate its performance.

The coefficient for FCN is positive and significant, indicating that in the long run (240 days), the more fund companies that cluster hold a stock, the higher its excess returns. The coefficient for ISt is not significant across all time frames, suggesting that in 2019, the proportion of institutional holdings did not significantly impact stock excess returns, possibly due to multicollinearity.

The coefficient for F_change is not significant across all timeframes, indicating that in 2019, the change in the number of funds holding the stock did not have a significant impact on its excess returns. This might be because in a bull market environment, investors' reactions to individual fund's buying and selling actions for a particular stock are attenuated, with a greater focus on the overall market trend.

The coefficient for PCA_flock is not significant across all timeframes, suggesting that in 2019, the "consensus" level of a stock did not significantly impact its excess returns. This could be attributed to the fact that in a bull market environment, stock prices are more influenced by overall market factors rather than specific circumstances related to individual stocks.

Below are the regression results for VP at t=5, 30, 120, and 240 for the years 2018 and 2019:

Table 4. Regression for VP in 2018

Year: 2018	Regression for VP Coef (P-Value)			
Variable	t=5	t=30	t=120	t=240
AVt	0 (0.911)	0 (0.918)	0 (0.78)	0 (0.339)
TRt	0.037 (0)	0.009 (0)	0.002 (0)	0.001 (0)
At	0.104 (0)	0.034 (0)	0.015 (0)	0.005 (0)
SLt	1.061 (0)	0.426 (0)	0.174 (0)	0.067 (0)
DLt	1.524 (0)	0.287 (0)	0.089 (0)	0.12 (0)
ROEt	0.055 (0)	0.009 (0.049)	0.003 (0.328)	0.001 (0.444)
Pet	0 (0.356)	0 (0.073)	0.001 (0.034)	0.002 (0)
DTAt	-0.002 (0.424)	-0.004 (0.003)	-0.001 (0.098)	-0.001 (0.057)
MVt	-0.118 (0.139)	-0.204 (0)	-0.111 (0.001)	-0.206 (0)
Holder_T10_t	-0.01 (0.011)	0.004 (0.034)	0.003 (0.096)	0.005 (0.001)
I_dummy	0.008 (0.479)	0.007 (0.249)	-0.005 (0.329)	-0.006 (0.136)
FN	-0.002 (0.712)	-0.005 (0.177)	-0.002 (0.539)	-0.006 (0.004)
FCN	0.008 (0.623)	0.02 (0.022)	0.006 (0.361)	0.023 (0)
ISt	0.005 (0.069)	0 (0.705)	0.002 (0.134)	0.001 (0.327)
F_change	-0.002 (0.878)	-0.004 (0.485)	0.006 (0.178)	0.006 (0.096)
PCA_flock	0.049 (0.104)	0.063 (0)	0.108 (0)	0.098 (0)
Constant	4.173 (0.021)	5.805 (0)	4.005 (0)	6.382 (0)

At t=5 and t=30, only one out of the three metrics measuring institutional clustering shows significance. For t=120, a combined significance test yields a p-value of 0.2860, suggesting institutional clustering doesn't notably influence stock volatility in the short and medium term.

The negative coefficient for FN suggests that, over longer time frames, there's an inverse relationship between stock volatility and the number of funds holding it. Higher fund ownership might stabilize the stock, reducing volatility.

Conversely, the positive coefficient for FCN indicates that as more fund companies hold a stock, its price volatility might increase. Multiple fund companies holding a stock might amplify market attention, hence increasing its volatility.

The metric of ISt is significant only at t=5, possibly due to multicollinearity.

For F_change, its coefficient becomes significant at t=240. Over the long term, the more a fund increases its holdings of a stock, the higher the stock's volatility. This can be attributed to funds being perceived as informed traders. Thus, their buying might be seen as positive, leading to increased volatility.

For PCA_cluster, its coefficient is significant and positive at t=30, 120, and 240. This suggests that, over time, a higher "consensus" level on a stock correlates with greater stock volatility. When there's a strong market consensus about a stock, reactions to its news might be more pronounced, causing significant price fluctuations.

Table 5. Regression for VP in 2019

Year: 2019	Regression for VP Coef (P-Value)			
Variable	t=5	t=30	t=120	t=240
AVt	0 (0.519)	0 (0.142)	0 (0.542)	0 (0.923)
TRt	0.021 (0.004)	0.001 (0.148)	0.001 (0)	0.001 (0)
At	0.08 (0)	0.049 (0)	0.008 (0)	0.003 (0)
SLt	2.326 (0)	0.392 (0)	0.097 (0)	0.061 (0)
DLt	0.614 (0.127)	0.282 (0)	0.289 (0)	0.123 (0)
ROEt	0.005 (0.798)	0.024 (0)	0.001 (0.7)	-0.001 (0.713)
Pet	0.002 (0.051)	0.002 (0)	0.001 (0.052)	0 (0.18)
DTAt	-0.002 (0.379)	0 (0.99)	-0.001 (0.122)	-0.002 (0.006)
MVt	-0.041 (0.589)	-0.134 (0.001)	-0.121 (0)	-0.083 (0.003)
Holder_T10_t	-0.007 (0.067)	-0.001 (0.552)	-0.001 (0.695)	-0.002 (0.288)
I_dummy	-0.01 (0.353)	-0.018 (0.001)	-0.013 (0.003)	-0.014 (0)
FN	0.001 (0.897)	0 (0.869)	-0.001 (0.743)	-0.001 (0.443)
FCN	0.004 (0.767)	0.002 (0.787)	0.004 (0.436)	0.004 (0.441)
ISt	0.005 (0.057)	0.002 (0.306)	0 (0.949)	0 (0.881)
F_change	0.021 (0.081)	0.007 (0.299)	0.007 (0.146)	0.01 (0.031)
PCA_cluster	0.057 (0.093)	0.071 (0)	0.069 (0)	0.078 (0)
Constant	2.387 (0.166)	4.473 (0)	4.665 (0)	4.116 (0)

It can be observed that at t=5, only the coefficient of ISt is significant among the three variables measuring institutional clustering level, and at t=30, 120, and 240, none of the three variables are significant in measuring institutional clustering level. Therefore, it is necessary to conduct a joint significance test, with p-values of 0.6274, 0.7582, and 0.8836, respectively. It can be seen that in a bull market, regardless of the time span, institutional clustering behavior does not have a significant impact on stock price volatility.

As for F_change, at t=30 and t=120, its coefficient is not significant, indicating that in the bull market environment of 2019, the short-term and medium-term changes in fund buying and selling behavior do not have a significant impact on stock price volatility (at t=5, the p-value is close to 0.1, so it is assumed that there may be data bias here, and there should be no impact in the short term). However, at t=240, its coefficient is significantly different from 0, indicating that in the long term (240 days), changes in fund buying and selling behavior for a certain stock are positively correlated with stock price volatility, meaning that the greater the changes in fund buying and selling behavior for a certain stock, the greater the potential for stock price volatility.

As for PCA_cluster, at t=5, 30, 120, and 240, its coefficient is significant and greater than 0. This result is consistent with the impact of PCA_flock on stock price volatility during a bear market.

6. Conclusion

In summary, this research reveals some core investment behaviors of institutional investors in China's A-share market through empirical analysis. Institutional investors' clustering behavior has different impacts on the excess returns of company stocks under different time scales and market conditions. However, the impact on stock price volatility is only different across time scales.

Research on excess returns indicates that whether in a bear market or a bull market, the short-term clustering behavior of institutions does not have a significant impact on excess returns. However, in the medium and long term, clustering behavior has a significant effect on stock excess returns, and the impact differs in different market environments.

Specifically, in a bear market, the more funds that cluster and hold a stock, and the fewer the number of fund companies, the higher the excess returns from the stock. In a bull market, the more funds that cluster and hold a stock, and the fewer the number of fund companies, the lower the excess returns. In these two different market conditions, FN and FCN's long-term impact on excess returns will be significantly different.

In a bear market, if institutions increase their holdings of a stock in the long run, the stock's excess return rate will rise. However, changes in institutional shareholdings in a bull market do not significantly affect excess returns.

In a bear market, the impact of the "consensus" level on stocks varies with time scale. In the medium term (120 days), its coefficient is significant and positive, suggesting that high consensus stocks may achieve higher excess returns in the medium term. This could be due to investors seeking safety during a bear market, leading to high consensus stocks gaining acceptance, demonstrating stability and price stickiness in a declining market, and thus achieving higher excess returns. In the short and long term, this coefficient is not significant. The short term may see market-wide declines and spreading panic, causing these "consensus" stocks to fall as well. But in the long term, the coefficient is negative, possibly due to the market transitioning from a bear market to a rebound or bull market, causing stocks that had previously fallen significantly to rebound, suppressing the relative excess returns of high consensus stocks. In a bull market, stock prices are more influenced by overall market factors than by the specifics of an individual stock.

Research on stock volatility shows that in a bear market, short-term clustering by institutions does not significantly affect stock price volatility. However, in the medium to long term, the more funds that hold a stock, and the fewer the number of fund companies, the higher the stock price volatility. It's also found that in the long term, institutions increasing their stock holdings will cause an increase in stock price volatility, and on all time scales, the stronger the "consensus" on a stock, the higher its price volatility. In a bull market, the degree of institutional clustering doesn't significantly impact volatility, but institutions increasing their stock holdings and the strengthening "consensus" on a stock still lead to higher stock price volatility. The positive correlation between the "consensus" level and stock price volatility could be because widespread market consensus makes the market react more sensitively to related company information, triggering violent stock price fluctuations. The impact of the studied variables on stock price volatility does not show significant differences across different market conditions.

References

[1] Fang C, Zhou S. Research on the impact of clustered investment by securities investment funds on the market [J]. Northern Economy and Trade, 2023, (02): 85 - 88.

[2] Su Z, He X, Zhang Y. Funds "clustering": net value growth and risk of sharp decline [J]. Management Review: 1 - 15.

[3] Li S. Fund clustering leads to a stampede causing a serious incident [J]. Stock Market Trend Analysis Weekly, 2021 (05): 17.

[4] Lakonishok J, Shleifer A, Vishny R W. The impact of institutional trading on stock prices [J]. Journal of financial economics, 1992, 32 (1): 23 - 43.

[5] Spyrou S. Herding in financial markets: a review of the literature [J]. Review of Behavioral Finance, 2013, 5 (2): 175 - 194.

[6] Xiang C, Feng L. Is fund clustering value investing or blind trend-following? — Research based on the impact on individual stock returns [J]. Review of Investment Studies, 2022, 41 (12): 67 - 90.

[7] Sun Y. The causes of fund conglomeration and its impact on stock price volatility [D]. Southwestern University of Finance and Economics, 2022

[8] Qu C. Research on the clustering behavior of institutional investors in China [J]. New Economy, 2022, (04): 83 - 90.

[9] Jiang J. Empirical research on the influence of institutional investors' holding group behavior on stock price[D]. Zhejiang University, 2022.

[10] Jiang S, Qian Y. Research on the impact of institutional investor clustering in the fund network on the stock market [J]. Finance and Economy, 2021, (10): 82 - 90.