

Research on Stock Quantification Strategy of Fund Flow Factor Based on Data Analysis

Lingwei Zhang^a, Xiaolei Ding^{b*}, Biyuan Yang^c

^ae-mail: 1195494097@qq.com, * Corresponding author: ^bklausding@zufe.edu.cn

^ce-mail: 2495636430@qq.com

Zhejiang University of Finance & Economics, Hangzhou, China

Abstract: At present, the portraits of the participants in the China stock market present relatively significant differences, which makes the capital market involves different types of capital inflows and outflows, thus causing fluctuations in the stock price. In recent years, the gradual rise of quantitative trading has promoted the continuous development and maturity of quantitative trading strategies. Within this context, how to combine the fund flow factor with the quantitative strategy to realize the accurate prediction of a future stock price increase and decrease further evolved into a difficult problem to that numerous scholars have devoted themselves. Based on Python, this paper firstly uses the median absolute deviation method to carry out data cleaning and effectiveness testing on various fund flow factors, applying the factors with higher effectiveness to the quantitative stock-selection strategy for stock screening. Furthermore, aiming at the investment portfolio formed under different factors, this paper applies the scientific calculation expansion libraries including NumPy, Pandas, and Matplotlib in Python to carry out backtest analysis to test the practical effect of different strategies, thereby exploring the feasibility and practicability of the stock quantitative strategy based on fund flow factor. Lastly, this paper proposes a host of pertinent recommendations on the application of the fund flow factor in stock quantitative strategy.

Keywords: Chinese Stock Market; Fund Flow; Quantitative Strategy; Median Absolute Deviation Method; Scientific Calculation Expansion Library; Backtest Analysis

1. Introduction

Under the background of the ongoing major global public safety events as well as the violent and turbulent economic environment, the global economy is inevitably facing increasing downward pressure. Despite this, the turnover of China's stock shows an upward trend. Moreover, with the change in people's awareness and way of investment and wealth management, the capital activity in China's securities market is also continuously improving. In different investment methods and strategies, a capital market composed of institutions, large, medium-sized, and retail investors has gradually formed. In essence, the current stock market in China is an emotional-oriented trading market, with changes in the inflow and outflow of funds largely reflecting the emotional fluctuations of the corresponding groups of traders.

At the present stage, some relevant studies regard the net increase of main funds as a bullish signal and the net inflow of retail funds as a bearish indicator. Additionally, some scholars found that the inflow and outflow of large funds can be used as one effective indicator of stock buying

and selling. The reason lies in that large fund flows in the China market typically represent some financial institutions and large-scale idle funds. Compared with retail investors, they are more likely to grasp the development trend and policy orientation of the current market and even control the short-term trend of the stock market. Nevertheless, current academic circles seldom research stock quantitative strategy from the perspective of comprehensively combining all fund flow factors, and have not put forward sufficient effective recommendations on the application of fund flow factors in stock quantitative strategy. Hence, this paper will carry out systematic research on the aforesaid issues, to improve the application system concerning fund flow factor in the field of stock quantification and to explore more effective stock quantitative strategies based on fund flow factor, thus providing a beneficial experience for further research by more scholars.

2. Data Processing and Factor Test

This paper takes the stocks of all A-share listed companies in China before 18 November 2022 as the research object, with the research interval of the sample ranging from 4 January 2007 to 18 November 2022. In this research interval of the sample, the China stock market has experienced repeated bull-bear market transitions, covering the entire trend of the stock market under different systemic risks. In addition, the data involved in this paper are from the Choice database of East Money Information, with the data analysis software as Python.

Specifically, this paper first judges the different types of fund flows and the impact of their proportions on the stock price increase based on the rising probability and frequency of the stock price on the 1st, 3rd, and 5th day in the future, thus determining the optimal proportion of different types of fund flows on the stock price increase. Furthermore, according to the optimal proportion of different types of funds, this paper constructs a quantitative stock-selection model based on single-factor fund flow, as well as the corresponding securities investment portfolio based on the ranking idea of market value from small to large and the stock-selection model of selecting three stocks every three days. Moreover, by comparing the yield curve of the strategy with the Shanghai and Shenzhen 300 Index, this paper further judges the feasibility of adopting a fund flow strategy for stock selection.

2.1 Selection of Variables

In the process of modeling, according to the number of entry orders, fund flows are further divided into four types for statistics, including institutional funds (> CNY 1 million), large funds (CNY 200,000-1 million), medium-sized funds (CNY 40,000-200,000), and retail funds (< CNY 40,000). Subsequently, this paper calculates the proportions of buying and selling of four types of funds, in turn, grouping them into 10 equal parts to observe their impact on the stock increase.

Table 1 Indicators in the Fund Flow Stock-selection Model

Classification of Fund Flow	Indicators Name
Institutional Funds	Purchases and Sales of Institutional Funds
Large Funds	Purchases and Sales of Large Funds
Medium-sized Funds	Purchases and Sales of Medium-sized Funds
Retail Funds	Purchases and Sales of Retail Funds

2.2 Standardized Processing of Data

Given the difference between the value range of indicators and the measurement unit, it is necessary to standardize the extracted data of each indicator before empirical analysis. In this connection, the data involved in this research are processed by the median absolute deviation (hereinafter referred to as "MAD") method, which first determines the Median of all factors, and calculates the absolute deviation value of each factor from the median, thus obtaining the absolute deviation value, namely, the $X - \text{Median}$. Further, MAD is calculated based on the absolute deviation value. Lastly, the range of $[\text{median} - n \times \text{MAD}, \text{median} + n \times \text{MAD}]$ can be obtained by setting the parameter $n=3 \times 1.4826$. The following adjustments are made to the factor values beyond the reasonable range:

$$U_A = \begin{cases} X_{\text{median}} + n\text{MAD} & \text{if } X_i > X_{\text{median}} + n\text{MAD} \\ X_{\text{median}} & \text{if } X_i < X_{\text{median}} - n\text{MAD} \\ X_i & \text{if } X_{\text{median}} - n\text{MAD} < X_i < X_{\text{median}} + n\text{MAD} \end{cases}$$

2.3 Factor Feasibility Test

Prior to modeling, Python is used to perform a factor test on the standardized capital flow data of all A-share listed companies from 2007 to 2022. Meanwhile, through the statistics of the forecast effect of each variable on the stock price increase and the probability of the stock price increase under different percentages, this research further judges whether to do further relevant factor analysis on them. The relevant results are as follows:

Proportion of Institutional Purchases	Frequency of occurrence	1st Day			3rd Day			5th Day		
		Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise
>0.1	1426137	690241	48.40%	0.20%	688769	48.30%	0.29%	683638	47.94%	0.31%
>0.2	409063	205150	50.15%	0.52%	202067	49.40%	0.80%	199103	48.67%	0.84%
>0.3	146758	79075	53.88%	1.09%	76145	51.88%	1.74%	74328	50.65%	1.90%
>0.4	61273	36511	59.59%	1.95%	34243	55.89%	3.36%	32997	53.85%	3.83%
>0.5	28458	19168	67.38%	3.23%	17527	61.59%	5.91%	16658	58.54%	6.96%
>0.6	15305	11602	75.81%	4.70%	10471	68.42%	9.08%	9813	64.12%	11.07%
>0.7	9744	8028	82.39%	5.86%	7220	74.10%	11.80%	6725	69.02%	14.62%
>0.8	6614	5772	87.27%	6.67%	5132	77.59%	13.67%	4795	72.50%	16.98%
>0.9	4129	3695	89.49%	7.14%	3287	79.61%	14.75%	3065	74.23%	18.48%

Figure 1 Proportion of Institutional Purchases

Proportion of Institutional Sales	Frequency of occurrence	1st Day			3rd Day			5th Day		
		Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise
>0.1	1874307	910014	48.55%	0.00%	892871	47.64%	-0.03%	892887	47.64%	-0.05%
>0.2	508350	242340	47.67%	-0.04%	238535	46.92%	-0.09%	239192	47.05%	-0.14%
>0.3	150910	70140	46.48%	-0.13%	69931	46.34%	-0.20%	70141	46.48%	-0.30%
>0.4	49522	22110	44.65%	-0.32%	22209	44.85%	-0.53%	22479	45.39%	-0.71%
>0.5	17522	7329	41.83%	-0.81%	7277	41.53%	-1.46%	7459	42.57%	-1.95%
>0.6	7050	2522	35.77%	-1.93%	2495	35.39%	-3.58%	2541	36.04%	-4.64%
>0.7	3506	953	27.18%	-3.33%	959	27.35%	-5.34%	1009	28.78%	-7.01%
>0.8	2014	394	19.56%	-4.64%	393	19.51%	-5.07%	435	21.60%	-6.11%
>0.9	1158	164	14.16%	-5.71%	157	13.56%	-7.76%	191	16.49%	-8.52%

Figure 2 Proportion of Institutional Sales

Proportion of Purchases from Large Investors	Frequency of occurrence	1st Day			3rd Day			5th Day		
		Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise
>0.1	8365149	4102390	49.04%	0.04%	4078614	48.76%	0.08%	4081093	48.79%	0.12%
>0.2	4851447	2358904	48.62%	0.07%	2358701	48.62%	0.10%	2352889	48.50%	0.13%
>0.3	1384739	683022	49.32%	0.18%	691908	49.97%	0.32%	690315	49.85%	0.40%
>0.4	173583	88145	50.78%	0.32%	89673	51.66%	0.70%	89925	51.81%	0.95%
>0.5	17600	9119	51.81%	0.52%	9265	52.64%	1.38%	9318	52.94%	2.10%
>0.6	3710	2079	56.04%	1.39%	2069	55.77%	4.05%	2182	58.81%	6.60%
>0.7	1730	1068	61.73%	2.44%	1103	63.76%	7.45%	1196	69.13%	12.51%
>0.8	1083	727	67.13%	3.19%	765	70.64%	10.10%	849	78.39%	16.94%
>0.9	778	520	66.84%	3.31%	553	71.08%	10.65%	619	79.56%	18.05%

Figure 3 Proportion of Purchases from Large Investors

Proportion of Sales from Large Investors	Frequency of occurrence	1st Day			3rd Day			5th Day		
		Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise
>0.1	8873190	4376397	49.32%	0.04%	4339424	48.90%	0.08%	4346769	48.99%	0.12%
>0.2	5992392	2943094	49.11%	0.03%	2902123	48.43%	0.03%	2903219	48.45%	0.05%
>0.3	2072970	1020359	49.22%	0.02%	1005280	48.49%	0.01%	1009230	48.69%	0.03%
>0.4	275604	134855	48.93%	0.00%	135550	49.18%	0.05%	137054	49.73%	0.12%
>0.5	26686	12445	46.63%	-0.16%	12937	48.48%	-0.18%	13268	49.72%	-0.16%
>0.6	4401	1703	38.70%	-0.89%	1862	42.31%	-1.70%	2048	46.53%	-2.00%
>0.7	1806	588	32.56%	-1.61%	673	37.26%	-3.42%	826	45.74%	-4.10%
>0.8	1028	301	29.28%	-2.00%	383	37.26%	-4.11%	494	48.05%	-5.04%
>0.9	721	219	30.37%	-2.04%	280	38.83%	-4.35%	365	50.62%	-5.07%

Figure 4 Proportion of Sales from Large Investors

Proportion of Purchases from Medium-sized Investors	Frequency of occurrence	1st Day			3rd Day			5th Day		
		Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise
>0.1	9853318	4881939	49.55%	0.05%	4853475	49.26%	0.12%	4868688	49.41%	0.19%
>0.2	9735421	4819762	49.51%	0.04%	4792741	49.23%	0.11%	4808039	49.39%	0.17%
>0.3	8427169	4176273	49.56%	0.03%	4149407	49.24%	0.07%	4163248	49.40%	0.14%
>0.4	3045790	1537481	50.48%	0.07%	1533258	50.34%	0.16%	1538323	50.51%	0.25%
>0.5	265329	137028	51.64%	0.19%	139149	52.44%	0.42%	139929	52.74%	0.60%
>0.6	22702	11022	48.55%	0.13%	11207	49.37%	0.33%	11438	50.38%	0.51%
>0.7	6332	2989	47.20%	0.11%	3012	47.57%	0.36%	3203	50.58%	0.87%
>0.8	1915	890	46.48%	0.31%	1016	53.05%	1.71%	1250	65.27%	3.91%
>0.9	1332	620	46.55%	0.40%	749	56.23%	2.28%	980	73.57%	5.16%

Figure 5 Proportion of Purchases from Medium-sized Investors

Proportion of Sales from Medium-sized Investors	Frequency of occurrence	1st Day			3rd Day			5th Day		
		Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise
>0.1	9860356	4889426	49.59%	0.06%	4860059	49.29%	0.14%	4874677	49.44%	0.21%
>0.2	9752946	4836874	49.59%	0.06%	4807140	49.29%	0.13%	4821181	49.43%	0.20%
>0.3	8451170	4200763	49.71%	0.06%	4173399	49.38%	0.12%	4184759	49.52%	0.18%
>0.4	3344821	1678951	50.20%	0.05%	1667703	49.86%	0.11%	1678146	50.17%	0.18%
>0.5	341300	170659	50.00%	0.05%	171061	50.12%	0.14%	172847	50.64%	0.21%
>0.6	25892	11957	46.18%	-0.02%	12258	47.34%	0.02%	12572	48.56%	0.03%
>0.7	6058	2644	43.64%	-0.21%	2735	45.15%	-0.40%	2966	48.96%	-0.39%
>0.8	1902	787	41.38%	-0.48%	916	48.16%	-0.67%	1142	60.04%	-0.33%
>0.9	1323	538	40.67%	-0.57%	671	50.72%	-0.58%	902	68.18%	0.38%

Figure 6 Proportion of Sales from Medium-sized Investors

Proportion of Purchases from Retail Investors	Frequency of occurrence	1st Day			3rd Day			5th Day		
		Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise
>0.1	9787231	4843928	49.49%	0.04%	4816878	49.22%	0.10%	4833090	49.38%	0.17%
>0.2	9143625	4520795	49.44%	0.02%	4493058	49.14%	0.07%	4512743	49.35%	0.14%
>0.3	7082797	3525611	49.78%	0.01%	3492273	49.31%	0.07%	3516182	49.64%	0.16%
>0.4	4027621	2041136	50.68%	0.03%	2016437	50.07%	0.12%	2040545	50.66%	0.27%
>0.5	1564840	810054	51.77%	0.07%	803508	51.35%	0.23%	818258	52.29%	0.44%
>0.6	418457	219633	52.49%	0.09%	219802	52.53%	0.30%	225388	53.86%	0.60%
>0.7	81252	42729	52.59%	0.07%	42852	52.74%	0.25%	44410	54.66%	0.62%
>0.8	11641	5837	50.14%	0.15%	6011	51.64%	0.18%	6530	56.09%	0.39%
>0.9	2502	1046	41.81%	0.89%	1187	47.44%	1.66%	1466	58.59%	1.45%

Figure 7 Proportion of Purchases from Retail Investors

Proportion of Sales from Retail Investors	Frequency of occurrence	1st Day			3rd Day			5th Day		
		Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise	Rising frequency	Rising Probability	Average Rise
>0.1	9780415	4853879	49.63%	0.07%	4824273	49.33%	0.15%	4838690	49.47%	0.23%
>0.2	8775881	4362686	49.71%	0.08%	4343361	49.49%	0.17%	4356050	49.64%	0.26%
>0.3	5957662	2979945	50.02%	0.10%	2985447	50.11%	0.25%	2999842	50.35%	0.39%
>0.4	2795795	1418117	50.72%	0.15%	1434478	51.31%	0.41%	1446818	51.75%	0.65%
>0.5	952657	492463	51.69%	0.24%	501325	52.62%	0.69%	507552	53.28%	1.10%
>0.6	256139	136807	53.11%	0.44%	138648	54.13%	1.27%	140837	54.98%	2.09%
>0.7	52856	29574	56.05%	0.88%	29553	55.91%	2.51%	30158	57.06%	4.14%
>0.8	9935	5992	60.31%	1.94%	5983	60.22%	5.52%	6334	63.75%	9.12%
>0.9	3345	2150	64.28%	3.21%	2247	67.17%	9.60%	2471	73.87%	15.98%

Figure 8 Proportion of Sales from Retail Investors

Based on the analysis of the results of the above-mentioned backtest, the proportions of the purchases from institutions, large, and medium-sized investors as well as the proportion of sales from retail investors have a good ability to forecast the stock price increase. Among them, if the purchases from large fund flow account for more than 0.6%, the stock market will have a 67.13% chance of rising the next day. This indicator performs best in the forecast effect of all types of fund flows, and it will still maintain a high probability of an increase in the next 3 and 5 days, with the average rise and fall of stocks increasing with the change of time. In contrast, the proportion of the purchases of medium-sized investors is more than 45% which can successfully predict the next day's rise of the stocks and has a positive correlation with future rise and fall. Regarding the proportion of institutional purchases and retail sales, the probability of a stock rising increases with the increase of their proportions.

3. Construction and Test of Different Fund Flow Models

Based on the analysis outlined above, this paper finds that the greater the proportion of factors, the greater the probability of future stock price increase. Nevertheless, it is worth noting that the frequency of such cases has decreased accordingly. From the perspective of a quantitative investment strategy of short-term operation, it is necessary to maintain more investment targets. Therefore, in the interest of ensuring that the investment portfolio can cover more targets, this paper makes a gradient test on the above factors, and further finds that when the factor value > 0.5, the investment portfolio contains more targets. Furthermore, when constructing the single-factor quantitative model, this paper takes the stock yield rate as the dependent variable, with the proportion of institutional purchases > 0.5, the proportion of purchases from large investors > 0.5, the proportion of purchases from medium-sized investors > 0.5 and the proportion of sales from retail investors > 0.5 as the independent variables. On the same note,

this paper selects three stocks and holds them for three days according to the order of market value from big to small, intending to construct the single-factor stock-selection model. On the other hand, aiming at the evaluation of the model, this paper analyzes whether the stock-selection model with different independent variables can perform well by counting a series of indicators from 2007 to 2022, such as yield rate and maximum drawdown, and comparing with the returns of Shanghai and Shenzhen 300 Index, testing its investment effect. The relevant results are as follows:

Table 2 Investment Performance of Institutional Fund Flow Proportion Strategy

Transaction Date	Strategy Yield	Basic Yield	Excess Yield
2007/12/31	47.43%	162.98%	-115.55%
2008/12/31	-42.35%	-65.84%	23.50%
2009/12/31	-2.77%	94.11%	-96.88%
2010/12/31	-58.60%	-13.90%	-44.70%
2011/12/31	-42.00%	-24.57%	-17.43%
2012/12/31	-53.22%	7.30%	-60.52%
2013/12/31	-30.52%	-7.28%	-23.24%
2014/12/31	-48.83%	50.36%	-99.20%
2015/12/31	42.21%	8.90%	33.31%
2016/12/31	-37.22%	-12.41%	-24.81%
2017/12/31	-34.32%	21.87%	-56.19%
2018/12/31	-82.12%	-25.59%	-56.53%
2019/12/31	-75.76%	36.49%	-112.25%
2020/12/31	-63.37%	25.29%	-88.65%
2021/12/31	-63.13%	-3.76%	-59.37%
2022/12/31	-62.57%	-22.41%	-40.16%

As can be seen from Table 2, the investment portfolio constructed by using the proportion of institutional purchases > 0.5 shows a poor yield effect, and simply successfully outperformed the Shanghai and Shenzhen 300 Index in two years. The backtest analysis of the 15-year data in the interval (excluding 2022) indicates that the annualized return and the maximum drawdown of the investment portfolio are both negative, with a maximum drawdown of 100%. It can be seen that the fund flow strategy based on the proportion of institutional purchases does not apply to the current A-share market. The possible reason lies in that institutions do not buy stocks and take a position at the lowest point of stocks. Moreover, its excessive purchases and slippage easily cause some institutions to be unable to buy related stocks thoroughly, which leads to institutions usually choosing to start buying stocks gradually while the stocks are still falling. However, the strategy of this paper only holds the selected stocks for three days, so it is incapable of fully capturing the rising opportunity of related stocks, which thus leads to the unsatisfactory investment effect of this strategy.

Table 3 Investment Performance of Large Fund Flow Proportion Strategy

Transaction Date	Strategy Yield	Basic Yield	Excess Yield
2007/12/31	79.30%	162.98%	-83.68%
2008/12/31	-12.28%	-65.84%	53.57%
2009/12/31	21.27%	94.11%	-72.84%
2010/12/31	123.88%	-13.90%	137.79%
2011/12/31	-5.21%	-24.57%	19.36%
2012/12/31	15.00%	7.30%	7.70%
2013/12/31	31.59%	-7.28%	38.87%
2014/12/31	43.37%	50.36%	-6.99%
2015/12/31	17.11%	8.90%	8.22%
2016/12/31	34.82%	-12.41%	47.24%
2017/12/31	19.18%	21.87%	-2.68%
2018/12/31	-33.48%	-25.59%	-7.89%
2019/12/31	44.40%	36.49%	7.91%
2020/12/31	19.49%	25.29%	-5.80%
2021/12/31	-20.88%	-3.76%	-17.12%
2022/12/31	-18.12%	-22.41%	4.29%

As can be seen from Table 3, the investment portfolio selected by using the proportion of purchases from large investors > 0.5 shows a relatively general yield effect. The yield of this strategy has not outperformed the Shanghai and Shenzhen 300 Index for 6 years. Moreover, the 15-year annualized yield in the interval was only 17.11%, whereas the maximum drawdown was -60.46%. Therefore, it is not suitable for investors with risk aversion preference. Meanwhile, the cumulative net value of the strategy is only 12.29, showing a relatively general effect.

Table 4 Investment Performance of the Proportion Strategy of Purchases from Medium-sized Funds

Transaction Date	Strategy Yield	Basic Yield	Excess Yield
2007/12/31	315.56%	162.98%	152.58%
2008/12/31	21.54%	-65.84%	87.39%
2009/12/31	252.03%	94.11%	157.92%
2010/12/31	154.34%	-13.90%	168.25%
2011/12/31	27.29%	-24.57%	51.86%
2012/12/31	43.26%	7.30%	35.96%
2013/12/31	88.32%	-7.28%	95.60%
2014/12/31	77.72%	50.36%	27.35%
2015/12/31	148.84%	8.90%	139.94%
2016/12/31	39.44%	-12.41%	51.85%
2017/12/31	-0.29%	21.87%	-22.15%

Transaction Date	Strategy Yield	Basic Yield	Excess Yield
2018/12/31	2.47%	-25.59%	28.06%
2019/12/31	64.20%	36.49%	27.71%
2020/12/31	15.68%	25.29%	-9.61%
2021/12/31	-3.52%	-3.76%	0.24%
2022/12/31	-23.20%	-22.41%	-0.79%

In contrast, the proportion strategy of the purchases from medium-sized investors has achieved quite excellent performance, with an average excess yield of 62.01% in the 15-year interval. In the meantime, its 15-year annualized yield in the interval was 57.64%, whereas the maximum drawdown was only -47.47%, both of which were significantly higher than the results of the backtest of the proportion factor of institutional purchases, as well as the proportion factor of purchases from large investors. Nevertheless, in recent years, the annualized yield of this strategy has shown a significant downward trend and has not yet outperformed the Shanghai and Shenzhen 300 Index.

Table 5 Investment Performance of the Proportion Strategy of Sales from Retail Funds

Transaction Date	Strategy Yield	Basic Yield	Excess Yield
2007/12/31	222.37%	162.98%	59.40%
2008/12/31	-34.86%	-65.84%	30.98%
2009/12/31	257.73%	94.11%	163.62%
2010/12/31	61.88%	-13.90%	75.78%
2011/12/31	4.84%	-24.57%	29.41%
2012/12/31	26.57%	7.30%	19.27%
2013/12/31	51.09%	-7.28%	58.37%
2014/12/31	45.03%	50.36%	-5.33%
2015/12/31	109.43%	8.90%	100.53%
2016/12/31	-0.92%	-12.41%	11.49%
2017/12/31	-20.70%	21.87%	-42.56%
2018/12/31	1.33%	-25.59%	26.92%
2019/12/31	60.41%	36.49%	23.93%
2020/12/31	16.99%	25.29%	-8.29%
2021/12/31	8.33%	-3.76%	12.09%
2022/12/31	-15.95%	-22.41%	6.46%

The proportion strategy of sales from retail investors outperformed the Shanghai and Shenzhen 300 Index, with an average 15-year excess yield of 35.13%, an annualized yield of 33.71%, and a maximum drawdown of -52.65%. Although the strategy had negative returns in some years of the interval, the drawdown was small and controllable. Therefore, the strategy presents a relatively excellent performance.

4. Conclusion

To sum up, by comparing the single-factor quantitative stock-selection model constructed year by year, this paper draws the following conclusions:

First of all, the short-term quantitative stock-selection strategy based on fund flow is feasible and can obtain a higher yield. Based on the inflow and outflow of different types of fund flows, it is possible to fully tap the stock targets that may rise in the next few days, thereby selecting the stocks with large short-term gains. The securities investment portfolio constructed year by year through the fund flow model of large, medium-sized, and retail investors can achieve higher profits in most market conditions, thus helping investors to obtain positive returns. It can be seen that the strategy based on different types of fund flows is feasible and effective in short-term stock trading in China.

Secondly, different market conditions will bring more significant impact and changes on the returns of quantitative investment strategies based on fund flows. The specific proportion of fund flows should be adjusted based on the changes in the current stock market. Meanwhile, it is necessary to increase the timing trading strategy, to ensure that the selected securities investment portfolio can obtain a more stable yield effect annually.

Thirdly, in the fund flow strategy, it is strongly recommended to focus on the proportion factor of the purchases from large and medium-sized investors as well as the proportion factor of the sales from retail investors, excluding the proportion factors of the institutional purchases and sales. This is because the first three factors could lead to a rise in stocks over the next three days. However, institutions, whether buying or selling, could cause the stock to fall in the next three days. This further indicates that most institutions have relatively large amounts of funds and relatively low operating frequency. In addition, the operation of institutions will have a greater impact on the stock market, which is mainly a negative impact. Consequently, in the development and improvement of a stock quantitative investment strategy based on fund flow, the role of institutional funds should be generally ignored.

Fourthly, the quantitative investment strategy based on different types of fund flows must reduce the proportion of each major fund to be bought and sold, to obtain more selection targets and give full play to the ranking ability of market value, thus selecting the securities investment portfolio with the potential advantage of short-term rise. Additionally, choosing excessively harsh conditions may result in an investment portfolio with fewer underlying stocks, thus making it difficult to grasp better investment opportunities.

References

- [1] S. M. Ahmed, "Quantification of investor emotion in financial news by analyzing the stock price reaction," 2017 International Conference on Information and Communication Technologies (ICICT), 2017, pp. 119-123.
- [2] J. Yang, "Stock Price Informativeness and Investment-Cash Flow Sensitivity," 2010 International Conference of Information Science and Management Engineering, 2010, pp. 510-514.
- [3] S.M. Fazzari R.G. Hubbard and B.C. Petersen "Investment-Cash Flow Sensitivities are Useful: A Comment on Kaplan and Zingales" The Quarterly Journal of Economics vol. 115 pp. 695-705 2000.

- [4] A. Ascioğlu, S. P. Hedge and J. B. McDermott "Information asymmetry and investment-cash flow sensitivity" *Journal of Banking& Finance* vol. 32 pp. 1036-1048 2008.
- [5] S. Agca and A. Mozumdar "The impact of capital market imperfections on investment-cash flow sensitivity" *Journal of Banking& Finance* vol. 32 pp. 207-216 2008.
- [6] H. Yang, C. Wang and Y. Zhao, "The Cross-section of Expected Stock Returns: Evidence from Chinese A-share Market," 2012 Fifth International Conference on Business Intelligence and Financial Engineering, 2012, pp. 303-307.
- [7] Y. Shi, "Analysis of Multi-factor Stock Market Choice Portfolio Model Based on Regression," 2021 5th Annual International Conference on Data Science and Business Analytics (ICDSBA), 2021, pp. 1-5.
- [8] C. M. Frans, P. A. Nigo and N. N. Qomariyah, "Stock Market Statistical Analysis: Investing Versus Trading Strategies," 2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE), 2022, pp. 33-38.
- [9] D. Fangfei, H. Ningning and L. Su, "Institutional Investors Heterogeneity, Future Earnings and the Informativeness of Stock Prices," 2013 Third International Conference on Intelligent System Design and Engineering Applications, 2013, pp. 887-890.
- [10] Isha, S. Dixit, M. K. Ahirwar, D. Sakethnath and M. Rakha, "Stock Prediction by Analyzing the Past Market Trend," 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2021, pp. 1-4.