Research on Stock Investment Strategy of Semiconductor Industry Based on PCA Modeling

Aoxue Mo*, Xia Lou, Weiwei Chen

Faculty of Economics, Hangzhou Normal University, Zhejiang, China, 311121

*Corresponding author: M.2023.02@outlook.com

Abstract. With the continuous innovation of information technology, in order to meet the needs of the development of high-precision industries, the global semiconductor industry's market demand is growing and expanding. In recent years, the western countries on China's semiconductor industry to chase and intercept, played to determine the future strategic dominance between the countries, the right to survive the industrial chain of the chip war. In order to take advantage of the opportunity to invest in well-developed semiconductor companies, this paper selects 134 stocks in the semiconductor industry as a stock pool from the Flush IFIND database, and adopts the PCA principal component analysis method, selects 15 fundamental indicators, and analyzes and sorts out the stocks from the five analytical indexes of profitability, development ability, profitability, solvency and operating ability, and conducts a comprehensive evaluation and analysis of stocks. Based on Markowitz theory, the optimal risk portfolio is finally determined. At the same time, this paper is based on the pyramid model for position management, combined with averaging trading strategy and high sell-low suck strategy real-time position adjustment, through 43 days of trading to observe the accuracy of this investment strategy.

Keywords: semiconductor industry, PCA model, financial statement analysis, value investment.

1. Introduction

With the deepening of the reform and opening-up process, China has become an important force driving the stable development of the world economy. However, out of the dual considerations of curbing the development of China's semiconductor industry and maintaining a huge share in China, the United States, in conjunction with Japan, the Netherlands and other countries to limit the export of advanced arithmetic chips and process equipment, etc., and the implementation of sanctions on related technologies. Consumption from the semiconductor field, China's consumption share far exceeds that of other countries, up to 40%, and the huge market demand deviates from is that most of the domestic semiconductor products self-sufficiency rate is less than 30%, and in the wafer manufacturing materials and other high-priority areas, the degree of dependence on imports is even greater. Based on this, the government has introduced a series of incentives, including tax breaks, subsidies, financial support and land concessions, and joint scientific research institutions and universities to promote the transformation of scientific and technological achievements. In order to invest in well-developed semiconductor companies through the stock market, Li Muyun (2023) based on the BL model of the multi-period portfolio optimization problem, through the inclusion of asset returns, risk costs, transaction costs, etc. into the performance indicators to describe, while taking into account the constraints in the trading process, the comprehensive formation of the optimization problem, and then use the rolling time-domain optimization way to solve to get the optimal trading strategy[1]. Ben Huang (2023) proposed to use K-Means algorithm to cluster and label the stock volatility, and the labeled stocks of each category are used as the stock pool of quantitative time-timing strategy, which is combined with the original quantitative time-timing strategy to construct the stock selection and time-timing portfolio investment strategy[2]. Min Zhu (2022) proposed to try to use LSTM neural network based on HS-IPS model to learn time series data such as stock historical trading data, decision-making process of information observers and momentum traders, investor's attention, and information dissemination speed, to predict the future short-term stock-related data, and to construct related momentum strategies[3]. Jixiang Sun (2022) constructed a multi-factor

investment trading system based on S-TabNet, the prototype system uses factor and A-share data as the basis, combines with S-TabNet to screen stocks and obtains a list of recommended stocks to buy and sell[4]. Lei Le (2022) used the entropy weight method to establish a system of characteristic indexes affecting the stock trend by considering multidimensional factors and explored their impact on the company's stock trend; according to the system of characteristic indexes, a stock selection model based on the CART decision tree was established to make a decision on stock selection[5] .

In this study, the paper cut through the semiconductor industry to identify stocks with potential for optimal portfolio construction. In terms of investment strategy, the paper form a stock pool with A-share listed companies in the semiconductor industry, and use PCA principal component analysis to select net profit margin of total assets (X11), net profit margin of fixed assets (X12), growth rate of income from main business (X13), capital preservation and appreciation rate (X21), growth rate of total assets (X22), growth rate of net profit (X23), net assets per share (X31), price-to-book ratio (X32), price-to-earnings ratio (X33), current ratio (X41), working capital-to-total-net-assets ratio (X42), shareholders' equity-to-fixed-assets ratio (X43), accounts-receivable turnover (X51), inventory turnover (X52), and total-assets turnover (X53), which are the 15 company fundamentals, are used to analyze and sort the stocks for comprehensive evaluation. Before data processing, the paper standardized the indicator data and constructed the correlation coefficient matrix to solve the covariance problem among the data. In the reliability test, the KMO value of this data is 0.645, and the approximate chi-square value in the Bartlett's test of sphericity is 846.844, which has good correlation and credibility. Then, the paper selected four feature roots according to the principle of feature root greater than or equal to 1. By identifying the contribution value of each principal component, we selected the top 4 stocks in terms of score ranking as investment objects. After screening the four stocks, the paper look for the tangent point between the portfolio line and the efficient frontier to construct the portfolio, manage the position based on the pyramid model, and adopt the mean trade strategy and the high sell-low suck strategy to keep the risk within the acceptable range, targeting the low cost and high spread. This paper constructs the optimal investment portfolio according to the investment strategy, and verifies the correctness of the strategy through empirical measurements, so as to provide strategic references for the investment in the semiconductor industry.

2. Model building and solving

2.1. Refers to the construction of the correlation coefficient matrix

After inputting the relevant index data of 134 semiconductor enterprises, the paper standardize them. Assuming that the value of the jth evaluation index of the ith evaluation object is Xij, the raw data can be arranged into the following matrix[6]:

$$\begin{bmatrix} X_{11} & \cdots & X_{13} \\ \vdots & \ddots & \vdots \\ X_{51} & \cdots & X_{53} \end{bmatrix} \tag{1}$$

The variables satisfy a certain correlation between them in order to carry out principal component analysis, therefore, after standardizing the data, the paper establish the covariance matrix R based on the standardized data matrix, and the matrix R is a statistical indicator reflecting the degree of correlation between the standardized data (Rudan)[7]. After the correlation coefficient test, this study found that the correlation coefficients of X13 and X23, X21 and X22, and X12 and X43 were large, 0.631, 0.926, and 0.803, respectively, and the correlation coefficients of X11 and X23, X21 and X31, X22 and X31, and X41 and X12 were close to 0.5, with a higher possibility of covariance.

2.2. KMO and Barlett’s test

In order to determine the suitability of the data for principal component analysis, KMO test and Barlett test were performed. According to Table 1, the KMO value was 0.645, which is greater than 0.5, indicating that there is some correlation between the indicators. In addition, the approximate chi-

square value obtained from the Barlett's test of sphericity was 846.844, with a p-value of less than 0.001, which is significant. Both test results indicate that the stock valuation model is suitable for principal component analysis and reflect the high credibility of the results of this this study.

Table 1 KMO and Bartlett's test

KMO Quantity of Sample Suitability		0.645
Bartlett's test of sphericity	Approximate cardinality (F)	846.844
	Degree of Freedom (DF)	105
	Significance (P)	<0.001

2.3. Identify the contribution value of each principal component

In the total variance explained principal components table, the paper base the extraction of the number of principal components on the eigenvalues and cumulative variance contribution ratio. The variance contribution ratio of the ith principal component is shown in the following equation:

$$f_i = \frac{\gamma_i}{\gamma_1+\gamma_2+\cdots+\gamma_p}, i = 1,2,\cdots,p \quad (2)$$

A total of 15 component factors were identified in this analysis, and these four components totaling more than 1 were sequentially named f1, f2, f3, and f4. the paper rotated the stock evaluation principal component matrix according to the maximal method to derive the rotated component matrix, and the rotated component matrix can make the interpretation clearer[8].

As can be seen from Table 2, the main factors determining the stock scores are categorized into four groups: component 1 factor loadings are 0.819, 0.767, 0.688, 0.674, and 0.650, representing the factor loadings for the growth rate of total assets, the capital appreciation rate, the current ratio, the net profit margin on fixed assets, and the ratio of stockholders' equity to fixed assets, respectively; component 2 factor loadings are 0.770, 0.760, 0.622 and 0.554, which represent the factor loadings of the growth rate of income from main business, the growth rate of net profit, the net profit margin of total assets and the price-to-capitalization ratio, respectively; the factor loadings of component 3 with a larger amount of 0.609 and 0.574, which represent the factor loadings of the turnover ratio of receivables and the net profit margin of fixed assets, respectively; and the factor loading of component 4 with a larger amount of 0.665, respectively. The other factor loadings are all relatively small, so this paper will use 0.665 to represent the factor loadings of the price-earnings ratio, ignoring the rest of the factors that have little or no significant influence. In this paper, components 1, 2, 3 and 4 are named F1, F2, F3 and F4 in turn.

Table 2 Component matrix

	components			
	1	2	3	4
Zscore(X11)	0.331	0.622	0.362	-0.062
Zscore(X12)	0.674	-0.031	0.574	0.04
Zscore(X13)	0.238	0.77	-0.158	-0.183
Zscore(X21)	0.767	-0.025	-0.412	-0.235
Zscore(X22)	0.819	-0.07	-0.352	-0.179
Zscore(X23)	0.261	0.76	-0.03	0.004
Zscore(X31)	0.591	-0.044	-0.343	-0.133
Zscore(X32)	0.135	0.554	-0.171	0.377
Zscore(X33)	0.005	0.084	-0.106	0.665
Zscore(X41)	0.688	-0.365	-0.083	0.119
Zscore(X42)	0.621	0.01	0.005	0.451
Zscore(X43)	0.65	-0.326	0.439	0.06
Zscore(X51)	0.303	-0.1	0.609	-0.201
Zscore(X52)	-0.088	0.124	-0.03	-0.511
Zscore(X53)	-0.11	0.379	0.444	-0.024

Extraction method: principal component analysis.
a. Four components were extracted.

2.4. Calculate the composite score of the stock

Based on the score calculation model $F_i = w_{i1}X_1 + w_{i2}X_2 w_{i3}X_3 + \cdots + w_{in}X_n$ for each principal component and the composite score calculation model $F = \alpha_1 F_1 + \alpha_2 F_2 + \alpha_3 F_3 + \alpha_4 F_4$, the paper can get the formula of the four principal components and the formula of the composite score as follows[9]:

$$F1 = [X_1, X_2, \cdots, X_{15}][0.171846638, 0.349923365, \cdots, -0.057109154]^T \quad (3)$$

$$F2 = [X_1, X_2, \cdots, X_{15}][0.411118943, -0.020489851, \cdots, 0.250504951]^T \quad (4)$$

$$F3 = [X_1, X_2, \cdots, X_{15}][0.277071471, 0.439334321, \cdots, 0.339833517]^T \quad (5)$$

$$F4 = [X_1, X_2, \cdots, X_{15}][-0.055454486, 0.035777088, \cdots, -0.021466253]^T \quad (6)$$

$$F = 0.34732F_1 + 0.25261F_2 + 0.14378F_3 + 0.08335F_4 \quad (7)$$

In the end, this article selects the top four companies to invest in, namely, Tomfoolery Microelectronics, Linkage Technology, Jingchen and Weir.

2.5. Position Management and Trading

In the portfolio management, in this paper, the paper utilize the data of the logarithmic returns of the trading days of these four companies in the last three months and the Markowitz optimal risk portfolio methodology, and the paper analyze the data in R language. According to the tangent point of the optimal portfolio line and the efficient frontier, the paper obtain the maximum Sharpe ratio and construct a specific portfolio, as shown in Table 3.

Table 3 Investment ratios under Markowitz's investment theory

Name of security Percentage of investment	Name of security Percentage of investment
Amlogic Limited	28%
linkage technology	35%
Weir Shares	16%
Tomfool Microelectronics	21%

At the same time, taking into account the high volatility of the semiconductor industry, as well as the existence of uncertainty in the early stage of the market grasp of the general market, this paper adopts the positive pyramid buying method and inverted pyramid selling method. And according to moving averages and Granville's Law initially develop a reasonable time to buy and sell, to take the average trading strategy and the strategy of high-pitch low-suction. SMA trading strategy refers to determine the timing of buying and selling based on whether there is a golden cross and a dead cross between the moving averages, buying when there is a golden cross and selling when there is a dead cross[10]; the high and low absorption strategy refers to the simultaneous attention to the deviation rate of low absorption and the number of high absorption, the MACD top divergence and bottom divergence two indicators. In the short market, the stock price continued to fall sharply, running below the averages, the negative deviation rate continues to decrease close to the negative value of the previous period, this time to buy, the opposite is to sell.

3. Trading results

In this paper, transactions are conducted from August 7 to September 21, 2023 and the results are shown in Table 4.

Table 4 Summary of basic information on securities transactions

serial number	stock code	Stock	Profit and loss	Amount	Number of days available for trading	Reasons for buying and selling	earnings yield
1	688099	Amlogic Limited	Profit	103886.6	46	Top rankings for principal component analysis, double averages, kdj golden crosses	8.19%
2	603501	Weir	Profit	10320	46	Top rankings for principal component analysis, double averages, kdj golden crosses	1.43%
3	002156	Tong Fu Microelectronics	Profit	8968	46	Top rankings for principal component analysis, double averages, kdj golden crosses	0.95%
4	301369	Linkage Technology	Loss	42488	46	Principal Component Analysis ranked high with double SMA bonding	-2.67%

4. Conclusion

In this paper, the selected positions of Tong Fu Microelectronics, Linkage Technology, Jing Chen shares and Weir shares in the process of manipulation by the demand recovery and industry boom re-rise factors, the rate of return is more in line with the expectations, of which three stocks have made a profit, the total rate of return of 0.74%. This shows that the top-down investment strategy adopted in this paper is highly practical and scientific.

In summary, this paper summarizes the following two operational summary: first, to analyze the overall trend of the industry have a handle on the overall trend of the industry, and timely understanding of the macro factors affecting the investment, such as monetary policy, industry cyclicality and so on. Clearly the main risks that may arise in the investment process, clear focus on the center of gravity, through the analysis of the reasonable operation to reduce the degree of risk to

the impact of investment returns. Secondly, the paper must pay attention to the trend of stock prices, according to the selected method of position adjustment, trading strategy and other theoretical principles, timely opening, adding positions, reducing positions, in the stock price trend is not yet clear when the prudent investment, investment operations in the early stages of not fully analyze the trend of a good position at a slightly higher level, the follow-up to avoid the funds were set can only be turned into a real loss of the floating loss.

References

[1] Li Muyun. Research on multi-period investment strategy based on rolling optimization and Black-Litterman model [D]. Donghua University,2023.

[2] Xin-Ben Huang. K-Means volatility agglomeration based stock selection and timing portfolio investment strategy[D]. Yunnan University of Finance and Economics, 2023.

[3] Zhu M. Research on stock price momentum investment strategy from the perspective of HS-IPS model[D]. Huzhou Normal College,2022.

[4] Sun Jixiang. Research on multi-factor stock selection trading strategy based on TabNet[D].Chongqing University of Posts and Telecommunications,2022.

[5] Lei Le,Li Ziyi,Yin Lianqian. Quantitative stock selection strategy based on entropy weighting method and CART decision tree[J]. SAR Economy,2022(05):114-118.

[6] Hu Meijuan. Evolutionary process, mechanism and path optimization of urban ecological welfare performance under the perspective of comprehensive welfare [D]. Nanjing Normal University,2020.

[7] Rudan. Research on Financial Agglomeration and Influencing Factors in Five Northwest Provinces and Regions[J]. Western Finance,2020(06):87-94+97.DOI:10.16395/j.cnki.61-1462/f.2020.06.017. Wang WT. Evaluation of textural quality and product development of potato whole flour-wheat flour mixed dough [D]. Shanghai Normal University,2017.

[8] Mei Xueping. Research on the evaluation of factors affecting economic development in Sichuan province based on principal component analysis[J]. China Collective Economy,2023(05):33-36.

[9] Xing Zhangfan. Empirical Research and Optimization of the Law of One [D]. Zhejiang University,2019.