Correlation Analysis between the Return of a Single Stock and the Market Index Based on the CAPM Model

Zhehao Chen[1,†], Mengqun Liu[2,*,†] Yuda Wu[3,†]

[1]Tianfu school

[2]Capital University of Economics and Business

[3]Shanghai United International School Jiaoke Campus

*Corresponding author: 32019090011@cueb.edu.cn

[†]These authors contributed equally.

Abstract. In recent years, with the rapid development of information technology, it has been difficult for people to make investment decisions due to the increasing amount of stock data information. Quantitative investment is a new investment strategy emerging in this situation. Studying quantitative investment strategy and using the CAPM model for data analysis can get a more objective and efficient investment strategy. This paper analyzes the correlation between stock return rate and market index by collecting daily data of 50 Stocks in Shanghai and Shenzhen A-share markets and carrying out an empirical test based on the CAPM model. Finally, it is concluded that the fluctuation of stock prices is highly correlated with the change of market indexes, especially in some industries such as materials and retail. Therefore, the instability of SSE and Shenzhen indexes can be taken into consideration to a large extent when selecting investment objects.

Keywords: CAPM model, Rate of return, Index rate of return.

1. Introduction

Electronic information is now widely used in a variety of industries depend on the rapid rise of big data and the ongoing development of information technology. The uses of various industries in society and the popularity of electronic information have caused the trading volume and trading data of the stock market to continue to rise in the age of big data. In this era, it has been challenging for the traditional investment strategy to manage such a massive amount of information and select the best investment strategy. A brand-new investment method called quantitative investment is better suited to the current financial climate. Quantitative investment is popular in recent years in China. Its fundamental concept is to employ computer technology and a specific digital model to put ideas and tactics for investors into practice [1]. In 1964, American academics combined asset portfolio theory and capital market theory to create the capital asset pricing (CAPM) model. It focuses primarily on the relationship between hazardous assets and the projected rate of return of assets on the securities market. It can also investigate how to produce balanced prices. The foundation of contemporary financial market price theory, which has significant research implications, is the CAPM model.

The study of quantitative investment strategies and the use of the CAPM model for data analysis can make investment methods swifter, disciplined, and objective as quantitative investing gains prominence in the world of financial investment. Following a thorough review of the literature, it is believed that China's current stock market is not ideal compared to that of other nations, and that the research on quantitative investment is insufficient compared to that of other nations. As a result, it is thought that further study of quantitative investment strategies and models will help to fully comprehend how quantitative investment is applied in the actual stock market and will highlight the objectivity, accuracy, and accuracy of this type of investment. The application of quantitative investment in the actual stock market as well as the benefits of quantitative investment can be completely understood through the study of the CAPM model in the stock market on the analysis of single stock data and its correlation analysis to the market index.

This paper aims to use CAPM model to analyze the correlation between stock return and market index, fully understand the advantages of quantitative investment compared with other strategies, and

how to select stocks with investment significance efficiently and objectively. The first part of the paper will summarize the literature and explain the theory. The second part will introduce the data. The third part will analyze the correlation between CAPM model and the market index through the data analysis and empirical test of 50 randomly intercepted daily data of Shanghai and Shenzhen A shares.

2. Literature review

2.1 Basic Concepts of CAPM Model

CAPM model is the abbreviation of Capital Asset Pricing Model in the financial market. The capital asset pricing model mainly studies the relationship between the expected return rate of assets and risky assets in the securities market and how the equilibrium price of stocks is formed in the securities market. Since its birth, capital asset pricing has been the focus of the modern finance research model.

CAPM model is a fundamental mathematical theoretical model in the field of finance, which uses the assumed mathematical formula to calculate the expected return rate of stocks. Domestic scholars have also carried out relevant research on the basic concept of CAPM. Hao believes that since future returns of capital assets such as stocks are uncertain, the essence of the CAPM model is to discuss the relationship between capital risks and returns. That is, high returns accompany high risks. CAPM model is derived based on certain assumptions [2]. CAPM model can show that there is a linear relationship between the expected returns of different assets in equilibrium. Liu believes that the CAPM model needs to be studied on the premise that the market reaches equilibrium to verify what conditions the expected return rate of assets needs to meet in the equilibrium market. Therefore, it is necessary to consider the relationship between asset price and expected return of asset and the relationship between average past return of asset and expected future return of asset [3]. Wang, Tao, Li, and Hou discuss the effectiveness of the CAPM model in the Chinese stock market and predict the accuracy of future return rates with the help of different listed companies in China, proving that the CAPM model is effective to a certain extent for the same amount of portfolio [4]. Zhang proposed that the CAPM model does not apply to China's securities market to some extent, the author analyzed the ineffectiveness of the application of this theory in China's securities market and put forward suggestions to expand the stock market scale, improve investor structure and strengthen information management to improve the applicability of this theory in China [5]. He and Hu do empirical research on Shanghai and Shenzhen stock market data, the risk coefficient of value and the relationship between stock and stock return are tested and think in Shanghai and Shenzhen stock market in China, the applicability of CAPM model is weak at the present stage, the system risk in asset pricing plays a vital role in [6].

2.2 The Importance of Studying the CAPM Model for Stock Return and Correlation Analysis

Similarly, the practical application of the CAPM model in China's A-share market and correlation analysis of stocks are also essential because the A-share market is the main body of China's stock market and is the principal place for domestic enterprises to finance equity and domestic people to invest. Yang and Zheng believe that studying the correlation analysis of the CAPM model in an A-share market is conducive to optimizing the allocation of market resources and promoting the development of the national economy to A certain extent [7]. Fan studied the applicability of THE TRADITIONAL Chinese medicine industry to CAPM model research and explored the industry's overall risk tolerance, which is of great significance to the study of the relationship between risk and income of the traditional Chinese medicine industry and the development of the traditional Chinese medicine industry [8]. Ban believes that with the further deepening of the economic system reform and the gradual improvement of the capital market system, the scale of Listed companies in China is further expanding. It is imperative to conduct empirical research through THE CAPM model to judge the degree of risk and to give investment suggestions and countermeasures [9]. Yin and Huang believe

that with the growth of the science and innovation board market, it is vital to explore its asset pricing methods and conduct an empirical study on the applicability of CAPM and its expansion model in the science and innovation board market [10]. Wu made an empirical analysis using the CAPM model to predict the development prospect of the transportation and logistics industry and provide a reference for investors [11].

3. Data and methodology

3.1 Methodology

A fundamental mathematical theoretical model for the financial industry, the CAPM model employs mathematical formulas to determine the expected return of equities on the presumption that the requirements are met. On the foundation of capital market theory and portfolio theory, it was created by American academics in 1964. The following is the CAPM model's calculation formula:

$$E(R_i) = R_f + \beta_i E\left[(R_m) - R_f\right] \quad (1)$$

Where Ri denotes the anticipated rate of return on asset i, Rf denotes the risk-free interest rate, βi denotes the β coefficient of stock i, and Rm denotes the anticipated yield of the market portfolio, E[(Rm) -Rf] denotes the market's anticipated rate of return less the risk-free interest rate, or the market risk premium. In order to confirm the relationship between stock returns and the Shanghai Stock Exchange and Shenzhen Stock Exchange indexes, this study will employ the CAPM model to conduct empirical analysis on the data of the selected sample stocks. Regression analysis is performed using Eviews software on the data of the chosen sample stocks, which is calculated using the CAPM model Value, in order to confirm the applicability of stocks in various industries on the A-share market as well as the relevance of stock returns to the Shanghai Stock Exchange and Shenzhen Stock Exchange indexes.

3.2 Data

The daily data of 50 A-share equities listed on the Shanghai Stock Exchange and the Shenzhen Stock Exchange from 2012 to 2021 is the data chosen for this article. The stock market trading module in the CSMAR database is where the data originates. Samples from the real estate, food and beverage, financial, and other sectors are chosen at random. To prevent singularity, choosing stocks from a variety of industries can help completely reflect and summarize the characteristics of the stock market as a whole. In order to guarantee an adequate data volume and boost the precision of model calculation, 10-year data of each sample are chosen at the same time.

Table 1. Descriptive statistics

Code	Minimum	Maximum	Mean	Variance	Skewness	Kurtosis
*000011	-10.0375%	10.0474%	0.0780%	10.1460	0.1440	2.0370
*000554	-10.0840%	10.1124%	0.0328%	8.7760	-0.0080	2.6910
*000570	-10.0346%	10.1031%	0.0466%	7.9210	0.0360	3.1990
*000702	-10.0423%	10.0840%	0.0474%	7.6420	-0.2290	3.4940
*000822	-10.0687%	10.1124%	0.0538%	8.3050	0.0580	2.4640
*000869	-23.6702%	10.0113%	-0.0199%	5.4480	-0.3440	7.1450
*000957	-55.1139%	10.0858%	0.0462%	10.8500	-1.7610	33.9560
*600004	-30.4061%	10.0356%	0.0544%	5.2420	-1.0940	16.1000
•600015	-28.8793%	10.0541%	-0.0106%	3.4440	-2.4480	41.1390
•600050	-10.0559%	10.1031%	0.0094%	4.9590	0.3950	5.5510
•600085	-10.0160%	10.0222%	0.0738%	5.0240	0.3400	4.7670
*600261	-34.9776%	10.0297%	-0.0176%	7.2790	-2.0160	27.0810
•600456	-10.0174%	10.0254%	0.1032%	9.0120	-0.0570	2.0090
*600563	-10.0119%	10.0212%	0.1533%	8.0720	0.1120	1.9270
*600642	-10.0446%	10.0719%	0.0380%	3.7090	-0.0100	7.9950
*600838	-10.3187%	10.0870%	0.0523%	8.4520	0.1050	3.2200

Table 1 displays the results of counting 50 representative stocks from each industry using SPSS. The average value is shown to be approximately 0.1. The small standard deviation values demonstrate the small size of their average variation. Their variances are very different, but they are all more than 3, which indicates that the data are widely dispersed. Four equities have kurtosis that is greater than ten percent, which shows that the extreme value disparities between them are very dissimilar.



Fig. 1 Shanghai Stock Exchange

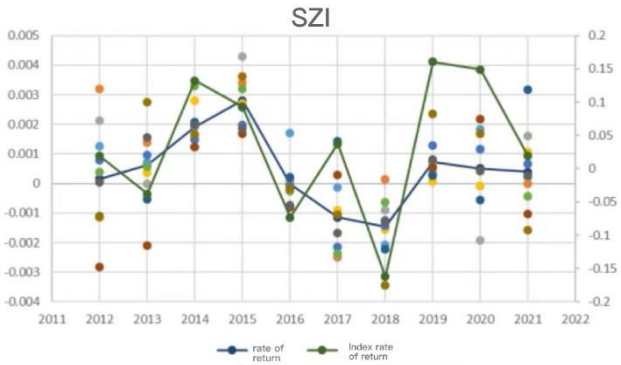


Fig. 2 Shenzhen Stock Exchange

With the use of two scatter figures constructed in accordance with the Shanghai Stock Exchange and the Shenzhen Stock Exchange, as well as a line figure of the mean value to illustrate the trend, we estimated the yield of 50 sample stocks. Then build a line figure, determine the index's average

return rate, and compare the results to the corresponding market indices. The annual average return broken line is shown in blue in the mean value data and figures for index return and mean value of return, and the annual average index return broken line is shown in green. Three years is a period of increase or reduction in return as compared to the scatter figure. The average return is roughly similar to the broken line trend of the average index return, and the Shanghai Stock Exchange rose to a higher index return of 17.9 percent in the first cycle, as well as the Shenzhen Stock Exchange, which also rose to a higher index return of 13.24 percent in the same year. We can infer a preliminary conclusion from this: the correlation coefficient has a tendency to be near to 1, indicating that there is some relationship between the return rate and the change in index return rate.

4. Empirical analysis

4.1 Firm Level

This paper first uses the unique analysis method of a single stock as an example in order to clearly demonstrate the data analysis process.

Establish the stock return of Jinmo shares using the CAPM model by importing the computed stock return of Jinmo shares and the return rate of the Shanghai stock index into the Eviews program. coefficients of a single-variable linear regression model.

$$K_i = a + \beta K_m \tag{2}$$

In the model, Ki stands for the stock return of Jinmo shares less the risk-free return (Ki = Ri – Rf), and Km for the return rate of the Shanghai Stock Exchange Index less the risk-free return (Km = Rm−Rf).

Using Eviews, it is possible to determine the slope of the univariate linear regression equation, which is the slope of the gold molybdenum stock coefficient, in order to derive the model parameters.

Table 2. Parameters of regression equation

		Value
β		1.2285
a		0.4394
R-squared		0.4857
F-statistic		2291.5060
t- statistic	β	47.8697
	a	6.8658
Prob.	β	0.0000
	a	0.0000

The significance of regression parameters is examined using the T-test. Since the significance level in this equation is set at 0.05, the effect of the return on the Shanghai stock index on the return on Jinmo shares is significant. It is assumed that there is no heteroscedasticity because the primary explanatory variables in the model are complete and there is no setting error or measurement error of sample data. The results of the regression analysis show that the model's determining coefficient R2 is 0.4857, which indicates that the goodness of fit is generally good and that the change in the Shanghai Stock Exchange index has a significant bearing on the yield. Between January 1, 2012, and December 31, 2021, Jinmo's stock yield was According to the coefficient of 1.2285, the stock yield of Jinmo shares will rise by 1.2285 percent for every one percent increase in the Shanghai index. his β If the coefficient is larger than 1, there is a significant link between the stock's return rate and the Shanghai Stock Exchange Index. The return rate fluctuation range of individual equities is wider than

the portfolio representing the entire market. The predicted return on investment in the asset is greater than the market's average return.

4.2 Industry Level

We categorize the 50 stocks by their industries and calculate the mean values of β coefficients and R-squared figures for each industry to analyze the difference in the degree to which different industries are affected by changes in the market index. The mean values of β coefficients and R-squared figures for each industry are shown in Table 3.

Table 3. Mean values of β coefficients and R-squared figures for each industry

Industry	β Coefficient	R-squared
Materials	1.2316	0.4122
Commercial Retail	1.1746	0.3474
Information Technology	1.1533	0.3088
Real Estate	1.1054	0.3505
Industry	1.0986	0.3888
Energy	1.0624	0.4027
Transportation Equipment	1.0448	0.3164
Household Appliances	1.0364	0.3069
Telecommunication Services	1.0243	0.4953
Chemicals	1.0204	0.3634
Transportation	1.0119	0.4550
Health Care	0.9991	0.3623
Finance	0.9956	0.4782
Food & Beverage	0.9691	0.3563
Agricultural Products	0.9627	0.3543

The results show that the β coefficients of the materials, commercial retail, and information technology sectors are large, ranging from 1.1 to 1.3, while the β coefficients of the health care, finance, food and beverage, and agricultural products sectors are smaller, all less than 1. It can be seen that there are indeed differences between the industries. Industries with large β coefficients are more likely to be investment stocks and share price movements are more susceptible to movements in the market index. Turning to stocks in the health care, finance, food and beverage, and agricultural products sectors, they are mostly stable and less susceptible to changes in the market index, with systemic risk accounting for a smaller share of total risk. The stock price movements are mostly caused by non-systemic risk, which means that they are more affected by their operating conditions. From the results, it is clear that investing in industries with high β coefficients, such as materials and commercial retail, is expected to gain greater investment returns when the market index moves.

4.3 Market Level

The regression results of all stock data are shown in Table 4, including 10 stocks in the A-share market of Shenzhen and 40 stocks in the A-share market of Shanghai. We calculate the β Coefficient and R-squared of these 50 stocks respectively. The p-values of all stocks obtained by t-test are 0.0000, which proves that there is a significant relationship between index rate of return and rate of return.

Table 4. Parameters of regression equation

Stock code	β Coefficient	R-squared	Stock code	β Coefficient	R-squared
000011	1.0374	0.3124	600269	0.9930	0.6345
000014	1.0576	0.3179	600302	1.1713	0.3878
000017	0.9505	0.2436	600377	0.7267	0.4154
000554	0.9546	0.3042	600456	1.3203	0.4050
000570	0.9817	0.3549	600519	0.8302	0.3207
000702	0.9627	0.3543	600563	1.1172	0.3230
000822	1.0204	0.3634	600594	0.9624	0.2676
000869	0.8878	0.4009	600642	0.9660	0.4974
000919	0.9345	0.4120	600650	1.2001	0.3778
000957	1.1001	0.3288	600683	1.1797	0.3190
600004	1.0448	0.4263	600692	1.1971	0.3558
600007	0.9803	0.3542	600742	1.1570	0.3523
600010	1.1151	0.3038	600776	1.1894	0.2945
600012	1.0210	0.3643	600808	1.1216	0.4152
600015	0.9153	0.4759	600838	1.1746	0.3474
600017	1.0926	0.4920	600897	0.9649	0.4244
600028	0.8551	0.4619	600971	1.2941	0.4327
600050	1.0825	0.4932	600993	1.1309	0.3899
600066	0.8774	0.2681	601333	1.0517	0.5053
600085	0.9686	0.3797	601600	1.2623	0.4541
600118	1.2235	0.3781	601601	1.0758	0.4805
600159	1.1800	0.4439	601618	1.1517	0.4511
600197	1.1893	0.3474	601668	1.0315	0.4627
600261	1.0364	0.3069	601808	1.1458	0.4121
600262	1.0318	0.2980	601958	1.2285	0.4857

The size of the R-squared calculated in the model indicates the goodness of fit of the model and determines the share of systematic risk in the total risk, which means how much of the movement of the stock price is caused by the movement of the market index. The stock with the largest R-squared is Jiangxi Ganyue Expressway Co., Ltd (600269), with an R-squared of 0.6345, indicating that the stock has a good fit, and the stock price movement is more influenced by the market index movement. Shenzhen China Bicycle Company (Holdings) Limited is the stock with the lowest R-squared, with the R-squared of 0.2436, indicating that it fits the model poorly and its stock price changes are mainly caused by the company's operating conditions rather than by changes in the Shenzhen securities component index. The overall R-squared of these 50 stocks is at a moderate level, indicating that most of the sample stocks have sufficient goodness of fit with the model.

The β coefficients of the 50 sample stocks selected in this paper are distributed between 0.7 and 1.3, with 33 companies' stocks having β coefficients greater than 1 and 17 stocks having β coefficients less than 1. The highest β coefficient of 1.3203 is for Baoji Titanium Industry Co., Ltd (600456) and the smallest β coefficient of 0.7267 is for Jiangsu Expressway Company Limited (600377). The results show that there is a significant positive correlation between the stock rate of return and the market index, and most of them are strongly correlated, which means the return of investing in these stocks is greater than the level of the market return. This also reflects that most of the stocks listed in China's A-share market are speculative stocks. When the market index rises, investing more in stocks with higher β coefficients and less in stocks with lower β coefficients can lead to a more excess return. In contrast, when the market index declines, reducing investment in stocks with large β coefficients can effectively reduce investment loss. Most of the stock price changes of the sample stocks selected

in this paper are susceptible to market index movements, indicating that the stock return has a strong correlation with the SSE and SZSE index.

5. Conclusions

According to the CAPM model, we analyze the correlation of single stock data and the market index in the stock market to understand its application in the actual stock market. From the analysis, it can be concluded that among the stocks listed in the A-share market of the Shanghai Stock Exchange and Shenzhen Stock Exchange in China, the movements of most of the stocks' share prices are highly correlated with the movements of the market index. The selection of investment objects can largely refer to the movement of SSE and SZSE indices. In particular, when investing in sectors with large β coefficients, such as materials, commercial retail, and information technology, it is more likely to gain more than 1% for every 1% rise in the index. Therefore, we can make more reference to the changes in the market index when investing in these industries as a way to gain more investment returns.

REFERENCES

[1] X.W. Li. Analysis of the investment model of technological innovation: the application of quantitative investment in the financial market. China Circulation Economy, (13) (2022)70-72. DOI: 10.16834/j.cnki.issn1009-5292.2022.13.003.

[2] X.Y. Hao. Discussion of Theory and Applying of Capital Assets Pricing Models. Journal of Zhongyuan University of Technology, (04) (2002)43-46.

[3] Y.K. Liu. Asset pricing problem solving based on CAPM model and model improvement. Commercial Economics Review, (07) (2021)174-175.

[4] W.J. Wang, S Tao, J.Y. Li, W.B. Hou. An Empirical Analysis of Capital Asset Pricing Model in Different Enterprises in China. Journal of Huaibei Normal University (Natural Sciences), 42(04) (2021)23-31.

[5] M.M. Zhang. An analysis of the non-effectiveness of CAMP theory in the Chinese securities market. Times of Fortune, (11) (2021)182-183.

[6] X.R. He, R.D. Hu. The Empirical Study of CAPM Model in Shanghai and Shenzhen Stock Market of China. Journal of Fujian Business University, (06) (2020)23-29. DOI: 10.19473/j.cnki.1008-4940.2020.06.005.

[7] S.H. Yang, Z.K. Zheng. Empirical Test of the Applicability of CAPM Model in A-Share Market. Finance Theory and Teaching, (06) (2021)39-44. DOI: 10.13298/j.cnki.ftat.2021.06.010.

[8] S.C. Fan. An empirical study based on CAPM model in Chinese medicine industry. China Collective Economy, (17) (2021)66-68.

[9] D.L. Ban. Risk Analysis of Luzhou Lao Jiao Co., Ltd Based on CAPM Model. Economic Research Guide, (08) (2022)114-116.

[10] G.W. Y, R. Huang. An empirical study on the applicability of CAPM and its extended model in the KCI market. Quality & Market, (04) (2021)138-140.

[11] X. Wu. Empirical Analysis of Transportation Logistics Industry Stock Based on CAPM Model. Logistics Engineering and Management,43(02) (2021)108-110.