The 2022 International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/6/20220198

Research on the Stock Investment Value of Hong Kong Stock Connect based on Factor Analysis

Guanran Hao [1,†], Kai Hou [2,†], Junwei Hu [3,†], and Yao Tan [4,a,*,†]

[1]*No.2 Middle School of Yantai Shandong Province, Yantai, 264000, Shandong, China*
[2]*Shanxi Xiaoyi Middle School, Xiaoyi, 032300, Shanxi, China*
[3]*Columbia international college-Hamiltion-L9C7V6-Ontario-Canada*
[4] *Saint Paul American School, Beijing, 100000, Beijing, China*
a. yao.tan@stpaulamerican.org
**corresponding author*
[†]*These authors contributed equally*

Abstract: Based on the Hong Kong Stock Connect, this paper introduces a stock investment pricing method based on factor analysis, and combines it with portfolio model. A statistical method known as factor analysis is used to separate common factors from collections of data. British psychologist Spearman was the one who initially put out the idea. He modified and enhanced the portfolio model after discovering a particular link between several probable common elements., so as to form a complete set of methods for establishing stock alternative asset allocation. This paper will first introduce a method to evaluate and compare the business performance of listed companies in the industry, and apply it to the analysis data. Finally got the investment plan. In order to explain and demonstrate the methods and models discussed in this paper, this paper will take the Hong Kong Stock Connect listed companies as an example, and analyse and summarize the results. This paper will have certain reference value for investors, and promote the theoretical research of related issues more in-depth.

Keywords: Stock Investment Value, Factor Analysis, Hong Kong Stock

1. Introduction

Nowadays, the economy is currently growing more quickly than ever, and income levels are rising as well, and people no longer only pursue material life, but more spiritual life. In today's world almost everyone has spare funds, and for these funds, many people choose to invest, among the many investment products, the most popular investment should be a stock investment [1]. Therefore, a lot of stock investment concepts followed, and rational investment and value investment gradually become mainstream. Then, how to make a reasonable stock analysis is also a problem for every investor to solve. Factor analysis has been applied in economics, sociology, management, and other research fields. Its main goal is to identify the essential hidden variables that are buried inside a group of observable variables but cannot be assessed explicitly.[2].Based on the study of the internal dependence of the original variable correlation distance matrix, the variables of the intricate relationship are summarized into a multivariate statistical method with a few comprehensive factors.

DOI: 10.54254/2754-1169/6/20220198

In the face of numerous indicators in the financial statements of listed companies, investors can use this method to replace the original indicators with three-factor scoring indicators [3]. These three indicators also contain information reflecting the differences between stocks, so that investors can have a clear understanding of the financial status of listed companies. Value investment can not only test people's ability to analyze data but also test the value orientation. In the stock as a value investment method, to select the listed companies with good economic benefits, a large number of data increase the difficulty of investors to analyze the listed companies, so the application of factor analysis will be a good solution to this problem [4]. By using the factor analysis method, investors can make a quantitative analysis of each index and extract the factors that affect the stock. In this way, they can make a relatively objective evaluation of the stocks of listed companies and provide an evaluation criterion for small and medium investors. Therefore, establishing an appropriate queuing model is conducive to helping achieve the above goals.

At present, the development of the Chinese stock trading market is becoming more and more mature, meanwhile, the mode of stock trading supervision is becoming more and more perfect [5]. The stock market has attracted a lot of money managers. Stock market risks are everywhere. In order to effectively avoid investment risks, stock investors will collect market information through various channels to form accurate investment judgments. The stock market is heavily influenced by a variety of factors, and stock movements are extremely difficult to judge. To participate in the stock investment, we should not only avoid the risk but also build the stock investment value evaluation system through the factor analysis method.

Now, all kinds of statistical analysis software, such as SPSS, and Eviews have been widely used. Researchers are also focusing more and more on the empirical examination of stock investment value, which has produced a vast body of research findings. Through the application of the analytic hierarchy process, the stock behavior is studied as a whole, and all kinds of influencing factors involved in stock value are quantified, so as to optimize the investment method for investors [6]. Through the method of principal component analysis, financial indicators of research value are reasonably extracted from the Hong Kong Stock Connect, and the most influential comprehensive financial indicators are selected to rank listed securities companies in terms of influence and competitiveness [7]. By using SPSS software, cluster analysis, and factor analysis, sample stocks are grouped and ranked, and their characteristics are analyzed. A new analysis method of stock fundamentals is constructed, which can realize the comprehensive consideration of several analysis indicators in a lower dimension.

Overview of the approach of factor analysis A statistical analysis technique that may be used with several variables is the factor analysis approach. This method carries out scientific research on the dependency relationships existing within variables, and scientifically summarizes the related variables with complex relationships to form a few comprehensive factors [8]. That is to say, variables with strong correlations belong to the same class, and each class variable becomes a factor. Through a few factors, multiple indicators or the relationship between multiple factors can be described, and most of the information contained in the original data can be reflected. The key factors can be extracted, that is, the factors that have the greatest impact can be found out accurately [9]. In this study, the Hong Kong Stock Connect's equities are categorized scientifically using the factor analysis approach., and the main constraints on their earnings growth quality are accurately identified according to the obtained analysis results, and the effective improvement path of the corresponding stock investment value of the Hong Kong Stock Connect is explored, making investing decisions more scientific, increasing investment, and serving as a reference for investors.

DOI: 10.54254/2754-1169/6/20220198

2. Methodology

These two techniques for factor analysis. One seems to be independent variables and dependent variable evaluation and the other is convergent validity and discriminant validity analysis. The link between factors and measurements is not assumed in exploratory factor analysis. The exploratory factor analysis approach is mostly used in this study [10]. Principal component analysis and cofactor analysis are frequent analytical methods [10]. There is frequently a connection between the variables. When two variables exhibit a certain degree of correlation, it can be inferred that the respondents' knowledge is somewhat reflected in both variables. In order to construct as few new factors as possible that are independent of pounds, principal component analysis is aimed to remove superfluous duplicate variables or closely similar variables for all previously suggested variables.

Factor analysis' major objective is to identify some of a set of measured variables' more fundamental but indirectly quantifiable hidden variables. The benefit of confirmatory factor analysis is that it enables researchers to unambiguously express the specifics of the theoretical model. We evaluate the measuring method's validity when many measurement methods are employed. The purpose of the validity test is to determine whether an index has a substantial load with the design factor and whether there is no significant load with the unrelated component. The variables are shown in Table 1.

Table 1: Table of market indicators.

variable	Variable declaration	Range
PreClosePrice	The previous closing price refers to the closing price of the previous trading day	[87.9, 98.5]
OpenPrice	The opening price of a stock on a specified trading day, calculated and published by the exchange	[87.9, 97.8]
ClosePrice	The closing price of a security on a specified trading day, calculated and published by the exchange	[87.9, 97.3]
HighPrice	The highest price at which a security is traded on a given trading day	[88.6, 97.9]
LowPrice	The minimum price at which a security is traded on a given trading day	[87.9, 96.8]
Volume	The volume of trading in a security during the course of a given trading day	[23.69, 97.73]
Amount	The amount of turnover generated by the entire trading of a security on a given trading day	[22.47, 89.51]
Change	(The closing price of the last trading day of the week - the previous week's closing price） x100/Weekly closing price	[-2.05, 1.3]
ChangeRatio	Designated daily trading volume/Latest outstanding share capital as of the specified date	[0.021, 0.04]

The bartlett test results show that the p-value is less than 0.05 after standardizing the data. The factor analysis of the data is indicated by the bartlett test's finding that the p-value is less than 0.05 after standardizing the data. After performing KMO testing on the data, the KMO value of 0.65

DOI: 10.54254/2754-1169/6/20220198

shows that there is some connection between the variables, and by using factor analysis, the common components can be identified. Next, the characteristic root of each component was utilized to construct the gravel map after the number of optimal factors was determined at any moment. Three factors are initially chosen because they have distinctive roots greater than one, as shown by the gravel map.

3. Results and Discussion

By using factor analysis, this paper come up with the following results, as Table 2 and Figure 1 show. The following tables analyse the case involving factor extraction and information on factor extraction quantity. The three components from the factor analysis are derived from the aforementioned table 1, and the characteristic root value is more than 1. The total variance interpretation rate after rotation is 99.053%, while the variance interpretation rates for the three components are, respectively, 55.941%, 22.540%, and 20.571%.

Table 2: Table of explained rates of variance.

Rate of variance interpretation table									
	characteristic root			Rate of variance interpretation before rotation			Rate of variance interpretation after rotation		
number	characteristic root	Rate of variance interpretation%	accumulate%	characteristic root	Rate of variance interpretation%	accumulate%	characteristic root	Rate of variance interpretation%	Accumulate%
1	5.322	59.135	59.135	5.322	59.135	59.135	5.035	55.941	55.941
2	2.300	25.560	84.695	2.300	25.560	84.695	2.029	22.540	78.481
3	1.292	14.358	99.053	1.292	14.358	99.053	1.851	20.571	99.053
4	0.085	0.947	100.000	-	-	-	-	-	-
5	0.000	0.000	100.000	-	-	-	-	-	-
6	0.000	0.000	100.000	-	-	-	-	-	-
7	0.000	0.000	100.000	-	-	-	-	-	-
8	0.000	0.000	100.000	-	-	-	-	-	-
9	-0.000	0.000	100.000	-	-	-	-	-	-

DOI: 10.54254/2754-1169/6/20220198

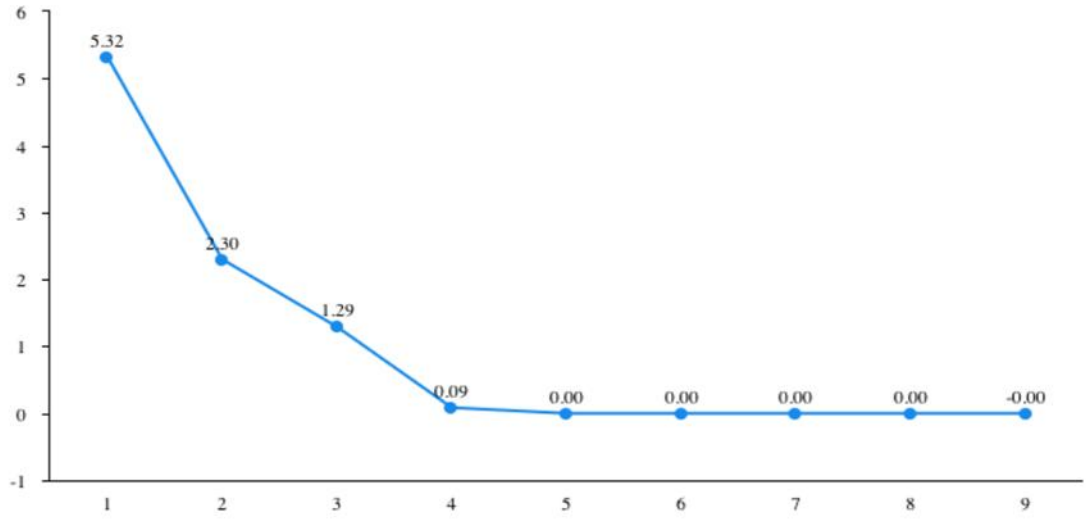

Figure 1: Gravel map.

The situation and volume of information used in factor extraction are analyzed in table 3 below. The three components from the factor analysis are derived from the aforementioned table 1, and the characteristic root number is much more above 1. The total variance translation ratio after rotation is 99.053%, while the variance interpretation rates for the three components are, correspondingly, 55.941%, 22.540%, and 20.571%.

Table 3: Table of the factor load coefficient after rotation.

Name	Factor load coefficient			Common degree
	Factor1	Factor2	Factor3	
PreClosePrice	0.990	-0.116	0.077	1.000
OpenPrice	0.990	-0.118	0.076	1.000
HighPrice	0.990	-0.116	0.079	1.000
Volume	-0.148	0.983	0.109	1.000
ClosePrice	0.990	-0.114	0.077	1.000
Amount	-0.148	0.983	0.109	1.000
Change	-0.063	0.156	0.965	0.960
LowPrice	0.990	-0.120	0.075	1.000
ChangeRatio	0.293	0.060	0.931	0.955

For the purpose of determining the data from this study were rotated using the maximum variance rotation method to determine the association between the variables and the research elements (varimax). The facts about factors' information extraction as well as the connection between factors and research items are shown in the table below. The following table shows that all research items have a common degree level of greater than 0.4, demonstrating a significant correlation between the research items and the factors as well as the factors' ability to extract information effectively. Check if the factor can effectively extract the majority of the data from the

DOI: 10.54254/2754-1169/6/20220198

study item, and then look at how closely the two correlate. (When the factor load coefficient's absolute value is higher, the factor's correspondence is more than 0.4.)

The final resulting data table 4 is shown below.

Table 4: Linear combination coefficient and weight results.

Linear combination coefficient and weight results					
Name	Factor1	Factor2	Factor3	Composite score	weight
Special roots	5.035	2.029	1.851		
Rate of variance	55.94%	22.54%	20.57%		
PreClosePrice	0.4413	-0.082	0.057	0.242	12.99%
OpenPrice	0.4412	-0.083	0.056	0.242	12.96%
HighPrice	0.4412	-0.082	0.058	0.243	13.00%
Volume	-0.0659	0.690	0.080	0.137	7.32%
ClosePrice	0.4414	-0.080	0.057	0.243	13.01%
Amount	-0.0659	0.690	0.080	0.137	7.32%
Change	-0.0281	0.110	0.711	0.157	8.39%
LowPrice	0.4411	-0.084	0.056	0.242	12.94%
ChangeRatio	0.1306	0.042	0.684	0.225	12.08%

The concept of dimension reduction is the foundation of factor analysis. In order to preserve or minimize the loss of the original data information, it condenses a large number of complex variables into a small number of independent common components that can capture the essence of the original large number of variables. It decreases the amount of variables while simultaneously reflecting how the variables are related internally. In general, factor analysis does three tasks: it reduces the dimensions of the factors being studied; it computes the factors' weights; and it computes the weighted factor summary scores.

Initially, we use KMO test to test whether the result of factor analysis is valid. The initial hypothesis is disproved by the KMO test, which reveals that the value of KMO is 0.65, indicating that there is correlation between the variables. The factor analysis result is accurate, and the result's dependability is medium.

Next, by drawing the gravel graph, we could decide how many principal components ought to be retained. In the gravel diagram, the number of principle components pertaining to the steep to steady transition on the broken line is equal to the amount of referrals retrieved primary components. It can be seen from the figure that, starting from the third principal component, the characteristic root value of the principal component starts to slowly decline. When the contribution of the factor cumulative interpretation reaches 90%, we can choose to retain the three principal components.

What is more, in order to drawing the rate of variance interpretation table, we calculated the characteristic root, rate of variance interpretation before rotation, and rate of variance interpretation after rotation. The primary emphasis of the total variance interpretation table is the contribution rate of factors to variable interpretation, or how many factors are necessary to interpret a variable as 100%. Generally, it should be expressed to more than 90%. Otherwise, the number of factors should be adjusted. The variance interpretation table illustrates that the cumulative interpretation contribution rate of the first three factors has reached 99.53%, far more than 90%. Therefore, the

DOI: 10.54254/2754-1169/6/20220198

first three factors can be used to evaluate the investment value of the Shanghai Hong Kong Stock Connect.

The importance of each factor hidden variable can be analyzed from the factor load coefficient table after rotation. The first factor is highly correlated with Preclose Price, Openprice, Highprice, Closeprice, and Lowprice, the five variables, which can be summarized as price factor; The second factor is closely related to two variables: Volume and Amount, which can be summarized as the amount factor; The third factor has a high correlation with Change and Changeratio, which can be defined as frequency factor.

According to the table of the factor load coefficient after rotation, the factor score function can be obtained as follows:

$$F_1 = 0.99x_1 + 0.99x_2 + ... + 0.239x_9 \quad (1)$$

$$F_2 = -0.116x_1 - 0.118x_2 + ... + 0.06x_9 \quad (2)$$

$$F_3 = 0.077x_1 + 0.076x_2 + ... + 0.931x_9 \quad (3)$$

Where x_1 is PreClosePrice, x_2 is OpenPrice, x_3 is HighPrice, x_4 is Volume, x_5 is ClosePrice, x_6 is Amount, x_7 is Change, x_8 is LowPrice and finally x_9 is ChangeRatio.

According to the calculations for the factor weight analysis, factor 1 has a weight of 55.941%, while factors 2 and 3 have weights of 22.54% and 20.571%, respectively. It is clear that the pricing element has the most influence on the stock value of Hong Kong Stock Connect.

The factor score is calculated. On this basis, the variance contribution rate is taken as the basis, and the ratio of the cumulative contribution rate of the two factors is combined to strengthen the summary score, and the following comprehensive score formula is obtained:

$$F = \frac{55.94\%F_1 + 22.54\%F_2 + 20.57\%F_3}{99.05\%} \quad (4)$$

Through the analysis of the above conclusions, the following suggestions are put forward for investors: first, investors should first consider the stock price factor, then the amount factor, and finally the transaction frequency factor when purchasing stocks. Using the ratio of variance contribution rate of the three principal components to the total contribution rate to calculate the comprehensive score F, and we can get the comprehensive score of the investment value of Hong Kong Stock Connect from 2016 to nowadays. Ranking of stock investment value can be obtained according to the comprehensive scores from large to small.

4. Conclusion

Above all, when selecting stocks, investors should not only consider the individual influencing factors of listed companies, but also comprehensively analyze all aspects of the company's indicators, and properly add industry related factors to accurately predict the internal value of the enterprise. In addition, when selecting stocks, we should pay attention to the volatility of the stock price. From the factor analysis results of the past seven years, the price factor is the biggest factor affecting the investment value of Hong Kong Stock Connect, and the stock price fluctuates in different periods of time, which indicates that stock buyers can properly choose a reasonable research period for analysis. In addition, investors should try their best to be "rational people" when investing in stocks. They should not take the temporary dividend level and stock price rise and fall of the enterprise as the evaluation criteria. They should maintain a rational and objective concept, and conduct a comprehensive analysis and comparison of the stocks they want to choose from both vertically and horizontally. At the same time, investors should receive investor education as much as possible, constantly improve their knowledge reserves, understand and pay attention to current events and national macro policies, establish rational investment concepts, and make appropriate investment decisions as much as possible to avoid risks and obtain higher investment income.

DOI: 10.54254/2754-1169/6/20220198

References

[1] *Floyd F J and Widaman K F 1995 Factor analysis in the development and refinement of clinical assessment instruments. Psychological Assessment, 7(3), 286-299.*

[2] *Han Q, Chen P and Ma T 2015 Influencing factor analysis of shale micro-indentation measurement. Journal of Natural Gas Science & Engineering, 27, 641-650.*

[3] *Kim J, et al. 2015 Risk factor analysis for the recurrence of chronic subdural hematoma: a review of 368 consecutive surgical cases. Korean Journal of Neurotrauma, 11(2).*

[4] *Buelow M T and Blaine A L 2015 The assessment of risky decision making: a factor analysis of performance on the iowa gambling task, balloon analogue risk task, and columbia card task. Psychol Assess, 27(3), 777-785.*

[5] *Wang X M 2004 The application of factor analysis in the stock evaluation. Application of Stats and Management.*

[6] *Wang J, Cui Q, Sun X and He M 2022 Asian stock markets closing index forecast based on secondary decomposition, multi-factor analysis and attention-based lstm model. Engineering Applications of Artificial Intelligence. The International Journal of Intelligent Real-Time Automation, 113.*

[7] *Lui S 2007 A study of stock volatility in the context of factor volatility models for large datasets: factor analysis and forecasting. Queen Mary University of London.*

[8] *Flad M and Jung R 2022 A common factor analysis for the us and the german stock markets during overlapping trading hours. Social Science Electronic Publishing.*

[9] *Hao J H, et al. 2007 Factor analysis for stock price fluctuation in shanghai stock market. Journal of Shenyang University of Technology.*

[10] *Khula A and Moroke N D 2017 The performance of maximum likelihood factor analysis on south african stock price performance. Journal of Economics and Behavioral Studies.*