Research on the Investment Value of Hybrid Funds in the Chinese Capital Market based on Factor Analysis

Runyang Lin[1, †], Yi Wang[2, *, †]

[1]School of Mathematics and Statistics, University of Sydney, Sydney, Australia

[2]Pennsylvania state University, Pennsylvania, United State

[*]Corresponding author: ykw5273@psu.edu

[†]These authors contributed equally.

Abstract. Factor analysis is a cutting-edge and effective data processing method that has a wide range of uses in the field of finance, as it can reduce the numerous indicators in finance into a few valid common factors in the form of factor analysis. The researcher can name the factors by the percentage contribution of each variable in the different factors. After that, the scores of financial products (e.g., stocks, funds) on different factors give a clear picture of how well the product performs in these aspects. The main research direction of this paper is to reduce the dimensionality of each index of hybrid funds in the Chinese fund market through the factor analysis technique to get reasonable public factors. And to rank these funds with factor scores to get the best funds among these ninety hybrid funds to make recommendations for investors. Finally, this paper successfully selects four excellent hybrid funds, which have excellent performance in terms of size, earnings, and debt-servicing ability, which can provide some reasonable suggestions for investors who want to invest in Chinese hybrid funds.

Keywords: Factor analysis; common factors; score rank.

1. Introduction

In recent years, China's fund industry has been booming with the development of banks, securities, and financial markets [1, 2]. Investment in the fund market is no longer limited to fund managers, and the amount of investment demand from individual investors for funds is also increasing. Since hybrid funds have both aggressive and conservative portfolio strategies, they have become the primary choice for most initial financial investors. The return and capital risk of hybrid funds are much lower than equity fund investments but at the same time higher than bond and market money market funds [3]. The current method of evaluating the performance of hybrid funds in China is to compare the NAV return of the fund with the return of an index [4]. The higher NAV return is usually considered the type of fund worth investing in the market [5, 6]. However, the significant disadvantage of this evaluation method is that the risk analysis of the fund is too one-sided and the time span of the fund analysis can have a great impact on the final prediction of the fund results.

To solve this problem, the CAPM model is used to analyze the market system analysis. The theory is based on the premise that the market is in stationery. In another word, the market has a traceable upward and downward trend. The CAPM model is shown below.

$$E_{ri} = R_f + \beta_i\left(ER_m - R_f\right) \tag{1}$$

Where E_{ri} =expected return of investment, R_f =risk-free rate, β_i =beta of the investment, $(ER_m - R_f)$market risk premium. In summary, the beta of the CAMP model is used to measure the riskiness of the fund with respect to the market. When the market is stationary, then the only factor that drives excess returns on equity values is the market systematic risk premium. When the asset takes on a more market systematic risk premium, it also earns more excess returns [7]. Therefore, according to this theory, the excess return is positively correlated with the beta coefficient. However, the limitations of CAMP, based index analysis for individual stock returns are as follows. Firstly, the data model must be based on the market being stationary, to carry out a value-weighted portfolio must meet the mean-variance validity. Based on the CAMP model, Jensen index maximization is also a common investment concept for maximizing fund investment returns. Based on the CAMP model, Jensen's measure is a risk-adjusted performance measure based on a visual representation of the average level of return for a mutual fund or an investment.

$$\alpha = R_p - \left[R_f + \beta * \left(R_m - R_f\right)\right] \tag{2}$$

Where R_p =Portfolio return, R_f = Risk free return, R_m = expected market return, β =portfolio beta. Unlike the CAMP model, which analyzes the market value of a stock, Jensen prefers to analyze the performance of an investment manager, i.e., a mixed stock or fund investment value analysis and uses an alpha measure to measure it. When the alpha value is higher, it reflects that the overall portfolio return is greater [8]. Therefore, in the fund market, investors need to pay attention not only to the overall portfolio return but also to the corresponding portfolio risk, in order to further understand whether the investment return can bear the corresponding market risk. Similar to the CAMP model, Jensen's alpha value is used as the only variable in the market regression line analysis for mutual funds. However, in real financial markets, the information that affects each stock is numerous and complex. Such as company size, assets and liabilities, company profits, and the ratio of book to market value all have an impact on the value of individual stocks [9]. The presence of many different variables will affect the performance of a hybrid fund in the financial market.

To address the characteristics of mixed fund market variables with a lot of information and heterogeneity, the multi-factor model can better explain the impact of the gap between mixed fund returns and common factors on the investment value of each mixed fund. The multi-factor analysis is a method to reduce the information of a large number of variables to a smaller number of factor associations, and therefore differs from dependent procedures such as CAMP and Jensen measure. Based on Kousmanen, Kopa and Post, the Handbook of Financial Econometric and Statistic proposed portfolio optimization and diversification using a standard test of stochastic dominance that would use factor analysis to achieve the relative weight of the criteria and the corresponding value of each criterion [10]. The combined utility value for mutual fund investment approach. The advantage lies in the depth of the investor's analysis of portfolio returns and risks. The core of asset allocation is based on the diversification effect to achieve risk minimization or maximize returns for a given level of risk.

2. Methodology

2.1 Data Description

All data in this paper are used in the public funds database of the CSMAR database. CSMAR is a famous database in China, which has all kinds of data on China's livelihood and economy. Many famous domestic and foreign journals have used its data. In this paper, 120 funds are randomly selected from the mixed funds in the Chinese fund market for analysis. After screening the random 120 funds for missing data, 90 valid funds were obtained. All data are used by the semi-annual data of the fund from Jan. 1, 2022 to Jun.30,2022 because of accuracy and timeliness.

2.2 Variables

For selecting variables for hybrid funds, a total of 110 variables describing hybrid funds in the database are treated in this paper after removing missing values from these variables and performing correlation tests to reduce highly repetitive variables [8]. Based on past cases of factor analysis funds, a total of 15 variables were selected in this paper. These 15 variables are BankDeposits(X_1), TotalLiability (X_2), TotalAsset (X_3), RetainedmentEarning (X_4), Profit (X_5), AVGProfit (X_6), ProfitAvailable (X_7), AVGNetProfitMargin (X_8), NAVGrowth (X_9), Insterestincome (X_{10}), InvestmentIncome (X_{11}), ManagementFee (X_{12}), TotalOperatingCost (X_{13}), ProportionOfFee (X_{14}), GearingRatio (X_{15}).

Table 1. Variables Introduction

Variables	rename	Intro
BankDeposits	X_1	The Fund's deposits with banks during the quarter
TotalLiability	X_2	Total of all liabilities of the Fund as of the quarter
TotalAsset	X_3	Total of all assets of the Fund as of this quarter
RetainmentEarning	X_4	Undistributed earnings at the beginning of the period + net income for the period - various surplus reserves appropriated - profit distributed + prior years' profit and loss adjustments.
Profit	X_5	Total profit of the Fund for the quarter
AVGProfit	X_6	Profit/amount
ProfitAvailable	X_7	The Fund's profits available for dividend distribution during the quarter.
AVGNetProfitMar	X_8	ROE=P/(E0+NP÷2+Ei×Mi÷M0-Ej×Mj÷M0)
NAVGrowth	X_9	(cumulative net value of shares - par value of units) ÷ par value of units
Insterestincome	X_{10}	The Fund's total income on interest for the quarter
Investmentincome	X_{11}	The Fund's total income on stocks and bonds for the quarter
ManagementFee	X_{12}	Total income of the fund manager for the period
TotalOperatingCost	X_{13}	Total Fund expenses for the quarter up to all
ProportionOfFee	X_{14}	Fund Custodian Fee Rate
GearingRatio	X_{15}	Total liability/Total assset

2.3 Factor Analysis

Factor analysis is a data analysis method that attributes information on multiple variables to a few new composite factors. The main way of this method is to group the variable information in terms of their correlation magnitude. In turn, the data with the higher correlation of variables in the data is obtained, and this data group represents basic variable information [7]. In factor analysis, this is called a common factor. It is assumed that the fund under observation has K evaluation indicators and n observed unit quantities. Therefore, the model of factor analysis represents each of the n observed unit quantities as p<K public factors, where p is fixed in advance. The public factors can be referred to as F1…Fn.

The general form of the factor benchmark model is:

$$R = r_f + b_1F_1 + b_2F_2 \ldots + b_nF_n + \varepsilon \tag{3}$$

R is the expected return of the factor model, r_f is the risk-free rate, b_n is the combination for the loading at the nth factor, F_n is the risk premium for the factor, $\mathcal{E}$ is the residual value.

Factor analysis is used as a form of dimensionality reduction for data analysis. Before processing the investment evaluation of mutual funds, each financial indicator of the fund needs to be screened. Firstly, missing and invalid indicators need to be removed. In order, to ensure the integrity of the shared factor data. Secondly, the KMO test and Bartlett test are also needed for the collated data. These two tests are used to determine whether the collated data are suitable for factor analysis. When the KMO value is closer to 1, the stronger the correlation between the variables and the more suitable for factor analysis. When the KMO value is less than 0.5, it is not suitable for factor analysis, while the Bartlett test is used to determine whether the variables are independent of each other, and when the p-value is less than 0.05, it proves that the variables are independent of each other and can be used for factor analysis. Based on the basic factor analysis process, the actual sample analysis and operation process will be explained next.

3. Results and Discussion

3.1 Factor Analysis Results

Table 2 shows after standardizing the data, the Bartlett test was performed using SPSS, and it was found that the p-value was less than 0.0001, which indicates that these fifteen variable indicators have a strong correlation and are not independent of each other and can be subjected to factor analysis. In the KMO test, the KMO value was 0.771, and this data set is suitable for factor analysis, and the common factors can be extracted.

Table 2. KMO and Bartlett test result

KMO Measure of Sapling Adequacy	0.771	
Bartlett's Test of Sphericity	Approx. Chi-Square	2169.110
	Df	105
	Sig.	0.000

Table 3. Common factor variance

Variable	Initial	Extraction
BankDeposits	1.000	0.749
TotalLiability	1.000	0.755
TotalAsset	1.000	0.955
RetainmentEarning	1.000	0.937
Profit	1.000	0.788
AVGProfit	1.000	0.743
ProfitAvailable	1.000	0.833
AVGNetProfitMargin	1.000	0.813
NAVGrowth	1.000	0.783
Insterestincome	1.000	0.605
Investmentincome	1.000	0.462
ManagementFee	1.000	0.969
TotalOperatingCost	1.000	0.976
ProportionOfFee	1.000	0.430
GearingRatio	1.000	0.795

From the Table 3 variance plot of the common factor, it can be found that the vast majority of variables are above 0.5, which indicates that the common factor can explain these variables very well. This suggests that there is a strong correlation between these variables. In addition, there are some variables with small variance contribution values (e.g., ProportionofFee is 0.430), and the common factor will not be able to explain these variables well.

After collating the variance of the common factors for each variable, this paper uses the gravel plot to find the optimal number of common factors. The screen plot method is a classic analysis method in factor analysis, which can quickly and clearly show the degree of interpretation of different numbers of common factors on the overall variables. From the Figure 1 screen plot, three factors were found to have characteristic roots greater than one, so three factors were initially selected as the number of common factors.

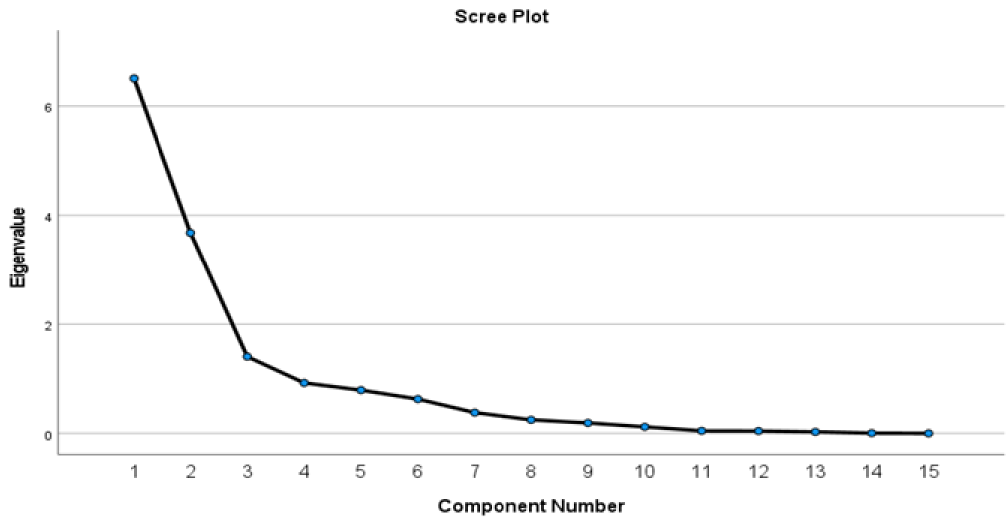

Fig. 1 Scree Plot

References are cited in the text just by square brackets [1]. (If square brackets are not available, slashes may be used instead, e.g. /2/.) Two or more references at a time may be put in one set of brackets [3, 4]. The references are to be numbered in the order in which they are cited in the text and are to be listed at the end of the contribution under a heading *References*, see our example below.

Table 4. Characteristic root shar and cumulative share

Component	1	2	3
1	0.987	-0.157	0.02
2	0.143	0.938	0.316
3	0.069	0.309	-0.949

Table 4 clearly shows that the cumulative percentage of the characteristic roots reaches 77.293 percent when the number of common factors reaches three, which proves that choosing three common factors is the optimal solution. Through the above method, this paper successfully extracted three common factors from 15 financial variables by using factor analysis. This represents the successful dimensionality reduction of the financial variables.

Table 5. Rotated Component Matrix

Variables	Factor1	Factor2	Factor3
TotalOperatingCost	0.98		
TotalAsset	0.976		
ManagementFee	0.97		
RetainedmentEarning	0.968		
ProfitAvailable	0.907		
BankDeposit	0.853		
Profit	-0.654	0.577	
InsterestIncome	0.572	0.443	
AVGNetProfitMargin		0.864	
NAVGrowth		0.845	
AVGProfit		0.832	
InvestmentIncome		0.629	
ProportionOfFee		-0.554	-0.331
GearingRatio			0.85
TotalLiability	0.5		0.695

Finally, the component matrix was rotated using the maximum variance method. The variables less than 0.3 in the factors were also screened out, making it easier to summarize the characteristics of each factor, and finally, the factor loading matrix plots were obtained. In turn, the relationship between the variables and the common factors can be analyzed by the rotated factor loadings in Table 4. The factors are named according to the characteristics of each variable indicator in the fund and the positive and negative factor loading values.

3.2 Fund Size Factor

As seen from the graph, factor one has a greater weight on the variables X_1, X_3, X_4, X_7, X_{12}, and X_{13}. These variables are all related to the size of a fund, such as total assets and total fees, which can directly reflect the volume of a fund. In addition, variables such as management fees are also sided indicators of a fund's size and performance. To sum up, this paper names factor one as the fund size factor. Funds that perform well on this factor tend to have excellent historical performance and experienced fund managers. This makes this category of funds the ability to manage vast amounts of money.

3.3 Earnings and Growth Factor

The second factor has a greater weight on the variables X_5, X_6, X_8, X_9,X_{10}, and X_{11}. These variables are all related to the profitability and growth of the fund. For example, investment and interest income, which are essential components of a hybrid fund's earnings, can be a good indicator of its profitability. There are also variables like AVGNetProfitMargin and NAVGrowth, which can describe the growth of a fund. In summary, Factor 2 is named Earnings and Growth Factor.

3.4 Debt Factor

Finally, factor three has this larger share on X_2 and X_{15}. These two variables are closely related to the debt profile of a fund. The debt total provides a clear view of the current debt profile of the fund and allows investors to gauge better the risk of buying the fund. The gearing ratio reflects a fund's ability to service its debt over time. Therefore, this paper names it the Debt Factor, which allows investors to quickly understand a fund's debt profile and solvency by looking at the Debt Factor.

Table 6. Component Score Coefficient Matrix

Variables	1	2	3
BankDeposit	0.13	-0.026	0.035
TotalAsset	0.157	0.047	-0.046
GearingRatio	-0.036	-0.035	0.539
TotalLiability	0.062	-0.03	0.43
InsterestIncome	0.112	0.19	-0.265
InvestmentIncome	0.047	0.227	-0.227
ManagementFee	0.148	-0.029	0.026
TotalOperatingCost	0.151	-0.015	0.021
ProportionOfFee	0.014	-0.126	-0.153
RetainedmentEarnin	0.153	0.02	-0.026
Profit	-0.083	0.183	-0.165
AVGProfit	-0.011	0.238	-0.019
ProfitAvailable	0.148	0.058	-0.08
AVGNetProfitMargi	0.014	0.235	0.061
NAVGrowth	0.015	0.229	0.066

The above three factors summarize and refine the individual characteristics of hybrid funds and can provide some reference value for investors. Table 5 shows the final factor score matrix, which can be used to calculate and rank the composite score of each factor. In this paper, we write the expressions for the factor 1 fund size factor as an example.

$$\text{Factor1} = 0.13X1 + 0.062X2 + 0.157X3 + 0.148X4 - 0.083X5 - 0.011X6 + 0.148X7 + 0.014X8 + 0.015X9 + 0.112X10 + 0.047X11 + 0.148X12 + 0.151X13 + 0.014X14 \quad (4)$$

3.5 Discussion

In this paper, using SPSS, the composite score ranking of 90 hybrid funds was obtained using these three factors. As table 7 shows, the top four hybrid funds with composite scores greater than 1 are screened out and recorded in Table 6. These four funds have very good performance in these three aspects and belong to the excellent investment targets among these ninety funds, which can give some reasonable suggestions to investors.

Table 7. Funds Rank

Funds	Factor1	Factor2	Factor3	Score
002190	6.273	0.690	-0.148	3.67
001316	3.858	3.855	-2.751	2.93
001475	3.985	-2.288	1.552	1.73
005730	1.548	0.544	0.452	1.09

Hybrid funds, which are medium to high-risk-return investments, have higher expected risk and expected return levels than money funds and bond funds and lower than stock funds. Through SPSS software, three effective common factors were finally extracted from 15 variables. They are the size factor, profitability factor, and debt service factor. The top four hybrid funds are NCBC Huili New Energy Theme Flexible Mixed Fund, Anxin Stable Value-added Mixed Fund A, Efunda Defense Military Mixed Fund A, and Guotai Jiang Yuan Advantage Select Flexible Mixed Fund A. The top two hybrid funds have a large proportion of the size and profitability factors, which means that their fund size and profitability are better than the other two hybrid funds. The first two hybrid funds have a large proportion of the size factor and profitability factor, which means that their fund size and profitability are better than the other two hybrid funds.

4. Conclusion

China's hybrid investment funds are in a period of massive expansion and development. In this thesis, based on factor analysis for the mixed funds in public funds, 120 random funds were trying to analyze. However, only 90 valid funds were obtained due to the missing information of some hybrid funds, and all data analysis is based on the semi-annual report data from 2022/01/01 to 2022/06/30. Due to the limited sample data. The variables were processed to obtain only three common factors, which are the return factor, growth factor, and debt factor. The data sample is too small compared to the whole number of mixed funds in the market. It is not possible to provide a very detailed data analysis. The final analysis obtained can only be used as a reference analysis for investment in China's mutual funds market and not as an investment guide for investors to use. If there is an opportunity to continue the data analysis in this direction in the future. Firstly, the number of data samples to be analyzed will be expanded to ensure that more accurate data analysis can be provided. Secondly according to the increase in the number of sample data. For a more systematic analysis that can be based on more granular public factors.

References

[1] Zhou Zhizhong, Guo Yan, Zhou Zhicheng. Analysis of market factors affecting the short-term performance of Chinese funds. Forecasting Magazine, 2002, (2): 29-33.

[2] Yang Jingjing, Xhi Jing, Young Martin. Mutual Fund Investment Strategies and Preferences. The Chinese Economy, 2014, 47:1, 5-37.

[3] Wang Shinyun, Lee Cheng-Few. Multi-criteria Decision Making for Evaluating Mutual Funds Investment Strategies. Handbook of Financial Econometrics and Statistics, 2015, 1857-1876.

[4] Chen Carl R., Steve Stockum. Selectivity Market Timming, And RandomBeta Behavior of Mutual Funds: A Generalized Model. Journal of Financial Research. Southern Finance Association; Southwestern Finance Association, 1986, 9(1), 87-96.

[5] Greene Jasong T., Ciccotello Conrad S. Mutual Fund Dilution from Market Timing Trades. Journal of Investment Management, 2006, 4: 31-54.

[6] Qiu Chengyu. Analysis on Application of CAPM Model in Chinese Securities Market and Its Deficiencies. Proceedings of the 2022 2nd International Conference on Enterprise Management and Economic Development (ICEMED 2022), 2022 online ISSN: 2352-5428.

[7] Wei, Jingru. A study on fund evaluation based on factor analysis. Statistics and Application, 2019, (1): 31-38.

[8] Chen Yan, Sun Xiaoguang. Performance evaluation of securities investment funds based on factor analysis. Special Economic Zone, 2005, (10): 108-110.

[9] Zhang Li. Research on comprehensive performance evaluation of funds based on factor analysis. Changsha: Central South University, 2007.

[10] Yu LP, Pan YT, Wu Yishan. Classification of academic journal evaluation indexes based on factor analysis. Library Information Work, 2009, 53(8): 146-149.