ISSN 2616-5902 Vol. 3, Issue 5: 107-117, DOI: 10.25236/AJBM.2021.030517

Stock Selection Model Based on Quality Factor: Evidence from Chinese A-share Market

Qing Ye

School of Management, Shanghai University, Shanghai, 200444, China
YQing6@outlook.com

Abstract: *This paper extracts four variables through the derivation of Gordon's dividend growth model, and constructs the evaluation index system of quality factors from 12 secondary indicators of four single factors of profitability, growth, safety and dividend distribution. We use integration method, mixed method and principal component analysis to construct Quality factor. This paper takes the A-share market as the research object to back-test the stock selection model, and compares and analyzes the robustness, yield and risk of the three models. The results show that the stock selection model of the quality factor under the integration method performs best; and the quality of the company has a greater impact on the excess return of the investment portfolio and less on the risk, but the hedging portfolio can effectively reduce the risk.*

Keywords: *Quality factor, Integrated method, Mixed method, Principal component analysis*

1. Introduction

With value investing becoming one of the mainstream concepts of stock market investment, investors prefer stocks with high intrinsic value when making investments. However, estimating the intrinsic value of a company is a relatively complex task, and the development of quantitative investment has led many scholars and investors to devote themselves to studying the quantitative operation of the value investment concept. At the same time, company quality as a reflection of the intrinsic value of the company, investors are paying more and more attention to it when making investment decisions, and it has become a consensus that high-quality companies have a significant premium in the long term. Recently, a number of scholars have studied the quality of listed companies. Asness et al. (2019) provided a model of firm quality scores with profitability, growth and safety as quality characteristics and formally defined quality factor using a quantitative approach. Subsequently, Libo Yin and Huiyi Liao (2020) proposed a quality score model constructed with four dimensions of profitability, growth, safety, and dividend distribution for the Chinese A-share market, and defined the quality growth factor using the year-over-year change in the quality score. All of the above quality factors are composite factors constructed from multiple single factors. The way in which single factors are compounded can have an impact on the stock selection model based on quality factors, but no scholars have yet explored the way in which quality factors are compounded.

Smart Beta funds, which also use composite factors for stock selection, have been widely used in practice. Smart Beta funds use one or more factors with a risk premium to determine a portfolio to achieve higher returns or lower risk. Smart Beta funds prefer to combine multiple factors to achieve more robust returns than to track single factor products. There are currently two main ideas for constructing composite factors: hybrid and integrated approaches. The hybrid method treats each factor as an independent entity, and each factor selects stocks independently without affecting each other; the integrated method considers the overall stock performance on multiple factors when selecting stocks, and selects stocks according to the high or low overall performance. In addition, a number of scholars have also introduced the principal component analysis (PCA) into the construction of composite factors by extracting principal components through dimensionality reduction of multiple single factors for the composite. Each of the three methods has its own characteristics: the hybrid method is simple and intuitive, the integration method is well integrated, and the PCA is of great interest for its simplicity and effectiveness. However, the industry has not yet formed a unified view on the advantages and disadvantages of the three approaches.

Therefore, we can compare and study the investment returns and risks associated with stock selection by the integration method and the hybrid method from the multi-factor portfolio construction approach,

ISSN 2616-5902 Vol. 3, Issue 5: 107-117, DOI: 10.25236/AJBM.2021.030517

and also construct a stock selection model based on the quality factor by the PCA.

2. Literature Review

Quality factor has been used in investment practice for a long time, and in recent years there has been a consensus among investors that high-quality companies have better long-term performance. In the academic field, although there is a large literature dedicated to the study of some specific aspects of firm quality, there are few theoretical perspectives to define and systematically study quality factors. There is a large literature on the performance of high quality stocks: Novy-Marx (2012), Hou et al. (2015), Ball et al. (2015) find that stocks with high profitability have higher investment returns. Frazzini and Pedersen (2014) define the beta factor and find that low beta is associated with high returns. George and Hwang (2010) argue that firms with low leverage have higher alpha. Campbell et al. (2008) conclude from an empirical analysis that firms with high credit risk tend to underperform. Mohanram (2005) argues that growth companies are better investments.

Unlike other risk factors that are measured using a single indicator, the quality factor is a combination of multiple types of factors related to a company's fundamentals. Moreover, the quality factor was not clearly defined in the early days of academia. The first person to propose a quality factor was Piotroski (2000), who constructed a company quality evaluation model through a combination of nine financial indicators in terms of company profitability, operating efficiency, capital structure and capital security, and constructed a quality factor through a scoring method. Jason Hsu et al. (2019) classified the indicators used to define quality factors into seven categories: profitability, earnings stability, capital structure, growth capacity, investment level, accounting quality, and dividends and dividend increases.

Asness et al. (2019), on the other hand, used the derivation of the formula of Gordon dividend model as the theoretical basis for constructing a firm quality evaluation index system, and constructed a firm quality score model by integrating the corresponding 22 secondary indicators using four dimensions of profitability, growth, safety and dividend distribution as quality characteristics. Asness et al. (2019) bought high quality companies' stocks and sell short low-quality companies' stocks thus constructing a quality factor and empirically study it in 24 developed countries' stock markets around the world, and the results prove that the quality factor investment strategy can achieve excess returns in both developed and emerging markets. Libo Yin and Huiyi Liao (2020) also defined the quality factor from the theory related to the dividend Gordon model and constructed a company quality evaluation model for the Chinese A-share market by integrating a total of 14 secondary indicators in four dimensions: profitability, growth, safety, and dividend distribution. They define quality growth as the year-over-year change in the quality factor to circumvent the influence of seasonal factors, and find that the quality growth factor performs better in stocks with declining quality after the empirical study.

Meanwhile, some scholars have also conducted research on the way composite factor stock selection models are constructed. Bender and Wang (2016) conducted a comparative study of four factors: value, low volatility, quality, and momentum through an integrated method and a hybrid method, and the empirical results showed that the integrated method performs better. Clarke et al. (2016) also support the integration method more, while Fitzgibbons et al. (2017) found that the integration method is superior to the mixed method when the number of factors is high or when there is a negative correlation between the factors. In contrast, Amenc et al. (2017) and Chow et al. (2018) argue that the integration approach sacrifices portfolio stability and that the hybrid approach yields better active management performance. Lester (2019) constructed a theoretical model and derived the ratio of factor exposures for the integrated and hybrid methods. He argued that the expected return of a multi-factor portfolio under the integration method increases with the number of factors while the portfolio risk remains the same, and the portfolio risk under the hybrid method decreases with the number of factors and the expected return remains the same.

In addition, many scholars have introduced machine learning to factor investment and empirical asset pricing, among which the PCA has attracted much attention for its simple and effective advantages. Giglio and Xiu (2019) and Kelly et al. (2019) used the PCA to construct implicit factor models that achieved accurate estimation of factor premiums despite the unobservability of the true factor. Rapach & Zhou (2019) downscaled 120 macroeconomic variables based on principal component analysis, extracted 10 principal components, and found that the 10 principal components could represent the three classical indicators of nominal rate of return, inflation rate, and output rate. They then combined the three factors with the market factor to construct a sparse macro four-factor model and verified that the model has strong explanatory power.

ISSN 2616-5902 Vol. 3, Issue 5: 107-117, DOI: 10.25236/AJBM.2021.030517

In summary, from the perspective of the construction of the quality factor stock selection model, the current literature focuses on stock selection based on relevant financial indicators that reflect the quality characteristics of the company. Since Asness et al. (2019) constructed the quality factor from financial logic, many scholars have carried out the test of applicability of the quality factor stock selection model for different stock markets. However, there are still some gaps in the optimization of this model, and the optimization of the quality factor stock selection model will be a key focus in the field of quality factor research in the future. Based on the application of multi-factor portfolio construction methods and principal component analysis methods in the field of factor stock selection model construction, we select the integration method, hybrid method and the PCA to study the stock selection models of quality factors in order to obtain better investment results.

3. Main Theory

3.1 Evaluation indicators of quality factors

The quality factor is a composite factor reflecting the quality of listed companies, which is essentially a combination of multiple types of evaluation indicators related to company fundamentals. The selection of quality factor evaluation indexes originates from the Gordon dividend model, through which Asness et al. (2019) conclude that the quality of stocks can be reflected by a combination of four dimensions: profitability, growth, safety and dividend distribution of listed companies. The Gordon dividend model assumes that when the dividend growth rate and discount rate are constant, the present value of the firm can be expressed as:

$$P = \frac{D_1}{r - g}$$

Where P is the present value of the stock, D_1 is the first period dividend cash flow, r is the discount rate, and g is the dividend growth rate. By dividing both sides of the equation by the book value of the stock at the same time , we can get:

$$\frac{P}{B} = \frac{Pro/B \times D_1/Pro}{r - g}$$

Where Pro/B is profit. Four variables are extracted from the right side of the equation: Pro/B, D_1/Pro, r and g, which denote the four dimensions of profitability, dividend distribution, discount rate (reflecting safety) and growth of the listed company, respectively.

In order to reflect each dimension more fully, this paper refers to Libo Yin & Huiyi Liao (2020) and collects 14 secondary indicators reflecting the four dimensions of profitability, growth, safety and dividend distribution of listed companies for metrics, which are shown in Table 1:

Table 1: Evaluation indicators of quality factor

Dimension	Secondary indicators
Profitability	1. Return on equity(ROE)= net income / equity of total shareholders 2. Return on assets(ROA)= net profit / total assets 3. Gross profits over assets(GPOA)= total profit / total assets 4. Gross profit margin(GMAR)= total profit / total sales
Growth	1. Growth in return on equity(DROE)= (net income in year t - net income in year t-1) / equity of total shareholders in year t-1 2. Growth in return on assets(DROA)= (net profit in year t - net profit in year t-1) / total assets in year t-1 3. Growth in gross profits over assets(DGPOA)= (total profit in year t - total profit in year t-1) / total assets in year t-1 4. Growth in gross profit margin(DGMAR)= (total profit in year t - total profit in year t-1) / total sales in year t-1
Safety	1. - Beta(-β)= - covariance of stock returns with market returns / variance of market returns 2. - Idiosyncratic volatility(-IVOL)= - residuals obtained using the FAMA-FRENCH three-factor pricing model 3. - Leverage(-LEV)= - total debt / total assers
Payout	1. – Net equity issuance(-EISS)= -ln (number of shares issued in year t / number of shares issued in year t-1) 2. – Net debt issuance(-DISS)= -ln (total debt in year t / total debt in year t-1) 3. Total net payout over profits(-NPOP)= cash dividends per share on common stock / earnings per share on common stock

Among the above secondary indicators, stocks with low beta, low idiosyncratic volatility, low

ISSN 2616-5902 Vol. 3, Issue 5: 107-117, DOI: 10.25236/AJBM.2021.030517

leverage, low stock issuance, and low debt increase are of higher quality, so a negative sign is added to the above indicators before calculation. After completing the calculation of each secondary indicator rank them:

$$R = rank(x)$$

Where x is a secondary indicator of the stock and R is the ranking value of the secondary indicator. Standardization of each secondary indicator based on ranking values:

$$Z_x = \frac{(R - \mu_R)}{\sigma_R}$$

Where u_R is the mean of R, σ_R is the standard deviation of R, and Z_x is the standardized score of the secondary indicator. The four single-factor combinations of profitability, growth, safety and dividend distribution are thus constructed:

$$Profitability = Z(Z_{ROE} + Z_{ROA} + Z_{GPOA} + Z_{GMAR})$$
$$Growth = Z(Z_{DROE} + Z_{DROA} + Z_{DGPOA} + Z_{DGMAR})$$
$$Safety = Z(Z_{-\beta} + Z_{-IVOL} + Z_{-LEV})$$
$$Payout = Z(Z_{-EISS} + Z_{-DISS} + Z_{NPOP})$$

3.2 Stock selection model with quality factor

The quality factor is compounded by multiple secondary indicators of four dimensions of listed companies: profitability, growth, safety and dividend distribution, and the compounding method of secondary indicators will have certain influence on the stock selection model of quality factor. According to the evaluation index system of quality factor, it can be known that the stock selection model of quality factor can be constructed from four single factors using integration method or hybrid method, or the stock selection model can be constructed directly by using 14 secondary indicators through the PCA.

The whole method is a portfolio approach that constructs a composite factor based on the composite performance of a stock on multiple single factors. The integration method maximizes the total exposure of the factors because it considers the composite performance of the stock on multiple single factors. To construct a stock selection model for the quality factor using the integration method, the stock scores on four single factors, namely, profitability, growth, safety and dividend distribution, are added together to obtain the stock scores on the quality factor:

$$Quality = Z(Profitability + Growth + Safety + Payout)$$

The stocks were divided into five groups according to the integrated scores from lowest to highest, noted as P1, P2, P3, P4, and P5, and the hedging portfolio P5-P1 was constructed as a control group (shown in Figure 1).

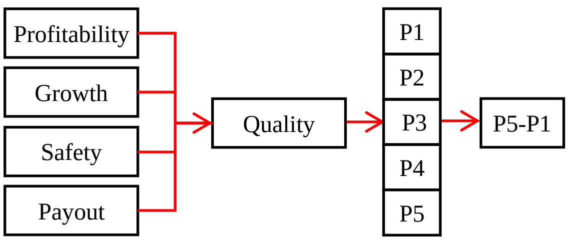

Figure 1: The integration method

The hybrid method is a method of constructing a portfolio by setting stock screening criteria based on the performance of stocks on each single factor individually and then allocating each single factor with a certain weighting. The hybrid method can present the portfolio construction logic in a more intuitive way since the performance of stocks on each single factor is considered separately. The stock selection model with quality factors is constructed using the hybrid method: stocks are divided into five groups according to their scores on the four single factors of profitability, growth, safety and payout from lowest to highest, respectively, and hedged portfolios are constructed, and then the groupings

ISSN 2616-5902 Vol. 3, Issue 5: 107-117, DOI: 10.25236/AJBM.2021.030517

corresponding to the four single factors are combined in equal proportions to construct portfolios P1, P2, P3, P4, P5 and P5-P1 (shown in Figure 2).

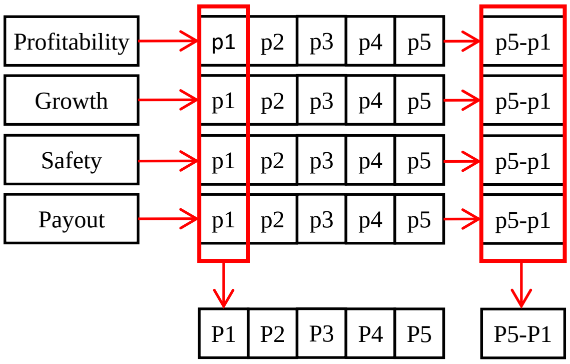

Figure 2: The hybrid method

The principal component analysis(PCA), on the other hand, constructs new factors by extracting the principal components on the covariance matrix of multiple indicators, and then uses the new factors for stock selection. In this paper, the quality factor evaluation index system is known, and since there is a certain correlation among the 14 two-dimensional indicators, the PCA can enhance the exposure of the quality factor by reducing the influence of correlation through dimensionality reduction techniques. The stock selection model of quality factor is constructed by using the PCA: firstly, the applicability test is conducted directly from 14 secondary indicators, and then p ($p \le 14$) principal component factors Y_i are selected by dimensionality reduction, and the quality factor PCA-Quality is constructed with its information contribution rate b_i as the weight:

$$PCA-Quality = \sum_{i=1}^{p} b_i Y_i$$

Based on the quality scores under the PCA, the stocks were likewise grouped from lowest to highest and a control group was set up (shown in Figure 3).

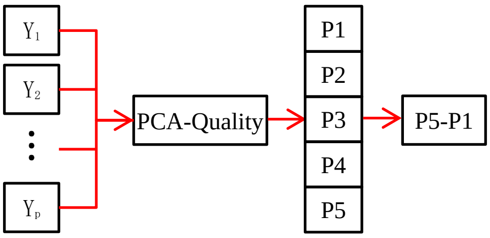

Figure 3: The PCA

We use market capitalization as the weight and adjust positions quarterly to backtest the sample with historical A-share data. Then we compare and analyze the average annualized return, return volatility and Sharpe ratio and other related indicators of the quality factor stock selection model under the integrated method, hybrid method and the PCA.

4. Data screening and pre-processing

We use all the nearly 4,000 A-shares in the Shanghai and Shenzhen markets in China as the research target, excluding abnormal stocks such as ST, sub-prime stocks listed for less than 6 months and stocks in the financial sector, then we collect and collate daily trading data and relevant financial data of each stock as the research sample, which are obtained from the CSMAR database. Considering the impact of the equity share reform, the sample interval is selected from January 2007 to August 2020, and the corresponding financial data are selected from quarterly reports from 2006 to 2020. Fourteen secondary indicators, including return on shareholders' equity, are calculated and collated for each stock according to the evaluation index system of quality factors, and then the sample data of the 14 secondary indicators are pre-processed, including depolarization, standardization and neutralization, among which the standardization process is the Z-value method, which has been explained in the main theory mentioned previously.

ISSN 2616-5902 Vol. 3, Issue 5: 107-117, DOI: 10.25236/AJBM.2021.030517

We use the MAD absolute deviation method for depolarization, which can robustly overcome the effect of sample dispersion. The method uses the median and absolute deviation MAD of the sample data for each indicator to detect outliers, where MAD is calculated as follows:

$$MAD = median(|x - median(x)|)$$

Where $median(x)$ is the median of the secondary index x. According to the MAD absolute deviation method we perform outlier detection on the sample data and considered data on the interval $[meidan(x) - n * 1.4826 * MAD,\ meidan(x) + n * 1.4826 * MAD]$ as normal. We generally consider $n = 3$, and outliers above or below this interval are replaced with the upper or lower line of the interval.

We neutralize the sample data after depolarization and standardization, and then remove the effects of market capitalization and industry factors. We run a linear regression on the log of market capitalization and industry dummy variables, and then extract the residuals from the regression as the pre-processed sample data, calculated as follows:

$$x' = \beta_M * \ln(MktVal) + \sum_{j=1}^{n} \beta_j * Industry_j + \varepsilon$$

Where x' is the sample data of a stock in a secondary indicator after going to extremes and normalization, $MktVal$ is the total market value of the stock, $Industry_j$ is the 0-1 dummy variable of the industry, which means that $\beta_j = 1$ if the stock is classified as industry j in the Shen Wan primary industry classification, otherwise $\beta_j = 0$. In addition, ε in the equation means the extracted residual value.

5. Empirical analysis

5.1 The PCA

Before conducting the robustness test and model backtesting, we need to refine the quality factor stock selection model under the PCA, which means we have to test the applicability of thr PCA on the sample data and construct the stock selection factor PCA-Quality.

5.1.1 Applicability test

For multi-level and multi-indicator quality factor evaluation index system, not all the index data are suitable for the PCA. Therefore, before conducting the PCA, suitability tests are performed. The common test methods are:

(1) KMO Test (Kaiser-Meyer-Olkin)

This test focuses on comparing the relative magnitudes of the simple and partial correlation coefficients. If the sum of squares of simple correlation coefficients among all variables is much larger than the sum of squares of bias correlation coefficients, the smaller the bias correlation coefficient among variables, the closer the KMO value is to 1, and the more suitable the original variable data is for the PCA. The value of KMO is between 0 and 1. Kaiser believes that a KMO value greater than 0.05 indicates that the sample data are suitable for thr PCA.

(2) Bartlett Spherical Test (Bartlett Test of Sphericity)

The null hypothesis of this test is that the correlation matrix is a unit matrix. If the null hypothesis is rejected, it is considered that the correlation matrix is not a unit matrix and there is correlation between the original variable data. It is generally considered that if its probability of companionship (Sig) is less than the significance level (usually 0.05), the null hypothesis is rejected and the original variable data are suitable for the PCA.

The results of the tests using KMO and Bartlett are shown in Table 2:

ISSN 2616-5902 Vol. 3, Issue 5: 107-117, DOI: 10.25236/AJBM.2021.030517

Table 2: Applicability test results

Kaiser-Meyer-Olkin metric when sampling is sufficient		0.780
Bartlett's test for sphericity	Approximate cardinality	85.314
	df	37.00
	Sig.	0.00

As shown in Table 2, the KMO value is 0.780, which is greater than 0.5, referring to the data in the table indicates that this group of indicators can be used for the PCA. The Bartlett's sphericity test results in a companion probability of 0.000, which is less than the significance level of 0.05, which means that the null hypothesis of Bartlett's sphericity test is rejected and therefore considered suitable for the PCA.

5.1.2 Calculate the correlation coefficient matrix

Because we have normalized the data of 14 secondary indicators $\tilde{x}$ after pre-processing, the correlation coefficient matrix is found as follows:

$$r_{ij} = \frac{1}{14-1}\sum_{k=1}^{14} \tilde{x}_{ki} * \tilde{x}_{ki}, \quad i, j = 1, 2, \cdots, 14$$

Where $r_{ij} = 1, \ r_{ij} = r_{ji}$, and is the correlation coefficient between the ith secondary index and the jth secondary index.

5.1.3 Calculate information contribution rate and cumulative contribution rate

We first calculate the eigenvalues λ_i of the correlation coefficient matrix and the corresponding normalized eigenvectors u_i, and then calculate the information contribution rate b_i and the cumulative contribution rate a_p according to the following equations:

$$a_p = \sum_{k=1}^{p} \lambda_k \Big/ \sum_{k=1}^{14} \lambda_k$$

Where $p = rank(b_i)$, i.e., the principal component Y_i is ranked by the magnitude of information contribution.

5.1.4 Select p (p≤14) principal components to construct PCA-Quality

When a_p is greater than 85%, the first p are selected as principal components instead of the original 14 secondary indicators, and a comprehensive analysis of the p principal components is performed to obtain the following table 3:

Table 3: Results of the PCA

Serial number	Contribution rate	Cumulative contribution rate	Serial number	Contribution rate	Cumulative contribution rate
1	29.027%	29.027%	8	4.734%	94.083%
2	22.992%	52.018%	9	2.149%	96.232%
3	10.125%	62.144%	10	1.975%	98.207%
4	7.909%	70.053%	11	0.860%	99.067%
5	7.165%	77.218%	12	0.531%	99.599%
6	6.698%	83.916%	13	0.251%	99.849%
7	5.434%	89.350%	14	0.151%	100.000%

It can be seen that the cumulative contribution of the first 7 principal components alone reached 85%, so the first 7 principal components were selected for the construction of PCA-Quality:

$$PCA-Quality = 29.027\%Y_1 + 22.992\%Y_2 + 10.125\%Y_3 + 7.909\%Y_4 + 7.165\%Y_5 + 6.698\%Y_6 + 5.434\%Y_7$$

ISSN 2616-5902 Vol. 3, Issue 5: 107-117, DOI: 10.25236/AJBM.2021.030517

5.2 Robustness tests

In order to test the effectiveness of the three stock selection models, we use IC and IR to test the robustness of the stock selection factors in the three stock selection models. IC, the information coefficient, is the correlation coefficient between the factor value of a stock and the stock's next period return after data preprocessing. The mean value of IC evaluates the ability of the factor value to predict the next period return: the larger the absolute value of the mean value of IC, the stronger the stock selection ability of the factor. IR, the information ratio, is the mean value of excess returns divided by its standard deviation. IR evaluates the stability of a factor's stock selection ability: the larger the IR, the more stable the factor's stock selection ability. A stepwise derivation from excess returns reveals that IR can be approximated by IC as follows:

$$IR \approx \frac{\overline{IC}}{std(\overline{IC})}$$

Where $\overline{IC}$ is the mean value of IC and $std(\overline{IC})$ is the standard deviation of IC. It is generally believed that when the absolute value of IC mean is greater than 0.03, the stock selection ability of the factor is stronger; when IR is greater than 0.5, the stock selection ability of the factor is more stable. The specific test results are shown in Table 4:

Table 4: Robustness test results

Factor	Mean value of IC	Standard deviation of IC	IR	Mean value of p-value	IC significant percentage
Quality	0.070	0.103	0.680	0.080	69.231%
Profitability	0.042	0.121	0.346	0.093	76.923%
Growth	0.060	0.057	1.038	0.193	30.769%
Safety	0.051	0.100	0.507	0.056	69.231%
Payout	0.028	0.061	0.456	0.230	53.846%
PCA-Quality	0.056	0.117	0.481	0.057	85.366%

In Table 4, Quality is the stock selection factor used in the quality factor stock selection model under the integration method; Profitability, Growth, Safety, and Payout are the stock selection factors used in the quality factor stock selection model under the mixed method; and PCA-Quality is the stock selection factor used in the quality factor stock selection model under the PCA. As can be seen from Table 4, the mean IC values of all six factors are greater than zero, which means that there is a positive correlation between the IC values of all six factors and the stock's next period return. Numerically, only the IC mean value of Payout factor of the hybrid method is slightly less than 0.03, which indicates that all three quality factor stock selection models have stronger stock selection ability, specifically the integration method is the strongest in terms of stock selection ability, and the PCA is slightly stronger than the hybrid method.

In terms of the stability of stock selection ability, the IR of Quality, Growth and Safety are all greater than 0.5, with Growth having the largest IR. Therefore, Growth has the most stable stock selection ability, while Profitability, which has the smallest IR, has the least stable stock selection ability. Overall, the integrated method is slightly stronger than the mixed method in terms of the stability of the quality factor stock selection model, and the PCA is the most unstable.

In addition, we conducted t-tests on the IC values of the factors to investigate the significance of the IC values. Since the sample interval is from January 2007 to August 2020 and the transfer period is one quarter, we made a total of 41 transfers. The p-value means in Table 4 are the means of the 41 IC values corresponding to the p-values. IC is considered significant when $p \leq 0.05$. In terms of the percentage of IC significant, the overall quality factor stock selection model under the PCA is the most significant.

5.3 Backtest results

We conduct an empirical study according to the three aforementioned quality factor stock selection models, using historical trading data from January 2007 to August 2020 for backtesting. Comparing and analyzing the return indicators and risk indicators of the three stock selection models, the backtest results are shown in Table 5:

The average return is the average return for each portfolio position; the return volatility is the quarterly volatility of the return for each position; the annualized Sharpe ratio is converted from the

ISSN 2616-5902 Vol. 3, Issue 5: 107-117, DOI: 10.25236/AJBM.2021.030517

quarterly excess return; and the total return at the end of the period is the final total return of the portfolio.

Table 5: Backtest result

	P1	P2	P3	P4	P5	P5-P1
The integration method						
Average return	0.020	0.018	0.026	0.036	0.055	0.017
Earnings volatility	0.235	0.232	0.233	0.233	0.206	0.041
Annualized sharpe ratio	0.147	0.137	0.192	0.269	0.461	0.735
Total return at period end	-0.094	-0.134	0.184	0.777	3.373	0.960
The hybrid method						
Average return	0.024	0.030	0.0337	0.037	0.045	0.011
Earnings volatility	0.231	0.230	0.231	0.226	0.210	0.026
Annualized sharpe ratio	0.181	0.222	0.253	0.286	0.373	0.695
Total return at period end	0.101	0.395	0.641	0.945	1.955	0.517
The PCA						
Average return	0.023	0.023	0.025	0.038	0.051	0.014
Earnings volatility	0.231	0.233	0.239	0.235	0.213	0.047
Annualized sharpe ratio	0.170	0.172	0.183	0.277	0.414	0.526
Total return at period end	0.049	0.058	0.125	0.883	2.562	0.707

Comparing each portfolio in the three stock selection models, we find that the P5 group has the highest scores for both the average return and the total end-of-period return, and there is an overall trend of increasing returns from P1 to P5, which implies that high-quality stocks are more likely to earn excess returns. In contrast, the control group P5-P1 has relatively lower scores for average return and total end-of-period return, indicating that the hedged portfolios under the three stock selection models do not obtain as much return as the highest quality portfolios. On the return volatility indicator, the difference between the scores of the P1 to P5 portfolios is small, compared to the lower score of P5-P1, indicating that the high quality has less impact on the return volatility of the portfolio; however, the portfolio formed by hedging using high and low quality is effective in reducing the volatility in returns. On the annualized Sharpe ratio metric, the scores of portfolios P1 to P5 show an increasing trend, but P5-P1 has the highest score, indicating that the hedged portfolio can achieve the highest return for the same unit of risk taken, and that high-quality stocks achieve higher returns than low-quality stocks. Overall as shown in the cumulative return curves in Figure 4, Figure 5 and Figure 6, portfolio P5 performs best in terms of return, while the cumulative return curve for portfolio P5-P1 is the smoothest, i.e. the hedged portfolio performs best in terms of risk.

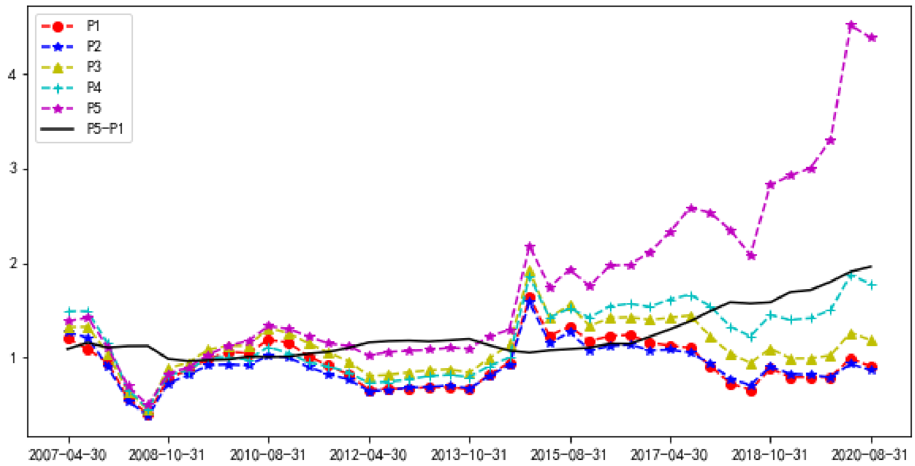

Figure 4: Cumulative return of the stock selection model by the integration method

ISSN 2616-5902 Vol. 3, Issue 5: 107-117, DOI: 10.25236/AJBM.2021.030517

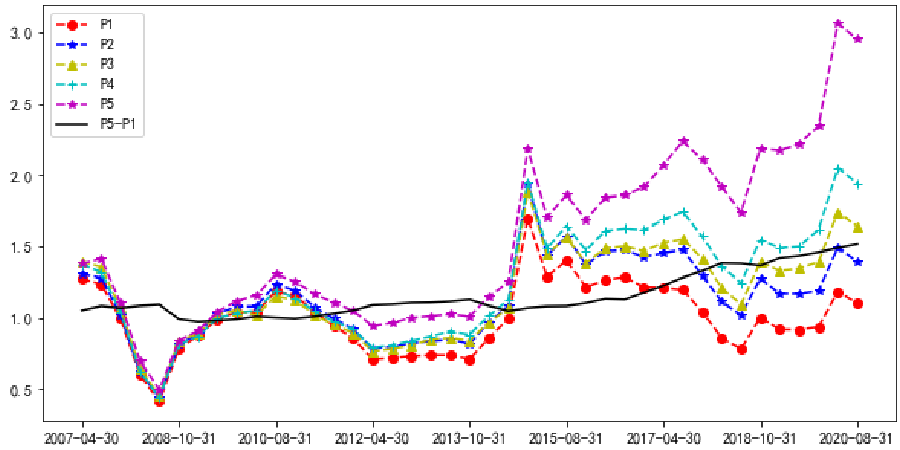

Figure 5: Cumulative return of the stock selection model by the hybrid method

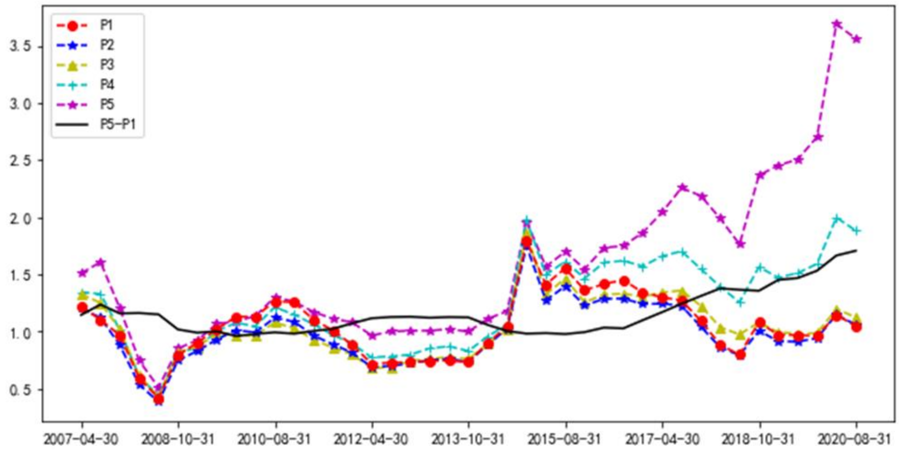

Figure 6: Cumulative return of the stock selection model by the PCA

Then, we compare and analyze the advantages and disadvantages of the three stock selection models in terms of return and risk to determine the best quality factor stock selection model. If the portfolio is to ensure the least risk while maximizing the return, we have to find the portfolio with the least risk among the three stock selection models in portfolio P5. From Table 5, it can be seen that the integrated method stock selection model in portfolio P5 has the lowest return volatility and the highest annualized Sharpe ratio, so the quality factor stock selection model under the integrated method is optimal in terms of risk control while maximizing returns. If the portfolio is to maximize the return with minimum risk, we have to find the combination of the three stock selection models with the highest return in the portfolio P5-P1. As shown in Table 5, the average return and the total return at the end of the period of the integrated method stock selection model in portfolio P5-P1 are the highest, so the quality factor stock selection model under the integrated method is still the best in terms of return acquisition with minimum risk.

6. Conclusion

We study the quality factor stock selection model from the construction method of quality factor. Firstly, the evaluation index system of quality factor is constructed according to the financial logic, and then the stock selection model of quality factor is constructed by using the integration method, the hybrid method and the PCA respectively. Taking Shanghai and Shenzhen A-shares as the research object, the sample interval of January 2007-August 2020 is selected for the empirical analysis of the stock selection models, and the three stock selection models are compared and analyzed in terms of robustness, return and risk.

Based on IC mean, IR and IC significant percentage to compare the stock selection ability, stock selection stability and stock selection significance of the three stock selection models, we find that the integration method is the strongest in stock selection ability and stock selection stability, but the PCA is the most significant in stock selection significance. The returns of the quality factor stock selection model

ISSN 2616-5902 Vol. 3, Issue 5: 107-117, DOI: 10.25236/AJBM.2021.030517

are reflected by average returns and total returns at the end of the period, specifically finding that high quality stocks are more likely to earn excess returns, while hedged portfolios do not earn as much as the highest quality portfolios. The risk of the quality factor stock selection model is reflected by the return volatility and the annualized Sharpe ratio. It is specifically found that high quality has less impact on the risk of the portfolio, while the hedged portfolio is effective in reducing the risk.

Combining the advantages and disadvantages of the three stock selection models in terms of return and risk, our analysis reveals that the stock selection model with the quality factor under the integration method performs best, whether it is to ensure minimum risk while maximizing return, or to maximize return while minimizing risk.

References

[1] Amenc, Ducoulombier, Frédéric, Esakia M, et al. Accounting for Cross-Factor Interactions in Multifactor Portfolios without Sacrificing Diversification and Risk Control [J]. The Journal of Portfolio Management, 2017, 43(5): 99-114.
[2] Asness C.S., Frazzini A., Pedersen L. H. Quality minus junk [J]. Review of Accounting Studies, 2019, 24(1): 34-112.
[3] Ball, R., J. Gerakos, J. T. Linnainmaa, and V. Nikolaev. Deflating profitability [J]. Journal of Financial Economics, 2015, 117(2): 225-248.
[4] Bender, J. and T. Wang. Can the whole be more than the sum of the parts? Bottom-up versus top-down multifactor protfolio construction [J]. The Journal of Portfolio Management, 2016, 42(5): 39-50.
[5] Campbell, J. Y., Hilscher, J. and Szilagyi, J. In search of distress risk [J]. Journal of Finance, 2008, 63(3): 2899-2939.
[6] Chow, T.-M., F. Li, and Y. Shim. Smart beta multifactor construction methodology: Mixing versus integrating [J]. Journal of Index Investing, 2018, 8(4): 47-60.
[7] Clarke, Roger, Harindra De Silva, and Steven Thorley. Fundamentals of Efficient Factor Investing [J]. Financial Analysts Journal, 2016, 72(6): 9-26.
[8] Fitzgibbons, Shaun, Jacques Friedman, Lukasz Pomorski, and Laura Serban. Long-only Style Investing: Don't Just Mix, Integrate [J]. Journal of Investing, 2017, 26(4): 153-164.
[9] Frazzini, A. and L. H. Pedersen. Betting against beta [J]. Journal of Financial Economics, 2014, 111(1): 1-25.
[10] George, T. J. and Hwang, C. Y. A resolution of the distress risk and leverage puzzles in the cross section of stock returns [J]. Journal of Financial Economics, 2010, 96: 56-79.
[11] Giglio, S. and D. Xiu. Asset pricing with omitted factors [J]. Chicago Booth Research Paper, 2019, 16-21.
[12] Hou, K., C. Xue, and L. Zhang. Digesting anomalies: An investment approach [J]. Review of Financial Studies, 2015, 28(3): 650-705.
[13] Jason Hsu, Vitali Kalesnik and Engin Kose. What Is Quality? [J]. Financial Analysts Journal, 2019, 75(2): 44-61.
[14] Kelly, B. T., S. Pruitt, and Y. Su. Characteristics are covariances: A unified model of risk and return [J]. Journal of Financial Economics, 2019, 134(3): 501-524.
[15] Lester, Ashley. On the Theory and Practice of Multifactor Portfolios [J]. Journal of Portfolio Management, 2019, 45(4): 87-100.
[16] Libo Yin, Huiyi Liao. Firm's Quality Increases and the Cross-Section of Stock Returns: Evidence from China [J]. International Review of Economics and Finance, 2020, 66: 228-243.
[17] Mohanram, P. Separating winners from losers among low book-to-market stocks using financial stateent analysis [J]. Review of Accounting Studies, 2005, 10(2-3): 133-170.
[18] Novy-Marx, R. Quality investing [R]. Working paper, Rochester, 2012.
[19] Piotroski, J. D. Value investing: The use of historical financial statement information to separate winners from losers [J]. Journal of Accounting Research, 2000, 38: 1-41.
[20] Rapac, D. E. and G. Zhou. Sparse macro factors [D]. Saint Louis University, Washington University in St. Louis, 2019.