Journal of Engineering, Project, and Production Management
2023, 13(1), 76-80

Risk Management of Enterprise Quantitative Investment Strategies through Data Modeling

Weizheng Wu
Lecturer, Huanghuai University, No. 76, Kaiyuan Avenue, Zhumadian, Henan 463000, China, E-mail: uwa8y2@163.com

Project Management
Received April 2, 2022; revised June 22, 2022; October 11, 2022; accepted October 12, 2022
Available online November 21, 2022

Abstract: Quantitative investment strategies have been increasingly used in the capital market. In order to help enterprises manage risks better and enhance the reliability of quantitative investment strategies, this paper designed a quantitative investment data model for Enterprise A. The constituent stocks of the Shanghai and Shenzhen 300 index were regarded as the stock pool. Then, factor screening was performed in the Uqer quantitative factor library. Two data models were established: a scoring model and a regression model. The two models were tested through backtesting. The return and risk were compared between the two models by taking the return rate, net value, Alpha, Beta, Sharpe ratio, maximum retracement and information ratio as the evaluation indicators. The backtest results showed that the data model established by the regression method had a higher return rate, annualized return rate, net value, larger α value, smaller β value, a Sharpe ratio of 0.76, a maximum retracement of 25.34%, and an information ratio of 2.42, which had better balance in return and risk compared with the scoring method. In addition, the larger the number of positions was, the smaller the frequency of position transfer was, and the less effective the model was in quantitative investment. The experimental results verify the reliability of the regression model in the formulation of the quantitative investment strategy of Enterprise A. The investment strategy of Enterprise A can be adjusted and determined by the regression model to promote the balance between enterprise benefits and risks. The research results provide some references for the theoretical research of quantitative investment.

Keywords: Risk management, data model, factor screening, yield rate, quantitative investment enterprise.

DOI 10.32738/JEPPM-2023-0008

1. Introduction

Quantitative investment refers to simulating an investor's philosophy through tools such as computers, statistics and artificial intelligence to obtain an executable data model. The traditional way of investment makes decisions based on the judgment of the profitability indicators of investment assets, relying more on the investor's ability and intuition, while quantitative investment constructs models using tools such as computers and then executes investments according to the models, which is more rationalized and objective (Ma, 2020). Quantitative investment strategies can be divided into three types: ① stock strategies: according to different interventions (human and computers), stock strategies are divided into two types: active equity investment (investors screen stocks after researching the market, subdivided into long/short strategy, short selling strategy, etc. (Kim, 2018)) and active quantitative investment (build portfolio models relying on computer big data analysis); ② macro strategy (commodity trading advisor strategy) (Sachs and Tiong, 2016): it analyzes the stock market based on behavioral finance theory to obtain high returns, e.g., the momentum inversion model; ③ arbitrage strategy (Wang et al., 2020): it obtains returns from fluctuations in market spreads, such as stock index future arbitrage (Biakowski and Perera, 2019) and option arbitrage (Laurini, 2015). Currently, two types of quantitative trading strategies are used in the capital market. One is the multi-factor stock selection model (Pan and Long, 2021). It considers that stock returns are determined by some factors. Factors are screened and combined to buy satisfactory stocks. The other is the high-frequency trading strategy (Chen et al., 2016): based on the iceberg algorithm, large orders are split into small orders to reduce large stock floats. Quantitative investment strategies have received more and more attention from researchers in recent years (Suhonen et al., 2017). Li et al. (2020) analyzed the evolutionary stability strategy (ESS) of quantitative investment strategy. They found through the study of Chinese stock index futures that a single strategy could not survive in the market, and the trading frequency should be appropriately reduced to cut transaction costs. Tang et al. (2019) analyzed the risk and return of the securities market with deep learning, trained and analyzed the model with the

Dow Jones Industrial Index and Shanghai Stock Exchange Index, and found that the model had good investment return and robustness. Wei and Watada (2016) established a T2 fuzzy stochastic support vector regression (T2 - FSVR) model to analyze market buying and selling behaviors for predicting effective trading strategies. Through testing on the MATLAB platform, they found that the model learned and automatically profited from a large amount of historical and real-time data and developed into a practical automated trading strategy. Wang et al. (2020) combined the wheel effect with random forest to build a stock selection model. They found through the experiment that the annualized return rate of the strategy was 3.6% higher than that of the single-round strategy. Marks and Shang (2019) have pointed out that the change, liquidity, and volatility of trading activities reflected by stocks can guide stock selection. Enterprise A is a quantitative investment enterprise, but its current quantitative investment strategy has some problems. In order to find out a more suitable quantitative investment strategy to help Enterprise A improve benefits and reduce risks, the multi-factor stock selection model has been extensively applied to predict the long-term stock price trend (Yuan et al., 2020). Pan and Long (2021) designed a multi-factor model for the food and drinks industry of A-share market, selected seven effective non-redundant factors from 20 candidate factors belonging to six classes to predict the future benefits of every stock in this industry, and performed an investment portfolio. However, the current quantitative investment strategies still have shortcomings, such as single strategy and strategy divergence. To study the multi-factor stock selection model deeper and find out the new quantitative investment method, based on the previous studies, this paper briefly analyzed the development status of Enterprise A, designed a multi-factor stock selection data model, and validated the return and risk of the model through backtesting to understand the effectiveness of the data model for risk management. Our paper makes some contributions to improve the quantitative investment strategy of enterprises, is beneficial to enriching the quantitative investment strategy, and further promotes the development of quantitative investment.

2. Research Method

2.1. Research Subject

Enterprise A adopts a linear management structure and integrates investment and research. Its investment deal relies on an automation system. It mainly uses computer and data models and guides its trading with investment models. The current investment strategies include stock index future arbitrage, commodity trading advisor strategies, stock selection strategies, etc. However, the multi-factor models used have high out-of-sample failure rates and large discrepancies between the real price and backtest result, leading to large investment risks. Moreover, the automated investment method has large operational risks. Therefore, in order to manage the investment risk of Enterprise A better, a stock selection data model was designed.

2.1. Candidate Factors and Stock Pools

Data used in the research came from the Uqer quantitative platform (https://uqer.datayes.com/). The Uqer quantitative platform was designed for researchers who study quantification, providing financial, factor, thematic, and macro industry-specific big data for all types of assets. It also provides 400+ quantitative factor libraries and can perform backtesting on strategies. Constituent stocks of the Shanghai and Shenzhen 300 index were used as the candidate stock pool in this study, and a quantitative investment data model of Enterprise A was established based on the historical data. The 244 factors in the Uqer quantitative factor pool were used as candidate factors, the sample interval was 2009.1.1-2020.12.31, and positions were transferred at the end of every month. Factors were screened using data between 2009.1.1 and 2010.12.31, and model testing was performed using data between 2011.1.1 and 2020.12.31. The initial capital was 10 million yuan. The number of stocks in the quantitative investment strategy was 100. From 2011 to 2020, the stock market experienced the initial public offerings (IPO) halt in 2013, the stock market crash in 2015, the COVID-19 shock in 2019; thus, the data were effective in testing the role of the model in risk management.

Factors with missing values greater than ten were eliminated. Then, missing values were filled using the median for the remaining 202 factors, abnormal values were processed using the mean-standard deviation method, and the data were normalized using the min-max method.

2.2 Factor Screening

Candidate factors were screened by factor information coefficient (IC) value ranking and factor information ratio (IR) value ranking.

(1) IC value: it reflects the influence of factors on the stock return rate in the next period. The larger the absolute value of IC value is, the better the prediction of the factor on the future return of the stock is. Generally, IC>3% is valid. Rank IC was used, and its expression is shown in Eq. (1):

$$\text{Rank IC} = \mathrm{corr}\left(\mathrm{order}_{t-1}^{f} - \mathrm{order}_{t}^{r}\right) \tag{1}$$

where $order_{t-1}^{f}$ is the ranking of stock factor value in period t-1 and $order_{t}^{r}$ is the stock return ranking in period t.

(2) IR value: it reflects the ability level of factors in capturing excess earnings, which can be obtained by approximately calculating the IC value. The relevant calculation formulas are shown in Eqs. (2) and (3):

$$\mathrm{IC} = \mathrm{corr}(\mathrm{f}_{t-1} - \mathrm{r}_{t}) \tag{2}$$

$$\mathrm{IR} \approx \frac{\overline{\mathrm{IC}_t}}{\mathrm{std}(\mathrm{IC}_t)} \tag{3}$$

where f_{t-1} is the value of the stock factor in period t-1, r_t is the rate of return in period t, and $std(IC_t)$ is the standard variance of the IC value in period t.

The factors in the union set of the top ten factors under the two rankings was taken as the valid factors for establishing the data model. There were 15 factors, as shown in Table 1.

3. Design and Analysis of Quantitative Investment Strategies Based on Data Models

3.1. Establishment of the Data Model

A data model was established using the following two methods in combination with the effective factors in Table 1.

Table 1. Effective factors

Factor code	Factor name
DDNSR	Downward fluctuations
HSIGMA	Historical fluctuations
REVS20	Stock's 20-day earnings
VOL5	5-day average turnover rate
VOL10	10-day average turnover rate
VOL20	20-day average turnover rate
VOL60	60-day average turnover rate
VOL120	120-day average turnover rate
ATR6	6-day average true range
BIAS20	10-day departure rate
BIAS60	60-day departure rate
ROC20	20-day rate of change
STM	ADTM factor intermediate variables
DIFF	MACD factor intermediate variables
MA10RegressCoeff 6	10-day moving average 6-day linear regression coefficient

(1) Scoring method: different factors of every stock in the stock pool were ranked in descending order and scored. The first to the last was given one point, the second to the last was given two points, and so on. The fifteen scores of the fifteen factors were totaled up.

Then, the stocks were ranked in descending order according to the total score, and the top 20% of stocks and the bottom 20% of stocks were taken to build the stock pool.

(2) Regression method: the future returns of the stocks were predicted using the forward stepwise regression algorithm, and stocks whose forecast value ranked in the top 20% and bottom 20% were taken to establish a stock pool.

3.2. Model Inspection Standards

The model test consisted of two parts, the test of the rate of return and the test of the risk profile. The test of the return rate includes the following criteria:

(1) Rate of return: the ratio of investment return to principal.

(2) Annualized return rate: the yield that would be earned if the portfolio were held for only one year.

(3) Net value: the actual market value of the portfolio held at a time point.

The test of the model risk profile includes the following criteria:

(1) Alpha: the excess return that can be earned on a portfolio's excess return minus compensation for systematic risk;

(2) Beta: the return earned by a portfolio for taking a systematic risk;

(3) Sharpe ratio (Kaplanski et al., 2016): the ratio of the portfolio's return rate to the market's risk-free return rate; the higher the value is, the more effective the investment strategy is;

(4) Maximum retracement: the limit of loss that an investor can tolerate;

(5) Information ratio: a measurement of the excess return brought by excess risk; higher values indicate higher excess returns.

4.3 Analysis of Backtest Results

The backtest results of the two data models are shown in Table 2.

Table 2. Backtest results of the data model

	Scoring method	Regression method
Return rate	368.79%	389.64%
Annualized return rate	12.67%	12.69%
Net value	4.126	4.978
A	0.0126	0.0132
B	1.135	1.128
Sharpe ratio	0.71	0.76
Maximum retracement	26.78%	25.34%
Information ratio	2.39	2.42

It was seen from Table 2 that the return rate, annualized return rate, and net value of the regression method were higher than the scoring method, the α value was larger than the scoring method, and the β value was smaller than the scoring method, indicating that the investment management ability of the regression method was better. The Sharpe ratio of the regression method was 0.76, which was slightly higher than that of the scoring method. In addition, the maximum retracement of the regression method was 25.34%, which was smaller than that of the scoring method. The information ratio of the regression method was 2.42, which was greater than the scoring method. Overall, the data model based on the regression method had a better performance.

A data model was established using the regression method to compare the effect of different numbers of positions held on the return and risk profile. The results are shown in Table 3.

Table 3. Model backtest results under different number of positions

Number of positions	20	60	100
Rate of return	497.65%	421.56%	389.64%
Annualized return rate	14.97%	13.21%	12.69%
Net value	6.452	5.215	4.978
A	0.0145	0.0138	0.0132
B	1.117	1.119	1.128
Sharpe ratio	0.91	0.87	0.76
Maximum retracement	23.64%	24.16%	25.34%
Information ratio	2.51	2.49	2.42

It was seen from Table 3 that when the number of positions held was 100, compared to the rate of return when it was 20, the rate of return of the model decreased by 108.01%, the annualized return rate decreased by 1.98%, the net value decreased by 1.664, the α value decreased, the β value increased, the Sharpe ratio decreased from 0.91 to 0.76, the maximum retracement increased by 1.7%, and the information ratio decreased by 0.09.

A model was established using the regression method. When one hundred positions were held, the effect of different position transfer frequencies on the return and risk profile was compared, and the results are shown in Table 4.

Table 4. Model backtest results under different position transfer frequencies

Frequency of position transfer	Monthly	Every three months	Every six months
Rate of return	389.64%	378.64%	365.28%
Annualized return rate	12.69%	11.64%	11.21%
Net value	4.978	4.872	4.123
α	0.0132	0.0131	0.0131
β	1.128	1.129	1.131
Sharpe ratio	0.76	0.55	0.46
Maximum retracement	25.34%	26.12%	26.89%
Information ratio	2.42	2.33	2.01

According to Table 4, the model initially adopted the monthly position transfer mode. With the decrease of the position transfer frequency, the rate of return of the model decreased. When the frequency of position transfer was once every six months, the rate of return of the data model decreased by 24.36%, the net value decreased by 0.855, and the Sharpe ratio also showed a significant decrease, from 0.76 to 0.46.

4. Discussion

This paper studied the quantitative investment strategy of Enterprise A on a Uquer quantitative platform. The backtest results of the two models established by the scoring method and the regression method were compared. It was found that the data model established using the regression model obtained more revenue compensations per unit of risk and achieved higher returns under lower risks, indicating that the model established by the regression method had a better backtesting effect and was more suitable for the quantitative investment risk management of Enterprise A. Then, the analysis of the specific strategy suggested that as the number of positions held increased, the rate of return of the model decreased, and the risk increased. The model backtest results showed that as the number of positions held increased, the possibility of making mistakes in stock selection also increased; therefore, the number of positions held should be controlled within a reasonable range in the quantitative investment process. It was also found from the results that extending the position transfer time was not beneficial to improving returns and reducing risks, which might be because the decreased frequency of position transfer declined the time-effectiveness of the model and made it unable to grasp market changes timely.

Compared with the previous studies, this paper further analyzed the multi-factor stock selection model, which expands the ideas for more extensive research on multi-factor models in the future. At the same time, it also analyzed the influence of specific strategies on the stock selection model and extended and optimized the models to provide some references for the decision-making of quantitative investment companies, which is conducive to promoting the further development of the quantitative investment industry. However, the model also has some shortcomings, such as the ignored consideration of transaction fees and incomprehensive selection of factors, which needs improvement in future work.

5. Conclusion

This paper studied the quantitative investment strategy for Enterprise A. Data models were established through screening factors using the scoring and regression methods to perform quantitative investment. The backtesting found that the data model established by the regression model obtained large returns with low risks, showing better performance in quantitative investment. In addition, the number of positions held and the frequency of position transfer also had impacts on the effectiveness of the model, so they need timely adjustments when a strategy is selected in reality. However, factor selection was not perfect in this paper, and the model was established based on historical data, which resulted in hysteresis. In future studies, there is a need to expand the selection of factors further and conduct experiments on more comprehensive data.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Institutional Review Board Statement

Not applicable.

References

Biakowski, J. and Perera, D. (2019). Stock index futures arbitrage: Evidence from a meta-analysis. *International Review of Financial Analysis*, 61, 284-294. doi: 10.1016/j.irfa.2018.09.002

Chen, Y. W., Guo, J. F., Wang, C., and Feng, N. D. (2016). High Frequency Trading Strategy Evaluation System Based on Grey Relational Analysis and Power Spectral Estimation. *Chemical Engineering Transactions*, 51, 661-666. doi: 10.3303/CET1651111

Kaplanski, G., Levy, H., Veld, C., and Veld-Merkoulova, Y. V. (2016). Past returns and the perceived Sharpe ratio. *Journal of Economic Behavior & Organization*, 123(Mar.), 149-167. doi: 10.1016/j.jebo.2015.11.010

Kim, R. (2018). Smart beta Strategy and Long-Short Factor Investing in Style Rotation. *Korean Journal of Financial Studies*, 47(5), 849-891. doi: 10.26845/KJFS.2018.10.47.5.849

Laurini, M. P. (2015). Imposing no-arbitrage conditions in implied volatilities using constrained smoothing splines. *Applied Stochastic Models in Business & Industry*, 27(6), 649-659. doi: 10.1002/asmb.877

Li, Y., Xiao, W., and Teng, Y. (2020). Evolutionarily stable strategy analysis of Quantitative investment strategy. *Journal of Physics Conference Series*, 1616, 012103. doi: 10.1088/1742-6596/1616/1/012103

Ma, S. (2020). Predicting the SP500 index trend based on GBDT and LightGBM methods. *E3S Web of*

Conferences, 214(5), 02019. doi: 10.1051/e3sconf/202021402019

Marks, J. M. and Shang, C. (2019). Factor crowding and liquidity exhaustion. *Journal of Financial Research*, 42(1), 147-180. doi: 10.1111/jfir.12165

Pan, H. and Long, M. (2021). Intelligent portfolio theory and application in stock investment with multi-factor models and trend following trading strategies. *Procedia Computer Science*, 187(1), 414-419. doi: 10.1016/j.procs.2021.04.116

Sachs, T. and Tiong, R. (2016). A liquid and investable benchmark index for the commodity trading advisor and managed futures industry. *Journal of Index Investing*, 6(4), 30-70. doi: 10.3905/jii.2016.6.4.030

Suhonen, A., Lennkh, M., and Perez, F. (2017). Quantifying backtest overfitting in alternative beta strategies. *The Journal of Portfolio Management*, 43(2), 90-104. doi: 10.2139/ssrn.2757113

Tang, C., Zhu, W., and Yu, X. (2019). Deep hierarchical strategy model for multi-source driven quantitative investment. *IEEE Access*, 7, 79331-79336. doi: 10.1109/ACCESS.2019.2923267

Wang, L., Xiong, X., and Xu, T. (2020). A study on pairing arbitrage strategy of stock passive structured fund in China under extreme market conditions. *Finance Research Letters*, 40, 101721. doi: 10.1016/j.frl.2020.101721

Wang, S., Li, Z., Zhu, J., Lin, Z., and Zhong, M. (2020). Stock selection strategy of A-share market based on rotation effect and random forest. *AIMS Mathematics*, 5(5), 4563-4580. doi: 10.3934/math.2020293

Wei, Y. and Watada, J. (2016). Building a type-2 fuzzy random support vector regression scheme in quantitative investment. *IEEJ Transactions on Electronics Information and Systems*, 136(4), 564-575. doi: 10.1541/ieejeiss.136.564

Yuan, X., Yuan, J., Jiang, T., Ain, Q.U. (2020). Integrated long-term stock selection models based on feature selection and machine learning algorithms for China stock market. *IEEE Access*, 8:22672-22685. doi: 10.1109/ACCESS.2020.2969293

Weizheng Wu was born in China, in 1988. He received a master's degree in economics from the Party School of the CPC Shaanxi Provincial Committee, in 2014. His current research interests include economic and financial, ideological and political education and other fields. He is currently a lecturer of Huanghuai University and deputy chief of the Party and Government Office of Huanghuai University. In recent years, more than ten scientific research projects and achievements have been completed by WeiZheng Wu. He once led the students to win the second prize in the 15th "Challenge Cup" Henan University Students' Extracurricular Academic Science and Technology Works Competition. He has one utility model invention patent and published more than ten papers.