

Portfolio Optimization for Cloud Computing, Biotechnology, Artificial Intelligence, Green Energy and Pharmaceutical Industries

Jisheng Duan

W.P. Carey School of Business, Arizona State University, Tempe, USA

jduan15@asu.edu

Abstract. Making reasonable portfolio allocation is a common problem nowadays. This paper analyzes asset allocation in cloud computing, biotechnology, medical engineering, green energy and artificial intelligence. These five assets are all fast-growing and suitable investments this year. This paper uses a time series approach to analyze these assets. Assets are analyzed using the mean-variance approach, Autoregressive Integrated Moving Average model (ARIMA), and FAME-French model. The mean variance model can better optimize the portfolio, while the ARIMA model is used to predict the selected ETF and judge its trend in the future period of time, and finally the Fama-French model is used to explain the portfolio. Finally, this paper evaluates the performance of the portfolio of these assets, including the specific proportion and the related return rate. The results of the analysis show that green energy and pharmaceutical industries account for the proportion of investment sent, the study is useful for potential investors in cloud computing, artificial intelligence, green energy, pharmaceutical industries, and biotechnology.

Keywords: Portfolio Optimization, Return and Risk, Investment.

1. Introduction

James Tobin, an American economist, won the Nobel Prize in economics in 1981 "for his study of financial markets and their relationships to spending choices, employment, production, and prices." He once said, "don't put all your eggs in one basket." This theory explains the importance of investment portfolios. The purpose of a portfolio is to help the investor achieve maximum return or minimize risk through the distribution of wealth, in short, to assist the investor in achieving their financial objectives. [1-2]

In recent years, with the development of science and technology and various industries, more new industries have gradually come into people's view. And as a result, more portfolios have been created. Therefore, how to better use investment risk to maximize returns and minimize risks has attracted more and more people's attention.

Most of the research on portfolio in the current market is aimed at the analysis of the current market situation. For instance, Wang Liang, Li Yuxuan, Liang Meiqi, Wang Yuanfei [3] studies the structure-based P2P product portfolio strategy under the risk reserve system. And Kraft Holger, Weiss Farina Demonstrated how to deploy a portfolio in the context of a pandemic, that is, a large event like COVID-19 [4]. Hedi Arouri Mohamed EI, Khuong Nguyen Duc [5], the oil price and the stock market as well as the portfolio related to securities were optimized. There's also a lot of research on what influences portfolios, like Jain Pankaj K, Kuvvet Emre, Pagano Michael S [6] research how corruption's impact on foreign portfolio investment This paper selects five industries with rapid growth at present, namely cloud computing, biotechnology, artificial intelligence, green energy and pharmaceutical to constructed financial portfolios with the commonly used mean-variance model [7-9].

2. Data

The data used in this paper came from Yahoo Finance (<https://finance.yahoo.com/>). This paper selected ETFs in five rapidly growing sectors namely First Trust Cloud Computing ETF (SKYY),

First Trust NYSE Arca Biotechnology Index Fund (FBT), Global × Robotics & artificial Intelligence ETF (BOTZ), iShares Green Clean Energy ETF (ICLN), VanEck Pharmaceutical (PPH). The reason why ETF is selected is that ETF is more stable than stocks, so it can better avoid market risks, ETF can better judge the trend of the industry, and ETF is composed of stocks of multiple industries, to avoid mutual interference between companies. All ETFs are dated from September 10, 2021, to September 9, 2022, with a total of one year of data. The information obtained from the relevant materials is shown in Table 1 and Table 2.

Table 1. Selected ETFs and related industries

	industry
SKYY	Cloud Computing
FBT	Biotechnology
BOTZ	Artificial Intelligence
ICLN	Green Energy
PPH	Pharmaceuticals

Table 2. Descriptive statistics of daily returns of the five ETFs

	Max	Min	Mean	Std Dev
ICLN	7.63%	-5.98%	0.42%	7.54%
PPH	3.21%	-2.98%	0.04%	3.47%
FBT	4.38%	-5.78%	-0.56%	5.90%
SKYY	6.31%	-6.90%	-1.72%	8.56%
BOTZ	4.40%	-5.74%	-2.61%	7.23%

As can be seen from the table 2, the ETF 'ICLN' the greatest average return 0.42%, and 'BOTZ' have the lowest profit -2.61%. And from the Std Dev the 'PPH' is the most stable, while 'SKYY' is the most volatile. Final 'ICLN' can provide max return, and 'SKYY' have the highest max return, 'SKYY' have the minimum Min return.

3. Method

3.1. ARIMA

The time series model is a more commonly used model for predicting stock prices. The model used in this article is the ARIMA model in the time sequence model. Model using an Autoregressive Integrated Moving Average (ARIMA model), the usual expression of ARIMA model is ARIMA (p, d, q). In this expression, AR represents the self-regression, p represents the number of self-regression terms, the last component, d, stands for the difference created to keep the time series stationary, and MA stands for the moving average, where q is the number of terms in the moving average. A generalization of an autoregressive moving average is called ARIMA (ARMA). The ARIMA model's formula can be written as:

$$(1 - \sum_{i=1}^p \phi_i L^i)(1 - L)^d X_t = (1 + \sum_{i=1}^q \theta_i L^i) \epsilon_t \quad (1)$$

Where L in the expression represents the lag factor.

3.2. Mean-variance analysis

Modern investment theory includes mean-variance analysis. Mean-variance analysis, in which the variance is expressed as the investment risk, is a method for calculating risk and expected return. Investors can evaluate how much risk they are ready to accept in exchange for various degrees of gain using this strategy. Investors can use this kind of analysis to determine the maximum return and the lowest risk for a specific level of return. The weights of each portfolio combined make up one in a portfolio.

$$\sum_i Weight_i = 1 \quad (2)$$

The Expected return is

$$E(R_p) = \sum_i \omega_i E(R_i) \quad (3)$$

R_i is the return on asset i , R_p is the return on the portfolio, and ω_i is the weight of composition of asset i in this calculation. Furthermore, the portfolio return variability is,

$$\sigma_p^2 = \sum_i \omega_i^2 \sigma_i^2 + \sum_i \sum_{j \neq i} \omega_i \omega_j \sigma_i \sigma_j \rho_{ij} \quad (4)$$

In this formula σ_i is the standard deviation of the periodic returns on an asset i , and ρ_{ij} is the correlation coefficient between the returns on asset i and j . And final the formula of portfolio return volatility is

$$\sigma_p = \sqrt{\sigma_p^2} \quad (5)$$

In this paper, this paper choose the portfolio with the maximum Sharpe rate, that is, how much excess return will be generated for each additional unit of risk of the portfolio. This is used as a benchmark to measure the test portfolio. The formula for the maximum Sharpe rate is as follows

$$SR = \frac{R_p - R_f}{\sigma_p} \quad (6)$$

3.3. Monte Carlo method

Because finance itself has uncertainty, it is suitable for Monte Carlo simulation. The fundamental tenet of Monte Carlo simulation is that the frequency of occurrence at a given time in numerous tests may be used to predict the likelihood of anything occurring. And the frequency of occurrence of a change event can be regarded as its probability when the sample size is large enough. In order to assess whether the structure fails, a large number of random variables that have an impact on its dependability can be randomly sampled. The sampling values can then be inserted into the functional function, and finally the likelihood of structure failure can be obtained.

In this study, the matrix of average stock return rate and covariance is used as its parameters, and the weight is randomly generated by each asset in the portfolio, and this method is used to determine the expected return and variance of the portfolio.

3.4. Fama-French three-factor model

An enhanced theory of capital asset pricing model is the three-factor model. This model is also the mainstream model at present, and there are related articles that show that the three-factor model is better than the five-factor model. For example Li Shuai and Zhang Qiang [10], it is stated that the three-factor model is superior to the five-factor model for the analysis of financial data. The market asset portfolio ($R_m - R_f$), the market capitalization factor (SMB), and the book-to-market factor, among others, are said to be able to explain the excess return of a portfolio (which includes a single stock) (HML).

$$r = R_f + \beta_3(R_m - R_f) + B_s * SMB + b_v * HML + \alpha \quad (7)$$

Where r is the predicted portfolio return, R_f is the market risk-free interest rate, and R_m is the market portfolio return, and the coefficient to be estimated beta of the three variables is the impact of the changes of risk premium, size premium and market to book ratio premium based on the anticipated rate of return. The beta coefficient in the CAPM model is similar to the idea of beta, which is the coefficient of market portfolio risk premium. SMB, or the company size variable, refers to the variation in return between a portfolio of firms with small market capitalizations and firms with big market capitalizations. The difference between the return on a portfolio of businesses with a high book-to-book ratio and the return on a portfolio of businesses with a low ratio is known as the price-to-book premium HML.

4. Result

Using the function to find the appropriate ARIMA function, the BIC method is used to find the appropriate p value and q value and obtain the applicable function. Table 3 shows the result.

Table 3. The result of the ARIMA related factors

Dep. Variable:	Adj Close	No. Observations:	252
Model:	ARIMA (3, 1, 1)	Log Likelihood	-144.917
Date:	Sun, 11 Sep 2022	AIC	299.834
Time:	02:49:02	BIC	317.461

Note: ETF 'ICLN' is chosen as the example to simulate the ARIMA model.

After determining the PDQ value of ARIMA model, this paper establish a model to predict the future trend of relevant ETFs and obtain the compound return rate of asset portfolio. Table 4 shows the result of the forecast.

Table 4. The result of the forecast for five ETFs

	ICLN	SKYY	FBT	BOTZ	PPH
2022/09/12	22.757344	68.492338	136.716567	20.586166	70.714133
2022/09/13	22.721761	68.492439	136.873476	20.586078	70.713686
2022/09/14	22.717485	68.492329	137.017667	20.586071	70.713623
2022/09/15	22.724105	68.493453	137.150168	20.585684	70.713614
2022/09/16	22.731138	68.492329	137.271929	20.585647	70.713623
2022/09/19	22.734750	68.494879	137.383819	20.585274	70.713625
2022/09/20	22.735550	68.495329	137.486639	20.585128	70.713613
2022/09/21	22.735074	68.495765	137.581124	20.584928	70.713635
2022/09/22	22.734439	68.496456	137.667950	20.584824	70.713637
2022/09/23	22.734067	68.497567	137.747737	20.585728	70.713724
2022/09/26	22.733959	68.495675	137.821057	20.586826	70.713736
2022/09/27	22.733986	68.496576	137.888433	20.587482	70.713834
2022/09/28	22.734042	68.496546	137.950347	20.588734	70.713724

After the prediction results are obtained, the results of all ETF predictions are simulated 10,000 times through Monte Carlo simulation. The resulting scatter plot is shown in Figure 1 below. The black marks represent positions formed by a single stock. The dark blue dashed line shows the efficient frontier. Each solid dot in the dark blue represents each portfolio under the given limit.

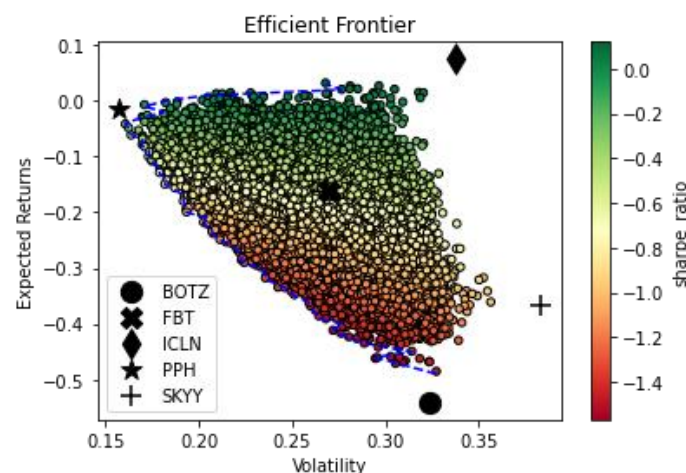


Figure 1. The result of Monte Carol

Then the best portfolio is selected according to the results obtained. Figure 2 marks two portfolios. The first one is the portfolio with the lowest volatility and highest Sharpe rate. In this paper, only the portfolio with the maximum Sharpe rate is selected for the experiment. The ratio between each ETF

that can be obtained should be BOTZ: 1.84%, FBT: 4.18%, ICLN: 71.28%, PPH: 22.32%, SKYY: 0.37%. The return one investor can get under this portfolio is 3.20% volatility is 26.79%, and the Sharpe rate of this portfolio is 11.96% (Details are shown in the Tables 5 and 6, respectively).

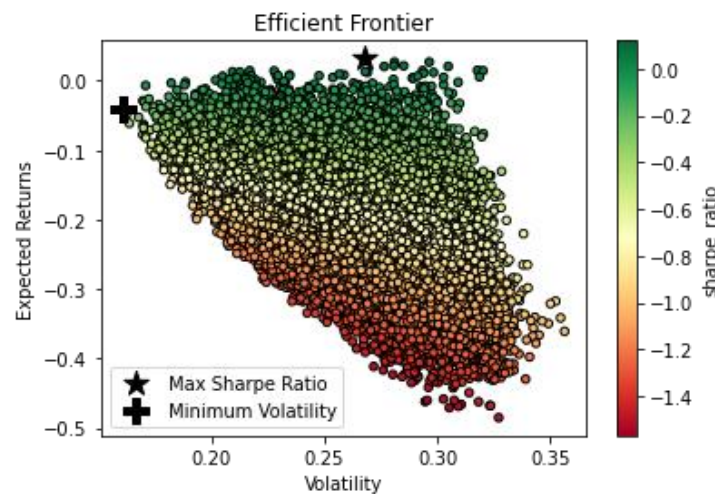


Figure 2. Max sharp rate and minimum volatility

Table 5. Weight of each stock in the two optimal portfolios (%)

	Max Sharpe Ratio
BOTZ	1.84%
FBT	4.18%
ICLN	71.28%
PPH	22.32%
SKYY	0.37%

Table 6. Return and volatility of the two portfolios

	Return	Volatility	Sharpe Ratio
Max Sharpe Ratio	3.20%	26.79%	11.96%

From a portfolio viewpoint, this is evident from the outcomes in the table, ETF 'ICLN' and 'PPH' account for the largest proportion of 'ICIN', accounting for 71.28%, and 'PPH' accounts for 22.32%. These two assets account for almost 95% of the total investment, and 'SKYY' accounts for the smallest proportion in the portfolio. Only 0.37% of total investment.

After that, the last step in this paper is to measure factor exposure. After the factor is defined, the factor exposure of the index can be obtained by weighting the factor score of the index component. The following table 7 reflects the relevant factor exposure mentioned in this paper.

Table 7. Correlation coefficient of factor exposure

	MKT	SMB	HML	Info Ratio	Alpha
BOTZ	0.002693	0.000335	0.000276	-112.495122	-7.912339
FBT	0.000347	-0.001330	-0.000178	59.919974	3.528971
ICLN	0.004334	-0.000818	-0.003115	-419.415378	-30.574519
PPH	0.000726	-0.001978	0.000179	1421.854228	48.254886
SKYY	0.003069	0.002009	-0.001345	99.918010	8.339106

From this table, it can be concluded that the SMB coefficient of BOTZ as well as SKYY is positive, which also indicates that the limited supply index will generate larger returns when small stocks outperform large stocks. The opposite is the case where the coefficient is negative which is the case of FBT, ICLN and PPH. From the value of HML, it can be concluded that the HML coefficients of BOTZ and PPH are positive, while the coefficients of FBT, ICLN and SKYY are negative. This shows that the excess return of ETF is affected by the HML coefficient, and a negative number means that during the study period, the chosen ETF prefers to invest development equities with a low book-

to-market ratio. Being regular means the opposite. The information ratio calculates how volatile a portfolio's returns are in relation to benchmark returns. A higher information ratio indicates the extent to which the fund exceeds the benchmark, while a lower one does the opposite. Investors generally choose ETFs based on their preferred risk profile. From the table, it can be said that PPH has the best information ratio, whereas ICLN has the worst.

5. Conclusion

Because the whole investment environment is a dynamic environment, with different industries, new and suitable industries are produced every year. The purpose of this paper is to analyze the investment portfolio analysis of industries that are suitable for investment in the present environment. To help potential investors provide reference. This paper analyzes the allocation of assets in the way of the time sequence and USES the mean variance analysis method to optimize the portfolio. In this article, the investment portfolio of the largest sharp rate is built, and the research finds that the amount of investment in the clean energy industry and the pharmaceutical industry is the largest if it is a portfolio of assets mentioned in this article.

However, there are also defects in this paper. For example, ARIMA model can only capture linear relations, but not nonlinear relations. Monte Carlo simulation also has some drawbacks such as the error is probabilistic error, and the issue of three-factor acquisition time for the Fama-French three-factor model.

References

- [1] Econlib. URL: <https://www.econlib.org/library/Enc/bios/Tobin.html> (Accessed on 2022-10-1).
- [2] He, J. A., Yang, X. Universal portfolio selection strategy by aggregating online expert advice. *Optimization and Engineering*, 2020, 1-25.
- [3] Wang, Liang, et al. Research on P2P Product Portfolio Strategy Based on Term Structure Under Risk Reserve System. *International Review of Economics & Finance*, 2023, 83: 124–38.
- [4] Kraft, Holger, Farina Weiss. Pandemic Portfolio Choice. *European Journal of Operational Research*, 2022.
- [5] Hedi Aroui, Mohamed El, Duc Khuong Nguyen. Oil Prices, Stock Markets and Portfolio Investment: Evidence from Sector Analysis in Europe over the Last Decade. *Energy Policy*, 2010, 38(8): 4528–4539.
- [6] Jain, Pankaj K., et al. Corruption's Impact on Foreign Portfolio Investment. *International Business Review*, 2017, 26(1): 23–35.
- [7] Markowitz, H. Foundations of portfolio theory. *The journal of finance*, 1991, 46(2): 469-477.
- [8] Dentcheva, Darinka, Gregory J. Stock. On the Price of Risk in a Mean-Risk Optimization Model. *Quantitative Finance*, 2018, 18(10): 1699–1713.
- [9] Mooney, Christopher Z. Monte Carlo Simulation. Sage Publications, 1997.
- [10] Li, S., Zhang, Q. Empirical research and regulatory thinking on the industry impact of the new crown epidemic on U.S. stocks—Based on the Fama-French three-factor and five-factor model. *Wuling Academic Journal*, 2021, (05): 48-57.