Portfolio Construction Based on Financial Indicators

Yuqi Qiao[1*,†], Jiaqi Yuan[2*,†], Zhiqi Wang[3*,†]
[1]*11810412@mail.sustech.edu.cn, [2]*zcahjy5@ucl.ac.uk, [3]*wzq2018050465@stu2018.jnu.edu.cn

[1]Department of Business, Southern University of Science and Technology Shenzhen, China
[2]Department of Mathematics, University College London London, United Kingdom
[3]Department of Economics, Jinan University Guangzhou, China
[†]These authors contributed equally.

Abstract—This paper focuses on the portfolio construction that determines the assets arranged in the portfolio based on some relatively novel and unique financial indicators. After that, the Monte Carlo simulation is used to find the efficient frontier and the allocation distribution of each asset. Subsequently, the ARIMA model is used to predict the overall trend of the price performance of the portfolio. Finally, there is a back test of the newly established asset portfolio to check the performance of our portfolio. In this paper, the historical data of 20 chosen blue-chip stocks is extracted from NYSE and NASDAQ. Each stock is assigned with an appropriate weight to achieve a smoother and more stable price trend by modeling. In the prediction section, there is a reasonable time prediction of the portfolio. The final back test also showed that the portfolio performs better than average market performance. In addition, the maximum sharp ratio proves that the research of this paper has made achievements on receiving extra returns under relatively stable risk and volatility. Altogether, the findings in this paper benefit certain investors in the related markets. It offers a brand-new quantitative investing strategy, which selects stocks scientifically and weighing them properly, decreasing the risk and receiving an acceptable return.

Keywords-Portfolio construction; Blue-chip stocks; Quantitative investment

1 INTRODUCTION

Risk and return are always the main concerns of investors in the quantitative investment area of the financial market. From this aspect, the portfolio construction process has been of significance for those who prioritized finding the asset allocation maximizing the returns under the accepted risk level as the ultimate goal. Since the pioneering work of Markowitz [1] and the advent of modern portfolio theory [2], institutional asset allocation as a practical field has evolved to include myriad techniques and perspectives [3].

By far, economists and other professionals did diversified investigations derived from the portfolio construction process. For instance, Shen [4] did the portfolio construction for the private market; the Russell company [5] focused on the assets allocation of the Emerging Market (EM) fixed income asset class, and Chattopadhyay [6] did the portfolio construction based on art and science-related assets. However, apart from these papers, the research combining the portfolio construction process with extra aspects, such as the financial indicators of assets and the time series prediction with back tests of the portfolio return, is relatively rare.

ICEMME 2022, November 18-20, Nanjing, People's Republic of China

DOI 10.4108/eai.18-11-2022.2326932

Therefore, to the best of our knowledge, this paper makes the following contributions to the literature. First, we evaluate and filtering the assets in the US stock market by using various financial indicators and build the efficient frontier of the representative assets by generating 50000 random simulations using the Monte Carlo method. For the following analysis, we decide to take the maximum sharp ratio asset allocation as the benchmark model, in which the maximum weight of 10.32% is contributed from the SBUX company. Second, we use the ARIMA model to make the time series prediction of our portfolio returns. The results of our analysis show that, in general, our portfolio performs better than the market in the prediction, and the extra return is simultaneously achieved. Third, we do the back test for our returns and find some important evaluation indicators such as sharp ratio, annual volatility, and a cumulative return of our portfolio. With these indicators, we evaluate the portfolio's performance and compare it with some stock market indexes.

This paper is organized as follows. Section 2 shows the data used in this paper. Section 3 and Section 4 present the methodologies and the related empirical results in this paper, respectively. Section 5 concludes this paper.

2 DATA

Pyramid algorithms assign different weights among the related indicators based on the financial data [7]. For example, the most critical metrics to investors, especially ROE and ROA, are given the highest importance. The second most crucial factor is the profitability, and then the cash position, the revenue, and the growth. Weights will be assigned to them in descending order. For the historical data range, the year 2020 is not included because of the passive, negative, and abnormal influences of the COVID-19. This means, consequently, the majority of industries and companies were affected by the COVID-19 epidemic, and the subsequent issuance of the Fed's massive amount of currency makes the financial data and stock prices of them not representative. However, there are good reasons to believe that the world will gradually return to normal after 2021, with many affected companies recovering. Therefore, the best 20 stocks were obtained according to the pyramid algorithm, and then the stock data from January 1, 2015, to December 31, 2019, were downloaded from yahoo finance (http://finance.yahoo.com). Some basic descriptive statistics are shown in the following Table 1.

TABLE 1. DESCRIPTIVE STATISTICS OF THE SELECTED ASSETS

	Mean	Std.	Min.	Max.	Skew.	Kurtosis
AAPL	0.3297	0.0176	-0.1049	0.0681	-0.5869	4.3167
ACN	0.1710	0.0130	-0.0756	0.0573	-0.6716	4.5583
BR	0.1532	0.0148	-0.1020	0.1058	-0.4323	11.1682
CBRL	0.0054	0.0132	-0.0707	0.0462	-0.7070	2.2741
CEPU	-0.6487	0.0511	-0.8191	0.1560	-8.4454	136.1701
FIZZ	-0.3818	0.0263	-0.1599	0.1139	-1.2343	7.5528
FTV	0.0103	0.0146	-0.0502	0.0468	-0.2920	1.1279
ITW	0.0561	0.0154	-0.0749	0.0704	-0.7102	3.2854
MA	0.2980	0.0161	-0.0643	0.0653	-0.2553	2.1159

NOC	0.0353	0.0157	-0.0676	0.0569	-0.5983	2.1175
NVDA	0.0008	0.0292	-0.2077	0.0895	-1.3243	7.3597
NVO	0.0958	0.0136	-0.0763	0.0374	-0.7467	3.0402
NVR	0.1149	0.0183	-0.0866	0.0591	-0.4250	1.7944
ROST	0.2147	0.0153	-0.0985	0.0695	-0.9491	5.7738
RYAAY	-0.1685	0.0188	-0.1712	0.0838	-1.6482	15.7669
SBUX	0.2584	0.0135	-0.0951	0.0926	0.2505	12.3180
SPGI	0.2356	0.0139	-0.0639	0.0786	-0.1296	3.9217
TJX	0.2509	0.0133	-0.0504	0.0672	0.1880	2.6207
TREX	0.2567	0.0254	-0.0997	0.1928	1.3433	11.4583
TXN	0.1148	0.0178	-0.0858	0.0717	-0.3614	2.8686

Detailed data of each stock are counted in the table. The combination of standard deviation skewness and abundance allows us to better measure the volatility and distribution of all stocks. Skewness is calculated to check whether the data distribution is symmetrical, and kurtosis is a measurement of the degree of data aggregation in the center. In a normal distribution, both skewness and kurtosis are 0. If the skewness is less than 0, it is a left skewness or negative skewness, which means that the mean value is to the left of the peak value, and there are more data of low stock prices than the high stock prices. If the skewness is greater than 0, it is a right skewness or positive skewness, which means that the mean is to the right of the peak, and there are more data of high stock prices than the negative low stock prices. A kurtosis less than 0 makes the distribution model flatter than the normal one, and a kurtosis greater than 0 makes the distribution model steeper than the normal one. Generally speaking, the stock market price typically has the characteristics of apparent left skew or right skew, and the kurtosis value is significantly greater than 0, which is shown as the tail after the peak [8].

3 METHOD

3.1 Mean Variance Model

In 1952, Markowitz published his paper "Portfolio Selection" in the Journal of Finance, which laid the foundation of portfolio theory and marked the beginning of modern portfolio theory [9]. The proposed mean-variance model creates the theory and method of the portfolio investment for rational investors under the condition of uncertainty. For the first time, it proves the advantages of diversified investment with accurate mathematical models. This model has become the mainstream direction of investment theory and practice. In the article, he pointed out that due to the correlation of various securities, the portfolio investment can reduce risks as long as there is an incomplete positive correlation between securities. In Markowitz's mean-variance model, the expectation represents the expected return of assets, and the variance represents the risk to study the selection and combination of assets. Specifically, the mean and variance are calculated by equations (1)-(2).

According to Markowitz [9], the expected return of the portfolio is calculated by equation (1),

$$R = x^T \mu \tag{1}$$

where R is the expected return of the portfolio, x is the weight for each asset, and is the expected return of each asset.

The corresponding variance of the expected return can be reached by the following equation (2),

$$V = D(R) = D(x^T\mu) = x^T\omega x \tag{2}$$

where V is the variance of the expected portfolio return and w is the co-variance matrix of the asset returns.

3.2 The Monte Carlo Simulation

Monte Carlo simulation (also known as the Monte Carlo Method) is a computerized mathematical technique to generate a large quantity of possible outcomes and assess the risk impact on them [10]. According to Carmona [11], the Monte Carlo simulation is feasible to be applied in portfolio construction in this paper. Thus, we chose this method to simulate the weights for the assets in the portfolio.

Monte Carlo simulation performs risk analysis by building models of possible results, substituting a range of values, which is a probability distribution, for any factor with inherent uncertainty [12]. In this paper, the normal distribution is applied, taking the mean and standard deviation values of the assets' returns as the parameters. By assigning random weights between 0 to 1 to twenty assets from the normal distribution function, with the total weight equaled to 1, the Monte Carlo method calculates the annualized return and variance repeatedly [12] for each sample set based on the following equations (3) - (4).

$$\mathrm{W} = (\mathrm{w}_1, \mathrm{w}_2, \dots, \mathrm{w}_\mathrm{i}, \dots, \mathrm{w}_{19}, \mathrm{w}_{20})^\mathrm{T} \tag{3}$$

$$\mathrm{w}_\mathrm{i} = \frac{\sum_{\mathrm{l}=1}^{20}(\delta_{\mathrm{il}}^2)^{-1}}{\sum_{\mathrm{k}=1,\mathrm{m}=1}^{20}(\delta_{\mathrm{km}}^2)^{-1}} \tag{4}$$

where δ_{ij} represents the standard deviation of the returns of i, j assets and is the$w_i$weight of the i asset in the portfolio which satisfies$\sum_{\mathrm{i}=1}^{20}\mathrm{w}_\mathrm{i} = 1$.

There are 50000 sample sets, and subsequently, 50000 calculations are generated and recorded.

3.3 The ARIMA Model

ARIMA is a natural extension of the class of ARMA models. The ARMA model is simply the merger between AR(p) and MA(q) models, in which AR(p) models try to capture (explain) effects of the momentum. It means reversion frequently observed in trading markets (market participant effects) and MA(q) models try to capture (explain) the shock effects observed in the white noise terms. These shock effects could be thought of as unexpected events affecting the observation process like Surprise earnings and terrorist attacks. ARIMA is used because it can reduce a non-stationary series to a stationary series by using a sequence of differences. Our data are usually not stationary; however, they can be made stationary by differentiating. The time series prediction is an auto-regressive integrated moving average model of order p, d, q, ARIMA(p,d,q). If the series is differentiated d times, it then follows an ARMA (p,q) process. We expand on the method described in the previous sheet. To fit data to an ARMA model, we use the Akaike Information Criterion (AIC) across a subset of values for (p,q) to find the model with minimum AIC. Then we apply the Ljung-Box test to determine if a good fit has been achieved for particular values of (p,q). If the p-value of the test is greater than the required significance, we can conclude that the residuals are independent and white noise. ARIMA (p,d,q) model expands ARMA (p,q) models. It can be shown as,

$$(1 - \sum_{i=1}^{p} \varphi_i L^i)\,(1 - L)^d X_t \;=\; (1 + \sum_{i=1}^{q} \theta_i L^i)\epsilon_t \tag{5}$$

where L is the lag operator, $d \in \mathbb{Z}, d > 0$.

4 Empirical Results

4.1 Portfolio Construction

In this paper, we implement the Monte Carlo simulation for 50000 times. Based on these asset weights, we calculate the return R and variance of the portfolio by using equations (6) and (7) and subsequently calculate the Sharpe ratio S of each sample sets based on the equation (8). In both calculations, we assume that there are 252 trading days per year.

$$R = \sum_{i=1}^{20} \mu_i w_i \cdot 252 \tag{6}$$

$$\delta^2 = W^T \cdot (Cov \cdot W) \cdot 252 \tag{7}$$

where $Cov_{i,j} = \frac{\sum_{x=1}^{n} E(Rix - E(Ri)) * E(Rjx - E(Rj))}{n-1}$ represents the co-variance value between i,j assets, n is the number of historical days, R_i represents the means of the returns of the i asset.

$$S = \frac{R - R_f}{\delta} \tag{8}$$

where R_f is the risk free ratio extracted from the US government website.

Correspondingly, we plot the efficient frontier and get the data of the maximum Sharpe ratio portfolio. The related results are shown in the following Figure 1 and Table 2.

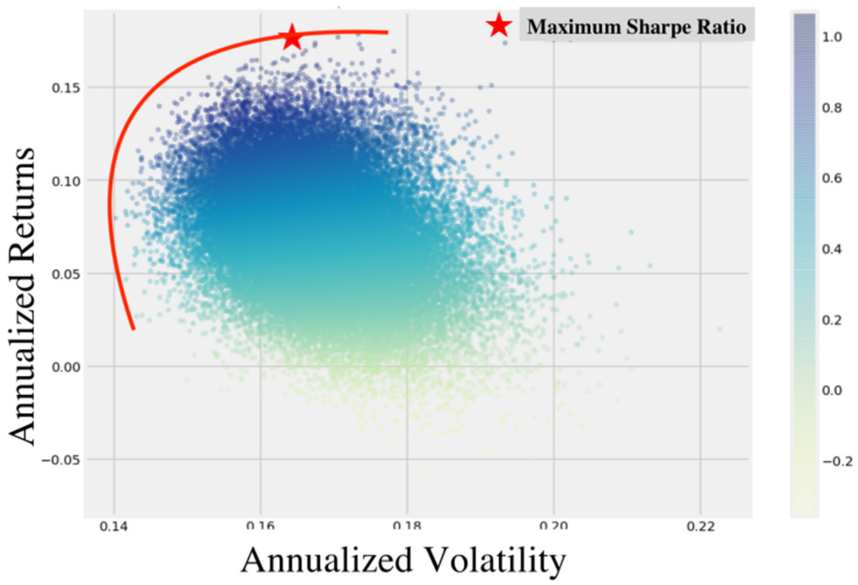

Figure 1. The efficient frontier

TABLE 2. PORTFOLIO WITH THE MAXIMIZED SHARPE RATIO

	AAPL	ACN	BR	CBRL	CEPU
Weight/%	4.54	7.71	4.21	0.35	0.11
	FTV	ITW	MA	NOC	NVDA
Weight/%	7.43	1.43	10.03	7.16	2.29
	NVR	ROST	RYAAY	**SBUX**	SPGI
Weight/%	2.58	6.66	0.46	**10.32**	7.51
	TREX	TXN	FIZZ	NVO	TJX
Weight/%	9.97	1.31	0.90	8.69	6.33
Portfolio Return/%	0.18		Portfolio Risk/%	0.16	

As shown in Table 2, it is clear that SBUX accounts for the largest proportion, with the magnitude of 10.32%, which may be accounted for because the SBUX is a large sized company with a small price-to-book ratio and relatively stable returns over the historical data set.

4.2 Times series Prediction

After the portfolio allocation is constructed, we made the time series prediction of our portfolio returns. For time series prediction, we choose the ARIMA model. First, we multiply the price sequence of all the selected stocks by their weights and add them together to get a portfolio price. As introduced before, the ARIMA model can be indirectly considered as a merger between AR(p) and MA(q) models, in which the AR(p) models capture market participant effects and MA(q) models try to capture the shock effects observed in the white noise terms. In our paper, ARIMA is used to reduce our non-stationary data series to a stationary data series by applying a sequence of differentiation. To fit data to an ARMA model, we use the Akaike Information Criterion (AIC) across a subset of values for (p,q) to find the model with minimum AIC. Then we apply the Ljung-Box test to determine if a good fit has been achieved, for particular values of (p,q). If the p-value of the test is greater than the required significance, we can conclude that the residuals are independent and white noise.

Following the above process, we found that the best order for our data is (2,0,1). The following figures show the Time Series Plots, Auto-correlation Coefficient and Partial Auto-correlation Coefficient data, the QQ Plot, and the Probability Plot. The related results are demonstrated in the following Figures (2)-(6).

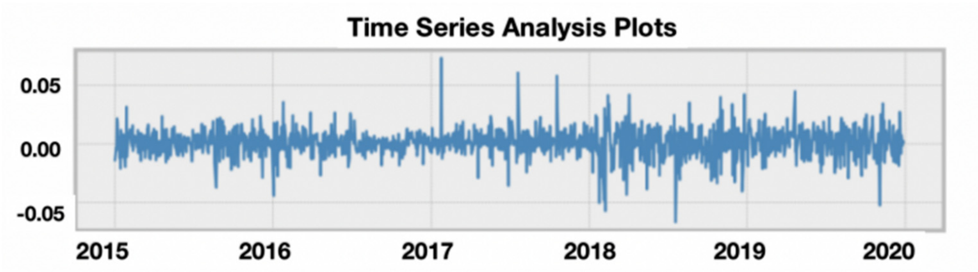

Figure 2. Time Series Analysis Plots

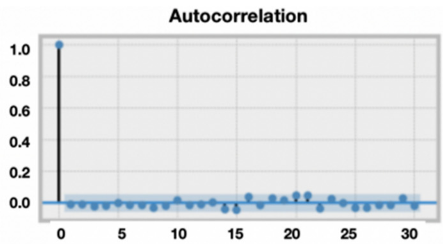

Figure 3. Autocorrelation

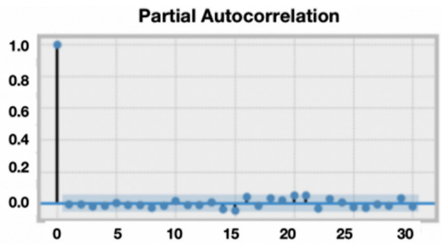

Figure 4. Partial Autocorrelation

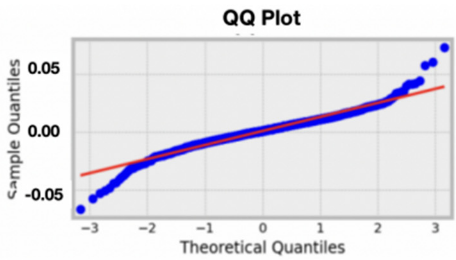

Figure 5. QQ Plot

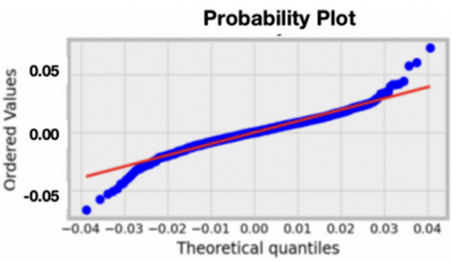

Figure 6. Probability Plot

We can see that after iterating, both the auto-correlation coefficient and the partial auto-correlation coefficient are in an acceptable range. It can be seen from the QQ plot that the residual basically meets the normal distribution. However, the bias of the curve's head and the tail is slightly larger. To reduce bias, we tried several other orders like (2,1,1) and (3,0,1), but it turns out that (2,0,1) is still the best choice.

After testing the normal distribution of the residuals, we predicted the return of the portfolio of early 2020. The result is shown in the following Figure 7. To make the result clearer, we also predicted our portfolio price, and the results are shown in figure (8).

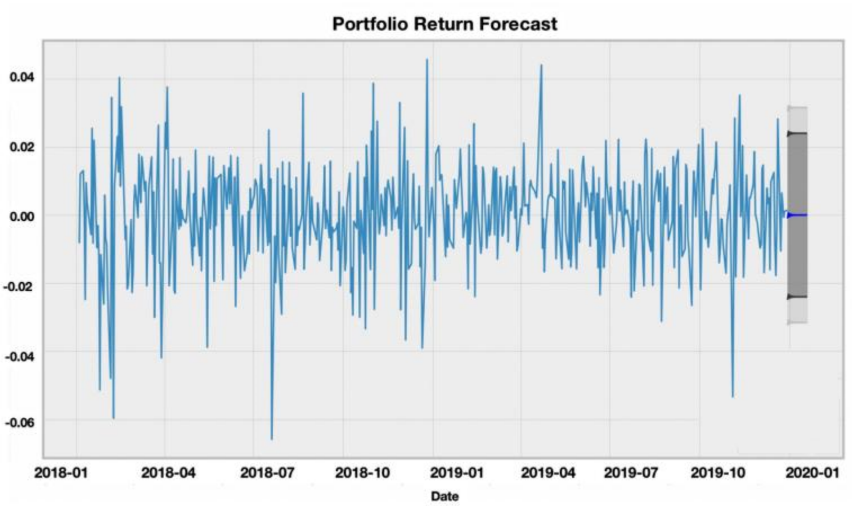

Figure 7. 21 Days Portfolio Return Forecast

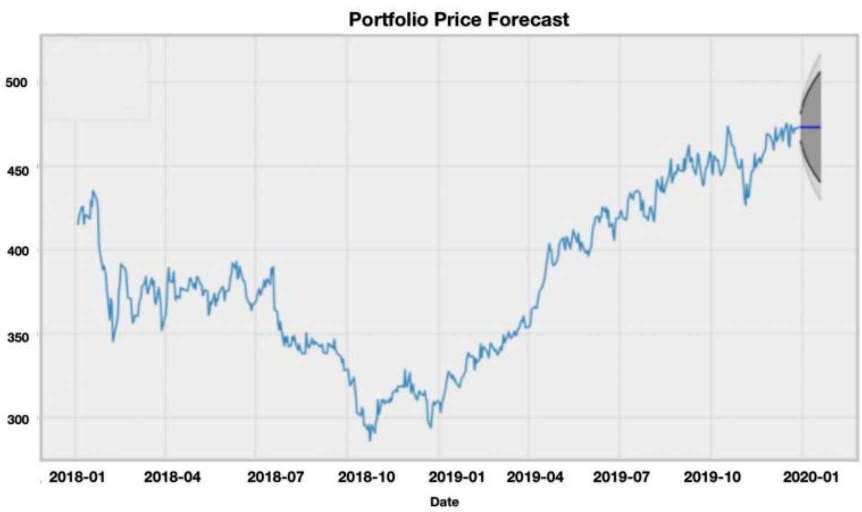

Figure 8. 21 Day Portfolio Price Forecast

4.3 Back Test

Figure Labels: After the time prediction, we do the back test for our portfolio. The results are shown in Table 3 and figures (9)-(11).

TABLE 3. BACK TEST

Start date	2015-01-02
End date	2019-12-30
Total months	59
Annual return	21.1%
Cumulative return	159.9%
Annual volatility	19.4%
Sharp ratio	1.08
Calmar ratio	0.59
Stability	0.86
Max drawdown	-36.0%
Omega ratio	1.21
Sortino ratio	1.58
Skew	-0.14
Kurtosis	3.83
Tail ratio	0.99
Daily value at risk	-2.4%

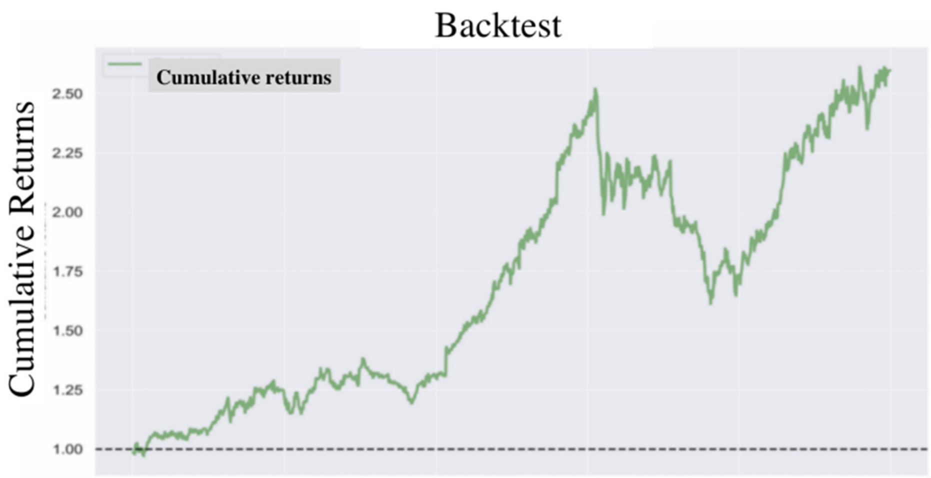

Figure 9. Cumulative returns

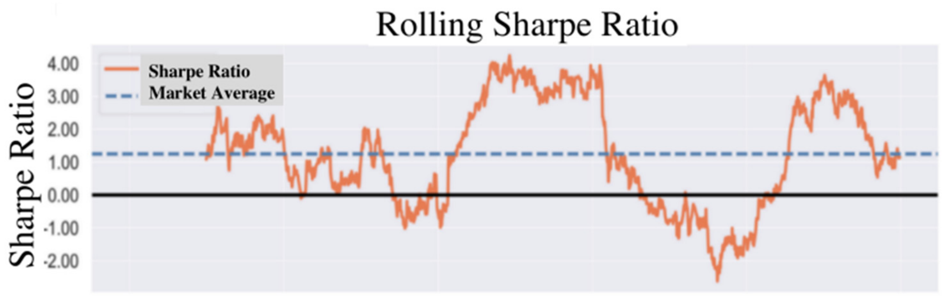

Figure 10. Rolling Sharp ratio(6-month)

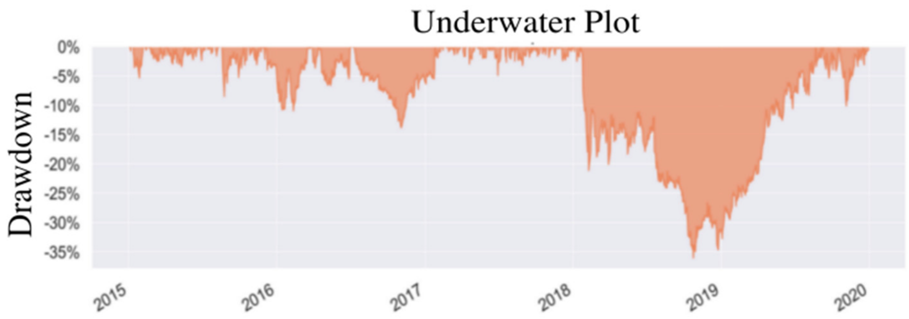

Figure 11. Underwater plot

From the table above: the annual return is 21.1%; the cumulative return is 159.9%; the annual volatility is 19.4%; the Sharp ratio is 1.09; the Calmar ratio is 0.59; the Stability is 0.86; the Max drawdown is -36.0%; the Omega ratio is 1.21; the Sortino ratio is 1.58; the Skew is -0.14; the Kurtosis is 3,83; the Tail ratio is 0.99; and the daily value at risk is -2.4%. We can also see the graph of Rolling Sharp Ratio (6-month), the underwater plot of drawdown, and the Cumulative Return.

5 Conclusion

This paper focuses on the portfolio construction area and applies various financial indicators in the asset allocation process.

First, we applied a new method in the stock selection process, introducing financial indicators to evaluate and rank all stocks in the US market. Then, by using the Monte Carlo Method to produce the efficient frontier, we find the allocation distribution of the maximum sharp ratio assets allocation. After that, we analyze our portfolio's returns by using the ARIMA models, and in the time series prediction, we successfully find that our portfolio performs better than the market. Finally, we evaluate the performance of our returns by using several important evaluation indicators such as sharp ratio, annual volatility, and a cumulative return of our portfolio.

Most of the results are not unexpected in our paper, as shown by our assets allocation' good performance over time. Nonetheless, there are spaces for improvements. For instance, we can choose more diversified dimensions to measure the value of stocks to select better stocks. Moreover, the inaccuracy of the Monte Carlo simulation is not eliminated. The allocation distribution may vary regarding changes in the number of simulations and different trial times. Although economists prove this method to be authentic in producing the efficient frontier, further investigation may still be worth considering. Finally, for the prediction part, the accuracy of the time series model will weaken with the increase of the forecast market. If we want to obtain more accurate forecast data, other methods might be needed.

REFERENCES

[1] Markowitz, H. Portfolio Selection. Journal of Finance 7, 1952, 77-91.

[2] Ben-Tal, A., Nemirovski, A. Lectures on modern convex optimization: analysis, algorithms, and engineering applications (Society for Industrial and Applied Mathematics).2001

[3] Eun, C.S., Bruce, G.R. Exchange Rate Uncertainty, Forward Contracts, and International Portfolio Selection. Journal of Finance 43, 1988, 197-215.

[4] Shen, J.Y. Asset Allocation and Private Market Investing. The Journal of Portfolio Management, 2020

[5] Russell, C. Emerging Market Fixed Income: Characteristics of the asset class. FTSE Russell, 2020

[6] Chattopadhyay, S. The Art and Science of Portfolio Construction. Carsongroup, 2019

[7] HOLD LE. URL: https://www.holdle.com/pyramid/index (Accessed on May 2, 2021)

[8] Jiang W.Y. An empirical study of Markowitz mean-variance model and RAROC model in Chinese securities market, ECONOMIC RESEARCH GUIDE 2020, 2, 103-106 (In Chinese)

[9] Haydock, Michael P.. Portfolio construction: The efficient diversification of marketing investments. Walden University, Beijing Zhong Ke I/E Company of China Science Publishing GroupProQuest Dissertations and Theses Full-text Search Platform database, 2010

[10] Kroese, D. P. Why the Monte Carlo method is so important today. WIREs Comput Stat, 2018, 386–392.

[11] Carmona, R.A. Numerical Methods in Finance. Springer Proceedings in Mathematics, 2012.

[12] Kroese, D. P. Handbook of Monte Carlo Methods. John Wiley & Sons, 2011