

Research and Analysis on Markowitz Model and Index Model of Portfolio Selection

Mingzhou Jin^{1,*}, ^{a, †}, Zexin Li^{2, *}, ^{b, †}, Shengkai Yuan^{3, *, c, †}

¹ College of Art & Science, Boston University, 02215, Boston United States

² School of Architecture & Fine Art, Dalian University Technology, 116024, Dalian, China

³ Social Science School, University of California, 92697, Irvine, United States

*Corresponding author. Email: ^amzjin@bu.edu, ^bAria.li.zexin@gmail.com, ^cshengky@uci.edu,

[†]These authors contributed equally.

ABSTRACT

With the development of stock portfolio theory, the research value of risk dispersion has become increasingly prominent. How investors construct portfolios has become an important research topic. In the article, we use recent 20 years of historical daily total return data for ten stocks, which belong in groups to three-four different sectors according to Yahoo Finance, the S&P 500 equity index, which include a total of eleven risky assets and a proxy for a risk-free rate, 1-month Fed Funds rate. We calculate all proper optimization inputs for the full Markowitz model, alongside the Index model. Using these optimization inputs for Markowitz model and Index model will need to the regions of permissible portfolios for the five different cases of the additional constraints. We present the results in both the tabular and graphical form with the objective to make inferences and comparisons between the sets of constraints for each optimization problem and between the Markowitz model and Index model. We hope to get an optimal portfolio by comparing the two models under different conditions. Also, we hope that the results of this project can lay a foundation for future on data analysis and investment portfolio creation.

Keywords: Markowitz model, Index model, risk, optimize the portfolio.

1. INTRODUCTION

As an international capital market, the US stock market has a long history and mature structure. It has completed listing and financing for many great companies in human history. However, most traders and investors still rely primarily on fundamental and technical analysis to make investment decisions. Today's stock market has moved away from the single investment stage and has become more focused on the risk and return of multiple stocks. Modern Portfolio Theory is a method to study how investors allocate their capital to different assets in order to find the optimal equilibrium relationship between overall risk and return under uncertainty. This paper introduces the mathematical analysis method and applies the Markowitz mean-variance analysis model (MM) and the Index model (IM) to the U.S. capital market to quantitative analysis and scientific decision making of stock portfolios, which can help improve investors' returns and reducing investment risks.

At the theory of relevant literature in the capital portfolio, most of the current studies are committed to the demonstration of the capital portfolio with a single Markowitz model, but lack of long-term dimension and demonstration, optimization, and comparison. Investment is the driving force of the continuous operation of the financial market. For investors, how to make a rational and dialectical investment is a problem worthy of study. Since the 1960s, portfolio theory has developed rapidly. American economist Markowitz used mathematical methods to apply the model to the study of the return risk relationship, concluded that the risk of the asset itself determines its expected return, and gave a method to calculate the optimal portfolio according to the mean and variance of each asset return, which promoted the development of modern portfolio theory [1, 2]. In 1963, Sharp simplified Markowitz's model and proposed a single index model, which linked the risk and return of the portfolio with the market portfolio, provided a new perspective for portfolio diversification, and greatly reduced the amount of calculation of the model [3]. Sharp, Lintner, and Mossin further developed the Markowitz

model, established the capital asset pricing model, and analyzed the sensitivity between securities return and market portfolio return, which has been widely recognized [4-6]. Ross conducted in-depth research on the capital asset pricing model, relaxed the requirements on assumptions, and proposed an arbitrage pricing model, which is more concise and easy to use [7].

Since then, a large number of scholars have widely used the model in various investment fields. Hwang T, Gao S, and Owen T taking the stocks in the UK stock market stocks from 1985 to 2012 as the research object, this paper analyzes the relationship between stock market value and expected return through Markowitz's portfolio effective frontier. The results show that the larger the size of the portfolio, the worse the performance of the portfolio, that is, the lower the return of the portfolio. The small-scale portfolio can not only obtain excess, but also spread risks well [8]. Kan and Zhou studied using the covariance matrix of the sample mean and return to estimate the optimal combination. It is found that the portfolio constructed by risk-free assets, tangent portfolio, and sample minimum variance portfolio is better than risk-free assets and tangent portfolio, which changes the theoretical suggestions of the two fund portfolios [9]. Hoevenaars and others studied the impact of parameter uncertainty on the risk of three types of assets in long-term investment: stocks, notes, and bonds. They found that the impact of the investment term on the optimal portfolio allocation is weaker than that of stock return parameter uncertainty. When the investment term exceeds 15 years, the risk term structure of stocks and bonds is relatively flat [10].

The aim of this research is the regions of permissible portfolios for the different situation accounting to existing historical data. Find the best portfolio for different risk preference with the best rate of return under the minimum risk. And methods maybe provide research and analysis ideas for our future portfolio.

The remainder of the paper is organized as follows; section 2 describes the sample and data analysis method; section 3 performs two venture capital models, Markowitz model and Index model; section 4 introduces the best portfolio in five different constraint; The last section presents our conclusions and portfolio examples in our real life.

2. DATA

We collect recent 20 years of daily data of total returns for the S&P 500, and for ten stocks such that there are three-four groups of stocks with stocks in each group belonging to one sector and an instrument representing risk-free rate, 1-month annual Fed Funds rate. Table of stock ticker symbols for our portfolios to work, as Table 1.

Table 1. Stock ticker symbols in portfolios

Portfolio	Full Name	Sector
1 QCOM	QUALCOMM Incorporated	Technology
2 AKAM	Akamai Technologies, Inc.	Technology
3 ORCL	Oracle Corporation	Technology
4 MSFT	Microsoft Corporation	Technology
5 CVX	Chevron Corporation	Energy
6 XOM	Exxon Mobil Corporation	Energy
7 IMO	Imperial Oil Limited	Energy
8 KO	The Coca-Cola Company	Consumer Defensive
9 PEP	PepsiCo, Inc.	Consumer Defensive
10 MCD	McDonald's Corporation	Consumer Cyclical

For our calculations, we use the full available historical data from 5/11/2001 to 5/12/2021. In order to reduce the non-Gaussian effects, we aggregate the daily data to the monthly observations. Details information and the stock develop trend of a listed company are as follow:

● QUALCOMM Incorporated (QCOM)

Qualcomm is a wireless communications and SoC research and development company located in San Diego, California, USA. The company produces semiconductors, software, etc. It is best known for working in 5G, CDMA2000, and WCDMA standards. The stock trend over the past 20 years is shown in the Figure1.

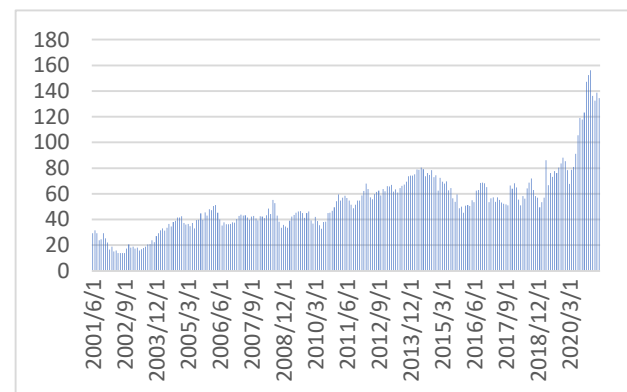


Figure 1 Stock trend in QCOM

● AKAM-Akamai Technologies

Akamai Technologies, Inc. is a global content delivery network, cybersecurity, and cloud service company, providing web and Internet security services. Akamai's Intelligent Edge Platform is one of the world's largest distributed computing platforms. The stock trend over the past 20 years is shown in the Figure2.

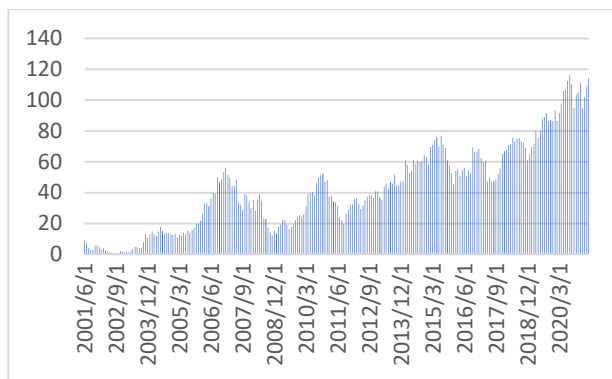


Figure 2 Stock trend in AKAM

- ORCL-Oracle

Oracle Corporation is an American company. It makes hardware systems and enterprise software (software for businesses) such as database management systems. Its headquarters are in Redwood City, California, United States. The founder is Lary Ellison. In 2010, it employed 105,000 people worldwide. It has increased in size through natural growth and by buying other companies. The stock trend over the past 20 years is shown in the Figure3.

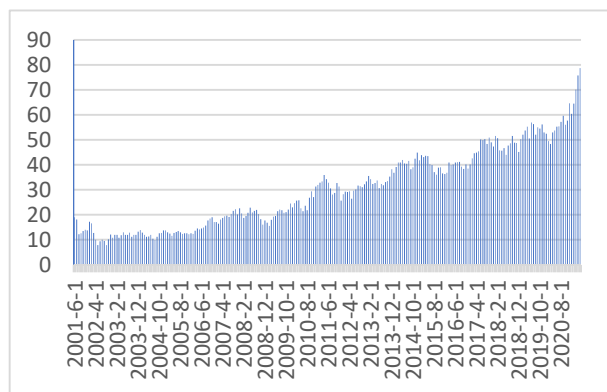


Figure 3 Stock trend in ORCL

- MSFT-Microsoft

Microsoft Corporation is an American multinational technology corporation that produces computer software, consumer electronics, personal computers, and related services. Its best known software products are the Microsoft Windows line of operating systems, the Microsoft Office suite, and the Internet Explorer and Edge web browsers. Its flagship hardware products are the Xbox video game consoles and the Microsoft Surface lineup of touchscreen personal computers. The stock trend over the past 20 years is shown in the Figure4.

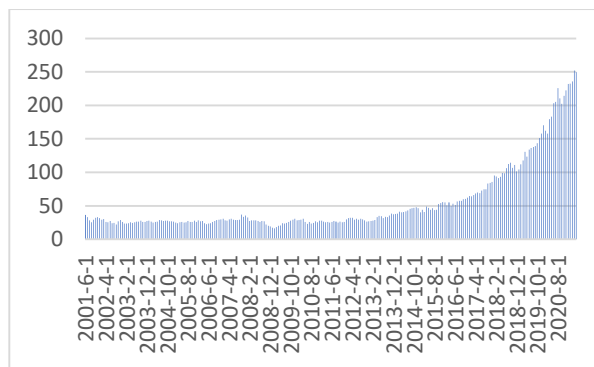


Figure 4 Stock trend in MSFT

- CVX-Chervron

Chevron Corporation is an American multinational energy corporation. It was founded in 1984 and is the third largest oil company in America. One of the successor companies of Standard Oil, it is headquartered in San Ramon, California, and is active in more than 180 countries. Chevron is engaged in every aspect of the oil and natural gas industries, including hydrocarbon exploration and production; refining, marketing and transport; chemicals manufacturing and sales; and power generation. The stock trend over the past 20 years is shown in the Figure5.

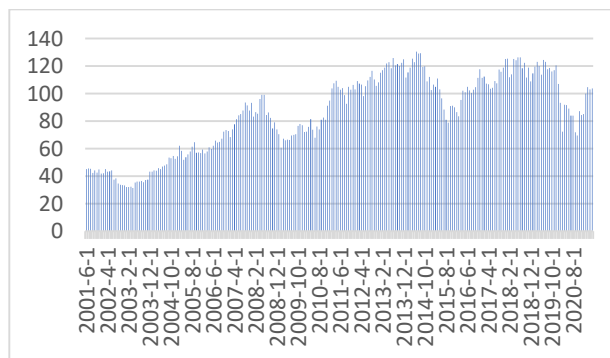


Figure 5 Stock trend in CVX

- XOM-Exxon Mobil

Exxon Mobil Corporation, stylized as ExxonMobil, is an American multinational oil and gas corporation headquartered in Irving, Texas. It is the largest direct descendant of John D. Rockefeller's Standard Oil and was formed on November 30, 1999, by the merger of Exxon and Mobil. ExxonMobil's primary brands are Exxon, Mobil, Esso, and ExxonMobil Chemical. The stock trend over the past 20 years is shown in the Figure 6.

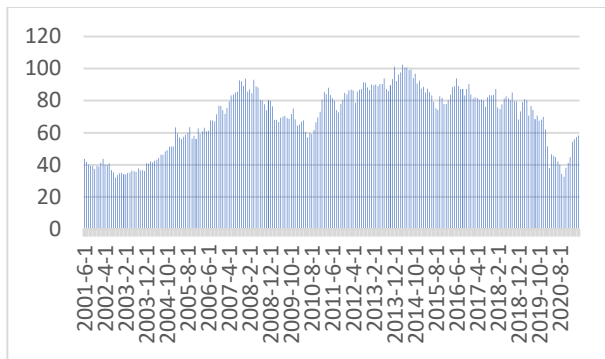


Figure 6 Stock trend in XOM

- IMO-Imperial Oil Limited

Imperial Oil Limited (French: Compagnie Pétrolière Impériale Ltée) is a Canadian petroleum company. It is Canada's second-biggest integrated oil company. ExxonMobil has a 69.6 percent ownership stake in the company. It is a significant producer of crude oil, diluted bitumen, and natural gas, Canada's major petroleum refiner, a key petrochemical producer, and a national marketer with coast-to-coast supply and retail networks. The stock trend over the past 20 years is shown in the Figure 7.

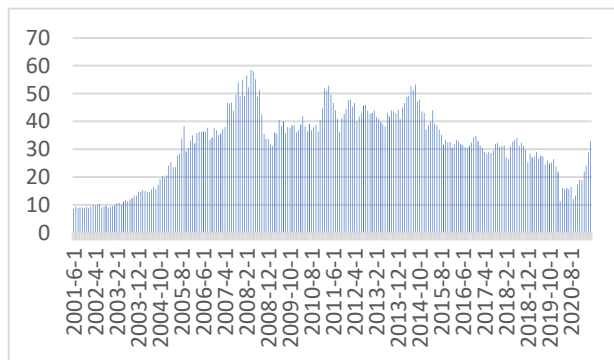


Figure 7 Stock trend in IMO

- KO-Coca Cola

The Coca-Cola Company produces concentrate, which is then sold to licensed Coca-Cola bottlers throughout the world. The bottlers, who hold exclusive territory contracts with the company, produce the finished product in cans and bottles from the concentrate, in combination with filtered water and sweeteners. The stock trend over the past 20 years is shown in the Figure 8.

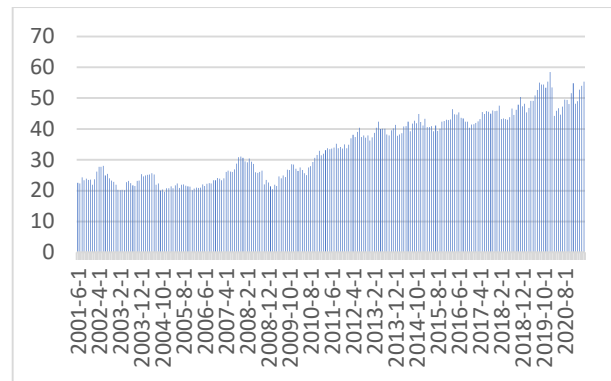


Figure 8 Stock trend in KO

- PEP- PepsiCo, Inc

PepsiCo, Inc. is an American multinational food, snack, and beverage corporation headquartered in Harrison, New York, in the hamlet of Purchase. PepsiCo's business encompasses all aspects of the food and beverage market. It oversees the manufacturing, distribution, and marketing of its products. PepsiCo was formed in 1965 with the merger of the Pepsi-Cola Company and Frito-Lay, Inc. PepsiCo has since expanded from its namesake product Pepsi Cola to an immensely diversified range of food and beverage brands. The stock trend over the past 20 years is shown in the Figure 9.

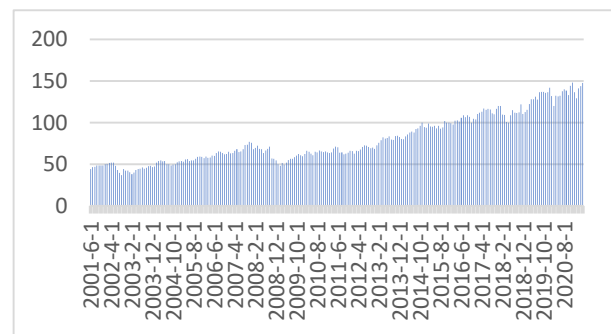


Figure 9 Stock trend in PEP

- MCD-McDonald's

McDonald's is an American fast food company, founded in 1940 as a restaurant operated by Richard and Maurice McDonald, in San Bernardino, California, United States. They rechristened their business as a hamburger, and later turned the company into a franchise, with the Golden Arches logo being introduced in 1953 at a location in Phoenix, Arizona. The stock trend over the past 20 years is shown in the Figure 10.

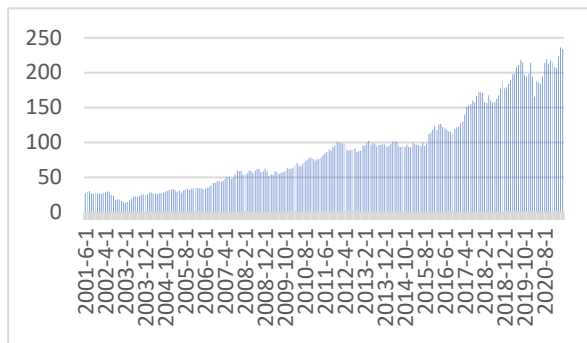


Figure 10 Stock trend in MCD

Based on those monthly observations, we prepare an Excel spreadsheet that makes all the necessary calculations to plot a Permissible Portfolios Region, which combines all proper optimization inputs for the full Markowitz model (MM), alongside the Index model (IM). Using these optimization inputs for MM and IM we find the Efficient Frontier, the Minimal Risk or Variance Frontier, and the Minimal Return Frontier for a given set of five different constraints. The Minimal Return Frontier and the Efficient Frontier together are forming the Minimal Risk or Variance Frontier.

To calculate a large numbers of multiple points on any of the required frontiers, the main tool that we use to solve the optimization problems for each point on the Minimal Risk or Variance Frontier is the Excel Solver.

We analyze all our results with the purpose of comparison of different constraints for each optimization problem (MM and IM) and the two optimization problem solutions between each other with the same constraints. And how to make a portfolio by using these parameters.

As an international capital market, the US stock market has a long history and mature structure. It has completed listing and financing for many great companies in human history. However, most traders and investors still rely primarily on fundamental and technical analysis to make investment decisions. Today's stock market has moved away from the single investment stage and has become more focused on the risk and return of multiple stocks. Modern Portfolio Theory is a method to study how investors allocate their capital to different assets in order to find the optimal equilibrium relationship between overall risk and return under uncertainty. This paper introduces the mathematical analysis method and applies the Markowitz mean-variance analysis model(MM) and the Index model(IM) to the U.S. capital market to quantitative analysis and scientific decision making of stock portfolios, which can help improve investors' returns and reducing investment risks.

3. METHOD

In this section, we will introduce two venture capital models, the Markowitz model and the Index model, that

help investors construct different types of investment portfolios, such as minimum variance portfolio and maximum Sharpe portfolio, due to investors' variant objectives.

3.1 Markowitz Mean-Variance model

Markowitz Mean-Variance model, also known as Markowitz model(MM), was introduced by H. M. Markowitz. In 1952, Markowitz proposed the venture capital model. Markowitz defined risk as the volatility of the rate of return, and for the first time, applied the method of mathematical statistics to the study of portfolio selection. This method makes the multi-objective optimization of profit and risk achieve the best balance effect. The investment in securities and other risky assets first needs to solve two core problems: expected return and risk. So how to measure the risk and return of portfolio investment and how to balance these two indicators for asset allocation are the problems that market investors urgently need to solve. It was against this background that Markowitz's theory emerged in the 1950s and early 1960s. In order to maximize the return of an investment and to minimize unexpected risk at the same time, the Markowitz model exerted two main variables, mean and variance, to find the optimal portfolio by following mathematical statistics logic.

3.1.1 Assumptions in Markowitz model

As I mentioned above, the Markowitz model was based on the method of mathematical statistics. In order to make the model as simplified as it could be and work correctly, there are quite a lot of prerequisites. Markowitz applied four assumptions to fully construct an effective and functional model.

- (1) When an investor considers each investment choice, it is based on the probability distribution of the returns on securities in a certain period of time.
- (2) Investors estimate the risk of a security portfolio according to the variance or standard deviation of the expected return rate of the security.
- (3) Investors' decisions are based solely on the risks and returns of the securities.
- (4) At a certain risk level, investors expect the greatest return; in contrast at a certain level of return.

3.1.2 Formulas

Based on the four assumptions, the Markowitz model can properly compute the target result of investors by two following formulas:

$$\text{Mean: } r_p = \sum x_i * r_i \quad (1)$$

For the part, r_p is the return of portfolio, r_i is the return of the i th stock, x_i and x_j are the investment ratio of stock i and j .

$$\text{Variance: } \min \sigma^2(r_p) = \sum \sum x_i * r_i * \text{Cov}(r_i, r_j) \quad (2)$$

In this formula, $\sigma^2(r_p)$ is the variance of portfolio investment (total risk of portfolio), and $\text{Cov}(r_i, r_j)$ is the covariance between the two securities. This model has laid the foundation for modern securities investment theory. The above formula shows that the yield of x_i securities can minimize the risk of $\sigma^2(r_p)$ under the restricted conditions, which can be obtained through the Langerian objective function.

Its economic significance is that investors can determine an expected return in advance, and the investment proportion (project capital allocation) of investors in each investment project (such as stocks) can be determined through the above formula to minimize the total investment risk. Different expected returns have different combinations of minimum variances, which form the set of minimum variances.

3.2 Index model

Single-index model (SIM), also known as the Index model, is a simple asset pricing model which is usually used in the financial industry to evaluate the risk and return of a stock. Single index model is also called the characteristic line and diagonal model. It is one of the portfolio analysis models. After the Markowitz portfolio analysis model was established, it was found that the Markowitz model had some defects. One important drawback is that it is very complicated and cumbersome to calculate. And amount of estimations reduced calculating efficiency and accuracy. For this reason, Professor W. Sharp created an analysis model that not only analyze the securities portfolio, be easy to use. This model is called the Single Index model.

3.2.1 Assumptions in Index model

To facilitate analysis, the single index model assumes that only one macro factor will cause stock return risk, which can be expressed as the return rate of a market index, such as the S&P 500.

The residual (random), which is assumed normally distributed with mean zero and standard deviation σ_i .

According to the assumptions of this model, the return of any stock can be decomposed into the expectation of the residual return of individual shares (expressed by a company-specific factor α), the return of macro events that affect the market, and the unpredictable composition of micro events that only affect the company.

3.2.2 Formulas

$$r_{it} - r_{ft} = \alpha_i + \beta_i(r_{mt} - r_{ft}) + \varepsilon_{it} \quad (3)$$

r_{it} is the return to stock i , r_{ft} is the risk-free rate, r_{mt} is the return to the market portfolio, α_i is the stock's alpha, or excess return, β_i is the stock's beta, or sensitiveness to the market return, ε_{it} is the residual (random) return, assumed normally distributed with mean zero and standard deviation σ_i .

3.2.3 Data

For the index model, we need to get different factors, for example, $\alpha_i + \beta_i$. Fewer factors are required for analysis because the Index model simplified the estimating process of the Markowitz model. So, we use the regression function to plot these three important points.

3.3 Constraint

In this project, we set 5 different constraints for each model. Basically, the constraint is different at the upper limit of weights and lower boundary of weights. Using the 5 constraints, we can mimic most of the real policy and regulations in the economic market and companies around the world. In this case, the comparison and contrast of the two models are more coherent and applicable.

1. This additional optimization constraint is designed to simulate Regulation T by FINRA (<https://www.finra.org/rules-guidance/key-topics/margin-accounts>), which allows broker-dealers to allow their customers to have positions, 50% or more of which are funded by the customer's account equity.

$$\sum_{i=1}^{11} |w_i| \leq 2 \quad (4)$$

2. This additional optimization constraint is designed to simulate some arbitrary "box" constraints on weights, which may be provided by the client.

$$|w_i| \leq 1, \text{ for } \forall i \quad (5)$$

3. A "free" problem, without any additional optimization constraints, to illustrate how the area of permissible portfolios in general and the efficient frontier, in particular, look like if you have no constraints;

4. This additional optimization constraint is designed to simulate the typical limitations existing in the U.S. mutual fund industry: a U.S. open-ended mutual fund is not allowed to have any short positions, for details, see the Investment Company Act of 1940, Section 12(a)(3)

$$w_i \geq 0, \text{ for } \forall i \quad (6)$$

5. Lastly, we would like to see if the inclusion of the broad index into our portfolio has a positive or negative

effect, for that, we would like to consider an additional optimization constraint.

$$w_1 = 0 \quad (7)$$

4. RESULTS ANALYSIS

Before going to deeper analysis, we need to allocate our original data. The data we collected was daily data of 10 stocks and 1 index's in the past 20 years.

Firstly, we need to convert daily data to monthly data or monthly observation. The reason for this conversion is that monthly data can reduce the amount of time that the computer consume to calculate. Moreover, monthly data can have better Gaussian distribution, one of the assumptions of both models is that all data are assumed to be Gaussian distributed. In this case, we have to convert our data at first. By using the filter function to pick out the end of each month and label them as EOM, we can find out one representative figure for each month. Copying and pasting these data in a new table, we have already finished our initial data processing. Then, the next step is to get some useful data from hundreds of figures. Calculating return, excess return, and residual helps us solve further problems. So, for each stock we simply use the number from this month to minus the number of last month to calculate return for each stock. Then, we need a figure called NRFR. Assuming the first month's NRFR is 1000.00, we can calculate next month's NRFR by using the formula

$$NRFR_{n+1} = NRFR_n * (1 + FEDR01_{n+1}/100/252) \quad (8)$$

In this formula, $FEDR01_{n+1}$ represents the federal reserve interest rate at $n+1$ month. NRFR stands for the

nominal risk free rate, which is identical to the federal interest rate here.

Following with NRFR, we still need to get NRFR increase rate by this formula

$$\Delta NRFR = (FEDR01_{n+1}/FEDR01_n) - 1 \quad (9)$$

Now we can get excess return by using each month's NRFR to minus each month's return. And we take the average return and standard deviation of each stock for analysis later. Also, co-variance is needed. So, we make a correlation matrix to construct co-variance for all stocks. Then, we can actually move into our two models. The return of portfolio in Markowitz model is follow by this formula

$$\text{Return rate} = \text{SUMPRODUCT}(\text{stock weights: average return}) \quad (10)$$

$$\text{SQRT}(\text{MMULT}(\text{MMULT}(\text{weights*standard deviation}, (\text{correlation matrix}), \text{TRANPOSE}(\text{weights* standard deviation})) \quad (11)$$

We randomly set a group of weights without any constraints to see how two different models behave in our experiment. From the table of weights in the portfolio, we assume that the weights of the Markowitz model and Index model are identical. The weights of the SPX index and ten stocks are showing in Table 2: SPX0.04, QCOM0.04, AKAM0.04, ORCL0.04, MSFT0.04, CVX0.04, XOM0.04, IMO0.04, KO0.04, PEP0.04, and MCD0.6. The total weight is 1.00.

From the Table 3, we get the results that Return 14.397%, StDev 16.511%, Sharpe Ratio 0.872 in Index model Portfolio. We get the results that Return 14.397%, StDev 16.758%, Sharpe Ratio 0.859 in Markowitz model Portfolio.

Table 2. Weights of SPX index

Sum	models	SPX	QCOM	AKAM	ORCL	MSFT	CVX	XOM	IMQ	KO	PEP	MCD
1.00	weight	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.60

Table 3. Results in MM & IM

Index model Portfolio			Markowitz model Portfolio		
Return	StDev	Sharpe	Return	StDev	Sharpe
14.397%	16.511%	0.872	14.397%	16.758%	0.859

Table 4. Minimum variance portfolio and maximum shape portfolio in MM

	SPX	QCOM	AKAM	ORCL	MSFT	CVX	XOM	IMO	KO	PEP	MCD	Return	StDev	Sharpe
Minimum Variance	0.169	0.007	-0.001	0.094	-0.018	-0.069	0.208	-0.062	0.205	0.374	0.094	9.193%	13.202%	0.696
Maximum Sharpe	0.880	0.126	0.067	0.247	0.262	0.230	-0.208	0.1271	0.053	0.422	0.551	18.298%	18.626%	0.982

Table 5. Minimum variance portfolio and maximum shape portfolio in IM

Index	SPX	QCO M	AKA M	OR CL	MSF T	CVX	XOM	IMO	KO	PEP	MCD	Retu rn	StD ev	Sharpe
Minimum Variance	0.02 4	- 0.02 9	- 0.01 7	0.01 8	0.00 9	0.015	0.093	- 0.036	0.307	0.425	0.189	35.6 26%	55.5 80%	0.641
Maximum Sharpe	- 0.81 9	0.08 8	0.07 2	0.11 3	0.27 8	0.071	-0.049	0.046	0.251	0.424	0.524	16.6 64%	16.1 89%	1.029

In this case, the two models do show almost identical results. Then, we need to find out the minimum variance portfolio and maximum shape portfolio. From the portfolio of Markowitz model and the portfolio of Index model, we get the results as Table 4 and Table 5.

Furthermore, we want to test that if two models can still have identical or similar results at different constraints.

Constraint 1: Both models have almost the same minimum variance portfolio, maximum Sharpe portfolio, and efficient frontier. The maximum of the Cal line in Markowitz is larger than the one in Index model. The maximum of the minimal variance frontier in Markowitz is larger than the one in Index model. The maximum of the minimum return frontier in Markowitz is larger than the one in Index model. The range of point dispersion of the portfolio is almost the same in both models.

Constraint 2: The maximum of the Cal line in Markowitz is larger than the one in Index model. The maximum Sharpe Portfolio in Markowitz model is larger than the one in Index model. The minimum of the minimal variance frontier in Markowitz is smaller than the one in Index model. The minimum variance portfolio, efficient frontier, and minimum return frontier are almost the same in both models. The range of point dispersion of the portfolio is almost the same in both models.

Constraint 3: The maximum of the Cal line in Markowitz is larger than the one in Index model. The minimum variance portfolio, maximum Sharpe portfolio, minimal variance frontier, efficient frontier, and minimum return frontier are almost the same in both models. The range of point dispersion of the portfolio is almost the same in both models.

Constraint 4: The maximum of the Cal line in Markowitz is larger than the one in Index model. The maximum Sharpe Portfolio in Markowitz model is larger than the one in Index model. The minimum variance portfolio, minimal variance frontier, efficient frontier, and minimum return frontier are almost the same in both models.

Constraint 5: The minimum variance portfolio, maximum Sharpe portfolio, Cal line, minimum return frontier, minimal variance frontier, and efficient frontier

are almost the same in both models. The range of point dispersion of the portfolio is almost the same in both models.

By calculating the minimum variance portfolio and maximum Sharpe portfolio on efficient frontier based on five constraints of two models, we find that both Markowitz model and Index model show a quite indifferent accuracy in estimating ten stocks' risks and returns. All five minimum variance frontiers of the two models present a similar shape which is bullet shaped. Also, the five frontiers show a similar tendency that the return of all that are minimum variance portfolio increases as the standard deviation increases; the return of all portfolios below minimum variance portfolio decreases as the standard deviation increases. Sometimes the boundary of efficient frontiers or minimum return frontiers differs between two models within 1% to 2% of return at the same standard deviation.

Furthermore, the difficulty of processing is different between the two models. There are $2N+N*(N-1)/2$ estimators for N stocks. In other words, there are 77 estimators for Markowitz model. On the contrary, Index model only need $3N+2$ estimators for N stocks. In this case, 35 estimators are enough to get a similar portfolio using Index model. As you can see from the comparison above, Index model only need half of the estimators of the Markowitz model to get the same result. Based on 10 stocks and 1 market index, the gap of the two models on using estimators is not that prominent. As the number of stocks increases to a larger scale, the difference in the number of estimators between the two models will grow exponentially. But the advantage of the Index model is prominent. The simplification process reduces the number of estimators and increases the inaccuracy of returns and risks. Even though there is a slight difference between the two models based on our research, as the number of stocks increase, there might be a big gap on the accuracy of the two models.

5. CONCLUSION

In this research, we want to get an investment portfolio through data analysis. At the beginning, we collected data on ten stocks. Then, we used tools in Excel for data analysis and drawing graphs. After obtaining two

models with five constraints, we compared their similarities and differences. In the end, we came up with a portfolio that we thought was optimal.

In general, two models plot almost the same graph for investor to determine the minimum variance portfolio frontier of ten stocks and 1 SPX index. So, investors can have the same investment portfolio by using either Markowitz model or Index model; however, the effort needed to get the same portfolio is different between the two models. Markowitz model needs far more estimators compared with Index model. To conclude, Index model, compared to Markowitz model, is more practical in the real situation of the market. The simplification of calculating co-variance helps Index model decrease the demand for amount of estimators. Hence, investors can frequently and largely use Index model to help themselves find optimal portfolios.

There are two main shortcomings in this project. Firstly, after the experiment and data analysis, we only came up with a basic investment portfolio. We haven't thought much about these 10 stocks just come to our conclusions based on the two data models and basic analysis methods. Although our conclusion is accurate, we think it really can be used as a reference for investment. The data is still insufficient, and we don't think it's representative of the entire industry. In the future, we will collect more raw data of stock companies in the same industry. Then, similar methods of this research were used for analysis. We will use a large number of stock companies to show how the industry has performed in the past, and we will try to predict both the short and long term future of the industry.

Secondly, our analytic steps are accurate, but it does not mean that our investment strategy is absolutely reasonable. Although we have collected a lot of data, it's all from the past. In recent two years, due to the sudden epidemic, the volatility of the stock market has been much more violent than that of the previous years. The data of the past history cannot be used to predict the future completely and accurately. Therefore, a qualified portfolio is not only based on data results but also on an investigation of the industry structure. We will collect more news and financial data in the future. We will handle this textual information and finally get a deeper understanding of stock companies and the health of industry.

REFERENCES

- [1] Portfolio Selection. Markowitz, H. Journal of Finance, The. 1952
- [2] Equilibrium in a capital asset market. Mossin, J. Econometrica. 1966
- [3] Sharpe, William F. A Simplified model for Portfolio Analysis. Management Science, 1963, 9(2): 277-293.
- [4] Sharpe W. Capital Asset Price: A Theory of Market Equilibrium under Conditions of Risk. Journal of Finance, 1964, 19(3): 425-442.
- [5] Lintner J. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. Review of Economics and Statistics, 1965, 47(1): 13-37.
- [6] Jan Mossin. Equilibrium in a Capital Asset Market. Econometrica, 1966, 34(4): 768-783.
- [7] Ross S. The Arbitrage Theory of Capital Asset Pricing. Journal Of Economic Theory, 1976, 13(3): 341-360.
- [8] Tienyu Hwang, Simon Gao, Heather Owen. Markowitz efficiency and size effect: evidence from the UK stock market. Review of Quantitative Finance and Accounting. 2014 (4)
- [9] Kan R, Zhou G. Optimal Portfolio Choice with Parameter Uncertainty. Journal of Financial and Quantitative Analysis, 2007, 42(3): 621-656.
- [10] Hoevenaars R, Molenaar R, Schotman P, Steenkamp T. Strategic Asset Allocation for Long-term Investors: Parameter Uncertainty and Prior Information. Journal of Applied Econometrics, 2014, 29(3): 353-376.