Empirical analysis in effective portfolio allocation through Markowitz and index models

Zhongquan(Bill) Zhang , Zixin Liu

New York University(NYU)-2023 Summer Academic Research Program(ASC)

Abstract
This study uses the Markowitz and index portfolio optimization models to analyze ten stocks’ returns in the past 20 years. Under the Markowitz and index models, we identify suitable input, such as returns and standard deviation. To optimize the inputs of the two models, we identify three realistic additional constraints and export-related forms. We compare the two models under the same constraint and then draw conclusions based on the differences.

Keywords: Markowitz Model; portfolio; optimization models; Index Model

1. Introduction
The paper aims to put into practice the concept of modern portfolio theory, such as Markowitz’s optimal portfolio selection and some basic index models. We then analyze the monthly total return data of 10 stocks. One is the stock index grade risk-free rate. We calculate the Markowitz model and other suitable optimization inputs from that data and use these within the three constraints we’ve given. For example, the minimum risk and optimal portfolios find additional constraints on the small return portfolio. Ultimately, we analyze all the results of the computational data to compare each optimization problem (MM/IM) and an optimal solution problem between multiple solutions to get the optimal portfolio of the selected stock data.

2. Theoretical Foundations
According to the Markowitz mean-variance Model, investments in risky assets such as securities must first solve two core problems: expected return and risk. Modern portfolio theory assumes that investors are risk-averse, meaning that given two portfolios with the same expected return, investors will gravitate toward the less risky portfolio. As a result, investors can only take on more risk if they get a higher expected return. Conversely, investors who want higher expected returns must take on more risk. The exact trade-offs are not the same for all investors. Investors will evaluate the trade-offs differently based on their risk aversion characteristics. In other words, a rational investor will not invest in a portfolio if the second portfolio has a more favorable expected return on risk; if an alternative portfolio with that level of risk exists, it will have a higher expected return. The following part is an introduction to the investment data model.

2.1 Markowitz Model(MM)
In 1952, Harry Markowitz published the paper “Portfolio Analysis,” which provides basic ideas for measuring the returns and risks of securities. He used the theory of probability and mathematical statistics to construct a model framework for analyzing securities prices. In his model, the price of a security is a random variable, and the mathematical expectation and variance of this random variable can measure the value and risk of the security. Starting from general psychological analysis, Markowitz assumed that economically rational individuals have a risk aversion tendency; that is, when the return is certain, the investment behavior with the least risk is adopted; that is, in his model, investors pursue the investment portfolio with the least variance when the return is certain. Although the model excludes the analysis of risk enthusiasts, there is no doubt that the vast majority of people, in reality, are risk averse, so its analysis is general. Based on a series of theoretical assumptions, Markowitz’s analysis of the stock market concludes that there is an efficient portfolio in the stock market. This theory provides a theoretical basis for financial practice to search for this combination, and its analytical framework has become the basis for the theoretical analysis of modern financial engineering. Meanwhile, the mean-variance model is the most important part of it. Mean-Variance Model Investors invest a given amount of money over a certain period. He buys some securities at the beginning of the period and then sells them at the end. So, at the beginning of the period, he has to decide which securities to buy and how to allocate...
the money among them. This means that the investor has to choose the optimal combination of all possible securities at the beginning of the period. At this time, investors have two decision-making goals: the highest possible rate of return and the lowest possible uncertainty risk. The best goal is to achieve the best balance between these two mutually restrictive goals. The investment model established is the mean-variance model. Investing in securities and other risky assets first must solve two core problems: expected return and risk. So, determining the risk and return of portfolio investment and balancing these two indicators for asset allocation is an urgent problem for market investors to solve. It was against this background, in the 1950s and early 1960s, that Markowitz’s theory came into being.

The mean-variance model is based on several assumptions:

1. When investors consider each investment choice, it is based on the probability distribution of security returns over a certain holding period.
2. Investors estimate the risk of a securities portfolio according to the variance or standard deviation of the expected return rate of securities.
3. The investor’s decision is based solely on the risks and returns of the security.
4. At a certain level of risk, investors expect the maximum return; in contrast, investors want the least risk at a certain level of return.

According to the above assumptions, Markowitz established the calculation method of expected return and risk of securities portfolio and the efficient boundary theory and established the mean-variance model of asset optimal allocation. The objective function of the model is the variance of the portfolio.

$$\sum_{i} x_i \sigma_i^2$$

The constraint condition is:

$$\sum_{i} x_i \mu_i \geq \mu, \quad \sum_{i} x_i = 1, \quad x_i \geq 0.$$  

If security allows short selling, the corresponding constraint of $x_i \geq 0$ can be removed. Here $x_i$, it represents the proportion of funds invested in security I, and the total proportion of all investments $\sum x_i \leq 1$ does not exceed the budget. The expected return of the portfolio is $\sum_{i} x_i E(r_i) \geq \mu$. To achieve the desired return $\mu$, the risk $\sigma^2$ can be minimized by adjusting the capital ratio $x_i$. Graph 1 shows individual assets that are distributed according to expected returns and standard deviation. The distribution of these assets consists of a curve; we call it an efficient frontier. In this curve, we can find the global minimum-variance portfolio, which represents the optimization of portfolio allocation.

2.2 Index Model(IM)

The index model assumes that only one macro factor will cause stock return risk, which can be represented by the return rate of a market index, such as the S&P 500 index (S&P 500). 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 here by a company-specific factor $\alpha$), the return of macro-events affecting the market, and the unpredictable micro-events affecting only the company composition. $\beta_i(r_m-r_f)$ represents the market movement under the influence of stocks, and $\beta_i(r_m-r_f)$ represents the market risk under the influence of corporate factors. Macro events, such as changes in interest rates and changes in labor costs, can cause systematic risks that affect the returns of the entire stock market. Company-specific events are micro-events that will cause changes in the earnings of a specific company, which will affect the company’s earnings but the impact the entire economy is negligible. In a portfolio, the unsystematic risk caused by company-specific factors can be reduced to zero by discretization. Graph 2 below explains how the Index Model simplifies stock risk.
This exponential model is based on the following assumptions:

1. Most stocks have positive covariance because they react similarly to macro events.
2. However, some firms are more sensitive to these factors than others, and the coefficient $\beta$ controls this sensitivity.
3. The covariance between bonds is due to different degrees of response to macro events. So, the covariance of each stock is equal to their beta multiplied by the market variance: $\text{Cov}(R_i, R_k) = \beta \sigma_i \sigma_k$.

The last equation greatly reduces the amount of covariance calculations. Otherwise, the covariance of bonds in the portfolio must be calculated using historical returns, and each bond must be calculated separately. With this equation, only $\beta$ and the variance of the market are required. So, the single exponential model greatly reduces the amount of calculation.

### 3. Research Design

#### 3.1 Stocks Description and Analysis

In the research, we chose 10 stocks from Yahoo! Finance and collected 20 years of historical daily total return data about them as the objectives we focus on. The ten stocks come from different sectors: technology, financial services, customer defense, and healthcare. We will briefly analyze the ten stocks from different sectors in the following.

1. **Technology–NVIDIA Corporation, Cisco Systems, Inc., and Intel Corporation**

   **NVIDIA Corporation** is an American multinational technology company based in Santa Clara, California. NVIDIA produces graphics processing units (GPUs) and application programming interfaces (APIs) for data science and high-performance computing. Recently, this firm has broadened its focus from traditional PC graphics applications to more complex and favorable opportunities, such as artificial intelligence and autonomous driving. According to a Morningstar report (July 2023), Nvidia might be the biggest winner in the rise of ChatGPT and generative AI. Figure 3 shows NVDA’s monthly returns in the past 20 years. This company has a lower standard deviation in its monthly returns and more constant positive returns.

   ![Monthly Return of NVDA within Recent 20 years](image)

   Cisco Systems, Inc., abbreviated by CSCO, is an American-based multinational digital communications technology company in San Jose, California. CSCO “designs, manufactures, and sells Internet Protocol-based networking and other products related to the communications and information technology industry” (Yahoo! Finance). In December 2021, CSCO had a market capitalization of around 267 dollars. Figure 4 shows CSCO’s monthly returns in the past 20 years. Generally, CSCO has had a small variance in its monthly and constant stock returns over the 20 years.

- $\alpha + \beta R_M + \epsilon$
- $R_M = R_m - r_f$
Intel Corporation, abbreviated by INTC, is a technology company headquartered in Santa Clara, California. As a multinational corporation, Intel has subsidiaries in India, Canada, France, and others. Intel designs, develops, manufactures, markets, and sells computing and related products worldwide (Yahoo! Finance). Intel is one of the world’s biggest semiconductor chip manufacturers in revenue. Figure 5 shows Intel Corporation’s monthly returns in the past 20 years. In the first ten years, INTC’s deviation in its monthly returns decreased, but it rose again in the last ten years. Returns were not constant in the past 20 years.


The Goldman Sachs Group, Inc. (GS) is a financial institution headquartered in New York. GS provides services for both corporations and individuals. These services include three segments: global banking & markets, asset & wealth management, and platform solutions. The global banking & markets segment provides financial advisory services and corporate defense activities. The asset & wealth management segment aims to transform insights into opportunities that satisfy investors’ needs across public and private markets (Goldman Sachs). The Platform Solutions segment offers credit cards and point-of-sale financing for purchasing goods or services (Yahoo! Finance). Figure 6 shows the monthly returns of the Goldman Sachs Group, Inc. (GS) in the past 20 years. The monthly returns of GS didn’t have large changes within 20 years. Most monthly returns are located between -10% to 10%.
U.S. Bancorp (USB) is a financial services holding company headquartered in Minneapolis, Minnesota. USB provides various financial services to individuals, businesses, institutional organizations, governmental entities, and other financial institutions in the United States. It operates in Corporate and Commercial Banking, Consumer and Business Banking, Wealth Management and Investment Services, Payment Services, and Treasury and Corporate Support segments (Yahoo! Finance). USB also provides depository services, lending services, corporate and purchasing cards, etc. Figure 7 below shows USB’s monthly returns in the past 20 years. USB’s stock had a small deviation in monthly returns, ranging from -10% to 10%. The largest deviation appeared near July 6th, 2009, and the returns were over 20%.

![Figure 7: Monthly Returns of USB within 20 years](image)

The Toronto-Dominion Bank (TD) is a financial company in Toronto, Canada. This company and its subsidiaries provide various financial products and services in Canada, the United States, and other nations (Yahoo! Finance). TD operates in three main segments: Canadian Retail, U.S. Retail, and Wholesale Banking. The company also provides personal deposits, credit cards and payments, capital markets, etc. The financial institution offers its products and services under the TD Bank and America’s Most Convenient Bank brand names, and it has a strategic alliance with Canada Post Corporation (Yahoo! Finance). Figure 8 shows TD CN’s monthly returns in the past 20 years. TD CN’s stocks had positive returns for the past 20 years. However, several big deviations were over 10% or below -10%.

![Figure 8: Monthly Returns of TD CN in the past 20 years](image)

The Allstate Corporation (ALL) is a financial company headquartered in Northbrook, Illinois, and it provides property and casualty, as well as other insurance products in the United States and Canada. The services ALL provides mainly include Allstate Protection, Protection Services, Allstate Health and Benefits, and Run-off Property-Liability segments (Yahoo! Finance). The Allstate Protection segment offers private passenger auto and homeowners insurance, while the Protection Services segment provides consumer product protection plans and related technical support for mobile phones, and consumer electronics (Yahoo! Finance). The Allstate Health and Benefits segment provides life, accident, critical illness, short-term disability, and other health insurance products. The Run-off property liability segment offers property and casualty insurance. We calculate monthly returns for the Allstate Corporation over the past 20 years in Figure 9. Stocks of ALL have larger deviations compared to other stocks we research. Especially on January 30, 2001, this stock had a -33.85% return, which was the smallest during the 20 years.

![Figure 9: Monthly Returns of Allstate Corporation](image)
III. Customer Defensive–The Procter & Gamble Company and Colgate-Palmolive Company

The Procter & Gamble Company (P&G) is an American multinational corporation in the customer defensive sector, and its headquarters is in Cincinnati, Ohio. P&G specializes in a range of personal health and personal care products. These products are used for baby care and home care in living rooms, kitchens, laundry rooms, and bathrooms (P&G). In 2014, P&G recorded $83.1 billion in sales. Figure 10 shows PG’s monthly returns within 20 years. PG’s stocks had dramatically large deviations in its stocks, which moved between -10% and 10% frequently.

Colgate-Palmolive Company (CL) is an American multinational consumer goods company headquartered in Manhattan, New York. The company sells health care, personal care, and veterinary products. This company offers toothpaste, toothbrushes, mouthwash, shampoos, and hand soaps. Figure 11 is a diagram that represents the monthly returns of CL stocks within 20 years. CL stock had a large deviation in its monthly returns; sometimes, its monthly returns were larger than 10% or less than -15%.
IV. Healthcare—Johnson & Johnson

Johnson & Johnson (JNJ) is a pharmaceutical industry company based in New Brunswick, New Jersey. Johnson & Johnson provides skin health/beauty products under different brands. This company also offers MedTech and pharmaceutical products. MedTech provides diverse healthcare expertise and purposeful technology (Johnson & Johnson). Some pharmaceutical medicines are also provided to treat devastating and complex diseases in our time. Figure 12 shows JNJ stocks’ monthly returns within 20 years. Overall, JNJ stock was constant over the past 20 years and didn’t have large deviations.

3.2 Data processing

In this section, we introduce how we use the ten stocks and their historical return rates, data we choose for our research, to calculate their changes in returns according to their risks.

I. Introduction:

Rate of return: rate of return refers to the return on investment of the fund, which is one of the important indicators for investors to judge the investment effect of the fund. Usually expressed as a percentage, the formula is (Net value per unit - Net value per unit at the time of purchase)/Net value per unit at the time of purchase *100%.

An investment fund’s ultimate purpose is to obtain income, so the rate of return is one of the core indicators of the fund. When choosing a fund, investors often use the historical rate of return as the evaluation criterion. However, the historical rate of return does not represent the future return situation (discussed above); investors also need to consider other factors, such as the fund’s investment strategy, the management team, the risk level, etc.

Risk indicators: risk indicators are used to measure the degree of risk of fund investment indicators; generally speaking, investment returns and risks are proportional; the higher the risk, the higher the return may be. Common risk indicators include volatility, maximum retracement, Sharpe ratio, etc. Among them, volatility refers to the volatility of the fund’s net worth in a certain period; the greater the volatility, the higher the fund’s risk. Maximum retracement refers to the maximum decline in the fund’s net value in a certain period. It is also an important indicator for measuring the risk of the fund. Investors need to pay attention to the maximum retracement of the fund to judge whether it aligns with their risk tolerance.

Sharpe ratio: The Sharpe ratio measures the risk-return ratio of the fund index; it can reflect the level of risk and return of the fund. The higher the Sharpe ratio, the better the risk-return ratio of the fund. It calculates the ratio of the return of an asset over the risk-free rate of return to the volatility of the asset. It is calculated by dividing the
excess return by the volatility of the asset. The higher the Sharpe ratio, the higher the excess return of each risk, which is one of the important indicators to measure the risk-return relationship of assets. Through the Sharpe ratio, investors can know the return performance of different assets under the same risk so as to better formulate investment strategies and choose asset portfolios. When investing in funds, investors must choose their risk indicators according to their risk tolerance to choose the right fund. Capital Allocation Line (CAL) represents all possible combinations of risk and return for the investor. The slope of the capital allocation line is denoted as S, which is equal to the expected return of the entire portfolio for every one-unit standard deviation increase. The slope S of CAL is called the reward-risk ratio or Sharpe ratio; it represents how fast one stock’s returns rise when its risk is higher. Figure below (figure 13) clearly explains how we get one stock’s expected returns and related risks using indifference curves and CAL tools.

II. Restriction condition

(1) We set up three constraints for our input data to avoid the possible randomness of one simple test. The first additional optimization constraint is designed to mimic the Financial Industry Regulatory Authority (FINRA) regulation T, which allows broker-dealers to allow their clients to hold positions. 50% or more of which is funded by the client’s account assets:

\[ \sum_{i=1}^{n} |w_i| \leq 2 \]

\[ |w_i| \leq 1, \text{ for all } i; \]

(2) The second additional optimization constraint is designed to simulate some arbitrary “box” constraints on weights, which may be provided by the client (typical for hedge funds):

\[ w_i \geq 0, \text{ for all } i \]

Due to this ‘box’ restriction, which needs to be implemented in the model for \( \forall i \), the purpose of this additional optimization restriction is to mimic some of the arbitrary weight” box restrictions that hedge fund investors often provide and to mimic the typical constraints of an SEC registered hedge fund or major asset manager. The unrestricted scenario corresponds to the third instance. The portfolio areas that are acceptable in general, and the valid boundary areas in particular, are shown as having no additional optimization constraints for the problem if you do not use constraints - the constraints of the “free” problem. The following constraint, \( \geq 0 \), is the so-called “box” constraint for \( \forall i \). U.S. open-end funds do not allow short positions, so this new optimization constraint mimics the typical U.S. mutual fund business restrictions. In particular, any registered investment enterprise is prohibited from short-selling any security unless it is involved in an underwriting business that does so.

The last constraint =0 corresponds to the fact that there are no indexes in our portfolio. This restriction is to see if adding a broad index to our portfolio has a good or negative effect.

III. Data processing

We obtain the overall data map after the daily data processing and analysis summary. We organized the disordered daily data into a monthly quantitative table in Excel to normalize it. Here are some images of the data (figure 14 and 15):
By utilizing the theory and tools discussed above, we finally export two graphs: one represents the Markowitz Model (MM), and another one represents the Index Model (IM). Standard deviation is on the x-axis, while returns are on the y-axis. The two graphs below (Figures 16 and 17) show the changes in standard deviation and returns under different constraints and how returns connect with standard deviation under these constraints.

(1) Markowitz Model (MM)
Analyzing the Markowitz Model, we can summarize that one stock may have two different returns under the same standard deviation. Under constraint 1, \( \sum_{i=1}^{11} w_i \leq 2 \), higher returns for one specific standard deviation are the efficient frontier, while lower returns are the inefficient frontier. The minimal variance frontier is a curve connecting all points representing the efficient and inefficient frontier.

When considering constraint 2, we notice that the frontier curve consisting of efficient and inefficient frontier points has a steeper slope. That means under constraint 2, the ten stocks have higher returns when the standard deviation is the same. Constraint 3 makes the frontier curve steeper than the curve under constraint 2. The frontier curve under constraint 4 has the flattest slope among these curves.

(2) Index Model(IM)
In the graph on IM, we notice that frontier curves have a similar shape to the ones in the graph of MM. Frontier curves under constraint 4 have the flattest slope. Frontier curves have steeper slopes from constraint 1 to 3. However, the Index Model makes covariance estimation easier and improves our analysis efficiency.

5. Conclusion

5.1 Comparisons on Minimum Variance and Maximum Sharpe Ratios

In this part, we have a closer look at the differences between the Markowitz and Index Model. We compared the minimum variance, which we named MinVar in our paper, and maximum Sharpe ratios, which we named MaxSharpe, between the two models. The two graphs below (Figures 18 and 19) show MinVar and MaxSharpe for each stock for both MM and IM.

Figure 18: MinVar and MaxSharpe under five constraints in the Markowitz Model

Figure 19: MinVar and MaxSharpe under five constraints in the Index Model
By comparing stocks and the returns and standard deviations that stocks bring, we notice that stocks have lower MinVar under the Markowitz Model. Lower MinVar means lower risks. Also, returns and standard deviation are lower under the Markowitz Model.

MaxSharpe compares the return on an investment with its risk. Also, it represents the risk premium of an investment versus a safe asset such as a Treasury bill or bond (Fernando, 2023). Higher MaxSharpe represents higher returns with the same risk. When we compare MaxSharpe between the two models, we don’t find unity among these data. Some MaxSharpe are larger in MM, and some MaxSharpe are not. Therefore, the MaxSharpe part may not have a clear answer. However, when we calculate MinVar, we find that stocks have higher returns and lower risks under the Markowitz Model.

5.2 Comparison between the Markowitz Model and Index Model

The Markowitz Model is built based on historical data but doesn’t work on current stocks and related data. Also, the model requires many estimates, including returns and variances. However, we do not have to worry about these problems in our analysis. We have used stock information in the past, and analyzing ten stocks doesn’t take a long time, even though we need estimates on returns and variances. We can use this model in our paper.

Compared to the Markowitz Model, the Index Model has an easier calculation process. However, the problem is that the model can’t reflect the real risk in the market. The index Model simplifies risks into systematic risks and firm-specific risks. If we hope to reflect the ten stocks’ real risk, the Markowitz Model is definitely a better choice.

6. Reference

(1) Exponential Model Baidu https://ms.mbd.baidu.com/r/14hSbYiPa6Q?f=cp&u=ba9b10970a963c7d
(3) Goldman Sachs https://www.goldmansachs.com/
(4) Markowitz Portfolio theory model Baidu https://mq.mbd.baidu.com/r/14hRQVvVqW4?f=cp&u=b283449b7a16b22a
(6) P&G https://us.pg.com/
(7) The three basic assumptions of technical analysis (July 15, 2020). zhuilan.zhihu.com/p/160664350