Diversification and Risk Assessment using Data Science: Downside Risk vs. Mean Variance Optimization

YASHASVI BORA
Maharaja Agrasen Institute of Technology

Abstract- Diversification and risk assessment are essential aspects of portfolio management. In this research paper, we explore the application of data science techniques to compare two popular portfolio optimization methods: Downside Risk and Mean Variance Optimization. The study begins by collecting historical financial data for a set of stocks. Using this data, we calculate the portfolio weights for both optimization methods. The Mean Variance Optimization technique aims to maximize returns while minimizing volatility, while the Downside Risk approach focuses on minimizing the potential for losses. To measure the diversification benefits of each strategy, we analyze the sector allocation of the optimized portfolios and compare the sortino ratios. By computing sector weights, we acquire insights into the concentration or dispersion of investments across different industries. This research allows us to evaluate the diversification obtained by each optimization method and compare their efficacy in distributing risk across sectors. The results generated from the data analysis and visualization illustrate the contrasting characteristics of the Downside Risk and Mean Variance Optimization approaches. The sector-wise analysis highlights variations in sector allocations, illustrating the disparities in diversification techniques adopted by each strategy. Additionally, the risk assessment analysis provides insights into the potential downside risks connected with each portfolio. This research contributes to the field of portfolio management by providing a comprehensive comparison between Downside Risk and Mean Variance Optimization methods. It demonstrates the potential of data science techniques in evaluating portfolio diversification and risk assessment. The findings can assist investors and financial professionals in making informed decisions regarding portfolio construction and risk management strategies.

Indexed Terms- Diversification, Risk Assessment, Portfolio Optimization, Data Science, Downside Risk, Mean Variance Optimization

I. INTRODUCTION

Diversification and risk assessment are essential parts of portfolio management that play a vital role in achieving optimal investment outcomes. With the introduction of data science approaches, investors now have tremendous tools at their disposal to make informed decisions based on advanced analyses. In this research article, we dig into the domain of portfolio optimization and risk assessment by comparing two famous methods: Downside Risk and Mean Variance Optimization.

In particular, we study a carefully selected collection of top 10 equities that represent various industries. By focusing on these well-established and significant companies, we want to assess the usefulness of the Downside Risk and Mean Variance Optimization approaches in managing portfolios consisting of high-profile stocks.

The basic purpose of portfolio optimization is to establish a balance between maximizing returns and avoiding risks. Mean Variance Optimization, pioneered by Harry Markowitz, tries to discover an allocation of assets that optimizes the portfolio’s projected return while minimizing its total risk [1]. This method assumes that returns are normally distributed and measures risk purely through the variance of returns.

On the other hand, Downside Risk optimization adopts a more cautious approach by prioritizing the reduction of potential losses. It lays higher emphasis on downside protection and examines the downside potential and downside risk of investments. This method tries to shield against poor market conditions and conserve capital during downturns.

To assess the performance of both optimization strategies, we take historical financial data of the selected stocks and employ data science techniques to generate the portfolio weights for both Downside Risk and Mean Variance Optimization. These weights describe the ideal allocation of investments across different assets, reflecting the composition of the portfolios. By comparing sector allocations and measuring downside risk metrics, we shed light on the strengths and weaknesses of each approach. Overall, this research attempts to expand our understanding of diversity and risk assessment in portfolio management and highlights the potential of data science in driving educated investing decisions.

II. LITERATURE REVIEW

The process of choosing the optimal portfolio (asset distribution) among all those under consideration in a portfolio is known as portfolio optimisation. The objective typically maximizes factors such as expected return, and minimizes costs like financial risk. Harry Markowitz first presented modern portfolio theory in his PhD dissertation in 1952. It is predicated on the idea that a portfolio's expected return should be maximized for any given level of risk. Investors are forced to choose between risk and expected return for portfolios that satisfy this requirement, known as efficient portfolios, because getting a greater expected return necessitates taking on more risk. The efficient frontier is a curve that graphically illustrates the risk-expected return connection of efficient portfolios. All efficient portfolios are well-diversified, with each represented as a point on the efficient frontier. [1] [2]

The portfolio optimization problem is specified as a con- strained utility-maximization problem. The predicted portfolio return (net of transaction and financing costs) minus a cost of risk is how portfolio utility functions are frequently formulated. The cost of risk, or unit price of risk, is the latter component and is defined as the portfolio risk times a risk aversion factor.

The" critical line method" was created by Harry Markowitz and is a general method for quadratic programming that can handle additional linear restrictions as well as upper and lower bounds on holdings [3]. This is how Mean Variance Optimization first came into existence.

The probability of financial loss is referred to as downside risk. In other words, it is the probability that the actual return will be lower than predicted, or the ambiguity about how much of a difference there will be [4]. The standard deviation is an example of a deviation risk measure, which assesses both the upside and downside risk. Risk measurements typically quantify the downside risk. Downside beta or lower semi- deviation measurements can specifically be used to quantify downside risk. The industry standard is known as the statistic below-target semi-deviation (BTSD), sometimes known as target semi-deviation (TSV) [5].

One research paper says that mean variance optimization is a better approach since minimizing the variance typically results in a lower downside deviation and a greater Sortino ratio because it can be measured more precisely, minimizing the semivariance is more in accordance with the genuine preferences of a rational investor. In actuality, the issue is with the idea of downside risk rather than with a precise measure of it. Though theoretically entirely fair, minimizing only losses rather than total variability requires the investor to estimate the necessary inputs using a small portion of the information. This translates into imprecise estimates, which lead to poor out-of-sample portfolio performance [6].

According to P. Cheng (2001) [7], the findings demonstrate that, in terms of real estate allocation, the Downside Risk model generates portfolio compositions that are more realistic and consistent with institutional investors' practices. Ex ante Downside Risk portfolio return distributions often have bigger median returns and narrower left tails than MV portfolios. These findings are positive for investors who are primarily concerned with downside risks because the Downside Risk method not only looks to improve portfolio performance with greater median returns, but it is also more consistent with investors' perceptions of risk.

This project aims to study both the methods under similar prerequisites and settings by keeping the stocks and their data same for both of them and solve both the optimization techniques using quadratic programming.

III. METHODOLOGY

A. Data

Since data is the backbone of this project, plenty of financial data is required. I have selected top 10 stocks in every sector of the US stock market for the project. Focusing on the top 10 stocks in each area provides for a sector-specific analysis. Each industry has its particular characteristics, performance metrics, and risk considerations. By selecting the top stocks within each sector, we receive insights into the leading firms and their performance within their respective industries. This sector-based study provides a more sophisticated view of market dynamics and helps identify potential investment opportunities and risks specific to each sector.

To get the necessary data, the yfinance library, a Python tool built for fetching historical market data from Yahoo Finance, was utilized. This library provided fast access to the historical stock price information for each of the selected equities. By choosing the necessary time range of 15 years, a large historical context was gathered, enabling a full examination. To visualize the collected findings, the matplotlib library was applied. This robust library offers a wide range of plotting functions and customization possibilities, enabling for the construction of insightful graphs and charts.

In addition to utilizing the yfinance library for data collection, several other Python libraries played a crucial role in the data analysis process. I used pandas library for data manipulation and analysis, enabling efficient handling of the collected stock price data. This powerful library provided capabilities for data alignment, merging, and computation of various statistical measures.

The numpy library proved essential in conducting numerical computations and array operations. With numpy, fundamental mathematical functions were applied to the data, allowing calculations of portfolio returns, covariance matrices, and predicted returns. The methodology adopted in this research focuses on the calculation of downside risk optimization for a particular portfolio. The target minimum return is set at 20%, signifying the desired amount of return. The risk-free rate applied in the analysis corresponds to the US Treasury Yield for a 10-year term, which currently at the time of authoring this article, sits at 3.76%.

B. Initial Steps for Downside Risk Optimization

To begin, the expected return for each stock in the portfolio is computed using the risk-free rate, along with beta values and volatility measures specific to each stock. This calculation is based on the formula by William Sharpe (1964) [8]:

$$expected\ return = risk\ free\ rate + beta * (volatility)^2$$

The expected return serves as a key input in the subsequent steps of the analysis.

Subsequently, the downside deviation and downside risk for each stock are determined. The downside deviation quantifies the variability of returns below the target minimum return. It is calculated by calculating the square root of the mean of the squared minimum of (returns - target return, 0). Here, returns represent the historical returns of each stock.

The downside risk is then determined, representing the risk-adjusted return for each stock. It is generated by dividing the difference between the expected return and the intended minimum return by the downside deviation. This measure gives an estimate of the risk associated with reaching the intended return.

The formula for the optimization problem can be represented as follows:

Minimize: $w' * P * w$
Subject to: $A * w = b$
$G * w <= h$

Where:

- w is the vector of weights representing the allocation to each asset
- P is the matrix of downside risk values (diagonal matrix)
- A is a matrix representing the equality constraint

(sum of weights = 1)

- b is the vector representing the target minimum return
- G is the matrix representing the inequality constraints
- h is the vector representing the upper bounds on the constraints
- w' denotes the transpose of the weight vector

To optimize the portfolio, a quadratic programming problem is developed utilizing the estimated downside risk metrics. The idea is to maximize the projected return while limiting downside risk. Matrices and vectors are defined to represent the objective function and constraints of the optimization problem. To achieve adequate weight allocations, the lower bound constraint on the weights is adjusted to allow for unfettered allocations to stocks. This improvement gives for greater flexibility in designing the ideal portfolio.

The optimization problem is solved using a specialized solver, specifically the solvers.qp function. This function employs the CVXOPT library and returns the solution to the quadratic programming problem.

Finally, the resulting portfolio weights are calculated, showing the optimal allocation for maximizing return while limiting downside risk. These weights represent the proportion of each stock that should be included in the portfolio to meet the appropriate risk-return trade off. Initially, any weights that are negative are filtered out, retaining just the positive weights. The equities connected with these positive weights are removed, creating a list of assets that contribute to the portfolio. The following stage requires determining the total weight of the positive weights, which is the sum of the weights of all assets with positive allocations. These weights are then translated into percentage values to illustrate the relative contribution of each asset in the portfolio as seen in the" Fig. 1".

According to one scholar, the primary distinction between the approaches comes to the notion of risk [13]. The mean variance technique uses the standard deviation to gauge risk, while the downside risk concept relies on several distinct measurements. However, the downside risk measures in this theory refer to the semivariance. The standard deviation estimates the departure from the mean return that includes upside and downside variances. The semistandard deviation exclude upside deviations and measures only the return deviation occurring below a certain return threshold level set by the perpetrator.

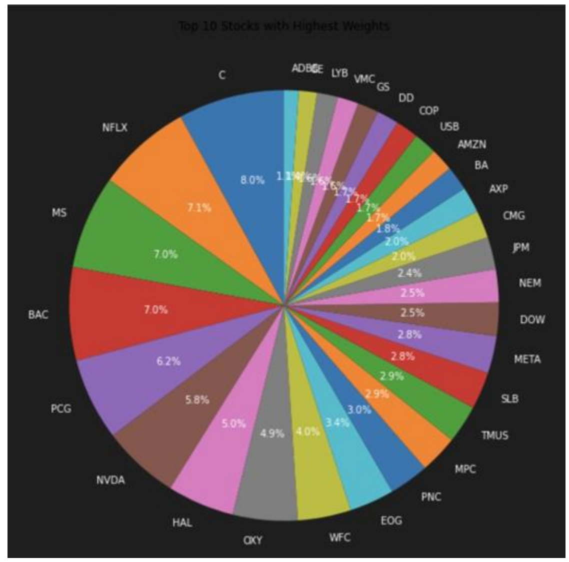

Fig. 1. % Contribution of top 30 stocks by weightage in the Downside Risk Optimization portfolio

C. Initial Steps for Mean Variance Optimization

Firstly, the returns of the portfolio members are determined using the historical price data. Log returns are computed by calculating the natural logarithm of the ratio of the current price to the previous price. This aids in capturing the relative price movements and removing any bias from absolute price levels.

The covariance matrix is then generated based on the log returns of the portfolio elements. This matrix evaluates the link and co-movement between multiple assets, providing useful insights into their joint risk. Next, the optimization issue is constructed using the mean-variance approach. The covariance matrix acts as the input for the problem. The predicted returns for each asset are computed using the risk-free rate and the beta values particular to each asset. These predicted returns represent the mean component of the optimization issue.

The equation for mean-variance optimization using quadratic programming can be represented as follows:

Maximize: $w' * P * w - q' * w\, G$
Subject to: $* \; w <= h$

Where:

- w is the vector of weights representing the allocation to each asset
- P is the covariance matrix of asset returns
- q is the vector of expected returns
- G is the matrix representing the inequality constraints
- h is the vector representing the upper bounds on the constraints
- w' denotes the transpose of the weight vector

Matrices and vectors are defined to represent the objective function and constraints for the optimization issue. The covariance matrix is utilized to generate the P matrix, whereas the negative anticipated returns are used to establish the q matrix. The constraints are established by the G and h matrices, combining the beta values, identity matrix, and lower bound restrictions. The lower bound constraint on weights is adjusted to allow for unfettered allocations to assets by setting the lower bounds to zero. The optimization issue is solved using the quadratic programming solver solvers.qp from the CVXOPT package.

The resulting portfolio weights, indicating the optimal allocation for the given risk-return tradeoff, are generated from the solution. These weights represent the proportion of each asset that should be included in the portfolio to maximize expected returns while considering the covariance among assets.

The weights for each asset are printed, offering a clear insight of the ideal portfolio allocation. The total weight of the portfolio is computed by summing the weights of all assets. The weights are then translated into percentages for each asset as indicated in the" Fig. 2".

IV. ANALYSIS

A. Sortino Ratios

The Sortino ratio is a variation of the Sharpe ratio that differentiates harmful volatility from total overall volatility by using the asset's standard deviation of negative portfolio returns—downside deviation—instead of the total standard deviation of portfolio returns. The Sortino ratio takes an asset or portfolio's return and subtracts the risk-free rate, and then

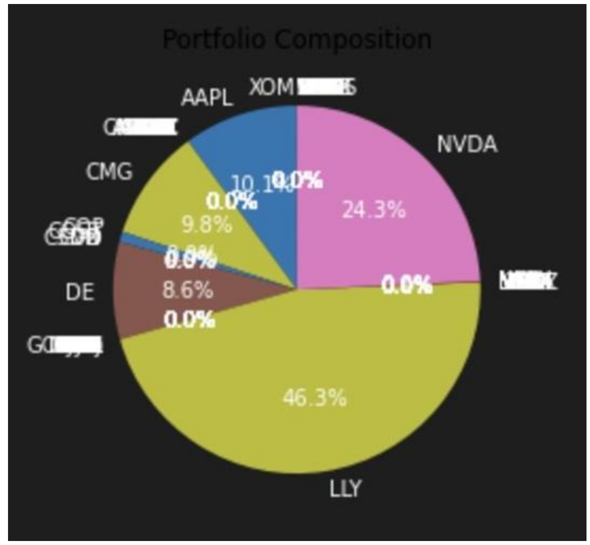

Fig. 2. % Contribution of each stock in the Mean Variance Optimization portfolio

divides that amount by the asset's downside deviation [9]. The ratio was named after Frank A. Sortino.

In this situation, the Sortino ratio for downside risk optimization is computed to be roughly 2.30, while the Sortino ratio for mean variance optimization is approximately 13.52. The higher Sortino ratio of 13.52 for mean variance optimization shows a superior risk-adjusted performance compared to the downside risk optimization.

The greater Sortino ratio generated through mean variance optimization shows that the portfolio's returns, after accounting for the risk-free rate, exhibit a more favorable trade-off with downside risk. This shows that the mean variance optimization strategy has been effective in designing a portfolio that generates a higher level of return compared to the downside risk incurred. It signifies that the portfolio has delivered higher risk-adjusted returns by efficiently minimizing downside risk.

On the other hand, the smaller Sortino ratio for downside risk optimization indicates a considerably inferior risk-adjusted performance compared to mean variance optimization. This shows that the downside risk optimization strategy may not have achieved as efficient risk management in the portfolio development process, leading to a smaller excess return per unit of downside risk.

These contrasted Sortino ratios provide insights into the relative performance of the two optimization strategies in controlling downside risk. The greater Sortino ratio for mean variance optimization suggests its potential advantage in producing better risk-adjusted returns by properly balancing risk and reward.

B. Risk-Return Tradeoff

Risk-return tradeoff states that the potential return rises with an increase in risk. Using this principle, individuals associate low levels of uncertainty with low potential returns, and high levels of uncertainty or risk with high potential returns [10].

In the case of downside risk optimization, the portfolio exhibits a decreased volatility of 3.78% and a moderate return of 9.35%. This suggests that the portfolio is designed to reduce the downside risk and preserve capital amid poor market conditions. The reduced volatility signifies a more consistent investment experience, but it also implies that the portfolio may surrender some potential upside profits. Therefore, the downside risk optimization strategy promotes capital preservation and risk mitigation over maximizing returns.

On the other hand, mean-variance optimization leads in a higher volatility of 19.83% but delivers a substantially larger return of 34.60%. This signifies that the portfolio is exposed to a higher level of overall risk but has the potential for bigger gains. The increased volatility means that the portfolio is more subject to market changes, including both upside and downside moves. The mean-variance optimization strategy seeks to establish a balance between risk and return, and in this situation, it stresses achieving larger returns at the expense of increasing volatility.

In contrast, the mean variance method showcases a higher volatility, indicating a broader range of potential fluctuations in the portfolio's value. This method considers both upside potential and downside risk, aiming to achieve an optimal balance between risk and return. The higher volatility suggests that the mean variance method may allow for greater exposure to market fluctuations and potentially higher returns. However, it also entails a higher level of risk, as reflected by the increased volatility.

C. Diversification

Firstly, the mean variance optimization method concentrates all of the weight on only five companies, indicating a high level of concentration and potential exposure to the performance of these particular stocks. This concentration strategy may lead to increased risk as the portfolio becomes more dependent on the performance of a few select assets. Also, mean variance optimization considers the covariance matrix of asset returns to determine the optimal allocation. If the covariance between the five selected companies is relatively low compared to other companies in the portfolio, the optimizer may find that allocating all the weight to those three companies leads to a more efficient risk-return profile based on the covariance structure.

On the other hand, the downside risk optimization method exhibits a more diversified portfolio allocation, spreading the weight across multiple stocks. This diversification

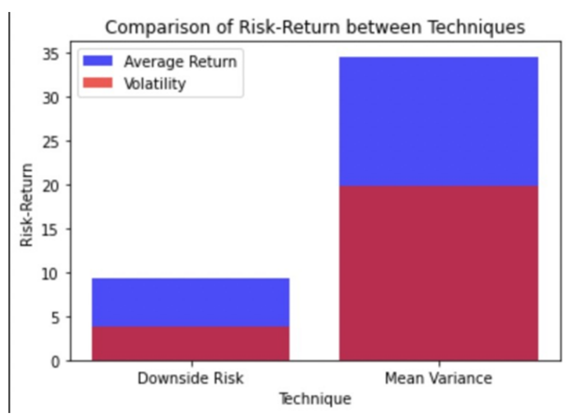

Fig. 3. Risk-Return Tradeoff for Downside Risk and Mean Variance Optimization

approach helps mitigate the concentration risk by reducing the dependence on the performance of a few individual stocks.

By diversifying the portfolio, the downside risk optimization method aims to achieve a more balanced risk exposure and potentially lower the overall risk of the portfolio. Consequently, the downside risk optimization method offers the benefit of risk mitigation through diversification, providing investors with a more robust and well-rounded portfolio allocation strategy.

D. Sector Allocation Comparison

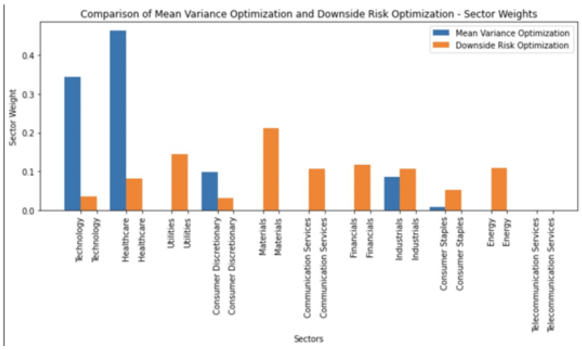

Fig. 4. Sector Allocation of Mean Variance and Downside Risk Optimization method

Mean Variance Optimization (MVO) and Downside Risk Optimization (DRO) result in different sector allocations. The weights assigned to each sector vary between the two optimization methods, indicating that they prioritize sectors differently.

Materials: DRO assigns a significantly higher weight to the Materials sector as compared to MVO. This suggests that DRO places more emphasis on the downside risk associated with the Materials sector as the downside risk volatility of the Materials sector stood at 19.84% while it grew by 136.48%. Energy: Compared to MVO, DRO gives Energy a weight that is comparatively higher. Commodity pricing and geopolitical variables might have an impact on investments in the energy sector.

Healthcare: Healthcare receives substantially more weight from both MVO and DRO. The defensive nature of the healthcare industry is well known for offering stability during market turbulence. The downside risk volatility of the top 10 stocks of the Healthcare sector stood at 15.5%. The cumulative return for the sector was at 124.92%.

Financials: Comparing DRO to MVO, DRO gives Financials a far higher weight. This shows that DRO concentrates on the financial sector's downside risk, which may be motivated by worries about economic cycles, interest rates, or regulatory variables that have an impact on the industry.

Communication Services: DRO gives Communication Services a substantially higher weight than MVO does.

Industrials: Numerous businesses engaged in manufacturing, building, and infrastructure make up the industrials sector. Both MVO and DRO assign relatively small weights to the Industrials sector, with the DRO weight being slightly higher. Utilities: The MVO and DRO weights for Utilities are both very small, suggesting a relatively low allocation to this sector in the portfolio.

Consumer Discretionary: The MVO weight for Consumer Discretionary is significant, indicating a relatively high allocation. The DRO weight is also relatively high, suggesting that the downside risk optimization approach maintains a similar emphasis on this sector.

Technology: Both MVO and DRO assign significant weights to the Technology sector. Healthcare receives substantially more weight from both MVO and DRO. It grew by 344.6%. The defensive nature of the healthcare industry is well known for offering stability during market turbulence.

Consumer Staples: The MVO weight for Consumer Staples is relatively small, while the DRO weight is higher, indicating a higher allocation to this sector when considering downside risk.

Telecommunication Services: MVO assigns a negligible weight to Telecommunication Services, while DRO assigns zero weight.

CONCLUSION

In this research paper, I explored the application of data science techniques in diversification and risk assessment by comparing two portfolio optimization methods: Downside Risk Optimization and Mean Variance Optimization. I utilized a dataset consisting of top stocks representing 11 sectors. Through the analysis, I observed notable differences in sector allocations between the two optimization methods. The Mean Variance Optimization exhibited higher weights in sectors such as Technology and Healthcare, while the Downside Risk Optimization allocated more weight to sectors like Materials and Utilities. This indicates that the two approaches prioritize different sectors based on their respective risk and return objectives.

According to the literature on portfolio risk measurement, a downside risk measure should be favoured over portfolio variance when there are non-normal returns and non- quadratic utility functions [11]. Additionally, Downside Risk Optimization method demonstrated its ability to identify sectors with lower downside potential and allocate higher weights to them. This highlights the effectiveness of the method in capturing downside risk and incorporating it into the portfolio construction process. Such risk-aware allocation can be particularly beneficial in volatile market conditions or during economic downturns, where preserving capital and managing downside risk become crucial objectives for investors.

By minimizing the downside risk ratio, investors can obtain the best downside risk adjusted return, similar to the situation where the investors can obtain the best risk adjusted return by maximizing the Sharpe ratio. The downside deviation (below-target semideviation) and the reward-to-semivariability ratio (R/SVt) are instruments for encapsulating the core of downside risk, according to Sortino and van der Meer [12]. Analysing the covariance matrix, correlation matrix, and the specific constraints used in the portfolio optimisation using quadratic programming will aid in a better understanding of the given weights. These variables may have an impact on allocation choices and help to explain why some sectors were allocated higher weights while having different return and negative risk characteristics.

In conclusion, our research sheds light on the benefits of data-driven diversification and risk assessment methods, emphasizing the importance of considering different optimization approaches to tailor investment strategies to individual needs.

ACKNOWLEDGMENT

I would like to express our gratitude to everyone who has contributed to the completion of this research paper on" Diversification and Risk Assessment using Data Science: Downside Risk vs. Mean Variance Optimization." First and foremost, I extend my heartfelt appreciation to my supervisor and mentor, Mr. Vibhor Sharma, for his guidance, support, and valuable insights throughout the research process. His expertise and feedback were instrumental in shaping the direction of my study. Furthermore, I acknowledge the academic community and researchers in the field of finance and data science for their contributions to the existing body of knowledge. Their studies and findings have laid the foundation for our research and have been essential in shaping our understanding of portfolio optimization techniques.

REFERENCES

[1] H. Markowitz," Portfolio selection," *Journal of Finance*, 7(1):77–91, 1952.

[2] H. Markowitz, "Portfolio selection: Efficient Diversification of Investments," OR, vol. 10, no. 4, p. 253, 1959. doi:10.2307/3006625

[3] H. Markowitz, "The optimization of a quadratic function subject to linear constraints," *Naval Research Logistics Quarterly*, vol. 3, no. 1–2, pp. 111–133, 1956. doi:10.1002/nav.3800030110

[4] T. Gwanoya, "Quantitative risk management: Concepts, techniques, tools. by Alexander J. McNeil, Ru¨diger Frey; Paul Embrechts (Princeton University Press, 2005)," *Annals of Actuarial Science*, vol. 2, no. 1, pp. 187–189, 2007. doi:10.1017/s1748499500000300

[5] D. N. Nawrocki, "A brief history of downside risk measures," *The Journal of Investing*, vol. 8, no. 3, pp. 9–25, 1999. doi:10.3905/joi.1999.319365

[6] A. Rigamonti, "Mean-variance optimization is a good choice, but for other reasons than you might think," *Risks*, vol. 8, no. 1, p. 29, 2020. doi:10.3390/risks8010029

[7] P. Cheng, "Comparing downside-risk and mean-variance analysis using bootstrap simulation," *Journal of Real Estate Portfolio Management*,

vol. 7, no. 3, pp. 225–238, 2001. doi:10.1080/10835547.2001.12089642

[8] W. F. Sharpe, “Capital asset prices: A theory of market equilibrium under conditions of risk,” *The Journal of Finance*, vol. 19, no. 3, p. 425, 1964. doi:10.2307/2977928

[9] F. A. Sortino and L. N. Price, “Performance measurement in a downside risk framework,” *The Journal of Investing*, vol. 3, no. 3, pp. 59–64, 1994. doi:10.3905/joi.3.3.59

[10] J. Chen, “Risk-return tradeoff: How the Investment Principle Works,” Investopedia, https://www.investopedia.com/terms/r/riskreturntradeoff.asp (accessed Jun. 21, 2023).

[11] Boudt, Kris & Peterson, Brian & Carl, Peter.” Hedge fund port- folio selection with modied expected shortfall,” January 2008, 10.2495/CF080101.

[12] F. A. Sortino and R. V. D. Meer, “Downside risk,” *Journal of Portfolio Management*, vol. 17, no. 4, pp. 27–32, 1991.

[13] Hamrin, “A Heuristic Downside Risk Approach to Real Estate Portfolio Structuring,” thesis, Stockholm, 2011