DELFT UNIVERSITY OF TECHNOLOGY

MASTERS THESIS

MULTI-PERIOD ROBUST MEAN-RISK PORTFOLIO OPTIMIZATION

Author:
Vangelis Nakos

Supervisor:
Fenghui Yu

*A thesis submitted in fulfillment of the requirements
for the degree of Master of Science*

in the

Applied Probability Group
Delft Institute of Applied Mathematics

September 2023

Vangelis Nakos: *Multi-period Robust Mean-Risk Portfolio Optimization,* Minimizing Risk and Enhancing Returns in Uncertain Market Environments.

MSc Thesis Applied Mathematics

Multi-period Robust Mean-Risk Portfolio Optimization

Vangelis Nakos (5462940)

Delft University of Technology

Supervisor:

Dr. Fenghui Yu

Other committee members:

Prof Dr. Antonis Papapantoleon

Dr. Nestor Parolya

Delft, The Netherlands

Diversification is the only free lunch in finance.

— Harry Markowitz

ABSTRACT

Portfolio optimization, a fundamental area of study in financial engineering, plays a crucial role in creating efficient portfolios. In this thesis, we consider a robust multi-period Mean-Variance portfolio optimization framework and apply it to real-world market data. The approach we look at incorporates a time-consistent structure that considers the variance of each period, aiming to minimize their sum, while ensuring that the expected return for each period exceeds a predefined threshold. Additionally, we introduce proportional transaction costs to simulate real-world market conditions. To account for uncertainties and increase robustness, we employ a distribution uncertainty set within a Wasserstein ball around the empirical distribution of historical data. This enables us to select worst-case portfolio scenarios for deriving robust optimal solutions. We aim to compare this method with other existing portfolio optimization methods, which we describe in depth in our analysis, to assess its effectiveness.

To achieve the aforementioned research objectives, we conducted an extensive review of portfolio optimization literature, exploring both Mean-Variance and Mean-CVaR portfolio optimization problems. Our research also included robust approaches on portfolio optimization including different distribution and parameter uncertainty sets. We proceeded to construct a comprehensive set of numerical experiments, evaluating portfolio optimization methods performance on real market data. Moreover, we included the S&P500 index to compare them against market performance. In these experiments, we randomly selected stock sets and evaluation period to work on, in order to ensure an unbiased assessment of the methods.

To evaluate the performance of each method, we used the Sharpe ratio of the realized portfolio returns. Our key findings indicate that, in most cases, at least one of the portfolio optimization models outperformed S&P500, suggesting that portfolio optimization problems perform relatively well in the real world. Furthermore, single-period models demonstrated better performance compared to multi-period having higher Sharpe ratio most of the times. Notably, robust optimization models exhibited superior performance compared to nominal models, underlying the significance of accounting for uncertainty. The implications of our research are twofold. Firstly, portfolio optimization problems, especially in the single-period context, demonstrated promising performance and should be embraced by financial practitioners seeking optimal risk-return investment strategies. Secondly, we recommend the preference for robust approaches over traditional models, as they offer improved flexibility to market uncertainties and potentially mitigate downside risks.

Keywords: Portfolio Optimization, Mean-Variance, Mean-CVaR, Robust Optimization, Single-Period, Multi-Period, Transaction Costs, Time-Consistency

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1

INTRODUCTION

The concepts of Portfolio Optimization (PO) and diversification have played a crucial role in the development and comprehension of financial markets and decision making. In 1952, Markowitz published a theory on PO that was a major breakthrough in this field (see Markowitz (1952)). His theory, which is now known as Modern Portfolio Theory (MPT), provided a solution to the fundamental question of how investors should distribute their funds among investments.

To find the best assets allocation of the portfolio, Markowitz set up a quantitative method for measuring the return and risk of an asset with the use of statistical tools like expected return and variance. He suggested that investors should consider both return and risk when determining the allocation of their funds among investment alternatives. Let $\boldsymbol{r}$ be the vector of uncertain returns from assets and $\boldsymbol{x}$ the decision vector, then, portfolio returns are $r_P = \boldsymbol{x}^\top \cdot \boldsymbol{r}$. Markowitz's theory can be formulated by the following optimization problem known as the Mean-Variance (MV) PO problem:

$$\min_{\boldsymbol{x}} \left\{ \mathrm{Var}(r_P) : \mathrm{E}(r_P) = \mu_0,\ \boldsymbol{x} \in X \right\},$$

where μ_0 is the desired expected return and X the set of all feasible decision vectors. This MV PO problem suggests that for any given level of expected return μ_0, investors should choose the portfolio with the minimum variance. In this way, they have the least risk for the same expected return. This idea of quantifying financial decision-making allowed for the evaluation of portfolio return and risk by considering asset returns, that is $\mathrm{E}(r_P)$, and their covariance, that is $\mathrm{Cov}(r_P)$. As a consequence, the principle of portfolio diversification emerged as a crucial concept, indicating that the risk level of a portfolio is influenced by the correlations among its components rather than solely the average risk level of each individual holding. In addition, financial decision-making has been formulated as an ongoing optimization process with numerous advancements made to date.

1.1 RISK MEASURES

While variance has been widely used as a risk measure in both theoretical and practical applications, it has its shortcomings. For instance, it treats both positive and negative deviations from expected returns equally as undesirable risk, despite the desirability of positive deviations for investors. To address this, downside risk measures that only consider negative deviations, such as the *semivariance* by Markowitz (1959), can be used. Moreover, since variance is a nonlinear risk measure, it can lead to more complicated formulations compared to linear risk measures such as the *mean absolute deviation* introduced by Konno and Yamazaki (1991). In terms of volatility risk measures, Sharpe (1966) and Bernardo and Ledoit (2000) introduced the *Sharpe ratio* and *Omega ratio*, respectively, to evaluate portfolio performance based on risk and return. Quantile-based risk measures such as Value-at-Risk (VaR) and Conditional-Value-at-Risk (CVaR)[1] are also known and widely used in risk management. In particular, VaR gained significant popularity starting from the mid-1990s, with its introduction in RiskMetrics by Morgan et al. (1996) and its adoption even being recommended as a standard for banking supervision by Basel Committee (1996). For some confidence level $\alpha \in (0,1)$, VaR of a portfolio with loss L at the confidence level α is given by the smallest number l such that the probability that the loss L exceeds l is no larger than $1-\alpha$. Mathematically, VaR is defined as:

$$\mathrm{VaR}_\alpha(L) = \min\{l : P(L > l) \leq 1-\alpha\}.$$

While VaR is a widely used measure of risk in finance, it has several downsides. Firstly, VaR is not subadditive in general distribution case and thus, it is not a coherent risk measure in the sense of Artzner et al. (1999). As a result, VaR does not conform to the principle that the risk of a portfolio should be less than the sum of its individual components. In practice, diversification across different assets should lead to a reduction in overall portfolio risk, but this principle is contradicted by VaR's limitations. Secondly, VaR informs us about the predicted maximum loss at a specified confidence level, but it fails to provide detailed insights into the nature of extreme losses beyond that level. This means that it may underestimate the potential impact of rare, catastrophic events that can have severe consequences on a portfolio.

On the other hand, CVaR is an alternative measure that overcomes the subadditivity limitation of VaR. CVaR is coherent (see Acerbi and Tasche (2002)), aligning with the diversification principle. Unlike VaR, CVaR provides information on the average magnitude of losses that exceed the VaR threshold. CVaR is more sensitive to the tail of the distribution, which makes it particularly well-suited to capture extreme events, such as financial crises or black swan events. The use of CVaR can lead to more conservative risk management strategies in portfolio optimization. For a loss L with $E(|L|) < \infty$, and cumulative distribution function F_L, CVaR at confidence level $\alpha \in (0,1)$ is defined as:

$$\mathrm{CVaR}_\alpha(L) = \frac{1}{1-\alpha}\int_0^\alpha u \cdot q_{F_L}(u)\, du$$

where $q_{F_L}(u) = F_L^{-1}(u)$ is the quantile function of F_L. For an integrable loss L with continuous cumulative distribution function F_L and for any $\alpha \in (0,1)$, we have:

$$\mathrm{CVaR}_\alpha(L) = \mathrm{E}(L | L \leq \mathrm{VaR}_\alpha(L)).$$

Moreover, the work of Rockafellar, Uryasev, et al. (2000) and Rockafellar and Uryasev (2002) demonstrated that minimizing CVaR can be achieved simul-

1 CVaR is also known as Expected Shortfall (ES).

taneously with determining the corresponding VaR, making it a more flexible and tractable optimization approach. The resulting formulation leads to a convex program, which is computationally efficient, making CVaR practical for portfolio optimization (as it will be shown in Section 2.2). Overall, CVaR has emerged as a valuable risk measure for investors and managers seeking a more comprehensive and robust measure of portfolio risk with Mean-CVaR emerging as a PO problem.

1.2 MULTI-PERIOD PORTFOLIO OPTIMIZATION

It is important to notice that Mean-Risk PO problems discussed consider only a single horizon ahead and they are called single-period PO problems. While single-period PO problems are useful for short-term decision-making, they fail to consider the impact of decisions over multiple periods. Given a specific exiting time, investors typically adjust their asset positions multiple times based on market conditions, making a multi-period selection model crucial for effective decision-making. This involves reallocating wealth at the start of each intermediate period with the aim of minimizing risk and maximizing return when exiting the market. One of the first extensions of MV PO to multi-period was by Li and Ng (2000), who provided the analytical solution for multi-period MV PO problem. A geometric solution was found by Leippold, Trojani, and Vanini (2004), while their solution has been successfully generalized by Bodnar, Parolya, and Schmid (2015a). Further advancement on the multi-period MV PO problem have been made using exponential utility (see Bodnar, Parolya, and Schmid (2015b)) or power utility (see Bodnar et al. (2023)). However, the focus remains on using the objective of simply variance, and concentrating on the time-consistent approach showed by Chen, Li, and Guo (2013). This approach will be later extended to the method used by Wu and Sun, 2023. Furthermore, using a multi-period approach in portfolio optimization provides a better way to account for transaction costs into the problem, which can significantly impact portfolio performance as shown by Arnott and Wagner (1990). In a single-period optimization, a greedy strategy that considers only the current period may be optimal, but it does not consider the impact of current holdings on future returns. However, in a multi-period optimization, one can analyze how current trades affect future trades, allowing for a more realistic model that incorporates the cost of moving in and out of positions. This approach ensures that the portfolio is positioned to trade profitably in future periods, while also taking into account the impact of transaction costs on returns.

1.3 ESTIMATION ERRORS IN MEAN-RISK PORTFOLIO OPTIMIZATION

Portfolios produced by the Mean-Risk PO problems often have unintuitive or extreme weights that cannot be realistically implemented in active trading. For instance, investment positions with significant negative weights. Although imposing additional constraints, such as no short positions, may resolve this issue, it can also result in the creation of portfolios that are very close to the boundary of the restrictions and are thus dependent on the allocation weight conditions as stated in Black and Litterman (1992). More importantly, PO problems have issues due to the random nature of asset returns, making it highly sensitive to inputs. Estimation errors in risk and return estimates are an issue as assets with large estimated returns, negative correlations, and small variances are over-weighted in the model, even though they may have large estimation errors. This problem arises due to the use of sample mean and variance, the ignorance of market factors, and the assumption of a single optimal portfolio (see Michaud (1989) and Broadie (1993)). Furthermore, Chopra and Ziemba (2013) studied the effects of estimation errors on portfolio performance and found that errors in estimating expected returns of assets are at least 10 times more important than errors in estimated variances or covariances.

The concept of robust optimization also concerned Ben-Tal and Nemirovski (1998), whose formulation allows for solving a robust formulation of the PO problem resulting in more stable allocations while being less sensitive to model parameter changes. This approach introduces uncertainty in the input parameters related to the mean vector and covariance matrix of asset returns, such as polytopic, box, and ellipsoidal uncertainty, or even uncertainty in the distribution of returns, such as those we will discuss in Chapter 4. Previous studies have investigated the use of worst-case VaR and CVaR in robust portfolio optimization when only partial information on the underlying probability distribution is available, like in Ghaoui, Oks, and Oustry (2003) and Zhu and Fukushima (2009). Fabozzi et al. (2007b) review the relationship between robust optimization and other methods for portfolio management, however, the focus in this thesis will be on worst-case optimization derived by distribution uncertainty set for the MV and Mean-CVaR PO problems.

1.4 THESIS SCOPE AND OUTLINE

In this thesis, we aim to address the challenges and limitations associated with traditional PO problems, such as MV and Mean-CVaR, by introducing a robust approach that emphasizes worst-case optimization from distribution uncertainty set. Additionally, we extend these problems to the multi-period setting, considering the impact of transaction costs while also incorporating a robust framework. The thesis is structured as follows. In Chapter 2 the problem formulation of the MV and Mean-CVaR PO problems are presented in the context of a single-period optimization. By solving these optimization problems, the efficient portfolios with the optimal balance between risk and expected return are found. Next, Chapter 3 is on multi-period portfolio optimization, the analysis of the previous Chapter is extended to a multi-period setting. The impact of transaction costs is also considered, which significantly influence portfolio performance. We develop a methodology to incorporate these costs into the optimization process. In addition, we introduce the concept of time-consistency and provide models that address this issue. Chapter 4 provides robust approaches in PO. It focuses on worst-case optimization accounting for uncertainty in input parameters using distribution uncertainty sets. Robust Mean-CVaR PO is presented in single-period, while robust MV PO is shown in multi-period. In Chapter 5, we conduct a comprehensive set of numerical experiments to evaluate the performance of the aforementioned PO problems in real market data. To ensure a rigorous assessment, we devise a method for randomly selecting a set of stocks and a specific evaluation period. Within this framework, we execute 20 independent numerical experiments, solving each PO problem for the chosen stocks set and evaluation period. For each experiment, we assess the performance of each method by computing the Sharpe ratio of the realized portfolio returns. The aim is to identify the model that consistently outperforms the others across different stocks sets and evaluation periods. Through these experiments, we gain insights into the strengths and weaknesses of each approach. In Chapter 6, we present the conclusions drawn from our research and discuss the implications of our findings while we provide recommendations for decision-makers. By exploring these topics, the thesis aims to contribute to the field of portfolio optimization by offering innovative approaches to address the limitations of traditional models and providing insights into their practical implications for financial decision-making.

2

PORTFOLIO OPTIMIZATION IN SINGLE-PERIOD

Portfolio optimization is a crucial task in financial management, as it involves selecting a combination of assets that maximizes returns while minimizing risks. Stochastic optimization models are particularly useful for this purpose, as they allow for uncertainty in asset returns and risk measures. This chapter focuses on single-period portfolio optimization problems, and presents two popular stochastic optimization models for portfolio optimization: MV and Mean-CVaR PO problem. We will explore the advantages and disadvantages of each model, and discuss how they can be used to construct optimal portfolios.

2.1 MEAN-VARIANCE PORTFOLIO OPTIMIZATION

MV PO problem was first introduced by Markowitz (1952) in his seminal work and have since become a foundation of MPT. The basic idea behind MV PO problem is to find the combination of assets that maximizes reward for a given level of risk, where Markowitz suggested that reward should be measured by the expected portfolio return and risk should be measured by the portfolio variance. In addition, the problem has been studied for the use of quadratic utility as in Bodnar, Parolya, and Schmid (2013), or even power utility as in Bodnar et al. (2020). Nevertheless, in this section, we will provide an overview of the plain MV approach from Fabozzi et al. (2007a), discuss its assumptions and limitations, and explain how to construct a MV optimal portfolio.

The MPT assumes that investors are rational and risk-averse, meaning that they prefer lower levels of risk for a given level of reward. Therefore, for any given level of expected return, they would choose the portfolio with the minimum variance and for any given level of variance, they would choose the portfolio with the maximum expected return. Different allocations of wealth in assets of the portfolio can yield different values of expected return and variance. The set of all possible portfolios that can be constructed is called *feasible set*, while the set of portfolios with minimum variance for any given level of expected return and minimum expected return for any given level of variance is called *efficient frontier*. The portfolio with the minimum variance overall is called *global minimum variance*.

To construct the MV PO problem, consider a portfolio consisting of N risky assets. Let $\boldsymbol{w} = (w_1, \dots, w_N)$ be the decision N-vector, which represents the proportion of initial wealth allocated to each asset. In other words, $\boldsymbol{w}$ is a vector of weights and so $\boldsymbol{w}^\top \mathbf{1} = 1$, for $\mathbf{1}$ a N-vector of ones. Let $\boldsymbol{r} = (r_1, \dots, r_N)$

be the N-vector of uncertain future returns and let $\boldsymbol{\mu} = (\mu_1, \dots, \mu_N)$ be the N-vector of means of $\boldsymbol{r}$ and Σ the $N \times N$ covariance matrix of the returns of all assets, that is

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_{NN} \end{pmatrix}$$

where $\sigma_{ii} = \sigma_i^2$ and $\sigma_{ij} = \sigma_{ji} = \rho_{ij}\sigma_i\sigma_j$ for $i \neq j$ and ρ_{ij} the correlation between assets i and j. Every covariance matrix is positive semidefinite and so $\boldsymbol{w}^\top \Sigma \boldsymbol{w} \geq 0$ for all $\boldsymbol{w}$. We will also assume that Σ is positive definite ($\boldsymbol{w}^\top \Sigma \boldsymbol{w} > 0$), that is, it is impossible to construct any of the assets in the portfolio using only the other assets in the portfolio. This assumption is in agreement with the reality while it ensures that Σ is an invertible matrix.

According to MPT, an investor seeks for the minimum variance given a specific expected return. So, for a target expected return μ_0, the MV PO problem is the following:

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^\top \Sigma \boldsymbol{w} \\ \text{s.t.} \quad & \boldsymbol{w}^\top \boldsymbol{\mu} = \mu_0, \\ & \boldsymbol{w}^\top \mathbf{1} = 1. \end{aligned} \tag{1}$$

Furthermore, an investor will also seek for the maximum expected return given a specific variance. So, for a target variance σ_0^2, the optimization problem becomes:

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^\top \boldsymbol{\mu} \\ \text{s.t.} \quad & \boldsymbol{w}^\top \Sigma \boldsymbol{w} = \sigma_0^2, \\ & \boldsymbol{w}^\top \mathbf{1} = 1. \end{aligned} \tag{2}$$

Or even, another formulation of the problem could be a utility function where expected return is rewarded and variance is penalized given a risk averseness parameter $\omega > 0$:

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^\top \boldsymbol{\mu} - \omega \cdot \boldsymbol{w}^\top \Sigma \boldsymbol{w} \\ \text{s.t.} \quad & \boldsymbol{w}^\top \mathbf{1} = 1. \end{aligned} \tag{3}$$

It is shown in Appendix A.1 that MV PO problems (1), (2) and (3) are equivalent given a relationship between the parameters μ_0, σ_0 and ω. It also provides the analytical solution of decision vectors $\boldsymbol{w}$ and the expected return and variance for each problem.

Under the problem (1) framework, the efficient frontier is defined as the set of portfolios with minimum variance for any given level of expected return. Therefore, it can be found by following the steps below:

Step-by-Step Process for MV Efficient Frontier.	
Step 1:	Solve problem (1) without the $\boldsymbol{w}^\top \boldsymbol{\mu} = \mu_0$ constrain to find the global minimum variance portfolio $\boldsymbol{w}_{min}$.
Step 2:	Attain the lowest expected return of the efficient frontier $\mu_{min} = \boldsymbol{w}_{min}^\top \boldsymbol{\mu}$.[1]
Step 3:	Solve the problem $\max_{\boldsymbol{w}} \boldsymbol{w}^\top \boldsymbol{\mu} \quad \text{s.t.} \quad \boldsymbol{w}^\top \mathbf{1} = 1$ to attain the highest expected return of the efficient frontier μ_{max}.[2]
Step 4:	Solve problem (1) for all $\mu_0 \in [\mu_{min}, \mu_{max}]$.

Following these steps for problem (1), and given the analytical solution provided at Appendix A.1, we can derive the decision vector $\boldsymbol{w}$ and the variance at the efficient frontier:

1. Global minimum variance portfolio: $\boldsymbol{w}_{min} = \dfrac{\Sigma^{-1}\mathbf{1}}{A}$.
2. Minimum expected return of the efficient frontier:
   $\mu_{min} = \boldsymbol{w}_{min}^\top \boldsymbol{\mu} = \dfrac{B}{A}$.
3. Maximum expected return of the efficient frontier:
   $\mu_{max} = \max\limits_{i=1,\dots,N} \mu_i$.
4. Efficient frontier:
   - Decisions: $\boldsymbol{w}^* = \dfrac{\lambda \Sigma^{-1}\boldsymbol{\mu} + \gamma \Sigma^{-1}\mathbf{1}}{2}$ with $\lambda = 2\dfrac{\mu_0 A - B}{\Delta}$ and $\gamma = 2\dfrac{C - \mu_0 B}{\Delta}$ for $\mu_0 \in [\dfrac{B}{A}, \max\limits_{i=1,\dots,N} \mu_i]$.
   - Variance: $\sigma_0^2 = \dfrac{A\mu_0^2 - 2B\mu_0 + C}{\Delta}$ for $\mu_0 \in [\dfrac{B}{A}, \max\limits_{i=1,\dots,N} \mu_i]$.

where $A = \mathbf{1}^\top \Sigma \mathbf{1}$, $B = \mathbf{1}^\top \Sigma \boldsymbol{\mu}$, $C = \boldsymbol{\mu}^\top \Sigma \boldsymbol{\mu}$ and $\Delta = AC - B^2$.

It is clear that the efficient frontier is a parabola on the expected return - variance space and on the expected return - standard deviation space. A plot of the efficient frontier and an indicator of the global minimum variance is presented in Figure 1:

1 This is the start point of the efficient frontier, as it represents the minimum expected return for any given variance. The rationale behind this choice is that expected returns below the global minimum variance would not fulfill the criteria of obtaining the highest expected return for a specific level of variance.

2 This is the end point of the efficient frontier. Since the problem is unbounded, an additional constraint $w_i \geq 0, i = 1, \dots, N$ is applied only for μ_{max}, and thus $\mu_{max} = \max\limits_{i=1,\dots,N} \mu_i$. This constraint is employed in all other models to be discussed later; however, it was excluded from the MV case for the sake of deriving an analytical solution.

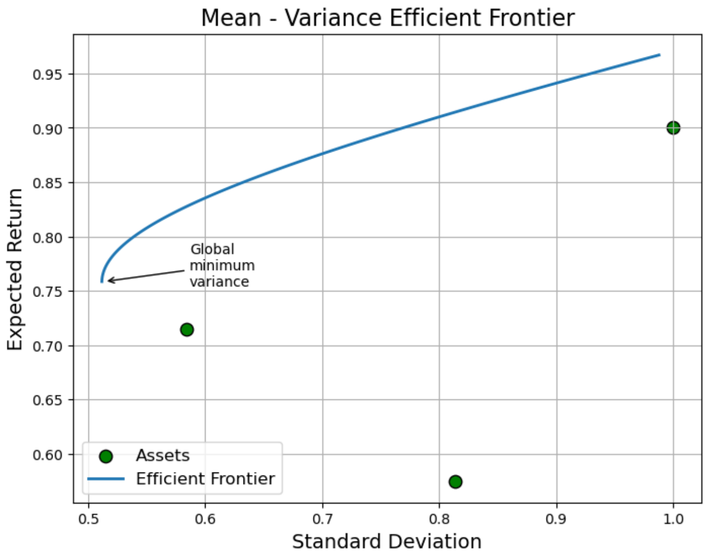


Figure 1: Plot of the efficient frontier on an expected return - standard deviation graph including the individual assets as points. The values A = 2.5, B = 1.77, and C = 1.31 are used in this example for the sake of making the graph.

2.1.1 *Mean-Variance Portfolio Optimization with the Addition of a Risk-free Asset*

In reality, investors also consider the amount of cash they hold in their portfolio. Holding cash, they get interest R as return which is considered as a no-risk investment. Therefore, it is essential that we include one risk-free asset (i.e. the holding cash) in the portfolio optimization problem.

Following the same set up, consider a portfolio consisting of N risky assets and 1 risk-free. Let $\boldsymbol{w} = (w_1, \ldots, w_N)$ be the decision N-vector for the risky assets and w_R the weight given at the risk-free asset and so $\boldsymbol{w}^\top \mathbf{1} + w_R = 1$, for $\mathbf{1}$ a N-vector of ones. The MV PO problem then is:

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^\top \Sigma \boldsymbol{w} \\ \text{s.t.} \quad & \boldsymbol{w}^\top \boldsymbol{\mu} + w_R R = \mu_0, \\ & \boldsymbol{w}^\top \mathbf{1} + w_R = 1. \end{aligned} \tag{4}$$

Following the 4 steps to find the efficient frontier given the results from Appendix A.2:

1. Global minimum variance portfolio: $\boldsymbol{w}_{min} = \mathbf{0}$ $(w_R = 1)$.
2. Minimum expected return of the efficient frontier: $\mu_{min} = R$.
3. Maximum expected return of the efficient frontier: $\mu_{max} = \max_{i=1,\ldots,N} \mu_i$.
4. Efficient frontier:

- Decisions: $\boldsymbol{w}^* = \frac{(\mu_0 - R)\Sigma^{-1}\boldsymbol{\mu}_D}{\boldsymbol{\mu}_D^\top \Sigma \boldsymbol{\mu}_D}$ for $\mu_0 \in [R, \max_{i=1,\ldots,N} \mu_i]$.
- Variance: $\boldsymbol{w}^{*\top}\Sigma\boldsymbol{w}^* = \frac{(\mu_0 - R)^2}{\boldsymbol{\mu}_D^\top \Sigma \boldsymbol{\mu}_D}$ for $\mu_0 \in [R, \max_{i=1,\ldots,N} \mu_i]$.

Taking a look at the optimal decisions

$$\boldsymbol{w}^* = \frac{\mu_0 - R}{\boldsymbol{\mu}_D^\top \Sigma \boldsymbol{\mu}_D}\Sigma^{-1}\boldsymbol{\mu}_D,$$

$\boldsymbol{w}^*$ is proportional to $\Sigma^{-1}\boldsymbol{\mu}_D$ which does not depend on μ_0. Therefore, the investor allocates his wealth among the risky assets in the same relative proportions, while changing μ_0 only changes the amount invested in that portfolio with $\boldsymbol{w}^*$ and in the risk-free asset via the $\frac{\mu_0 - R}{\boldsymbol{\mu}_D^\top \Sigma \boldsymbol{\mu}_D}$ proportionality constant. This portfolio with $\boldsymbol{w}^*$ consisting of risky assets is called the *tangency portfolio*.

Making the graph of the efficient frontier in a expected return - standard deviation plot (see Figure 2), it can be seen that the efficient frontier of the portfolio with the risk-free asset is a straight line, tangent to the efficient frontier of the risky assets. Their common point is the tangency portfolio.

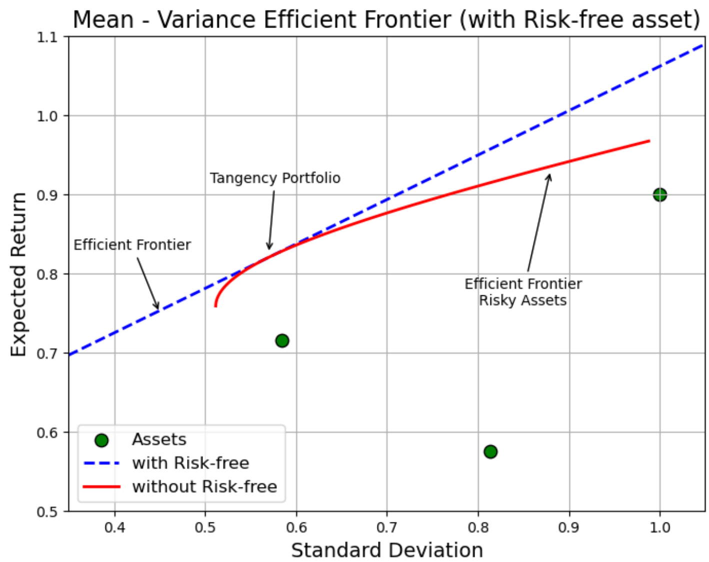
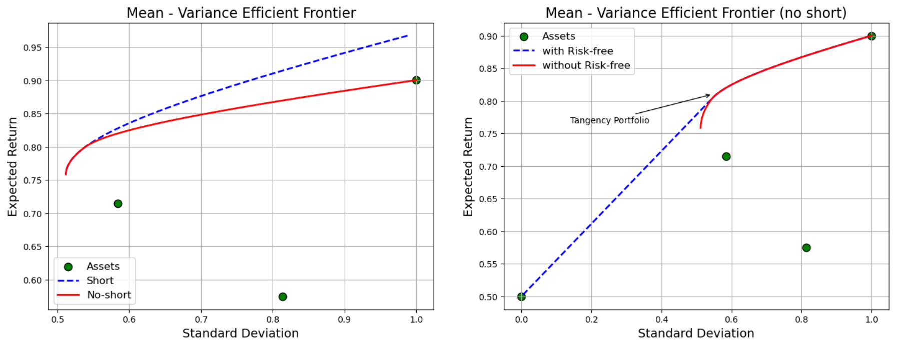


Figure 2: The efficient frontier of both with and without the risk-free asset on a expected return - standard deviation plot including the individual assets as points.The values A = 2.5, B = 1.77, C = 1.31 and R = 0.5 are used in this example for the sake of making the graph.

To find the exact decomposition $\boldsymbol{w}_T^*$ of the tangency portfolio, one should take the case where $w_R = 0$ and so, solve the $\boldsymbol{w}_T^{*\top}\mathbf{1} = 1$:

$$\frac{\mu_0 - R}{\boldsymbol{\mu}_D^\top \Sigma \boldsymbol{\mu}_D}\boldsymbol{\mu}_D \Sigma^{-1}\mathbf{1} = 1 \Rightarrow \mu_0 = \frac{\boldsymbol{\mu}_D^\top \Sigma \boldsymbol{\mu}_D}{\boldsymbol{\mu}_D \Sigma^{-1}\mathbf{1}} + R.$$



(a) Comparison of MV efficient frontiers with and without constraints.

(b) Efficient Frontier with and without Risk-free asset under no-short constraints.

Figure 3: Incorporating the no-short constrain in the MV optimization problem.

And so, the tangency portfolio is:

$$\boldsymbol{w}_T^* = \frac{\Sigma^{-1}\boldsymbol{\mu}_D}{\boldsymbol{\mu}_D\Sigma^{-1}\mathbf{1}}.$$

2.1.2 *Mean-Variance Portfolio Optimization with the Addition of No Short-Selling Constrain*

In our previous theoretical calculations, we did not limit the portfolio weights except for the requirement that they sum up to one. Specifically, they were permitted to have both positive and negative values without any restriction on short selling. However, in reality, some portfolio managers may not be able to sell assets short due to a variety of reasons and thus, they only take long positions. To reflect this constrain, the MV PO problem becomes:

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^\top \Sigma \boldsymbol{w} \\ \text{s.t.} \quad & \boldsymbol{w}^\top \boldsymbol{\mu} + w_R R = \mu_0, \\ & \boldsymbol{w}^\top \mathbf{1} + w_R = 1, \\ & w_R \geq 0, \\ & \boldsymbol{w} \geq 0. \end{aligned} \tag{5}$$

This problem can be solved for both considering the risk-free asset or not. We set $w_R = 0$ in case the risk-free asset is not included. The new efficient frontiers in both cases will be bellow the unconstrained ones since we are restricting the range of possible choices by constraining all the weights to be positive. In Figures (3a, 3b) below, the impact of prohibiting short selling is observed. Figures are generated using the quadratic programming solver from CVXOPT library in Python built by Andersen, Dahl, and Vandenberghe (2022) which uses methods from Sra, Nowozin, and Wright (2011).

After examining Figure 3, several conclusions can be drawn. Figure 3a demonstrates that the efficient frontier with constraints is positioned below the unconstrained one. The reason behind this is that even though allowing short selling may lead to higher expected returns, it also leads to higher risks, which most investors want to avoid. Additionally, Figure 3b displays that the efficient frontier with the risk-free asset is a straight line connecting the risk-free asset and the tangency portfolio. This occurs because the weight w_R of the risk-free asset and weight w_T of the tangency portfolio are constrained to be in $[0,1]$ and $w_R + w_T = 1$. Therefore, the lowest expected return (and risk) for the portfolio is for $w_R = 1$, while the highest expected return (and risk) for the portfolio is for $w_T = 1$. For expected returns higher than that of the tangency portfolio, the portfolio only includes risky assets since a combination of risk-free asset and tangency portfolio cannot achieve them. Thus, it is identical to the efficient frontier without the risk-free asset. Finally, in the absence of constraints (see Figure 2), higher expected returns could be attained by having $w_R \leq 0$ (borrowing cash) and $w_T \geq 1$ (investing borrowed cash) such that $w_R + w_T = 1$, thus allowing the efficient frontier straight line to continue beyond the tangency portfolio. In the remainder of this chapter, we will consistently incorporate the no-short constraints into every problem formulation, aligning with common practice in real-world scenarios.

2.1.3 *Mean-Variance Portfolio Optimization with the Addition of Transaction Costs*

If an investor wants to apply this model in real market data, then he will have to take into account for the transaction costs. Transaction costs usually come from the amount traded, so, an investor who reallocates his current portfolio to the new portfolio suggested by the MV PO solution should incur the transaction costs. Let TC be the estimated transaction costs, incorporating them into the MV PO problem looks like this:

$$
\begin{aligned}
\min_{\boldsymbol{w}} \quad & \boldsymbol{w}^\top \Sigma \boldsymbol{w} \\
\text{s.t.} \quad & \boldsymbol{w}^\top \boldsymbol{\mu} + w_R R \geq \mu_0, \\
& \boldsymbol{w}^\top \mathbf{1} + w_R = 1 - TC, \\
& w_R \geq 0, \\
& \boldsymbol{w} \geq 0,
\end{aligned}
\tag{6}
$$

where the expected return constrain is now relaxed to an inequality and normalized transaction costs are added to the weights constrain. The estimation of transaction costs is presented in detail in Section 3.1.2.1.

2.2 MEAN-CVAR PORTFOLIO OPTIMIZATION

When it comes to portfolio optimization, the choice of risk measure to be optimized is crucial. Although variance has been widely used since Markowitz (1952) seminal work, it has some limitations. One of the main drawbacks of variance is that it considers both positive and negative deviations around the expected return as equally undesirable, while in reality, investors may find positive deviations desirable. To address this issue, downside risk measures that only consider negative deviations can be used. VaR and CVaR are the most popular quantile-based risk measures. VaR quantifies the maximum loss at a specific confidence level, while CVaR represents the expected value of losses greater than VaR at a confidence level. However, VaR lacks of coherence as a risk measure, and it is unable to provide information about extreme losses beyond the specified probability level. Therefore, this section focuses on the Mean-CVaR PO problem.

Following the same structure as before, to construct the Mean-CVaR PO problem, consider a portfolio consisting of N risky assets. Let $\boldsymbol{w} = (w_1, \ldots, w_N)$ be the decision N-vector, which represents the percentage of initial wealth allocated to each asset, that is $\boldsymbol{w}^\top \mathbf{1} = 1$. In addition, we will include another constrain where no short-selling is allowed, that is, $\boldsymbol{w} \geq 0$ for the reasons mentioned in 2.1. Lastly, $\boldsymbol{r} = (r_1, \ldots, r_N)$ is the N-vector of uncertain future returns.

The corresponding loss function is $L = -\boldsymbol{w}^\top \boldsymbol{r}$ and the probability of L not exceeding a threshold l is given by the integral

$$P(L \leq l) = \int_{L \leq l} p(\boldsymbol{r}) d\boldsymbol{r},$$

where $p(\boldsymbol{r})$ is the probability density function of $\boldsymbol{r}$. VaR at a confidence level β is defined as

$$\text{VaR}_\beta(L) = \min\{l | P(L \leq l) \geq \beta\},$$

while the corresponding CVaR is given by

$$\text{CVaR}_\beta(L) = \frac{1}{1-\beta} \int_{L \geq \text{VaR}_\beta(L)} -\boldsymbol{w}^\top \boldsymbol{r} \cdot p(\boldsymbol{r}) d\boldsymbol{r}.$$

Due to the complexity of the CVaR, the Mean-CVaR PO problem cannot be solved analytically, and the use of software is necessary, which will be discussed in detail later in this chapter. Hence, the Mean-CVaR PO problem can be expressed as follows:

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \text{CVaR}_\beta(-\boldsymbol{w}^\top \boldsymbol{r}) \\ \text{s.t.} \quad & \boldsymbol{w}^\top \boldsymbol{\mu} = \mu_0, \\ & \boldsymbol{w}^\top \mathbf{1} = 1, \\ & \boldsymbol{w} \geq 0. \end{aligned} \tag{7}$$

Theorem 1 in Rockafellar, Uryasev, et al. (2000) states that the computation of CVaR can be accomplished by minimizing the auxiliary function

$$F_\beta(\boldsymbol{w}, \alpha) = \alpha + \frac{1}{1-\beta} \int_{r \in \mathbb{R}^N} -\boldsymbol{w}^\top \boldsymbol{r} \cdot p(\boldsymbol{r}) d\boldsymbol{r}, \tag{8}$$

with respect to the variable $\alpha \in \mathbb{R}$. Therefore, the formula for $\text{CVaR}_\beta(-\boldsymbol{w}^\top \boldsymbol{r})$ is given by

$$\text{CVaR}_\beta(-\boldsymbol{w}^\top \boldsymbol{r}) = \min_{\alpha \in \mathbb{R}} F_\beta(\boldsymbol{w}, \alpha). \tag{9}$$

And so, as proved in Theorem 2 of Rockafellar, Uryasev, et al. (2000), problem (7) becomes:

$$\begin{aligned} \min_{\boldsymbol{w},\alpha} \quad & F_\beta(\boldsymbol{w},\alpha) \\ \text{s.t.} \quad & \boldsymbol{w}^\top \boldsymbol{\mu} = \mu_0, \\ & \boldsymbol{w}^\top \mathbf{1} = 1, \\ & \boldsymbol{w} \geq 0, \\ & \alpha \in \mathbb{R}. \end{aligned} \tag{10}$$

The computation of $F_\beta(\boldsymbol{w},\alpha)$ presents a significant challenge due to the need to calculate the integral of a multivariate and generalized function. Nonetheless, this challenge can be overcome using an appropriate approximation method. Monte Carlo simulation is among the most effective methods for computing high-dimensional integrals. In fact, Rockafellar, Uryasev, et al. (2000) employed this method to approximate F_β as follows:

$$\tilde{F}_\beta(\boldsymbol{w},\alpha) = \alpha + \frac{1}{q(1-\beta)} \sum_{k=1}^{q} [-\boldsymbol{w}^\top \boldsymbol{r_k} - \alpha]^+,$$

where r_k denotes the k^{th} sample out of the q samples taken by the distribution $p(\cdot)$ and $[\cdot]^+ = \max\{\cdot, 0\}$. Since $L = -\boldsymbol{w}^\top \boldsymbol{r}$ is linear to $\boldsymbol{r}$ and if the feasible set of $\boldsymbol{w}$ is convex, then $\tilde{F}_\beta$ is a linear program that can be solved efficiently. Minimizing $\tilde{F}_\beta$ is an approximation to minimizing F_β, while it is also equivalent to minimizing the linear expression

$$\alpha + \frac{1}{q(1-\beta)} \sum_{k=1}^{q} u_k,$$

subject to the linear constrains $u_k \geq 0$ and $u_k \geq -\boldsymbol{w}^\top \boldsymbol{r_k} - \alpha$ for $k = 1, \ldots, q$. Consequently, the optimization problem becomes:

$$\begin{aligned} \min_{\boldsymbol{w},\boldsymbol{u},\alpha} \quad & \alpha + \frac{1}{q(1-\beta)} \sum_{k=1}^{q} u_k \\ \text{s.t.} \quad & \boldsymbol{w}^\top \boldsymbol{r_k} + \alpha + u_k \geq 0 \quad \text{for } k = 1, \ldots, q, \\ & u_k \geq 0 \quad \text{for } k = 1, \ldots, q, \\ & \boldsymbol{w}^\top \boldsymbol{\mu} = \mu_0, \\ & \boldsymbol{w}^\top \mathbf{1} = 1, \\ & \boldsymbol{w} \geq 0, \\ & \alpha \in \mathbb{R}. \end{aligned} \tag{11}$$

To effectively solve this convex optimization problem, we used the linear programming solver from the CVXOPT library in Python, as developed by Andersen, Dahl, and Vandenberghe (2022), which employs the methods described by Sra, Nowozin, and Wright (2011). To find the efficient frontier, we first need to solve (11) for all possible μ_0, that is $\mu_0 \in [\min_{i=1,\ldots,N} \mu_i, \max_{i=1,\ldots,N} \mu_i]$, given the returns $\boldsymbol{r_k}$ from the input data and then follow the aforementioned steps:

Step-by-Step Process for Mean-CVaR Efficient Frontier.	
Step 1:	Global minimum CVaR portfolio: Find $\boldsymbol{w}_{min}$ such that $\min_{\boldsymbol{w}} \text{CVaR}_\beta(L)$.
Step 2:	Minimum expected return of efficient frontier: $\mu_{min} = \boldsymbol{w}_{min}^\top \boldsymbol{\mu}$.
Step 3:	Maximum expected return of efficient frontier: $\mu_{max} = \max_{i=1,\dots,N} \mu_i$.
Step 4:	Efficient frontier: Get the CVaR values for all $\mu_0 \in [\mu_{min}, \mu_{max}]$.

Following these steps, and using the same inputs as in the MV case, the following graph for the efficient frontier of the Mean-CVaR portfolio optimization problem is derived (see Figure 4).

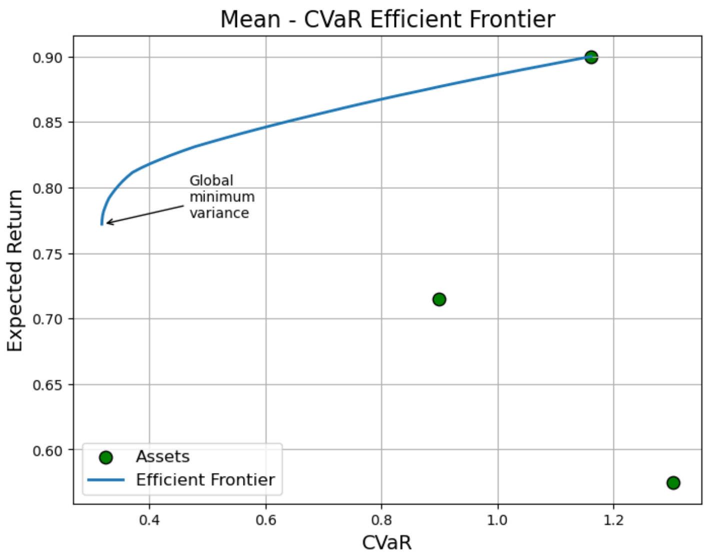


Figure 4: The efficient frontier Mean-CVaR on a expected return - CVaR plot including the individual assets as points. The confidence level of $\beta = 0.95$ is used in this example.

The Mean-CVaR efficient frontier looks a lot like the MV efficient frontier, with CVaR values similar to standard deviation in Figure 3. It is important to note, however, that these visual similarities can vary substantially depending on the input data.

2.2.1 *Mean-CVaR Portfolio Optimization with the Addition of a Risk-free Asset*

As with the MV PO problem, we will also add a case where there is a risk-free asset, that is, there is an asset with constant returns R. In this case, the optimization problem formulation is the same as (11) and we are using the same methods to solve it. Setting $R = 0.5$ (as the risk-free rate in MV case) and following the same 4 steps to get the efficient frontier, we plot both efficient frontiers with and without the risk-free asset (see Figure 5).

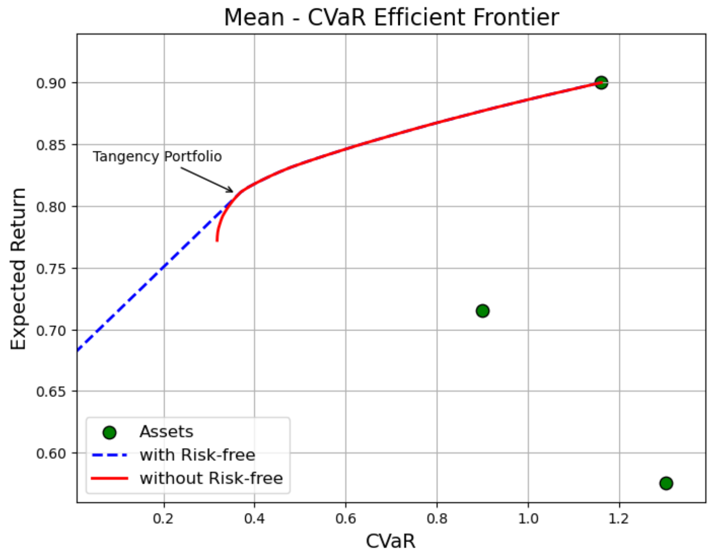


Figure 5: The efficient frontier Mean-CVaR on a expected return - CVaR plot with and without the risk-free asset including the individual assets as points. The confidence level of $\beta = 0.95$ and risk-free return $R = 0.435$ are used in this example.

When the risk-free asset is included, the Mean-CVaR efficient frontier takes the form of a straight line, indicating the presence of a tangency portfolio. Below the expected return of the tangency portfolio, the efficient frontier is a linear combination of the risk-free asset and the tangency portfolio. Above this expected return, the efficient frontier is equivalent to the one without the risk-free asset. The logic behind that is the same as in 2.1.1.

2.2.2 *Mean-CVaR Portfolio Optimization with the Addition of Transaction Costs*

If an investor decides to implement this model using real market data, they will need to consider transaction costs, which typically depend on the trading volume. When an investor reallocates their current portfolio to the new portfolio suggested by the Mean-CVaR PO solution, they will incur transaction costs represented as TC. To incorporate these costs, we use the following problem formulation:

$$
\begin{aligned}
\min_{\boldsymbol{w},\boldsymbol{u},\alpha} \quad & \alpha + \frac{1}{q(1-\beta)} \sum_{k=1}^{q} u_k \\
\text{s.t.} \quad & \boldsymbol{w}^\top \boldsymbol{r_k} + \alpha + u_k \geq 0 \quad \text{for } k = 1, \dots, q, \\
& u_k \geq 0 \quad \text{for } k = 1, \dots, q, \\
& \boldsymbol{w}^\top \boldsymbol{\mu} + w_R R \geq \mu_0, \\
& \boldsymbol{w}^\top \mathbf{1} + w_R = 1 - TC, \\
& w_R \geq 0, \\
& \boldsymbol{w} \geq 0, \\
& \alpha \in \mathbb{R}.
\end{aligned} \tag{12}
$$

where the expected return constrain is now relaxed to an inequality and normalized transaction costs are added to the weights constrain. The estimation of transaction costs is further explained in Section 3.1.2.1.

3 PORTFOLIO OPTIMIZATION IN MULTI-PERIOD

In the previous chapter, we discussed single-period portfolio optimization techniques that are effective for short-term decision-making. However, investors generally make decisions over multiple periods, adjusting their asset positions based on evolving market conditions. To better capture the dynamics of long-term investment strategies, it is essential to consider multi-period portfolio optimization. The focus of this chapter will remain on discrete time. In the multi-period context, given a time horizon, wealth is reallocated at the start of each period with the aim of minimizing risks and maximizing returns when exiting the market in the end of the time horizon. This chapter digs into the multi-period extensions of MV and Mean-CVaR PO problems.

Multi-period portfolio optimization offers several advantages over its single-period counterpart. Firstly, it provides a more realistic framework that accounts for the impact of transaction costs on returns, which can significantly affect portfolio performance. Secondly, multi-period optimization allows for a more comprehensive analysis of how current trades influence future trades, ensuring that the portfolio is well-positioned to profitably navigate future periods in discrete time. The insights and methodologies presented in this chapter will enable investors to develop efficient strategies for optimizing their portfolios in a multi-period setting.

3.1 MEAN-VARIANCE PORTFOLIO OPTIMIZATION

The MV PO framework, initially developed for single-period settings, has been successfully extended to multi-period scenarios. The multi-period MV PO approach considers the impact of investment decisions across multiple periods, capturing the trade-offs between expected returns and risks more accurately.

In this section, the focus will be on the formulation and implementation of multi-period MV PO. We begin by exploring the analytical solution proposed by Li and Ng (2000), which has been further improved by Cui et al. (2012) to ensure time consistency in the optimization process. To enhance the practical applicability of the model, the no-short constraint, as discussed by Cui et al. (2014), will be incorporated into the model. The analytical solution for this enhanced model will be presented. Moreover, we integrate a time-consistent approach within the constrained optimization problem, ensuring a consistent and realistic portfolio allocation strategy over multiple time periods. Finally, we address the impact of transaction costs by incorporating them into the optimization framework, thereby considering the practicality and feasibility of executing portfolio adjustments in real-world scenarios.

3.1.1 *Mean-Variance Portfolio Optimization - Analytical Solution*

Consider a portfolio consisting of $N+1$ risky assets. The investor joins the market at time 0 with an initial wealth x_0, which he can allocate among the assets, and exits the market at time period T. The wealth can be re-allocated among the $N+1$ assets at the beginning of each of the following $T-1$ consecutive time periods. Let $\boldsymbol{r_t} = (r_t^0, \ldots, r_t^N)$ be the $(N+1)$-vector of uncertain future rates of returns[1] at time period t. It is assumed that vectors $\boldsymbol{r_t}, t = 0, \ldots, T-1$, are statistically independent with known mean $\boldsymbol{\mu}_t = (\mu_t^0, \ldots, \mu_t^N)$ and covariance matrix Σ_t, that is

$$\Sigma_t = \begin{pmatrix} \sigma_{t,11} & \sigma_{t,12} & \cdots & \sigma_{t,1N} \\ \sigma_{t,21} & \sigma_{t,22} & \cdots & \sigma_{t,2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{t,N1} & \sigma_{t,N2} & \cdots & \sigma_{t,NN} \end{pmatrix}$$

where $\sigma_{t,ii} = \sigma_{t,i}^2$ and $\sigma_{t,ij} = \sigma_{t,ji} = \rho_{t,ij}\sigma_{t,i}\sigma_{t,i}$ for $i \neq j$ and $\rho_{t,ij}$ the correlation between assets i and j, all at time periods t. Let $\boldsymbol{u_t} = (u_t^1, \ldots, u_t^N)$ be the wealth allocation vector at time period t, where u_t^i is the wealth allocated at asset i at time period t. We take asset 0 as a reference, and the investment amount at time period t for this asset is given by $x_t - \sum_{i=1}^N u_t^i$. This formulation simplifies the process of incorporating a risk-free asset, which will be demonstrated later.

An investor seeks the optimal investment strategy, $\boldsymbol{u_t} = (u_t^1, \ldots, u_t^N), t = 0, \ldots, T-1$, to maximize the expected value of the terminal wealth x_T, subject to the constraint that the variance of the terminal wealth does not exceed a pre-selected risk level[2]. This can be represented by the following optimization problem:

$$\begin{aligned} \max_{\boldsymbol{u_t}} \quad & \mathrm{E}(x_T) \\ \text{s.t.} \quad & \mathrm{Var}(x_T) = \sigma_0^2, \\ & x_{t+1} = \sum_{i=1}^N r_t^i u_t^i + \Big(x_t - \sum_{i=1}^N u_t^i\Big) r_t^0, \quad t = 0, \ldots, T-1. \end{aligned} \tag{13}$$

As indicated by Li and Ng and also showed in Appendix A.1 for single-period, problem (13) is equivalent to minimizing $\mathrm{Var}(x_T)$ for $\mathrm{E}(x_T)$ not lower

1 Note that returns here are just $\frac{p_{t+1}}{p_t}$ (without subtracting 1).

2 The efficient frontier is constructed in such a way so that portfolios have the minimum risk for a given expected return and the maximum expected return for a given risk. As a result, expected return and risk (variance) are positively correlated on the efficient frontier. Therefore, maximizing the expected value subject to the constraint that the variance does not surpass a predetermined risk level is equivalent to employing an equality constraint. The same applies for the other way around. In this way, we are using equality in these constrains in the optimization problems while Li and Ng are using inequalities.

than a pre-selected minimum desired expected return, or maximizing $\mathrm{E}(x_T) - \omega\,\mathrm{Var}(x_T)$ for $\omega \geq 0$, which represents the risk aversion parameter.

$$\begin{aligned} \min_{\boldsymbol{u_t}} \quad & \mathrm{Var}(x_T) \\ \text{s.t.} \quad & \mathrm{E}(x_T) = \mu_0, \\ & x_{t+1} = \sum_{i=1}^{N} r_t^i u_t^i + \Big(x_t - \sum_{i=1}^{N} u_t^i\Big) r_t^0 = r_t^0 x_t + \boldsymbol{P_t}^\top \boldsymbol{u_t}, \\ & t = 0, \ldots, T-1, \end{aligned} \tag{14}$$

where $\boldsymbol{P_t} = \Big[P_t^1, \ldots, P_t^N\Big]^\top = \Big[(r_t^1 - r_t^0), (r_t^2 - r_t^0), \ldots, (r_t^N - r_t^0)\Big]^\top$. Assuming that

$$\mathrm{E}(\boldsymbol{r_t}\boldsymbol{r_t}^\top) = \mathrm{Cov}(\boldsymbol{r_t}) + \mathrm{E}(\boldsymbol{r_t})\mathrm{E}(\boldsymbol{r_t}^\top) > 0 \quad \forall t = 0, \ldots, T-1,$$

it follows (see equations 4-6 in Li and Ng (2000)) that

$$\mathrm{E}(\boldsymbol{P_t}\boldsymbol{P_t}^\top) > 0 \quad \forall t = 0, \ldots, T-1.$$

For the rest of the analysis, the formulation involving the risk aversion parameter ω will be employed, as it is more suitable for investment scenarios where an investor can determine their preferred balance between anticipated terminal wealth and the corresponding risk:

$$\begin{aligned} \max_{\pi} \quad & \mathrm{E}(x_T) - \omega \cdot \mathrm{Var}(x_T) \\ \text{s.t.} \quad & x_{t+1} = r_t^0 x_t + \boldsymbol{P_t}^\top \boldsymbol{u_t}, \quad t = 0, \ldots, T-1, \end{aligned} \tag{15}$$

where π is the multi-period portfolio policy $\{\boldsymbol{u_0}, \ldots, \boldsymbol{u}_{T-1}\}$. The constraint expressed in (15) is known as the *self-financing constraint*. This condition is crucial in multi-period portfolio optimization as it encapsulates the realistic financial situation of many investors, where additional external cash inflows into the portfolio are not typically possible. The self-financing constraint requires that all transactions within the portfolio, including purchases of new assets or payments of incurred costs (e.g. transaction costs, which we will later discuss), must be financed using the portfolio's existing assets. This means that any increase in one asset's holdings must be offset by a decrease in another asset's holdings or from the cash account at the case where we have a risk-free asset.

The analytical solution of problem (15), as derived from Li and Ng (2000, Section 4),is specified by the following analytical form:

$$\begin{aligned} \boldsymbol{u}_t^* = & -\mathrm{E}^{-1}(\boldsymbol{P_t}\boldsymbol{P_t}^\top)\mathrm{E}(r_t^0 \boldsymbol{P_t}) x_t \\ & + \frac{1}{2}\Big(b x_0 + \frac{\nu}{2\omega\alpha}\Big)\Big(\prod_{k=t+1}^{T-1} \frac{A_k^1}{A_k^2}\Big)\mathrm{E}^{-1}(\boldsymbol{P_t}\boldsymbol{P_t}^\top)\mathrm{E}(\boldsymbol{P_t}) \\ & \forall t = 0, \ldots, T-2, \end{aligned}$$

$$\begin{aligned} \boldsymbol{u}_{T-1}^{*} = & -\mathrm{E}^{-1}(\boldsymbol{P}_{T-1}\boldsymbol{P}_{T-1}^{\top})\mathrm{E}(r_{T-1}^{0}\boldsymbol{P}_{t})x_{T-1} \\ & + \frac{1}{2}\left(bx_0 + \frac{\nu}{2\omega\alpha}\right)\mathrm{E}^{-1}(\boldsymbol{P}_{T-1}\boldsymbol{P}_{T-1}^{\top})\mathrm{E}(\boldsymbol{P}_{T-1}), \end{aligned} \tag{16}$$

where all new parameters are defined in Li and Ng (2000, Section 3). The equivalent solutions for problems (13) and (14) can be obtained by setting:

$$\omega = \begin{cases} \frac{\nu}{2\sqrt{\alpha(\sigma_0^2 - cx_0^2)}} & \text{for problem (13),} \\ \frac{\nu^2}{2\alpha[\mu_0 - (\mu + b\nu)x_0]} & \text{for problem (14).} \end{cases}$$

The efficient frontier for all problems (13) - (15) is:

$$\mathrm{Var}(x_T) = \frac{\alpha}{\nu^2}[\mathrm{E}(x_T) - (\mu + b\nu)x_0]^2 + cx_0^2, \quad \text{for } \mathrm{E}(x_T) \geq (\mu + b\nu)x_0.$$

With the analytical solution (16), implementing the optimal multi-period portfolio policy for problems (13)-(15) becomes simple. The optimal multi-period portfolio policy is composed of two terms and displays a separation property between the investor's risk attitude and their present wealth. The second term in u_t^* depends on the investor's risk attitude and is independent of their current wealth. It can be calculated before the actual investment process starts, without the need for real-time data. The first term in u_t^* depends on the current wealth and is not influenced by the investor's risk attitude. It is computed during each time period as the present wealth becomes available, requiring real-time data.

Deriving the efficient frontier requires the following 4 steps:

Step-by-Step Process for Multi-Period Mean-Variance Efficient Frontier.	
Step 1:	Minimum expected return is $\mu_{min} = (\mu + b\nu)x_0$.
Step 2:	Find the maximum risk averse parameter ω_{max} from μ_{min}.
Step 3:	Find the minimum ω_{min} from $\mu_{max} = \max_{i=1,\ldots,N} \mu_i^T$.
Step 4:	Efficient frontier: Get the variance values for $\omega \in [\omega_{min}, \omega_{max}]$.

3.1.1.1 *Mean-Variance Portfolio Optimization with the Addition of a Risk-free Asset*

Investment scenarios that include a risk-free asset can be considered a special case within the general multi-period MV PO framework discussed earlier. Let the 0^{th} asset be risk-free, so we now examine a capital market consisting of N risky assets and a risk-free asset that offers a guaranteed rate of return s_t. In this situation, r_t^0 is equal to a constant s_t, and $\mathrm{Cov}(r_t^0, r_t^i) = 0$, for $i = 0, 1, \ldots, n$

and $\forall t = 0, 1, \ldots, T-1$. The optimal portfolio policy for problem (15) when a risk-free asset is included in the investment scenarios is given as:

$$\begin{aligned} \boldsymbol{u}_t^* &= -s_t \mathrm{E}^{-1}(\boldsymbol{P}_t \boldsymbol{P}_t^\top)\mathrm{E}(\boldsymbol{P}_t) x_t \\ &\quad + \Big(\prod_{k=0}^{T-1} s_k x_0 + \frac{1}{2\omega\big(\prod_{k=0}^{T-1}(1-B_k)\big)}\Big)\Big(\prod_{k=t+1}^{T-1} \frac{1}{s_k}\Big)\mathrm{E}^{-1}(\boldsymbol{P}_t \boldsymbol{P}_t^\top)\mathrm{E}(\boldsymbol{P}_t) \\ &\quad \forall t = 0, \ldots, T-2, \end{aligned}$$

$$\begin{aligned} \boldsymbol{u}_{T-1}^* &= -s_{T-1}\mathrm{E}^{-1}(\boldsymbol{P}_{T-1}\boldsymbol{P}_{T-1}^\top)\mathrm{E}(\boldsymbol{P}_t) x_{T-1} \\ &\quad + \Big(\prod_{k=0}^{T-1} s_k x_0 + \frac{1}{2\omega\big(\prod_{k=0}^{T-1}(1-B_k)\big)}\Big)\mathrm{E}^{-1}(\boldsymbol{P}_{T-1}\boldsymbol{P}_{T-1}^\top)\mathrm{E}(\boldsymbol{P}_{T-1}). \end{aligned} \tag{17}$$

The equivalent solutions for problems (13) and (14) can be obtained by setting:

$$\omega = \begin{cases} \frac{\nu}{2\sqrt{\alpha\sigma_0^2}} & \text{for problem (13)}, \\ \frac{\nu^2}{\alpha[2\mu_0 - b x_0]} & \text{for problem (14)}. \end{cases}$$

The efficient frontier for all problems (13) - (15) is:

$$\mathrm{Var}(x_T) = \frac{\alpha}{\nu^2}[\mathrm{E}(x_T) - \frac{b x_0}{2}]^2, \quad \text{for } \mathrm{E}(x_T) \geq \frac{b x_0}{2}.$$

Note that $\boldsymbol{u}_t^*$ is proportional to $\mathrm{E}^{-1}(\boldsymbol{P}_t \boldsymbol{P}_t^\top)\mathrm{E}(\boldsymbol{P}_t)$ for $t = 1, 2, \ldots, T-1$. This suggests that every investor will allocate his wealth among the risky assets while maintaining the same relative proportions. On the other hand, the proportion of investment in risky assets to investment in the risk-free asset is determined at each time period by considering the realized value of the investor's wealth and taking into account the investor's risk preferences. The steps for deriving the efficient frontier in this case are the same as before.

3.1.1.2 *Time Consistency in Multi-Period Mean-Variance Optimization Problem*

In multi-stage portfolio optimization, time consistency of risk measures is vital. As Boda, Filar, et al. (2006) describe time consistency, if a decision-maker minimizes a certain risk measure in an T-stage problem, the same policy, from any t^{th} stage onward, should minimize the risk measure for the remaining $(T-t+1)$-stage problem, for every $t = 1, \ldots, T$. This ensures the decision-making process remains coherent over time, aligning present and future decisions. Boda, Filar, et al. (2006) propose two requirements to formalize this concept, thus enhancing the robustness of multi-stage portfolio optimization strategies.

In this section's context, Cui et al. (2012) revised the optimal portfolio policy after identifying phenomena of time inconsistency in the dynamic MV PO problem. They discovered that during a certain period k between 1 and $T-1$, if the wealth x_k at the beginning of that period exceeds a threshold x_k^* (see Cui et al. (2012, Equation 3.1)), the investor's behavior deviates from rationality (ω becomes negative). In this situation, the investor's approach is falsely aiming to maximize both expected return and variance, with the trade-off being dependent on the value of x_k. Therefore, they showed that when $x_k > x_k^*$, then the investor should withdraw the amount $2(x_k - x_k^*)\left(1 - \prod_{j=k}^{T-1}(1-B_J)\right)$ out of the market (adding a positive cash-flow in contrast to the self-financing constrain). That is, at the start of the period k, the investor should use $\hat{x}_k$ as wealth to invest, where

$$\hat{x}_k = \begin{cases} x_k, & \text{if } x_k \leq x_k^*, \\ x_k - 2(x_k - x_k^*)\left(1 - \prod_{j=k}^{T-1}(1-B_j)\right), & \text{if } x_k > x_k^*. \end{cases}$$

By choosing $\hat{x}_k$ over x_k, the investor achieves the same expected return and variance in both cases (see Cui et al. (2012, Theorem 5.2)). However, since x_k is larger than $\hat{x}_k$, the investor can achieve the same amount of expected return and variance with less invested wealth. Thus, the investment policy associated with x_k is considered inefficient. Opting for $\hat{x}_k$ enables the investor to withdraw a specific amount, which will ultimately contribute to the total expected return at time T. Nonetheless, an aspect not addressed by Cui et al. is the potential for the investor to allocate the withdrawn amount to the risk-free asset, thereby achieving an even higher expected return without any additional risk. This approach highlights the benefits of adopting a more efficient investment strategy.

3.1.1.3 *Mean-Variance Portfolio Optimization with the Addition of No-Short Constrain*

Incorporating the no short-selling constraint in a multi-period setting proves to be more challenging than in a single-period context. Several years after deriving the analytical solution for the multi-period MV PO problem, Cui et al. (2014) developed a semi-analytical expression for the piece-wise quadratic value function under the no-short constraint. The problem formulation is as follows:

$$\begin{aligned} \min_{\boldsymbol{u_t}} \quad & \mathrm{Var}(x_T) \\ \text{s.t.} \quad & \mathrm{E}(x_T) = \mu_0, \\ & x_{t+1} = s_t x_t + \boldsymbol{P_t}^\top \boldsymbol{u_t}, \\ & \boldsymbol{u_t} \geq \boldsymbol{0}, t = 0, \ldots, T-1, \end{aligned} \tag{18}$$

where s_t is the rate of return of the risk-free asset, $\mu_0 \geq x_0 \prod_{t=0}^{T-1} s_t$ and $\mathrm{Cov}(\boldsymbol{e_t}) > 0$, which implies $\mathrm{E}(\boldsymbol{P_t}\boldsymbol{P_t}^\top) > 0$ and $\mathrm{E}(\boldsymbol{P_t}) > 0$.
Cui et al. showed that the optimal investment policy of problem (18) is expressed by (see Theorem 3.1):

$$\boldsymbol{u_t^*} = s_t \boldsymbol{K_t^*} y_t \mathbb{1}(y_t > 0),$$

where

$\boldsymbol{K}_t^* = \arg\min_{\boldsymbol{K}_t \geq \boldsymbol{0}} EM_t,$

$EM_t = E[C_{t+1}(1-\boldsymbol{P}_t^\top \boldsymbol{K}_t)^2 \mathbb{1}(P_t^\top K_t < 1) + (1-\boldsymbol{P}_t^\top \boldsymbol{K}_t)^2 \mathbb{1}(P_t^\top K_t \geq 1)],$

$C_t = \min_{\boldsymbol{K}_t \geq \boldsymbol{0}} EM_t, C_T = 1,$

$y_t = \frac{\mu_0 - \mu^*}{\rho_t} - x_t,$

$\rho_t = \prod_{l=t}^{T-1} s_l,$

$\mu^* = \frac{\mu_0 - \rho_0 x_0}{1 - C_0^{-1}}.$

The MV efficient frontier is

$$\mathrm{Var}(x_T) = \frac{(E[x_T] - \rho_0 x_0)^2}{C_0^{-1} - 1}, \quad \text{for } \mathrm{E}[x_T] \geq \rho_0 x_0.$$

Some observations regarding this solution:

1. The solution is considered semi-analytical, as the optimal parameter vector $\boldsymbol{K}_t^*$ can only be computed using a numerical approach.

2. The optimal parameter vector $\boldsymbol{K}_t^*$ depends on the distribution of $\boldsymbol{P}_t$, which necessitates information about the distribution rather than just the first two moments.

3. In cases where $y_t \leq 0$, implying that the wealth x_t exceeds the threshold $\frac{\mu_0 - \mu^*}{\rho_t}$, the optimal investment policy $\boldsymbol{u}_t$ is zero. This means that all the wealth is allocated to the risk-free asset. Since it offers constant returns, the wealth remains invested in it until time T and will never fall below that threshold again.

3.1.2 *Time Consistent Approach of Mean-Variance Portfolio Optimization*

Recognizing the imperative nature of time consistency for multi-period risk measures, Chen, Li, and Guo (2013) formulated a dynamic MV PO problem that obeys this requirement. They accomplish that by taking as a risk measure the sum of the single-period conditional risks. In their PO problem, their objective is to minimize the sum of the single-period conditional variances, offset by the product of the conditional expected returns and a risk aversion parameter ω for each respective period. They proved that the optimal investment policy of the problem with that objective satisfies both the time consistency requirements as defined in Boda, Filar, et al. (2006). The analytical solutions to this problem are expressed in Chen, Li, and Guo (2013, Theorems 4 and 5). Specifically, Theorem 4 provides the solution for a portfolio comprised solely of risky assets, while Theorem 5 extends the analysis to include a risk-free asset in the portfolio.

Transforming this problem to the equivalent one where we minimize the sum of conditional variances given that the expected return in each period is higher than a pre-selected minimum desired expected return, the optimization problem can be formulated as:

$$\begin{aligned} \min_{\boldsymbol{u_t}} \quad & \sum_{i=1}^{T} \mathrm{Var}(x_i) \\ \text{s.t.} \quad & \mathrm{E}(x_{t+1}) \geq x_t \mu_0, \\ & x_{t+1} = \boldsymbol{r_t}^\top \boldsymbol{u_t}, \\ & x_t = \mathbf{1}^\top \boldsymbol{u_t}, \\ & t = 0, \ldots, T-1, \end{aligned} \tag{19}$$

where we are using $\boldsymbol{u_t} = (u_t^0, u_t^1, \ldots, u_t^N)$ for this model only (risk-free asset allocation is included in the variable $\boldsymbol{u_t}$). $\mathrm{Var}(x_i)$ and $\mathrm{E}(x_i)$ can now be expressed in terms of the wealth allocation $\boldsymbol{u_i}$ at period i:

$$\mathrm{Var}(x_i) = \boldsymbol{u_i}^\top \Sigma_i \boldsymbol{u_i} \quad \text{and} \quad \mathrm{E}(x_i) = \boldsymbol{\mu}_i^\top \boldsymbol{u_i},$$

where $\boldsymbol{\mu}_i$ is the expected return and Σ_i the covariance at period i. Moreover, we can add the no-short constrains and the problem becomes:

$$\begin{aligned} \min_{\boldsymbol{u_t}} \quad & \sum_{i=0}^{T-1} \boldsymbol{u_i}^\top \Sigma_i \boldsymbol{u_i} \\ \text{s.t.} \quad & \boldsymbol{\mu}_t^\top \boldsymbol{u_t} \geq x_t \mu_0, \\ & x_{t+1} = \boldsymbol{r_t}^\top \boldsymbol{u_t}, \\ & x_t = \mathbf{1}^\top \boldsymbol{u_t}, \\ & \boldsymbol{u_t} \geq 0, \\ & t = 0, \ldots, T-1, \end{aligned} \tag{20}$$

This is a tractable convex optimization problem which can be efficiently solved with dedicated computational tools. To this end, the CVXPY library (see Diamond and Boyd (2016) and Agrawal et al. (2018)) is employed, which is a domain-specific library for convex optimization embedded in Python.

3.1.2.1 *Mean-Variance Portfolio Optimization including Transaction Costs*

In the portfolio allocation process, the primary goal is to achieve an optimal balance between return and risk. In the previous sections, we discussed this balance without considering transaction costs. This oversight can result in suboptimal target portfolio holdings that may incur significant trading costs, potentially impacting realized risk-adjusted returns negatively. By incorporating transaction costs directly into the portfolio allocation process, the resulting portfolios become more cost-effective and demonstrate improved realized risk-adjusted returns. Accounting for transaction costs complicates the portfolio optimization problem, as it necessitates tracking the amount traded in each period to accurately determine the associated costs. Dantzig and Infanger

(1993) introduced one of the most widely-recognized multi-period portfolio optimization problems with transaction costs. Their formulation employs three types of decision variables: x_j^t, y_j^t, and z_j^t, which represent the amounts of asset j held, bought, and sold by investors at period t, respectively. The new variables y_j^t and z_j^t are used to estimate the transaction costs for buying and selling an asset. A more contemporary and comprehensive approach has been proposed by Boyd et al. (2017). They are only using one extra variable y_j^t representing the amount transacted of asset j at period t. A positive value of y_j^t indicates that asset j was bought by that amount, while a negative value indicates that asset j was sold by that absolute of that amount. Following a similar way and adopting the notation from problem (14), in this section we include transaction costs in the multi-period MV PO problem without introducing any new variable.

Consider the vector $\boldsymbol{y_t} \in \mathbb{R}^n$, which signifies the dollar values of trades at current prices. In a simplified adaptation of the transaction costs model proposed by (Connor (2000)), the transaction cost of a risky asset i at period t is estimated as $TC_t^i = \gamma|y_t^i|$, interpreting the transaction cost as a proportion of the transaction amount, where γ is the transaction costs parameter. Given the assumption of separable transaction costs, the total transaction cost at period t is $TC_t = \sum_{i=1}^N TC_t^i$.
Next, we examine the investment procedure at an intermediate period $k \in \{0, \ldots, T-1\}$. Here, the investor holds a portfolio $\boldsymbol{u}_{k-1}^+$, with a total value of $x_k = \mathbf{1}^\top \boldsymbol{u}_{k-1}^+ + h_{k-1}^+$. The terms $\boldsymbol{u}_{k-1}^+$ and h_{k-1}^+ represent the post-trade wealth allocated to risky assets and risk-free asset respectively.
The investor proceeds to reallocate wealth to $\boldsymbol{u_k}$, subject to the constraints $\boldsymbol{u_k} \geq 0$ and $h_k \geq 0$. The latter constrain results in a condition that $x_k - \mathbf{1}^\top \boldsymbol{u_k} - TC_k \geq 0$, where $TC_k^i = \gamma|u_k^i - u_{k-1}^{i+}|$. Therefore, $TC_k = \sum_{i=1}^N TC_k^i = \gamma\|\boldsymbol{u_k} - \boldsymbol{u}_{k-1}^+\|_1$ accounts for the transaction costs of the trade from period $k-1$ to k. The term $\|\boldsymbol{u_k} - \boldsymbol{u}_{k-1}^+\|_1$ denotes the total amount traded. Consequently, the wealth allocated to the risk-free asset is now $h_k = x_k - \mathbf{1}^\top \boldsymbol{u_k} - TC_k$.
Having completed the investment for period k, the investor proceeds to period $k+1$. At this stage, the post-trade portfolio $\boldsymbol{u_k^+}$ is evaluated after realizing returns $\boldsymbol{r_k}$, yielding $\boldsymbol{u_k^+} = \boldsymbol{u_k} \circ \boldsymbol{r_k}$ and $h_k^+ = h_k \cdot s_k$, where $\circ$ denotes Hadamard (elementwise) multiplication of vectors. A visual representation of this process can be seen in Figure 6.
These post-trade allocations serve as inputs for period $k+1$, and this process is iteratively carried out up until period $T-1$. The final wealth at period T is consequently denoted as x_T.

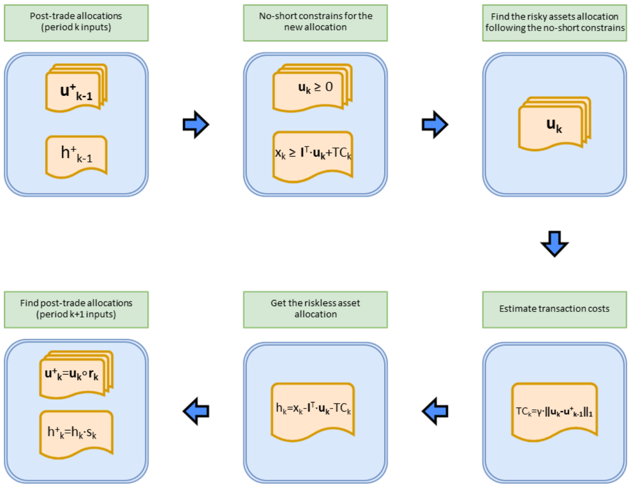


Figure 6: Schematic diagram of the multi-period portfolio optimization process with transaction costs at some intermediate period k.

Therefore, the new formulation of problem (20) using the notation from problem (14) is:

$$
\begin{aligned}
\min_{\boldsymbol{u}_t} \quad & \sum_{i=0}^{T-1} \boldsymbol{u}_i^\top \Sigma \boldsymbol{u}_i \\
\text{s.t.} \quad & \boldsymbol{u}_t^\top \boldsymbol{\mu} + R(x_t - \boldsymbol{u}_t^\top \mathbf{1} - TC_t) \geq x_t \mu_0, \\
& x_{t+1} = \boldsymbol{u}_t^\top \boldsymbol{\mu} + R(x_t - \boldsymbol{u}_t^\top \mathbf{1} - TC_t), \\
& x_t \geq \mathbf{1}^\top \boldsymbol{u}_t + TC_t, \\
& \boldsymbol{u}_t \geq 0, \\
& t = 0, \ldots, T-1,
\end{aligned}
\tag{21}
$$

where $TC_t = \gamma \| \boldsymbol{u}_t - \boldsymbol{u}_{t-1} \circ \boldsymbol{r}_{t-1} \|$ the estimated transaction costs at period t and u_{-1} is just zeros as in the beginning $(t = -1)$ we have not invested on anything.

Illustrative Example

Consider the case study in Li and Ng (2000, Section 7), assuming a stationary multi-period process with $T = 4$. An investor, starting with ten units of wealth at the beginning of the planning horizon ($x_0 = 10$), is trying to find the optimal allocation of his wealth among three risky assets, A, B, and C, in order to minimize the sum of variances over all periods, subject to expected return in each period not exceeding 1.1335, that is, $\mu_0 = 1.1335$.
The expected returns for the risky securities A, B, and C are $E(r_{A_t}) = 1.162$, $E(r_{B_t}) = 1.246$, and $E(r_{C_t}) = 1.228$ for $t = 0, 1, 2, 3$. The covariance of $r_t = [r_{A_t}, r_{B_t}, r_{C_t}]$ is given by

$$\text{Cov}(r_t) = \begin{bmatrix} 0.0146 & 0.0187 & 0.0145 \\ 0.0187 & 0.0854 & 0.0104 \\ 0.0145 & 0.0104 & 0.0289 \end{bmatrix},$$

for $t = 0, 1, 2, 3$. In addition to the three risky assets A, B, and C, there exists a risk-free asset with a guaranteed return rate of 1.04, that is, $E(r_{0_t}) = 1.04$. Transaction costs parameter is set to $\gamma = 1\%$ and using the notation u_t for all assets and $u_t^{(1:N)}$ for risky assets, here are the results from (21):

Results

Period 0:
Initial wealth: $x(0) = 10.0$
Risky assets allocation: $u^{(1:N)}(0) = [0.517, 1.081, 3.738]$ (Total: 5.335)
Transaction Costs: $TC(0) = 0.053$
Risk-free asset allocation: $u_0(0) = 4.611$
Wealth allocation for period 0: $u(0) = [4.611, 0.517, 1.081, 3.738]$ (Total: 9.947)
Post-trade wealth allocation for period 0: $u^+(0) = [4.796, 0.601, 1.346, 4.59]$ (Total wealth: 11.333)

Period 1:
Initial wealth: $x(1) = 11.333$
Risky assets allocation: $u^{(1:N)}(1) = [0.601, 1.144, 4.059]$ (Total: 5.804)
Transaction Costs: $TC(1) = 0.007$
Risk-free asset allocation: $u_0(1) = 5.521$
Wealth allocation for period 1: $u(1) = [5.521, 0.601, 1.144, 4.059]$ (Total: 11.326)
Post-trade wealth allocation for period 1: $u^+(1) = [5.742, 0.698, 1.426, 4.985]$ (Total: 12.851)

Period 2:
Initial wealth: $x(2) = 12.851$
Risky assets allocation: $u^{(1:N)}(2) = [0.698, 1.291, 4.584]$ (Total: 6.574)
Transaction Costs: $TC(2) = 0.005$
Risk-free asset allocation: $u_0(2) = 6.272$
Wealth allocation for period 2: $u(2) = [6.272, 0.698, 1.291, 4.584]$ (Total: 12.845)
Post-trade wealth allocation for period 2: $u^+(2) = [6.523, 0.811, 1.609, 5.63]$ (Total: 14.572)

Period 3:
Initial wealth: $x(3) = 14.572$
Risky assets allocation: $u^{(1:N)}(3) = [0.811, 1.451, 5.198]$ (Total: 7.461)
Transaction Costs: $TC(3) = 0.006$
Risk-free asset allocation: $u_0(3) = 7.105$
Wealth allocation for period 3: $u(3) = [7.105, 0.811, 1.451, 5.198]$ (Total: 14.566)
Post-trade wealth allocation for period 3: $u^+(3) = [7.389, 0.943, 1.809, 6.383]$

(Total: 16.524)

The expected final wealth is $E(x_T) = 16.524$, and the variance of the final wealth is $\text{Var}(x_T) = 3.755$.
In addition to the scenario with transaction costs, the problem was also solved without considering transaction costs (20). The figure below illustrates the desired expected return per period against the sum of variance for both cases (see Figure 7):

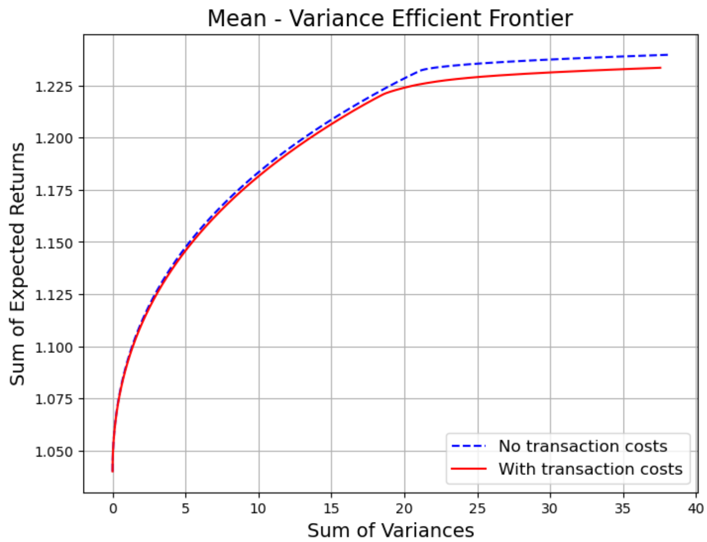


Figure 7: Plot of the desired expected return per period against the sum of variance for both cases: with and without transaction costs.

3.2 MEAN-CVAR PORTFOLIO OPTIMIZATION

In the field of portfolio optimization, the multi-period setting introduces an additional layer of complexity to the problem. Similar to the single-period setting examined in Chapter 2, the choice of an appropriate risk measure to optimize remains a key decision. Variance maintains its limitations, not only in single-period but also in multi-period scenarios as it was also mentioned in the conclusions of Cui et al. (2022). To overcome this issue, we turn our attention to downside risk measures that only focus on negative deviations. VaR and CVaR are two such quantile-based risk measures. While VaR defines the maximum loss at a certain confidence level, CVaR estimates the average value of losses exceeding the VaR at the same confidence level. However, VaR's shortcomings as a risk measure, such as its lack of coherence and inability to capture extreme losses beyond the specified probability level, make it less suitable for multi-period optimization. Given this, this section will concentrate on the Mean-CVaR PO problem in a multi-period context, emphasizing its benefits in managing downside risk over multiple periods.

Building upon the problem formulation in (21), where we included a risk-free asset, no-short constraints, and transaction costs, we will now explore the Mean-CVaR optimization problem. This problem aims to minimize the sum of CVaR values under the condition that the expected return in each period exceeds a predefined minimum expected return. As detailed in section 2.2, the computation of single-period $\text{CVaR}_\beta(L_t)$ at period t for $L_t = -\boldsymbol{u_t}^\top \boldsymbol{r_t}$ can be achieved through the following equation:

$$\text{CVaR}_\beta(L_t) = \min_{\alpha_t \in \mathbb{R}} F_\beta(\boldsymbol{u_t}, \alpha_t), \tag{22}$$

where $F_\beta(\boldsymbol{u_t}, \alpha_t) = \alpha_t + \frac{1}{1-\beta}\int_{\boldsymbol{r_t} \in \mathbb{R}^N} -\boldsymbol{u_t}^\top \boldsymbol{r_t} \cdot p(\boldsymbol{r_t}) d\boldsymbol{r_t}$. However, we can approximate $F_\beta(\boldsymbol{u_t}, \alpha_t)$ as follows:

$$\tilde{F}_\beta(\boldsymbol{u_t}, \alpha_t) = \alpha_t + \frac{1}{q(1-\beta)} \sum_{k=1}^{q} [-\boldsymbol{u_t}^\top \boldsymbol{r}_{t,k} - \alpha_t]^+,$$

where $r_{t,k}$ denotes the kth sample out of the q samples taken by the distribution $p(\cdot)$ at period t and $[\cdot]^+ = \max\{\cdot, 0\}$.

Given the formulations of the single-period Mean-CVaR problem in (11) and the multi-period problem in (21), we can define the multi-period Mean-CVaR PO problem as follows:

$$\begin{aligned}
\min_{\boldsymbol{u_t}, \boldsymbol{y_t}, \boldsymbol{\alpha}} \quad & \sum_{i=0}^{T-1} \alpha_i + \frac{1}{q(1-\beta)} \sum_{k=1}^{q} y_{i,k} \\
\text{s.t.} \quad & \boldsymbol{u_t}^\top \boldsymbol{\mu} + R(x_t - \boldsymbol{u_t}^\top \mathbf{1} - TC_t) \geq x_t \mu_0, \\
& \boldsymbol{u_t}^\top \boldsymbol{r}_{t,k} + \alpha_t + y_{t,k} \geq 0 \quad \text{for } k = 1, \ldots, q, \\
& y_{t,k} \geq 0 \quad \text{for } k = 1, \ldots, q, \\
& x_{t+1} = \boldsymbol{u_t}^\top \boldsymbol{\mu} + R(x_t - \boldsymbol{u_t}^\top \mathbf{1} - TC_t), \\
& x_t \geq \mathbf{1}^\top \boldsymbol{u_t} + TC_t, \\
& \boldsymbol{u_t} \geq 0, \\
& t = 0, \ldots, T-1, \\
& \alpha_t \in \mathbb{R}.
\end{aligned} \tag{23}$$

where $TC_t = \gamma \|\boldsymbol{u_t} - \boldsymbol{u}_{t-1} \circ \boldsymbol{r}_{t-1}\|$ the estimated transaction costs at period t.

Illustrative Example

Considering the same inputs as in section 3.1.2.1, the Multivariate Normal distribution is used to generate 1000 scenarios with vector mean = [0.162, 0.246, 0.228] and the covariance from that example. Taking $T = 4$, an investor starting with ten units of wealth at the beginning of the planning horizon ($x_0 = 10$), is trying to find the optimal allocation of his wealth among three risky assets and one risk-free with a guaranteed return rate of 1.04, in order to minimize the sum of CVaR values over all periods, subject to the expected return in each period not exceeding 1.1335, that is, $\mu_0 = 1.1335$. Transaction costs parameter is set to $\gamma = 1\%$ and for using the notation u_t for all assets and $u_t^{(1:N)}$ for risky assets, here are the results from (23):

Results

Period 0:
Initial wealth: $x(0) = 10.0$
Risky assets allocation: $u^{(1:N)}(0) = [2.471, 0.472, 3.178]$ (Total: 6.121)
Transaction Costs: $TC(0) = 0.061$
Risk-free asset allocation: $u_0(0) = 3.817$
Wealth allocation for period 0: $u(0) = [3.817, 2.471, 0.472, 3.178]$ (Total: 9.939)
Post-trade wealth allocation for period 0: $u^+(0) = [3.97, 2.871, 0.588, 3.903]$ (Total: 11.333)

Period 1:
Initial wealth: $x(1) = 11.333$
Risky assets allocation: $u^{(1:N)}(1) = [2.871, 0.457, 3.341]$ (Total: 6.669)
Transaction Costs: $TC(1) = 0.007$
Risk-free asset allocation: $u_0(1) = 4.657$
Wealth allocation for period 1: $u(1) = [4.657, 2.871, 0.457, 3.341]$ (Total: 11.326)
Post-trade wealth allocation for period 1: $u^+(1) = [4.843, 3.336, 0.569, 4.103]$ (Total: 12.851)

Period 2:
Initial wealth: $x(2) = 12.851$
Risky assets allocation: $u^{(1:N)}(2) = [3.336, 0.473, 3.767]$ (Total: 7.577)
Transaction Costs: $TC(2) = 0.004$
Risk-free asset allocation: $u_0(2) = 5.27$
Wealth allocation for period 2: $u(2) = [5.27, 3.336, 0.473, 3.767]$ (Total: 12.847)
Post-trade wealth allocation for period 2: $u^+(2) = [5.481, 3.877, 0.589, 4.626]$ (Total: 14.573)

Period 3:
Initial wealth: $x(3) = 14.573$
Risky assets allocation: $u^{(1:N)}(3) = [3.877, 0.51, 4.238]$ (Total: 8.625)
Transaction Costs: $TC(3) = 0.005$
Risk-free asset allocation: $u_0(3) = 5.944$
Wealth allocation for period 3: $u(3) = [5.944, 3.877, 0.51, 4.238]$ (Total: 14.569)
Post-trade wealth allocation for period 3: $u^+(3) = [6.181, 4.505, 0.636, 5.204]$ (Total: 16.526)

The expected final wealth is $E(x_T) = 16.526$, and the CVaR of the final wealth is $\text{CVaR}(x_T) = 2.285$.
The figure below illustrates the desired expected return per period against the sum of CVaR values for this example (see Figure 8):

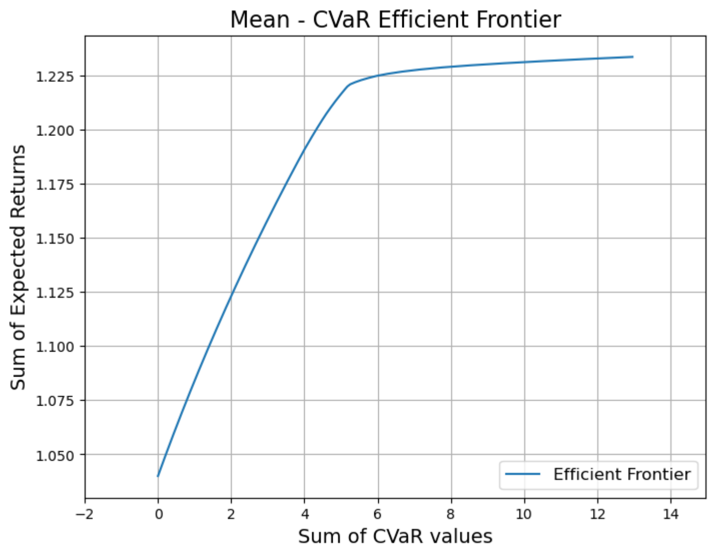


Figure 8: Plot of the desired expected return per period against the sum of CVaR values for this example.

4 ROBUST APPROACHES IN PORTFOLIO OPTIMIZATION

In this chapter, we explore robust approaches in portfolio optimization to address the limitations of traditional Mean-Risk models. Portfolios generated by Mean-Risk PO problems often exhibit extreme or impractical weights, making them difficult to implement in active trading. Robust optimization offers a more stable alternative by incorporating uncertainty in input parameters such as mean vectors and covariance matrices. This approach allows for more resilient allocations, reducing sensitivity to changes in model parameters and estimation errors. Uncertainty can be capture by constructing parameter uncertainty sets like box or ellipsoidal sets. Throughout this chapter, we focus on worst-case optimization derived from distribution uncertainty sets for the MV and Mean-CVaR PO problems, offering insights into its practical implications for portfolio decision-making.

4.1 ROBUST APPROACH IN SINGLE-PERIOD

In this section, we will address input parameter and distribution uncertainty for the returns. Consider the sets of parameters: θ and θ_0, where θ represents the true parameters, and θ_0 represents the estimated (nominal) parameters. For instance, in the MV model, θ represents the mean μ and the covariance Σ, while in the Mean-CVaR model, it refers to the distribution of portfolio return. θ is unknown, however, it is believed to belong to a specific uncertainty set U which is generated from the estimated parameter θ_0. In other words, $\theta \in U_{\theta_0}$, where the creation of U_{θ_0} will be specified later in this section. The goal of this approach is to construct a portfolio that minimizes the risk with respect to the worst-case scenario of the uncertain parameters in U_{θ_0}. In this section, we examine a distinct type of uncertainty that is linked to the distribution of the portfolio return instead of its moments. To be more precise, we assume that the density function of the portfolio return is only recognized to be part of a particular set $\mathscr{P}$ of distributions, that is $p(\cdot) \in \mathscr{P}$. To account for this uncertainty, Zhu and Fukushima (2009) introduced the concept of worst-case CVaR (WCVaR) for a given $\boldsymbol{w} \in W$ with reference to $\mathscr{P}$. The WCVaR is defined as

$$\text{WCVaR}_\beta(L) = \max_{p(\cdot)\in\mathscr{P}} \text{CVaR}_\beta(L).$$

Zhu and Fukushima (2009) address uncertainty related to the distribution of portfolio returns by assuming it belongs to a set of distributions, where each distribution in the set is a mixture of predetermined distribution scenarios.

The set can be expressed as the following mixture distribution uncertainty set:

$$p(\cdot) \in \mathscr{P} = \Big\{ \sum_{l=1}^{L} \lambda_l p^l(\cdot) : \sum_{l=1}^{L} \lambda_l = 1,\ \lambda_l \geq 0,\ l = 1, \dots, L \Big\},$$

where $p^l(\cdot)$ represents the density function of the l^{th} distribution scenario, and L denotes the total number of possible scenarios.

Using Theorem 1 from Zhu and Fukushima (2009) and the definition of CVaR from (8), WCVaR becomes:

$$\text{WCVaR}_\beta(L) = \min_{\alpha \in \mathbb{R}} \max_{p(\cdot) \in \mathscr{P}} F_\beta(\boldsymbol{w}, \alpha).$$

They also proved in Proposition 1 that WCVaR preserves coherence. Making the following reformulation, WCVaR can become:

$$\begin{aligned}
\max_{p(\cdot)\in\mathscr{P}} F_\beta(\boldsymbol{w},\alpha) &= \max_{\boldsymbol{\lambda}\in\Lambda} \Big\{ \alpha + \frac{1}{1-\beta} \int_{\boldsymbol{r}\in\mathbb{R}^N} -\boldsymbol{w}^\top \boldsymbol{r} \cdot p^{\boldsymbol{\lambda}}(\boldsymbol{r}) d\boldsymbol{r} \Big\} \\
&= \alpha + \frac{1}{1-\beta} \int_{\boldsymbol{r}\in\mathbb{R}^N} -\boldsymbol{w}^\top \boldsymbol{r} \cdot \max_{\lambda\in\Lambda} \sum_{l=1}^{L} \lambda_l p^l(\boldsymbol{r}) d\boldsymbol{r} \\
&= \alpha + \frac{1}{1-\beta} \int_{\boldsymbol{r}\in\mathbb{R}^N} -\boldsymbol{w}^\top \boldsymbol{r} \cdot \max_{l\in\mathscr{L}} p^l(\boldsymbol{r}) d\boldsymbol{r} \\
&= \max_{i\in\mathscr{L}} F^i_\beta(\boldsymbol{w},\alpha),
\end{aligned}$$

where $\Lambda = \{\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_L) : \sum_{l=1}^{L} \lambda_l = 1,\ \lambda_l \geq 0,\ l = 1, \dots, L\}$, $\mathscr{L} = \{1, \dots, L\}$ and $F^i_\beta(\boldsymbol{w},\alpha) = \alpha + \frac{1}{1-\beta} \int_{\boldsymbol{r}\in\mathbb{R}^N} -\boldsymbol{w}^\top \boldsymbol{r} \cdot p^i(\boldsymbol{r}) d\boldsymbol{r}$. Therefore, following the same logic as in equations (10) and (11), the optimization problem is:

$$\begin{aligned}
\min_{\boldsymbol{w},\boldsymbol{u},\boldsymbol{\alpha}} \quad & \alpha + \frac{1}{q_i(1-\beta)} \sum_{k=1}^{q_i} u_{k_i} \\
\text{s.t.} \quad & \boldsymbol{w}^\top \boldsymbol{r}_{k_i} + \alpha + u_{k_i} \geq 0 \quad \text{for } k = 1, \dots, q_i \text{ and } i = 1, \dots L, \\
& u_{k_i} \geq 0 \quad \text{for } k_i = 1, \dots, q_i \text{ and } i = 1, \dots L, \\
& \boldsymbol{w}^\top \boldsymbol{\mu} = \mu_0, \\
& \boldsymbol{w}^\top \mathbf{1} = 1, \\
& \boldsymbol{w} \geq 0, \\
& \alpha \in \mathbb{R},
\end{aligned} \tag{24}$$

where $\boldsymbol{r}_{k_i}$ is the k_i^{th} sample return from distribution $p^i(\cdot)$ and q_i is the number of samples from $p^i(\cdot)$.

Distributions $p^i(\cdot)$ for $i = 1, \dots, L$ should be determined as an input to the optimization problem. One way to find which distributions to include in the uncertainty set is to first attempt fitting a set of different distributions to the input data. Then, by conducting statistical tests, one can determine whether a distribution is representative of the data. The distributions that are a good fit to the data and pass the test are included in the uncertainty set. The procedure is described in the following steps:

Steps for deriving the Uncertainty Set.	
Step 1:	Fit distribution to data using maximum likelihood estimation.
Step 2:	Generate a sample from the fitted distribution of the same size as data.
Step 3:	Perform Hotelling's T-squared test as in Pituch and Stevens (2015) between data and sample and find the p-value.
Step 4:	Accept distribution in uncertainty set if p-value is greater than the usual significance level of 0.05.

The final step in the process involves the careful selection of candidate distributions to be included within the uncertainty set. However, there are not many distributions that can capture well the behavior of the returns. Distributions from the elliptical family are mostly used since they are easy to calibrate and generate data from, with Gaussian distribution being the most commonly used multivariate case. Nonetheless, Gaussian distribution has some disadvantages, such as its symmetry, which implies equal probabilities of losses and gains, and its use of linear correlation as a measure of dependence, which may not adequately capture non-linear dependencies observed in financial markets during times of crisis. To overcome these issues, we are using the approach implemented by Kakouris and Rustem (2014), where they use copulas to model the distribution of the data. The advantages of copulas include their flexibility in modeling the dependency between marginal distributions of random variables, allowing for the selection of the multivariate dependency separately from the univariate distributions. Copulas are also invariant under monotonic transformations and associated with many measures of dependence that measure the monotonic dependencies between two random variables, which are themselves invariant under monotonic transformations. This makes copulas a powerful tool for modeling complex dependence structures.

To obtain the efficient frontier in this case, we have to first define the uncertainty set containing mixture of copulas and then follow the necessary steps as shown earlier, to construct the efficient frontier. Thus, we proceed as follows: Following these steps, the following graph for the efficient frontier of the robust Mean-CVaR PO problem is derived (see Figure 9).

Upon analyzing Figure 9, it is evident that the robust Mean-CVaR efficient frontier retains a parabolic shape, where the expected returns increase concavely with WCVaR. This indicates that the balance between risk and reward is similar to other efficient frontiers discussed earlier. Moreover, the robust Mean-CVaR efficient frontier comprises a wider range of CVaR as it is de-

1 The empirical quantiles where found using statistical methods in Hyndman and Fan (1996).

Step-by-Step Process for Robust Mean-CVaR Efficient Frontier (Copulas).	
Step 1:	Define initial set of copulas $\mathscr{C}_0 = \{$Clayton, Gumbel, Frank, Gaussian, t-Student$\}$.
Step 2:	Perform the aforementioned four steps to get the fitted copulas that are representative to the input data and get the final uncertainty set $\mathscr{C}$.
Step 3:	Generate $q_i = S$ samples from each copula, that is for $i = 1, \ldots, L$.
Step 4:	Use the empirical distributions from the input data for each asset and apply its quantile function in the sample data[1].
Step 5:	Solve (24) for $\mu_0 \in [\min_{i=1,\ldots,N} \mu_i, \max_{i=1,\ldots,N} \mu_i]$ given the returns $\boldsymbol{r}_{k_i}$ from the sample data.
Step 6:	Global minimum CVaR portfolio: Find $\boldsymbol{w}_{min}$ such that $\min_{\boldsymbol{w}} \text{CVaR}_\beta(L)$.
Step 7:	Minimum expected return of the efficient frontier: $\mu_{min} = \boldsymbol{w}_{min}^\top \boldsymbol{\mu}$.
Step 8:	Maximum expected return of the efficient frontier: $\mu_{max} = \max_{i=1,\ldots,N} \mu_i$.
Step 9:	Efficient frontier: Get the CVaR values for $\mu_0 \in [\mu_{min}, \mu_{max}]$.

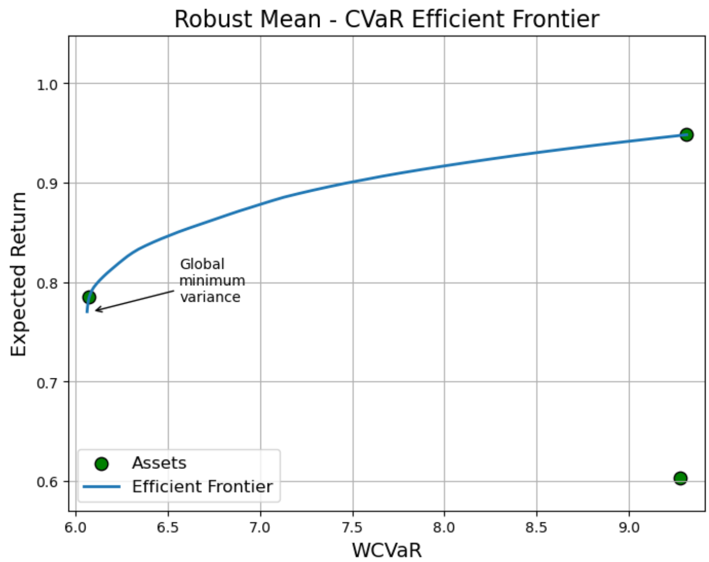


Figure 9: The efficient frontier Robust Mean-CVaR on a expected return - WCVaR plot (note that WCVaR is taken from the worst-case scenario), including the individual assets as points. Clayton, Gaussian and t-Student copulas are used in this example with confidence level of $\beta = 0.95$.

termined by sampling from fitted distributions and selecting the worst-case CVaR. However, it should be noted that the robust Mean-CVaR efficient frontier and the Mean-CVaR efficient frontier discussed earlier cannot be compared directly since the former uses WCVaR and the latter uses CVaR.

4.1.1 *Robust Mean-CVaR Portfolio Optimization with the Addition of a Risk-free Asset*

By incorporating the risk-free asset, the problem formulation and solution procedure remains the same, with the addition of a new component w_R to the weight vector $\boldsymbol{w}$ and a new constant component R to the mean vector $\boldsymbol{\mu}$. Thus, the first four steps in the solution procedure remain unchanged, while the remaining steps are as follows:

Step-by-Step Process for Robust Mean-CVaR Efficient Frontier (Copulas).	
Step 5:	The global minimum CVaR portfolio is achieved for $w_R = 1$ where CVaR is zero.
Step 6:	The minimum expected return of the efficient frontier is R, since w_{min} is for $w_R = 1$.
Step 7:	The maximum expected return of the efficient frontier is $\max\limits_{i=1,\ldots,N} \mu_i$.
Step 9:	The efficient frontier is obtained by calculating the CVaR values for μ_0 in the interval $[R, \mu_{max}]$.

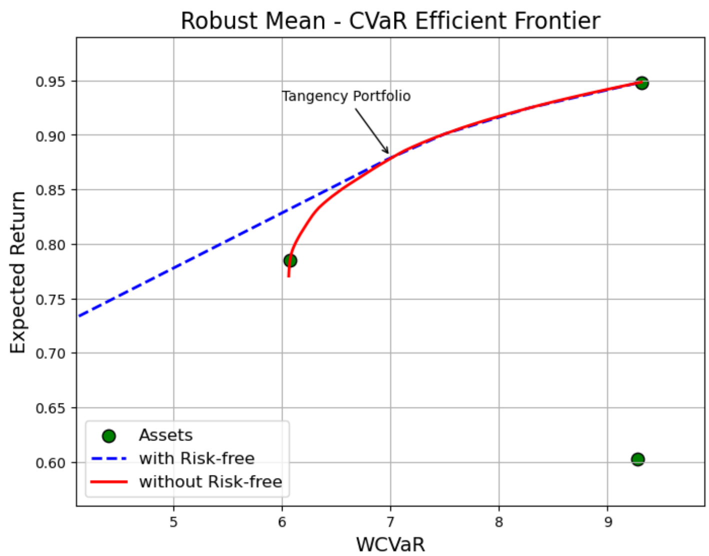


Figure 10: The efficient frontier Robust Mean-CVaR on a expected return - WCVaR plot with and without the Risk-free asset, including the individual assets as points.

Looking at Figure 10, once again, the robust Mean-CVaR efficient frontier partially takes the form of a straight line, indicating the presence of a tangency portfolio. Below the expected return of the tangency portfolio, the efficient frontier is a linear combination of the risk-free asset and the tangency portfolio, and above this return, the efficient frontier remains unchanged from the one without the risk-free asset.

4.1.2 *Robust Mean-CVaR Portfolio Optimization with the Addition of Transaction Costs*

If an investor chooses to apply this model to actual market data, they must account for transaction costs, which are usually related to the trading volume. When the investor adjusts their current portfolio to the new portfolio suggested by the Mean-CVaR PO solution, they will face transaction costs denoted as TC. To incorporate these costs, we use the following problem formulation:

$$
\begin{aligned}
\min_{\boldsymbol{w},\boldsymbol{u},\alpha} \quad & \alpha + \frac{1}{q_i(1-\beta)} \sum_{k=1}^{q_i} u_{k_i} \\
\text{s.t.} \quad & \boldsymbol{w}^\top \boldsymbol{r}_{k_i} + \alpha + u_{k_i} \geq 0 \quad \text{for } k = 1,\ldots,q_i \text{ and } i = 1,\ldots L, \\
& u_{k_i} \geq 0 \quad \text{for } k_i = 1,\ldots,q_i \text{ and } i = 1,\ldots L, \\
& \boldsymbol{w}^\top \boldsymbol{\mu} + w_R R \geq \mu_0, \\
& \boldsymbol{w}^\top \mathbf{1} + w_R = 1 - TC, \\
& w_R \geq 0, \\
& \boldsymbol{w} \geq 0, \\
& \alpha \in \mathbb{R},
\end{aligned}
\tag{25}
$$

where, $\boldsymbol{w}$ is the weights allocation and u_{k_i} is the k^{th} sample generated from the i^{th} copula from the copulas set. Here again μ_0 is the desired expected return of the portfolio. The estimated transaction costs are $TC = 0.01\|\boldsymbol{w} - \boldsymbol{w}_{prev} \circ \boldsymbol{r_t}\|$, where $\boldsymbol{w}_{prev}$ is the weights allocation of the previous portfolio.

4.2 ROBUST APPROACH IN MULTI-PERIOD

The previous sections in multi-period have primarily focused on portfolio optimization techniques that rely on the known return rates of invested assets. However, the characterization of uncertainty poses a significant challenge in practice. In reality, the underlying distribution of random variables, such as the rate of return of investment assets, remains unknown. Addressing this inherent uncertainty requires the utilization of efficient tools such as distributionally robust optimization.

Distributionally robust optimization provides a valuable approach for managing uncertainty in portfolio optimization. In particular, it assumes that the uncertain distribution resides within an ambiguity set, which is constructed based on incomplete distribution information of the random variables. One approach, described by Delage and Ye (2010) in their paper on distributionally robust optimization, involves utilizing generalized moment information. In their work, the authors introduce two parameters that serve as effective tools for controlling the size of the moment-based ambiguity set. Another method involves constructing the ambiguity set based on a metric, such as the Kullback-Leibler divergence, as demonstrated by Hu and Hong (2013) in their study. In the context of this research, the construction of the ambiguity set is achieved through the utilization of the Wasserstein metric.

The Wasserstein metric offers several advantages in the construction of ambiguity sets. Firstly, it facilitates the formation of natural confidence sets for unknown distributions through the use of Wasserstein balls. Moreover, the Wasserstein distance enables the measurement of distribution differences, even when their support sets are entirely dissimilar (see Kuhn et al. (2019)), making it a more flexible alternative to the most commonly used KL divergence. By adjusting the radius of the ambiguity set, one can modulate the level of conservativeness in the optimization problem. When the radius is set to zero, the model reduces to an ambiguity-free stochastic program. Consequently, the Wasserstein-based model exhibits improved out-of-sample performance in numerical experiments.

Motivated by the aforementioned advantages, we make the deliberate choice to incorporate an ambiguity set constructed using the Wasserstein metric as in Wu and Sun (2023). We also extend it by integrating transaction costs in the PO problem.

4.2.1 *Portfolio Optimization Problem Formulation in Robust Approach*

Consider a portfolio consisting of N risky assets and one risk-free with a constant rate of return $s_t = R$. The investor joins the market at time 0 with an initial wealth x_0 which he can allocate among the assets, and exits the market at time period T. The wealth can be reallocated among the $N+1$ assets at the beginning of each of the following $T-1$ consecutive time periods. Let $\boldsymbol{r_t} = (r_t^1, \ldots, r_t^N)$ be the N-vector of uncertain future rates of returns at time period t. It is assumed that vectors $\boldsymbol{r_t}, t = 0, \ldots, T-1$, are statistically independent. Let $\mathbb{P}_t$ be the underlying distribution of the random return vector $\boldsymbol{r_t}$, then the expected return and variance are $\mathrm{E}_{\mathbb{P}_t}(\boldsymbol{r_t})$ and $\mathrm{Var}_{\mathbb{P}_t}(\boldsymbol{r_t})$ respectively.

Given the worst-case scenario robust approach, an investor seeks the optimal investment strategy, $\boldsymbol{u_t} = (u_t^1, \ldots, u_t^N), t = 0, \ldots, T-1$, to minimize the worst-case sum of variances in each time period, subject to the constraint that the worst-case sum of the expected values in each time period does not exceed a pre-selected expected return level. This can be represented by the following optimization problem:

$$\begin{aligned}
\min_{\boldsymbol{u_t}} \quad & \sum_{i=0}^{T-1} \max_{\mathbb{P}_t \in \mathcal{Q}_{\theta_t}(\mathbb{P}_{t,N_t})} \mathrm{Var}_{\mathbb{P}_i}(x_i) \\
\text{s.t.} \quad & \max_{\mathbb{P}_t \in \mathcal{Q}_{\theta_t}(\mathbb{P}_{t,N_t})} \mathrm{E}_{\mathbb{P}_t}(x_t) \geq x_t \mu_0, \\
& x_{t+1} = \boldsymbol{u_t}^\top \mathrm{E}_{\mathbb{P}_t}(\boldsymbol{r_t}) + s_t(x_t - \mathbf{1}^\top \boldsymbol{u_t} - TC_t), \\
& x_t \geq \mathbf{1}^\top \boldsymbol{u_t} + TC_t, \\
& \boldsymbol{u_t} \geq 0, \\
& t = 0, \ldots, T-1,
\end{aligned} \tag{26}$$

where $TC_t = \gamma \|\boldsymbol{u_t} - \boldsymbol{u}_{t-1} \circ \boldsymbol{r}_{t-1}\|$ the estimated transaction costs at period t and $\mathcal{Q}_{\theta_t}(\mathbb{P}_{t,N_t})$ the ambiguity set.

4.2.2 *Wasserstein Metric and Ambiguity Set*

The construction of the ambiguity set involves the utilization of the Wasserstein metric in combination with the empirical distribution, denoted as $\mathbb{P}_{t,N_t}$, derived from the sample of size N_t for $\boldsymbol{r_t}$. This empirical distribution is mathematically represented as:

$$\mathbb{P}_{t,N_t} = \frac{1}{N_t} \sum_{i=1}^{N_t} \delta_{\boldsymbol{r_t},i}.$$

Here, the term $\delta_{\boldsymbol{r_t},i}$ represents the Dirac point measure positioned at $\boldsymbol{r_t}, i$.
From the definition given in Wu and Sun (2023), the Wasserstein metric plays a crucial role in quantifying the dissimilarity between probability distributions. In their work, the authors provide the following insights regarding the Wasserstein metric: The Wasserstein metric is defined on the space $\mathscr{M}(\Omega)$, which encompasses all probability distributions $\mathbb{Q}$ supported on Ω, satisfying the condition $\mathrm{E}_{\mathbb{Q}}(\|r\|) = \int \Omega \|r\| \mathbb{Q}(dr) < \infty$.
According to Definition 2 presented by Wu and Sun (2023), the Wasserstein metric, denoted as $D_W : \mathscr{M}(\Omega) \times \mathscr{M}(\Omega) \to \mathbb{R}^+$, is formulated as follows:

$$D_W(\mathbb{Q}_1, \mathbb{Q}_2) := \inf_{\Pi \in \Gamma(\mathbb{Q}_1, \mathbb{Q}_2)} \int_{\Omega^2} \|r - r'\| \, \Pi(dk, dk'),$$

where $\Gamma(\mathbb{Q}_1, \mathbb{Q}_2)$ represents the set of joint distributions Π with marginal distributions $\mathbb{Q}_1$ and $\mathbb{Q}_2$. Notably, when $p = 2$, the Wasserstein metric corresponds to a distance of order two.

It is important to note that while D_W does not guarantee a real distance, it can be interpreted as the minimum transportation cost required to move mass from distribution $\mathbb{Q}_1$ to distribution $\mathbb{Q}_2$. Based on this concept, the ambiguity set at time t, denoted as $\mathcal{Q}_{\theta_t}(\mathbb{P}t, N_t)$, can be defined as:

$$\mathcal{Q}_{\theta_t}(\mathbb{P}_{t,N_t}) = \{\mathscr{P}_t : D_W(\mathbb{P}_t, \mathbb{P}_{t,N_t}) \leq \theta_t\},$$

where $\mathcal{Q}\theta_t(\mathbb{P}_{t,N_t})$ represents the Wasserstein ball of radius θ_t centered at the empirical distribution $\mathbb{P}_{t,N_t}$. Notably, in mild situations, $\mathcal{Q}_{\theta_t}(\mathbb{P}_{t,N_t})$ comprises the true distribution $\mathbb{P}_t$ with high probability (see Mohajerin Esfahani and Kuhn (2018)). This property highlights the effectiveness and reliability of utilizing the Wasserstein metric in constructing the ambiguity set for robust portfolio optimization.

4.2.3 *Portfolio Optimization Problem Using Wasserstein Ambiguity Set*

In this section, we propose a reformulation of the portfolio optimization problem (26) using the Wasserstein ambiguity set, as introduced by Wu and Sun (2023) in Lemma 1 and Theorem 1. Before presenting the reformulation, we need to address the estimation of transaction costs, which involves the uncertainty of returns. To simplify the problem, we assume a small error in transaction cost estimation due to returns uncertainty and use the empirical distribution to estimate expected return and thus the transaction costs. Thus, the transaction cost TC_t can be expressed as follows:

$$TC_t = \gamma \| \boldsymbol{u_t} - \boldsymbol{u}_{t-1} \circ \mathrm{E}_{\mathbb{P}_{t,N_t}}(\boldsymbol{r}_{t-1}) \|.$$

By using Lemma 1 from Wu and Sun (2023), we can equivalently rewrite the second constraint in problem (26) as follows:

$$\boldsymbol{u_t}^\top \mathrm{E}_{\mathbb{P}_{t,N_t}}(\boldsymbol{r_t}) + s_t(x_t - \mathbf{1}^\top \boldsymbol{u_t} - TC_t) - \sqrt{\theta_t}\|\boldsymbol{u_t}\| \geq x_t \mu_0.$$

Furthermore, Theorem 1 from the same paper allows us to reformulate the objective function in problem (26) as:

$$\sum_{i=0}^{T-1} \sqrt{\boldsymbol{u_i}^\top \mathrm{Var}_{\mathbb{P}_{i,N_i}}(x_i) \boldsymbol{u_i}} + \sqrt{\theta_t}\|\boldsymbol{u_t}\|.$$

Consequently, the primal problem (26) can be reformulated into the following dual problem:

$$\begin{aligned}
\min_{\boldsymbol{u_t}} \quad & \sum_{i=0}^{T-1} \sqrt{\boldsymbol{u_i}^\top \mathrm{Var}_{\mathbb{P}_{i,N_i}}(x_i) \boldsymbol{u_i}} + \sqrt{\theta_t}\|\boldsymbol{u_t}\| \\
\text{s.t.} \quad & \boldsymbol{u_t}^\top \mathrm{E}_{\mathbb{P}_{t,N_t}}(\boldsymbol{r_t}) + s_t(x_t - \mathbf{1}^\top \boldsymbol{u_t} - TC_t) - \sqrt{\theta_t}\|\boldsymbol{u_t}\| \geq x_t \mu_0, \\
& x_{t+1} = \boldsymbol{u_t}^\top \mathrm{E}_{\mathbb{P}_{t,N_t}}(\boldsymbol{r_t}) + s_t(x_t - \mathbf{1}^\top \boldsymbol{u_t} - TC_t) - \sqrt{\theta_t}\|\boldsymbol{u_t}\|, \\
& x_t \geq \mathbf{1}^\top \boldsymbol{u_t} + TC_t, \\
& \boldsymbol{u_t} \geq 0, \\
& t = 0, \ldots, T-1.
\end{aligned} \tag{27}$$

The reformulated problem (27) preserves the optimal solutions and optimal value of the original problem (26). It takes advantage of the Wasserstein ambiguity set to ensure robustness in portfolio optimization by considering the uncertainty in the distribution of returns. Parameters θ_t are estimated using the bootstrap method proposed in Kang et al. (2019). To solve this optimization problem, the CVXPY library was utilized once again (refer to Diamond and Boyd (2016) and Agrawal et al. (2018)), which is a specialized Python library for convex optimization, to formulate and solve the portfolio optimization problem with the Wasserstein-based robust approach. This library provided the necessary tools and functions to express the problem in a mathematically rigorous manner and efficiently find the optimal solution.

In the next chapter, we will present empirical results and performance comparisons to demonstrate the effectiveness of the Wasserstein-based robust portfolio optimization approach in managing uncertainty and achieving superior risk-adjusted returns.

5 NUMERICAL EXPERIMENTS

In this chapter, we present a comprehensive analysis of portfolio optimization in a multi-period setting through a series of experiments using historical data. While many existing papers in the field typically analyze a specific set of stocks, our approach takes a broader perspective. We carefully selected a diverse wide range of stocks, considering factors such as market capitalization, industry diversification, and historical performance, to ensure a representative mix of companies across various sectors and market segments. To introduce variability and randomness into the selection process, we employ a randomized approach by randomly choosing a subset of the stocks to work with. Furthermore, to simulate realistic investment scenarios, we randomly select a starting date for the investor's initial investment. Having a random set of stocks and a random starting date, we use historical data from the chosen starting date and evaluate the performance of different portfolio optimization models over a specific time horizon.

Throughout the simulation, portfolio rebalancing occurs periodically, ensuring adaptability to changing market conditions. For the single-period models, the investor maintains the same weights for the duration of the rebalancing period, while for the multi-period models, the investor dynamically allocates wealth based on model's output using total number of periods equal to the rebalancing period. To capture the variability resulting from different random selections, we repeat this random process multiple times and track the expected returns, variance and CVaR for each portfolio model.

As a key performance metric, we evaluate the Sharpe ratio by Sharpe (1994) for each PO problem. We highlight the model with the highest Sharpe ratio in each case, as it serves as an important indicator of risk-adjusted returns and can provide valuable insights into the relative performance of different models. Additionally, we include the equal weights model and the S&P 500 as benchmark strategies for comparison. This comprehensive set of experiments provides valuable insights into the performance of the considered portfolio optimization methods and their relative strengths compared to benchmark strategies.

5.1 RANDOM NUMERICAL EXPERIMENT PROCESS DESCRIPTION

In this section, we will provide a detailed illustration of the methodology employed in our random numerical experiment. We will outline the step-by-step process of the experiment, including the selection of stocks, determination of the investment start date, application of portfolio optimization models, rebalancing intervals, and tracking of performance metrics. For the numerical experiments we use the models discussed in Chapters 2, 3 and 4.

In our random numerical experiment, we carefully selected a wide set of 23 stocks representing diverse sectors and market segments. These stocks include AAPL, AMZN, JPM, XOM, GE, INTC, DIS, IBM, MMM, HD, WMT, ORCL, MRK, BA, WFC, GS, BMY, GILD, COP, CAT, SBUX, F, PG. Next, we randomly selected a subset of stocks from this wide set, ranging from 5 to 15 stocks. Thereafter, we choose a random date from 2015 to 2023 to start our analysis. The historical data of daily returns for the 1000 days before the starting date are used as an input to our models. Each model has as an initial wealth of $x_0 = 10000$, and no additional external cash inflows are considered. To make the numerical experiments realistic, we apply transaction costs at a rate of 1% as a proportion of the transaction amount.
To evaluate the performance of the models, they are tested on the next 100 days of realized returns. The investor's objective is to minimize risk while ensuring that the expected return at each period exceeds a pre-specified expected return level. The expected return level is determined based on the means of the selected stocks from the historical data. Specifically, it is set equal to:

$$\mu_0 = \frac{\overline{\mu} + \mu_{max}}{2},$$

where $\boldsymbol{\mu}$ is the vector of empirical expected value of the returns from the historical data, $\overline{\mu}$ is the average of the expected return vector $\boldsymbol{\mu}$ and μ_{max} is the maximum expected return. The logic behind this is to get a higher pre-specified expected return level than the mean of the expected returns but not as high as the maximum value.
During the testing period, the investor reevaluates the portfolio allocation every 5 days. This means that after 5 days, the investor considers a new set of historical data comprising the 1000 previous days starting from that date. His current wealth at the evaluation date serves as the initial value while the portfolio allocation will also be used as initial allocation in order to calculate the transaction costs. In single-period models, the investor will rebalance each day the allocation to the weights specified in the beginning applying the transaction costs. In multi-period models, the investor will dynamically allocate wealth based on five-period multi-period model.
In the next subsections, we provide a detailed description of each model and its application within this experiment. All models include the risk-free asset whose return constant R is given by the yield of 3-months Treasury Bill (ˆIRX) at the starting date. The risk-free return stays a constant for the rest of the analysis. Although this is not the case for ˆIRX, in order to better utilize the idea of a risk-free asset we keep its value constant.

5.1.1 *Portfolio Optimization Models*

Single-Period Mean-Variance model

The first model used is the MV model for single-period as in problem (6) where the estimated transaction costs are $TC = 0.01\|\boldsymbol{w} - \boldsymbol{w}_{prev} \circ \boldsymbol{r_t}\|$, and

$\boldsymbol{w}_{prev}$ is the weights allocation of the previous portfolio.

Model's inputs are:

- $\boldsymbol{w}_{prev}$: the current weights allocation of the portfolio
- Parameters from historical data: Mean $\boldsymbol{\mu}$, Covariance Σ and Minimum desired expected return μ_0.

Single-Period Mean-CVaR model

The second model used is the Mean-CVaR model for single-period as in problem (12) where the estimated transaction costs are $TC = 0.01\|\boldsymbol{w} - \boldsymbol{w}_{prev} \circ \boldsymbol{r_t}\|$, and $\boldsymbol{w}_{prev}$ is the weights allocation of the previous portfolio.

Model's inputs are:

- $\boldsymbol{w}_{prev}$: the current weights allocation of the portfolio
- Parameters from historical data: Mean $\boldsymbol{\mu}$, Historical data $\{u_k, \quad k = 1.\dots,q\}$ and Minimum desired expected return μ_0.

Robust Single-Period Mean-CVaR model

Third and last model for single-period is the robust approach of Mean-CVaR as in (25).

Model's inputs are:

- $\boldsymbol{w}_{prev}$: the current wealth allocation of the portfolio
- Parameters from historical data: Mean $\boldsymbol{\mu}$, Historical data $\{u_k, \quad k = 1.\dots,q\}$ and Minimum desired expected return μ_0.

Multi-Period Mean-Variance model

For the multi-period, first model we use is the MV with $T = 5$ and following the time consistent approach as in (21). The estimated transaction costs are $TC_t = 0.01\|\boldsymbol{u_t} - \boldsymbol{u}_{t-1} \circ \boldsymbol{r_t}\|$, where $\boldsymbol{u}_{-1}$ is the wealth allocation of the previous portfolio. The wealth allocation of the risk-free asset is $h_t = x_0 - \boldsymbol{u_t}^\top \mathbf{1} - TC_t$.

Model's inputs are:

- $\boldsymbol{u}_{-1}$: the current wealth allocation of the portfolio
- x_0: the total wealth of the portfolio at that moment (it is also equal to $\boldsymbol{u}_{-1}^\top \mathbf{1}$).
- Parameters from historical data: Mean $\boldsymbol{\mu}$, Covariance Σ and Minimum desired expected return μ_0.

Multi-Period Mean-CVaR model

For the multi-period, first model we use is the MV with $T = 5$ and following the time consistent approach as in (23). The estimated transaction costs are $TC_t = 0.01\|\boldsymbol{u_t} - \boldsymbol{u}_{t-1} \circ \boldsymbol{r_t}\|$, for $\boldsymbol{u}_{-1}$ the wealth allocation of the previous portfolio. The wealth allocation of the risk-free asset at period t is $h_t = x_0 - \boldsymbol{u}_t^\top \mathbf{1} - TC_t$.

Model's inputs are:

- $\boldsymbol{u}_{-1}$: the current wealth allocation of the portfolio
- x_0: the total wealth of the portfolio at that moment (it is also equal to $\boldsymbol{u}_{-1}^\top \mathbf{1}$).
- Parameters from historical data: Mean $\boldsymbol{\mu}$, Historical data $\{y_k, \quad k = 1\ldots, q\}$ and Minimum desired expected return μ_0.

Robust Multi-Period Mean-Variance model

Last multi-period model is the robust approach of the MV using the Wasserstein Ambiguity Set discussed in (27). The estimated transaction costs are $TC_t = 0.01\|\boldsymbol{u_t} - \boldsymbol{u}_{t-1} \circ \mathrm{E}_{\mathbb{P}_{t,N_t}}(\boldsymbol{r}_{t-1})\|$, for $\boldsymbol{u}_{-1}$ the wealth allocation of the previous portfolio. The wealth allocation of the risk-free asset at period t is $h_t = x_0 - \boldsymbol{u}_t^\top \mathbf{1} - TC_t$.

Model's inputs are:

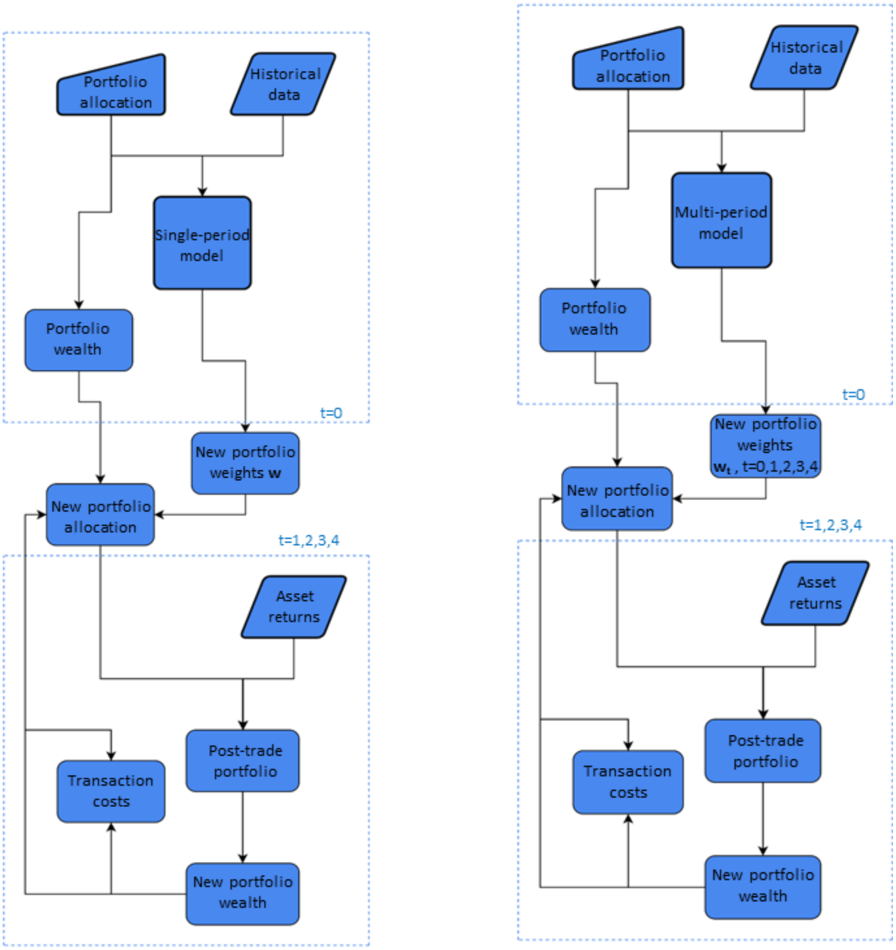
- $\boldsymbol{u}_{-1}$: the current wealth allocation of the portfolio
- x_0: the total wealth of the portfolio at that moment (it is also equal to $\boldsymbol{u}_{-1}^\top \mathbf{1}$).
- Parameters from historical data: Mean $\boldsymbol{\mu}$, Covariance Σ, Historical data $\{y_k, \quad k = 1\ldots, q\}$ to estimate $\boldsymbol{\theta}_t$ and Minimum desired expected return μ_0.

5.1.2 *Reevaluation Process*

After selecting a random starting date, we evaluate the performance of the portfolio models over the next 100 days. Throughout this period, the portfolios are reevaluated every 5 days based on updated historical data. The process of reevaluation differs between the single-period and multi-period models, we show and explain the process in Figure 11.

Single-Period Reevaluation Process

During the reevaluation process in the single-period models, the current portfolio allocation and historical data from the previous 1000 days serve as in-



(a) Single-period models. (b) Multi-period models.

Figure 11: Flow chart of the reevaluation process in the single-period models.

puts. The current allocation will provide the initial wealth, while it will be involved in the estimation of transaction costs. The historical data will contribute to the estimation of expected rewards and risks. Using these inputs, the single-period model provides the weights that the investor follows for the next day. Over the remaining four days of the reevaluation period, the investor rebalances the portfolio allocation to align with the weights provided by the model. Transaction costs are applied in this case. The 5-days process is presented in 11a.

Multi-Period Reevaluation Process

For the multi-period models, the inputs are the same, that is, the current portfolio allocation and historical data from the previous 1000 days. The current allocation provides the initial wealth and transaction costs, while the historical data contributes to the estimation of expected rewards and risks. Using these inputs, the multi-period model provides the weights that the investor will follow for the next 5 days. In this way, for the remaining four days of the reevaluation period, the investor rebalances the portfolio allocation to align with the weights provided by the model at that time period. Once again, transaction costs are applied in this case. This 5-days process can be seen in 11b.

5.2 NUMERICAL EXPERIMENT RESULTS

In this section, we present the results of the numerical experiments conducted to evaluate the performance of aforementioned portfolio optimization models. The experiments aim to assess the expected return, variance, CVaR, and Sharpe ratio of the portfolios generated by these models. The analysis is based

Portfolio Expected Return [1e-4]

	S&P500	EW	SP-MV	SP-MC	R-SP-MC	MP-MV	MP-MC	R-MP-MV
1	0.45	4.67	17.54	14.45	19.94	17.4	15.2	17.5
2	2.31	3.03	-1.71	-1.36	-0.23	-1.37	-0.95	-1.60
3	4.90	6.87	13.5	14.61	13.42	13.26	14.66	11.49
4	2.53	2.84	8.03	6.12	9.19	8.52	7.45	8.61
5	2.70	6.27	8.88	8.26	8.70	9.39	8.58	10.64
6	8.93	-5.69	-3.11	-1.86	-7.17	-2.99	-1.46	-3.27
7	9.73	12.35	3.12	1.55	-0.57	3.00	1.69	4.21
8	4.08	4.74	7.16	6.65	6.38	7.15	6.54	5.68
9	-2.96	-4.07	-4.59	-4.40	-3.82	-4.56	-4.45	-3.90
10	4.03	6.93	10.4	10.67	8.61	10.37	10.17	8.45
11	10.57	10.28	9.15	7.57	10.92	9.30	7.66	11.37
12	7.30	9.18	6.55	7.08	10.33	6.51	7.2	10.04
13	1.26	3.54	3.15	2.86	1.10	2.98	2.89	5.48
14	1.26	-1.92	-1.89	-1.20	2.20	-1.33	-0.76	-1.16
15	6.18	-0.94	-1.97	-1.68	-6.32	-2.04	-1.95	-1.39
16	-3.79	-5.27	0.07	2.34	1.68	-0.40	1.13	-2.52
17	5.76	6.45	-1.75	-2.40	3.45	-1.82	-2.51	1.84
18	-2.74	-1.95	-3.11	-3.46	-6.41	-2.86	-3.18	-3.89
19	0.37	-4.44	2.51	3.86	2.71	2.65	2.90	0.47
20	-9.62	-9.66	-1.04	-1.57	-1.04	-1.05	-1.77	-4.21

Table 1: The portfolio expected returns for each model per experiment.

on custom Python code developed specifically for this research, which can be found in code link. The code implementation allows for the systematic evaluation of various portfolio optimization strategies and provides reliable and reproducible results. The findings are presented in four tables, each presenting the respective metrics for the evaluated models. The labels for each model are defined in Table 6. We want to find the model that performed the best in each experiment and then count how many best performances each model had. To

Portfolio Variance [1e-5]

	S&P500	EW	SP-MV	SP-MC	R-SP-MC	MP-MV	MP-MC	R-MP-MV
1	11.42	15.00	18.77	17.82	14.53	19.14	18.44	20.23
2	18.24	20.35	35.49	35.12	39.15	35.47	35.01	38.76
3	2.16	2.88	6.79	7.05	6.87	6.76	7.04	6.31
4	4.52	4.94	9.88	10.19	9.40	9.98	10.48	9.21
5	5.03	5.85	9.10	9.39	10.75	9.24	9.51	9.29
6	4.30	5.12	8.23	8.87	8.55	8.27	8.85	8.25
7	6.88	11.62	8.38	8.99	9.42	8.49	9.24	8.75
8	1.86	1.49	2.67	2.87	3.65	2.67	2.87	2.6
9	6.11	5.97	8.38	8.60	8.23	8.41	8.62	8.44
10	1.85	2.03	1.57	2.02	2.36	1.58	1.95	2.02
11	3.68	4.62	3.76	4.08	5.66	3.82	4.16	4.09
12	2.60	2.56	3.20	3.35	3.22	3.20	3.35	3.22
13	14.99	13.06	18.8	18.50	18.03	18.87	18.45	19.67
14	14.99	13.48	15.04	15.71	15.62	15.13	15.72	15.46
15	4.36	4.56	6.00	7.00	6.63	5.94	7.01	5.75
16	14.10	11.81	13.41	13.46	13.32	13.46	13.45	14.72
17	5.17	6.82	6.87	7.72	6.38	6.96	8.09	7.44
18	16.23	12.89	17.92	17.49	18.02	17.9	17.39	18.63
19	13.46	11.18	14.47	14.22	13.37	14.63	14.65	14.78
20	7.31	5.44	4.03	4.10	4.38	4.06	4.21	4.67

Table 2: The portfolio variance for each model per experiment.

find the best-performing model, we are looking for the highest Sharpe ratio. We define Sharpe ratio as the following expression:

$$S_P = \frac{\mu_P}{\sigma_P},$$

where μ_P the empirical expected return and σ_P the volatility of the portfolio. However, when the empirical expected returns are negative, maximizing the Sharpe ratio alone may not accurately identify the best model, as higher volatility would yield higher Sharpe ratio. In such cases, we will look at the expected return and volatility of the highest Sharpe ratio model, if they are the highest and lowest respectively from all other models, then this model is the one performed better. Otherwise, we will employ an alternative approach, which is derived by its definition, as proposed by McLeod and Vuuren (2004):

$$\begin{aligned}\max \mathrm{P}(r_P \geq 0) &\Leftrightarrow \max \mathrm{P}\Big(\frac{r_P - \mu_P}{\sigma_P} \geq \frac{-\mu_P}{\sigma_P}\Big) \\ &\Leftrightarrow \max \mathrm{P}\Big(Z \geq \frac{-\mu_P}{\sigma_P}\Big) \text{ where } Z \sim \mathrm{N}(0,1) \\ &\Leftrightarrow \min\Big(\frac{-\mu_P}{\sigma_P}\Big) \\ &\Leftrightarrow \max\Big(\frac{\mu_P}{\sigma_P}\Big) \\ &= \max(\text{Sharpe ratio}).\end{aligned}$$

Portfolio Sharpe ratio [1e-3]

	S&P500	EW	SP-MV	SP-MC	R-SP-MC	MP-MV	MP-MC	R-MP-MV
1	4.24	38.16	128.00	108.24	<u>165.40</u>	125.80	111.93	123.04
2	17.13	21.21	-9.07	-7.26	-1.16	-7.26	-5.06	-8.12
3	105.48	128.14	163.91	174.04	161.92	161.21	<u>174.70</u>	144.69
4	37.68	40.37	80.80	60.65	<u>94.80</u>	85.26	72.80	89.71
5	38.07	81.92	93.14	85.23	83.89	97.73	88.00	<u>110.35</u>
6	<u>136.12</u>	-79.55	-34.30	-19.74	-77.59	-32.85	-15.48	-35.98
7	<u>117.23</u>	114.57	34.10	16.31	-5.91	32.54	17.56	44.99
8	94.55	122.93	<u>138.60</u>	124.12	105.49	138.31	122.20	111.50
9	-37.89	-52.73	-50.19	-47.47	-42.15	-49.74	-47.96	-42.45
10	93.67	153.72	<u>262.73</u>	237.62	177.21	261.06	230	188.12
11	174.25	151.19	149.19	118.54	145.19	150.58	118.75	<u>177.71</u>
12	143.13	181.44	115.77	122.24	<u>182.04</u>	115.22	124.36	176.97
13	10.29	30.99	23.00	21.04	8.22	21.67	21.31	<u>39.05</u>
14	10.29	-16.53	-15.45	-9.58	17.62	-10.80	-6.03	-9.31
15	93.59	-13.88	-25.42	-20.13	-77.56	-26.51	-23.34	-18.26
16	-31.92	-48.50	0.56	20.20	14.59	-3.44	9.72	-20.81
17	<u>80.16</u>	78.03	-21.16	-27.34	43.15	-21.82	-27.91	21.30
18	-21.27	-17.13	-23.21	-26.17	-47.77	-21.38	-24.13	-28.47
19	3.20	-42.01	20.86	32.38	23.43	21.87	23.99	3.87
20	-112.5	-130.0	-16.31	-24.46	-15.77	-16.55	-27.35	-61.68

Table 3: The portfolio Sharpe ratio for each model per experiment.

The values $\mathrm{P}\Big(Z \geq \frac{-\mu_P}{\sigma_P}\Big)$ for this alternative Sharpe ratio calculation are in Table 4. Looking at Table 3, experiments 1, 3, 4, 5, 7, 9, 10, 11 and 12 have positive Sharpe ratios and the best-performing model can be determined. In addition, in experiments 6 and 7 there are some negative Sharpe ratios, however, the highest Sharpe ratio model has both highest expected return and lowest variance and there is no need for further investigation. Experiments 2, 9, 14, 15, 16, 18, 19 and 20 have some negative Sharpe ratios and the best-performing model is determined by the alternative metric in Table 4. Finally, Portfolio Alternative Sharpe Ratio [1e-2]

	S&P500	EW	SP-MV	SP-MC	R-SP-MC	MP-MV	MP-MC	R-MP-MV
1	50.17	51.52	55.09	54.31	56.57	55.00	54.46	54.90
2	50.68	50.85	49.64	49.71	49.95	49.71	49.80	49.68
3	54.20	55.09	56.51	56.91	56.43	56.41	56.94	55.75
4	51.50	51.61	53.22	52.42	53.78	53.40	52.90	53.57
5	51.52	53.27	53.71	53.40	53.34	53.89	53.51	54.40
6	55.42	46.83	48.63	49.21	46.91	48.69	49.38	48.56
7	54.67	54.56	51.36	50.65	49.77	51.30	50.70	51.79
8	53.77	54.89	55.51	54.94	54.21	55.50	54.86	54.43
9	48.49	47.90	48.00	48.11	48.32	48.02	48.09	48.31
10	53.73	56.11	60.35	59.38	57.03	60.29	59.11	57.46
11	56.92	56.01	55.93	54.72	55.77	55.98	54.73	57.06
12	55.69	57.20	54.61	54.87	57.22	54.58	54.95	57.02
13	50.41	51.24	50.92	50.84	50.33	50.87	50.85	51.56
14	50.41	49.34	49.39	49.62	50.70	49.57	49.76	49.63
15	53.73	49.44	48.99	49.20	46.91	48.94	49.07	49.27
16	48.73	48.07	50.02	50.80	50.58	49.86	50.39	49.17
17	53.19	53.11	49.16	48.91	51.72	49.13	48.89	50.85
18	49.15	49.31	49.07	48.96	48.10	49.15	49.04	48.86
19	50.13	48.33	50.83	51.29	50.93	50.87	50.96	50.15
20	45.52	44.79	49.35	49.02	49.37	49.34	48.91	47.54

Table 4: The portfolio Sharpe ratio alternatives for each model per experiment.

a bar graph illustrating the distribution of maximum Sharpe ratio achieved by each model is provided in Figure 12.

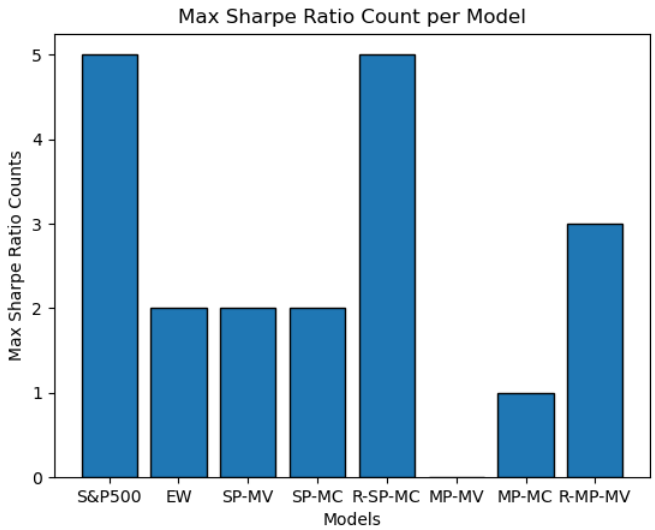


Figure 12: Bar graph with the counts of best-performing models for all experiments.

Based on the analysis of the bar graph in Figure 12, several key observations can be made. Firstly, the benchmark S&P500 was outperformed by at least one of the tested models in the majority of cases, indicating that models can actually perform in a similar or higher level as the benchmark. Additionally, the single-period models revealed relatively stronger performance, appearing as the best-performing models in 9 out of the 20 experiments, while the multi-period models did so only in 4 experiments. Moreover, the robust approaches demonstrated better performance compared to the nominal cases, highlighting the importance of considering robustness in PO problem . Remarkably, the robust single-period Mean-CVaR model stood out as the overall top-performing model, suggesting its effectiveness in achieving desirable risk-return trade-offs.

6 CONCLUSIONS AND DISCUSSION

In this chapter, we present the conclusions drawn from our research on MV, Mean-CVaR, and robust PO problems in both single-period and multi-period contexts. We then discuss the implications of our findings while we provide recommendations for decision-makers who want to use the models in practice. Finally, we discuss the limitations of our study and consider future research directions.

6.1 SUMMARY OF FINDINGS

We begin by summarizing the main findings from our research. We demonstrated the equivalence of PO problem formulations for a desired level of expected return, risk, and risk parameter. In addition, we provided a clear illustration of the efficient frontiers of the PO problems while also exploring the impact of incorporating risk-free assets and no-short constraints. By examining these variations, we gained valuable insights into the behavior of the efficient frontiers under different scenarios.

In the multi-period setting, we extended single-period PO problems by incorporating transaction costs without using any extra variable. We also formulated a time-consistent PO problem that includes transaction costs, making the time-consistent approach more realistic. Transaction costs were included in the robust optimization approach as well. We found out that the transaction cost models can work well and provide more realistic results.

Through a series of numerical experiments, we evaluated the performance of these models and compare their effectiveness in achieving optimal portfolio allocations in real data. Our experiments were designed to provide an unbiased assessment of each model's performance by incorporating random selection of some of the input parameters. Specifically, we randomly selected stocks from a predefined set while also the selection of the evaluation period was randomly chosen. We implemented all the models on market data, and then assessed their performance by calculating the Sharpe ratio in each experiment. In this way, we found that PO models perform well compared to market standards set by S&P500 as at least one of the models outperformed S&P500 most of the times. Single-period models outperformed multi-period models by having the highest Sharpe ratio 9 out of the 20 experiments in contrast to the 4 counts by all multi-period models. Moreover, robust models seemed to surpass the nominal models since they typically had better results. Overall, robust single-period Mean-CVaR PO problem had the best performance compered to the other PO problems.

6.2 IMPLICATIONS AND PRACTICAL RECOMMENDATIONS

Our research has yielded significant implications for PO practices. Our numerical experiments demonstrate that PO models, exhibit great performance across various scenarios competing the S&P500 index. As a result, practitioners should have confidence in utilizing these models to optimize their investment portfolios. Furthermore, the incorporation of robust optimization techniques has proven to be beneficial in the numerical experiments. Robust approaches consider uncertainties and extreme scenarios, resulting in portfolios that are better protected against potential market disruptions. Practitioners are encouraged to favor robust optimization models over traditional nominal models to protect their portfolios against uncertainties and achieve more stable performance.

6.3 LIMITATIONS AND FUTURE RESEARCH

Although our research has provided valuable insights into portfolio optimization, it is essential to recognize its limitations and consider potential directions for future research. To begin with, the models developed in our research are designed to construct the efficient frontier for any given desired expected return (or for any other varying parameter). However, an alternative approach could focus on determining a single portfolio with the highest Sharpe ratio. This alternative formulation would eliminate the need for additional parameters and align more closely with the performance assessment conducted in our numerical experiments.

Secondly, the numerical experiments conducted in our research were limited to a total of 20 experiments. While this number allowed for the presentation of results in manageable and readable tables, increasing the number of experiments could provide a more robust analysis and reduce the potential noise resulting from randomness. Future research could consider expanding the number of experiments to capture a wider range of scenarios and increase the statistical significance of the findings.

Lastly, the scope of the robust approach in our research was limited to one worst-case approach for both single-period and multi-period models. The former involved a distribution uncertainty set using different copulas while the latter involved an uncertainty ball constructed by the Wasserstein metric. However, there are various other robust optimization techniques that could be explored and compared. For instance, alternative distributions or metrics could be used to construct the uncertainty set or there could also be the inclusion of parameter-based uncertainty sets, such as box or ellipsoidal uncertainty sets. Investigating and comparing these additional approaches would provide a more comprehensive analysis of robust optimization techniques.

A

PROOFS

In this appendix chapter, detailed proofs of the theorems and propositions presented in MV PO problem are provided.

A.1 OPTIMIZATION PROBLEMS EQUIVALENCY

In this section, the equivalency between the following 3 problems will be shown:

Problem (A1)	Problem (A2)	Problem (A3)
$\min_{\boldsymbol{w}} \boldsymbol{w}^\top \Sigma \boldsymbol{w}$	$\max_{\boldsymbol{w}} \boldsymbol{w}^\top \boldsymbol{\mu}$	$\min_{\boldsymbol{w}} \boldsymbol{w}^\top \boldsymbol{\mu} - \omega \cdot \boldsymbol{w}^\top \Sigma \boldsymbol{w}$
s.t. $\boldsymbol{w}^\top \boldsymbol{\mu} = \mu_0$,	s.t. $\boldsymbol{w}^\top \Sigma \boldsymbol{w} = \sigma_0^2$,	s.t. $\boldsymbol{w}^\top \mathbf{1} = 1$.
$\boldsymbol{w}^\top \mathbf{1} = 1$.	$\boldsymbol{w}^\top \mathbf{1} = 1$.	

A.1.1 *Solution to Problem (A1)*

To solve the minimization problem (A1) for a given μ_0, the methodology as in chapter 4 of Focardi, Fabozzi, et al. (2004) is used, which originated from Merton (1972). This methodology uses Lagrange multipliers to convert the constrained optimization problem into an unconstrained one as in chapter 7 of Ingersoll (1987). The Lagrangian function for this problem is:

$$L(w,\lambda,\gamma) = \boldsymbol{w}^\top \Sigma \boldsymbol{w} - \lambda(\boldsymbol{w}^\top \boldsymbol{\mu} - \mu_0) - \gamma(\boldsymbol{w}^\top \mathbf{1} - 1),$$

where λ and γ are the Lagrange multipliers associated with the constraints $\boldsymbol{w}^\top \boldsymbol{\mu} = \mu_0$ and $\boldsymbol{w}^\top \mathbf{1} = 1$, respectively.
To find the solution, one needs to solve the system of equations:

$$\nabla_w L = 2\Sigma \boldsymbol{w} - \lambda \boldsymbol{\mu} - \gamma \mathbf{1} = 0,$$

$$\boldsymbol{w}^\top \boldsymbol{\mu} = \mu_0,$$

$$\boldsymbol{w}^\top \mathbf{1} = 1.$$

Taking the derivative of L with respect to $\boldsymbol{w}$ and setting it equal to zero:

$$2\Sigma \boldsymbol{w}^* - \lambda \boldsymbol{\mu} - \gamma \mathbf{1} = 0.$$

Solving for $\boldsymbol{w}^*$:

$$\boldsymbol{w}^* = \frac{\lambda \Sigma^{-1} \boldsymbol{\mu} + \gamma \Sigma^{-1} \mathbf{1}}{2}. \tag{28}$$

Substituting this expression for $\boldsymbol{w}$ into the constraints $\boldsymbol{w}^\top\boldsymbol{\mu} = \mu_0$ and $\boldsymbol{w}^\top\mathbf{1} = 1$, we obtain two solutions for the Lagrange multipliers:

$$\lambda = 2\frac{\mu_0 A - B}{\Delta},$$

$$\gamma = 2\frac{C - \mu_0 B}{\Delta},$$

where $A = \mathbf{1}^\top\Sigma\mathbf{1} > 0$, $B = \mathbf{1}^\top\Sigma\boldsymbol{\mu} > 0$, $C = \boldsymbol{\mu}^\top\Sigma\boldsymbol{\mu} > 0$ and $\Delta = AC - B^2$ with $\Delta > 0$ by the Cauchy-Schwarz inequality.

Therefore, substituting λ and γ in $\boldsymbol{w}^*$, the optimal variance is:

$$\boldsymbol{w}^{*\top}\Sigma\boldsymbol{w}^* = \frac{A\mu_0^2 - 2B\mu_0 + C}{\Delta}. \tag{29}$$

A.1.2 *Solution to Problem (A2)*

To solve the maximization problem (A2) for a given σ_0, again, we can use Lagrange multipliers to convert the constrained optimization problem into an unconstrained one. The Lagrangian function for this problem is:

$$L(w, \nu, \kappa) = \boldsymbol{w}^\top\boldsymbol{\mu} - \nu(\boldsymbol{w}^\top\Sigma\boldsymbol{w} - \sigma_0^2) - \kappa(\boldsymbol{w}^\top\mathbf{1} - 1),$$

where ν and κ are the Lagrange multipliers associated with the constraints $\boldsymbol{w}^\top\Sigma\boldsymbol{w} = \sigma_0^2$ and $\boldsymbol{w}^\top\mathbf{1} = 1$, respectively.
To find the solution, we need to solve the system of equations:

$$\nabla_w L = \boldsymbol{\mu} - \nu 2\Sigma\boldsymbol{w} - \kappa\mathbf{1} = 0,$$

$$\boldsymbol{w}^\top\Sigma\boldsymbol{w} = \sigma_0^2,$$

$$\boldsymbol{w}^\top\mathbf{1} = 1.$$

Taking the derivative of L with respect to $\boldsymbol{w}$ and setting it equal to zero:

$$\boldsymbol{\mu} - \nu 2\Sigma\boldsymbol{w}^* - \kappa\mathbf{1} = 0.$$

Solving for $\boldsymbol{w}^*$:

$$\boldsymbol{w}^* = \frac{\Sigma^{-1}\boldsymbol{\mu} - \kappa\Sigma^{-1}\mathbf{1}}{2\nu}. \tag{30}$$

Substituting this expression for $\boldsymbol{w}$ into the constraints $\boldsymbol{w}^\top\Sigma\boldsymbol{w} = \sigma_0^2$ and $\boldsymbol{w}^\top\mathbf{1} = 1$, we obtain two solutions for the Lagrange multipliers:

$$\nu = \frac{1}{2}\sqrt{\frac{\Delta}{\sigma_0^2 A - 1}},$$

$$\kappa = \frac{B - \sqrt{\frac{\Delta}{\sigma_0^2 A - 1}}}{A},$$

for $\sigma_0^2 A - 1 > 0$.

Therefore, substituting ν and κ in $\boldsymbol{w}^*$, the optimal mean is:

$$\boldsymbol{w}^{*\top}\boldsymbol{\mu} = \frac{\sqrt{\Delta(\sigma_0^2 A - 1)} + B}{A}. \tag{31}$$

Setting (31) equal to μ and solving to σ_0^2, we derive the following variance:

$$\sigma_0^2 = \frac{A\mu^2 - 2B\mu + C}{\Delta},$$

which is the exact same variance as the one in equation (29) for $\mu = \mu_0$ and so the optimal $\boldsymbol{w}^*$ are also the same.

As a result, problems (A1) and (A2) are equivalent for $\sigma_0^2 = \dfrac{A\mu_0^2 - 2B\mu_0 + C}{\Delta}$.

A.1.3 *Solution to Problem (A3)*

To solve the maximization problem (A3) for a given ω, we again use Lagrange multiplier. The Lagrangian function for this problem is:

$$L(w,\phi) = \boldsymbol{w}^\top\boldsymbol{\mu} - \omega\boldsymbol{w}^\top\Sigma\boldsymbol{w} - \phi(\boldsymbol{w}^\top\mathbf{1} - 1),$$

where ϕ is the Lagrange multiplier associated with the constraint $\boldsymbol{w}^\top\mathbf{1} = 1$. To find the solution, the system of equations must be solved:

$$\nabla_w L = \boldsymbol{\mu} - \omega \cdot 2\Sigma\boldsymbol{w} - \phi\mathbf{1} = 0,$$

$$\boldsymbol{w}^\top\mathbf{1} = 1.$$

Taking the derivative of L with respect to $\boldsymbol{w}$ and setting it equal to zero, we obtain:

$$\boldsymbol{\mu} - \omega \cdot 2\Sigma\boldsymbol{w}^* - \phi\mathbf{1} = 0.$$

Solving for $\boldsymbol{w}^*$:

$$\boldsymbol{w}^* = \frac{\Sigma^{-1}\boldsymbol{\mu} - \phi\Sigma^{-1}\mathbf{1}}{2\omega}. \tag{32}$$

Substituting this expression for $\boldsymbol{w}$ into the constraint $\boldsymbol{w}^\top\mathbf{1} = 1$, we obtain the solution for the Lagrange multiplier:

$$\phi = \frac{B - 2\omega}{A},$$

where $A = \mathbf{1}^\top\Sigma\mathbf{1}$ and $B = \mathbf{1}^\top\Sigma\boldsymbol{\mu}$.

Therefore, substituting ϕ in $\boldsymbol{w}^*$, the optimal mean is:

$$\boldsymbol{w}^{*\top}\boldsymbol{\mu} = \frac{\Delta + 2\omega B}{2\omega A}, \tag{33}$$

while the optimal variance given $\boldsymbol{w}^*$ is:

$$\boldsymbol{w}^{*\top}\Sigma\boldsymbol{w}^* = \frac{\Delta + 4\omega^2}{4\omega^2 A}. \tag{34}$$

So first, we will choose an ω, such that the mean (33) is equal to μ_0 (as in problem (A1)) and then look at what the variance will be:

$$\frac{\Delta + 2\omega B}{2\omega A} = \mu_0 \Rightarrow \omega = \frac{\Delta}{2A\mu_0 - 2B}.$$

And so, substituting this ω in the variance from equation (34):

$$\boldsymbol{w}^{*\top}\Sigma\boldsymbol{w}^* = \frac{\Delta + 4\omega^2}{4\omega^2 A} = \frac{A\mu_0^2 - 2B\mu_0 + C}{\Delta},$$

which is the exactly the same variance as in equation (29) and so the optimal $\boldsymbol{w}^*$ are also the same. Thus, problems (A1) and (A3) are equivalent for $\omega = \dfrac{\Delta}{2A\mu_0 - 2B}$.

We will now choose an ω, such that the variance (34) is equal to σ_0^2 (as in problem (A2)) and then look at what the mean will be:

$$\frac{\Delta + 4\omega^2}{4\omega^2 A} = \sigma_0^2 \Rightarrow \omega = \sqrt{\frac{\Delta}{4(\sigma_0^2 A - 1)}}.$$

And so, substituting this ω in the mean from equation (33):

$$\boldsymbol{w}^{*\top}\boldsymbol{\mu} = \frac{\Delta + 2\omega B}{2\omega A} = \frac{\sqrt{\Delta(\sigma_0^2 A - 1)} + B}{A},$$

which is the exactly the same mean as in equation (31) and so the optimal $\boldsymbol{w}^*$ are the same as well and so problems (A2) and (A3) are equivalent for $\omega = \sqrt{\dfrac{\Delta}{4(\sigma_0^2 A - 1)}}$.

To sum up, in the following table 5, the optimal allocation $\boldsymbol{w}^*$, expected return and variance for each problem is shown given their parameters $(\mu_0, \sigma_0^2, \omega)$ and for $A = \boldsymbol{I}^\top\Sigma\boldsymbol{I}$, $B = \boldsymbol{I}^\top\Sigma\boldsymbol{\mu}$, $C = \boldsymbol{\mu}^\top\Sigma\boldsymbol{\mu}$ and $\Delta = AC - B^2$:

Problem (A1)	Problem (A2)	Problem (A3)
$\frac{\mu_0 A - B}{\Delta}\Sigma^{-1}\boldsymbol{\mu} + \frac{C - \mu_0 B}{\Delta}\Sigma^{-1}\mathbf{1}$	$\dfrac{\Sigma^{-1}\boldsymbol{\mu} - \frac{B - \sqrt{\frac{\Delta}{\sigma_0^2 A - 1}}}{A}\Sigma^{-1}\mathbf{1}}{\sqrt{\frac{\Delta}{\sigma_0^2 A - 1}}}$	$\dfrac{\Sigma^{-1}\boldsymbol{\mu} + \frac{B - 2\omega}{A}\Sigma^{-1}\mathbf{1}}{2\omega}$
μ_0	$\dfrac{\sqrt{\Delta(\sigma_0^2 A - 1)} + B}{A}$	$\dfrac{\Delta + 2\omega B}{2\omega A}$
$\dfrac{A\mu_0^2 - 2B\mu_0 + C}{\Delta}$	σ_0^2	$\dfrac{\Delta + 4\omega^2}{4\omega^2 A}$

Table 5: Summary of optimal allocation $\boldsymbol{w}^*$, expected return and variance of each problem.

A.2 MEAN-VARIANCE OPTIMIZATION PROBLEM SOLUTION WITH A RISK-FREE ASSET

In this section the analytical solution of (4) will be provided. Using the second constrain in the first one, the problem becomes:

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^\top \Sigma \boldsymbol{w} \\ \text{s.t.} \quad & \boldsymbol{w}^\top \boldsymbol{\mu} + (1 - \boldsymbol{w}^\top \mathbf{1})R = \mu_0. \end{aligned} \tag{35}$$

To make computations easier and the results clearer, let $\boldsymbol{\mu}_D = \boldsymbol{\mu} - \mathbf{1}R$. In practice, this is the discounted return given the rate R for the single period time horizon. So, problem (35) becomes:

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \boldsymbol{w}^\top \Sigma \boldsymbol{w} \\ \text{s.t.} \quad & \boldsymbol{w}^\top \boldsymbol{\mu}_D = \mu_0 - R. \end{aligned} \tag{36}$$

To solve this minimization problem for a given μ_0, we will use Lagrange multipliers as before to convert the constrained optimization problem into an unconstrained one. The Lagrangian function for this problem is:

$$L(w,\lambda) = \boldsymbol{w}^\top \Sigma \boldsymbol{w} - \lambda(\boldsymbol{w}^\top \boldsymbol{\mu}_D - \mu_0 + R),$$

where λ is the Lagrange multiplier associated with the constraint $\boldsymbol{w}^\top \boldsymbol{\mu}_D = \mu_0 - R$.
To find the solution, we need to solve the system of equations:

$$\nabla_w L = \boldsymbol{w}^\top \Sigma \boldsymbol{w} - \lambda(\boldsymbol{w}^\top \boldsymbol{\mu}_D - \mu_0 + R) = 0,$$

$$\boldsymbol{w}^\top \boldsymbol{\mu}_D = \mu_0 - R.$$

Taking the derivative of L with respect to $\boldsymbol{w}$ and setting it equal to zero:

$$2\Sigma \boldsymbol{w}^* - \lambda \boldsymbol{\mu}_D = 0.$$

Solving for $\boldsymbol{w}^*$:

$$\boldsymbol{w}^* = \frac{\lambda \Sigma^{-1} \boldsymbol{\mu}_D}{2}. \tag{37}$$

Substituting this expression for $\boldsymbol{w}$ into the constraint $\boldsymbol{w}^\top \boldsymbol{\mu}_D = \mu_0 - R$, we obtain solution for the Lagrange multiplier:

$$\lambda = 2\frac{\mu_0 - R}{E},$$

where $E = \boldsymbol{\mu}_D^\top \Sigma^{-1} \boldsymbol{\mu}_D \neq 0$.

Therefore, substituting λ in $\boldsymbol{w}^*$, the optimal solution is:

$$\boldsymbol{w} = \frac{(\mu_0 - R)\Sigma^{-1} \boldsymbol{\mu}_D}{E}, \tag{38}$$

and the optimal variance is:

$$\boldsymbol{w}^{*\top} \Sigma \boldsymbol{w}^* = \frac{(\mu_0 - R)^2}{E}. \tag{39}$$

B

NUMERICAL EXPERIMENTS SUPPLEMENT

B.1 LABELS PER MODEL

In this table, we show the labels used for each model.

Label	Model
EW	Equal Weights
SP-MV	Single-Period Mean-Variance
SP-MC	Single-Period Mean-CVaR
R-SP-MC	Robust Single-Period Mean-CVaR
MP-MV	Multi-Period Mean-Variance
MP-MC	Multi-Period Mean-CVaR
R-MP-MV	Robust Multi-Period Mean-Variance

Table 6: The labels used for each model.

B.2 NUMERICAL EXPERIMENTS PLOTS

In this section, we will provide two plots per numerical experiment. The first plot is the portfolio wealth for each model over the 100-day testing period. The initial wealth is 10000 while the starting period is random. In addition, the number as well as the selection of the stocks is also random. The number of stocks used is displayed in the title of each plot. Next to that, the desired expected return μ_0 used in the models is displayed. The second plot includes 8 histograms, one for each model. These histograms represent the portfolio wealth returns for the testing period. The analysis is based on custom Python code developed specifically for this research, which can be found in code link.

Experiment 1

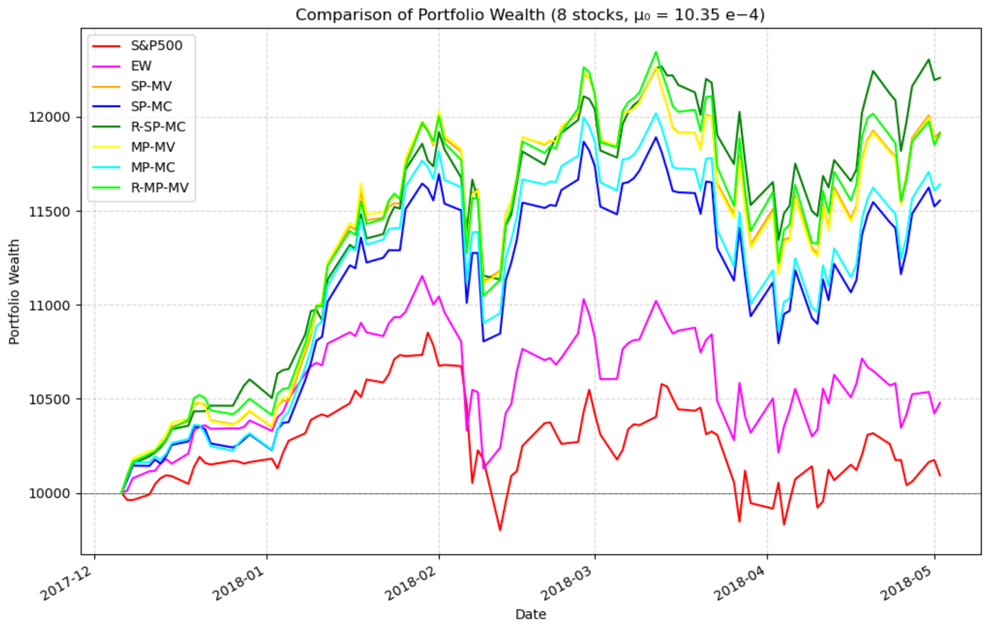


Figure 13: Portfolio wealth over 100-day testing period for experiment 1.

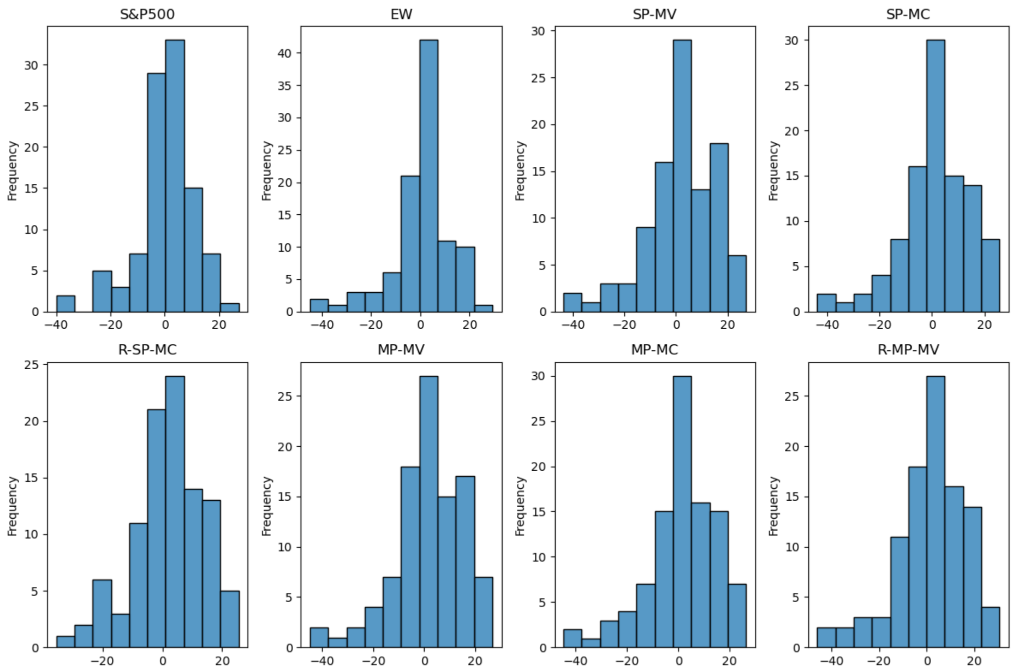


Figure 14: Distribution of portfolio wealth returns for experiment 1.

Experiment 2

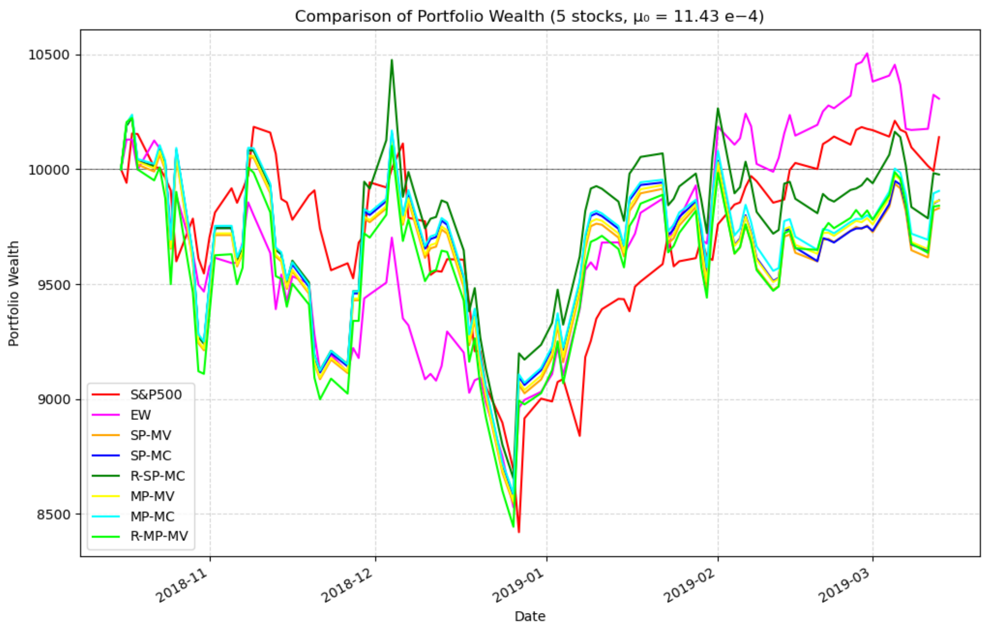


Figure 15: Portfolio wealth over 100-day testing period for experiment 2.

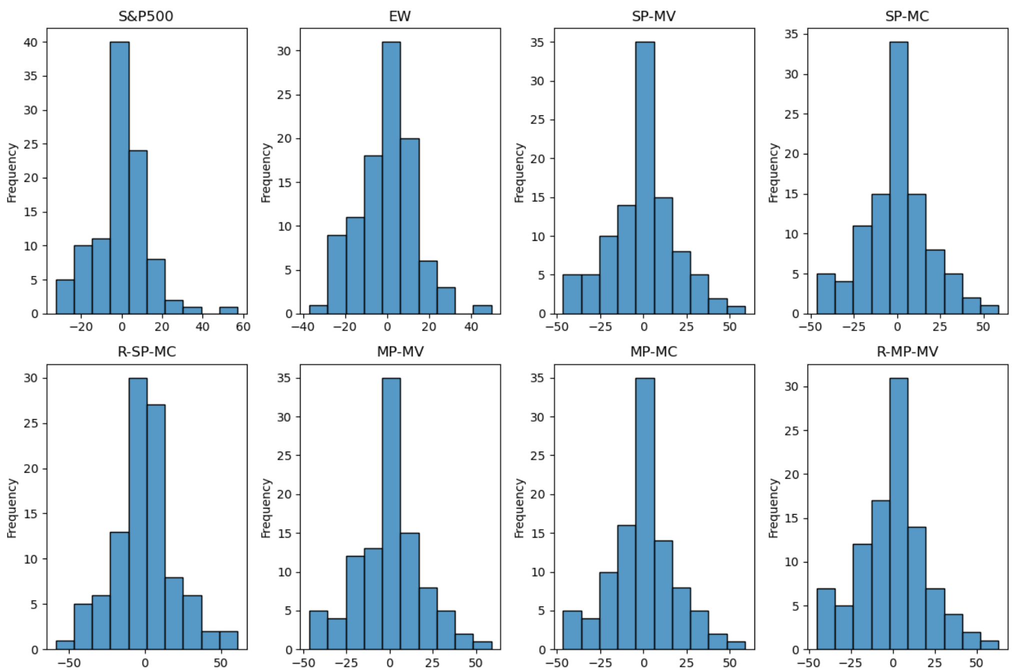


Figure 16: Distribution of portfolio wealth returns for experiment 2.

Experiment 3

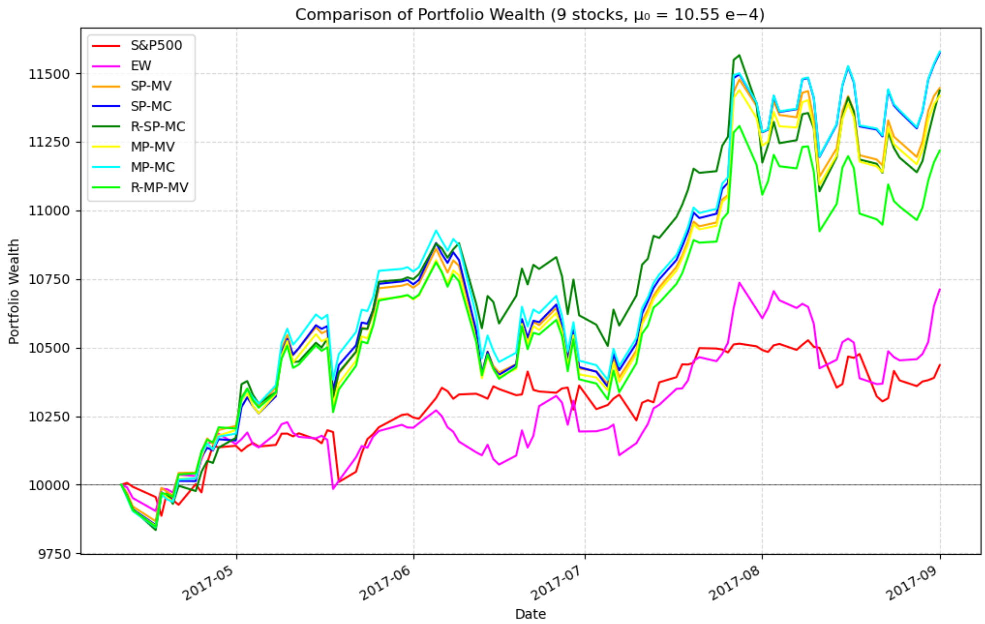


Figure 17: Portfolio wealth over 100-day testing period for experiment 3.

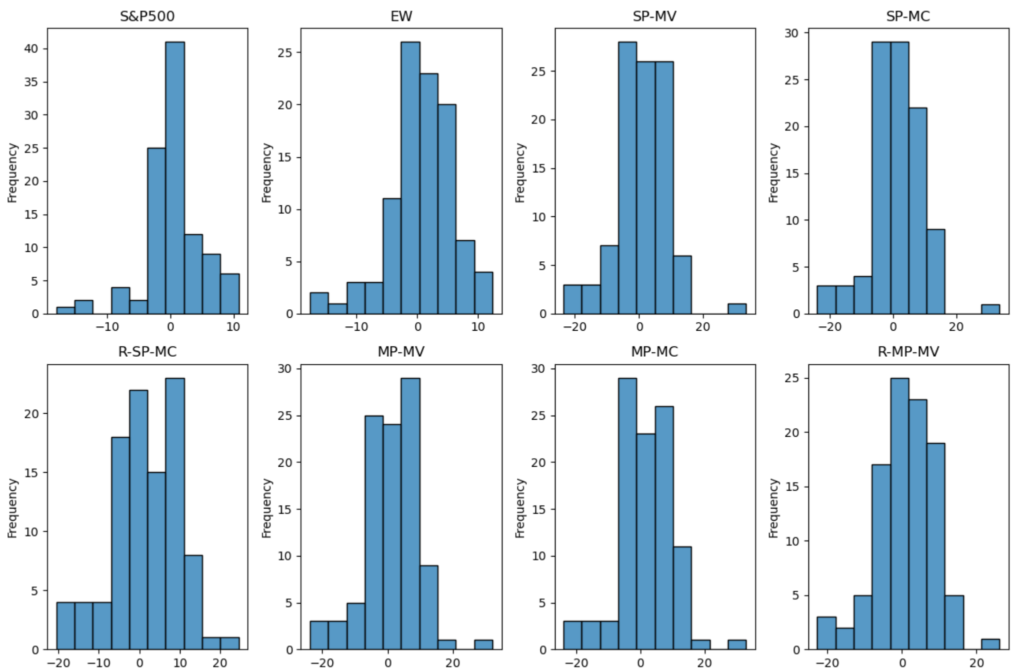


Figure 18: Distribution of portfolio wealth returns for experiment 3.

Experiment 4

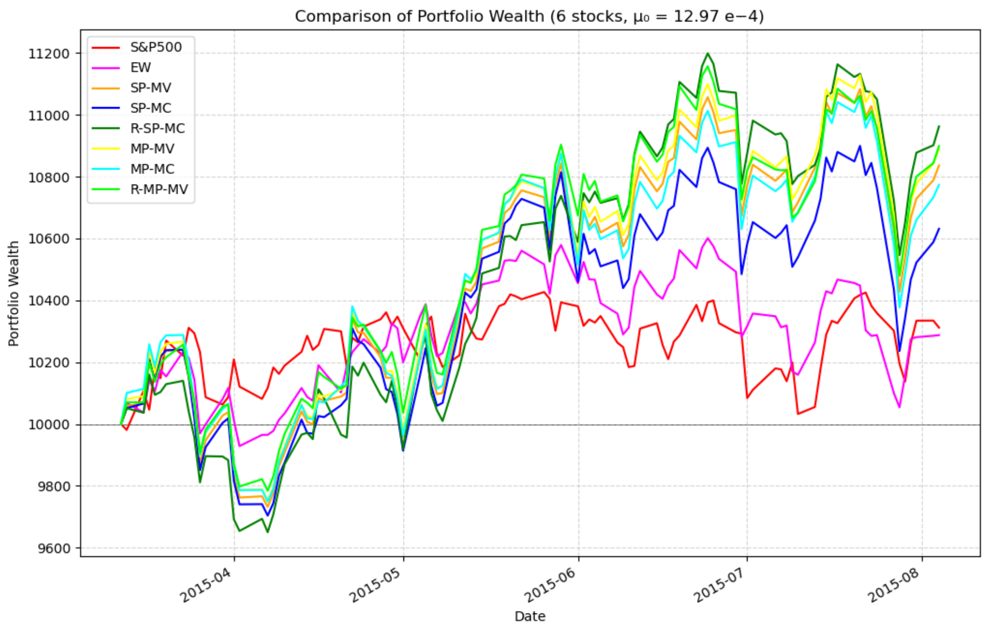


Figure 19: Portfolio wealth over 100-day testing period for experiment 4.

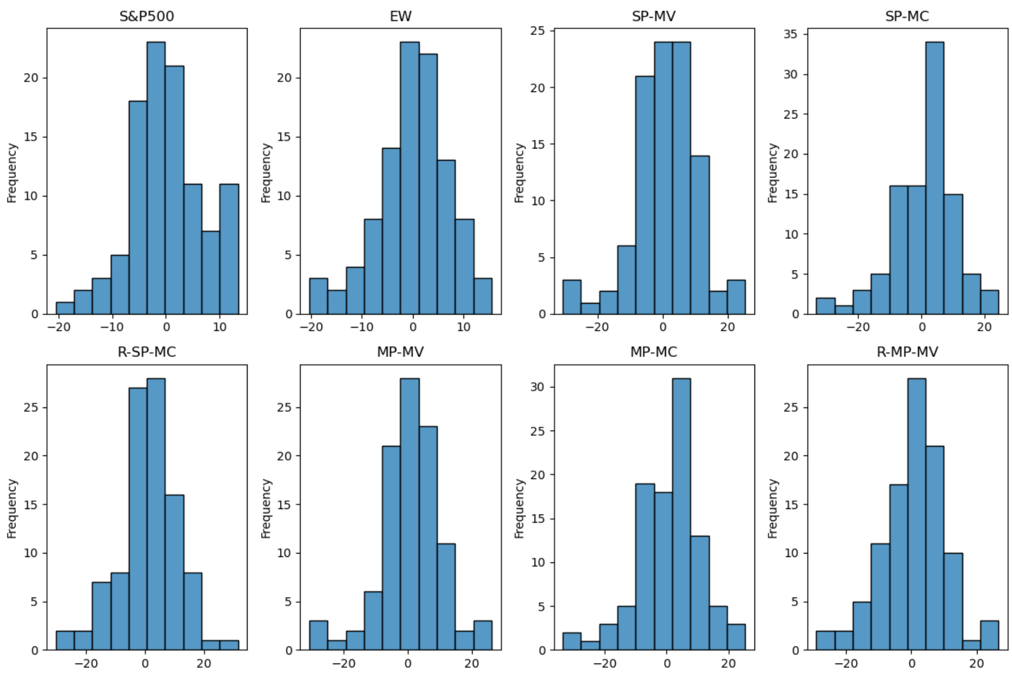


Figure 20: Distribution of portfolio wealth returns for experiment 4.

Experiment 5

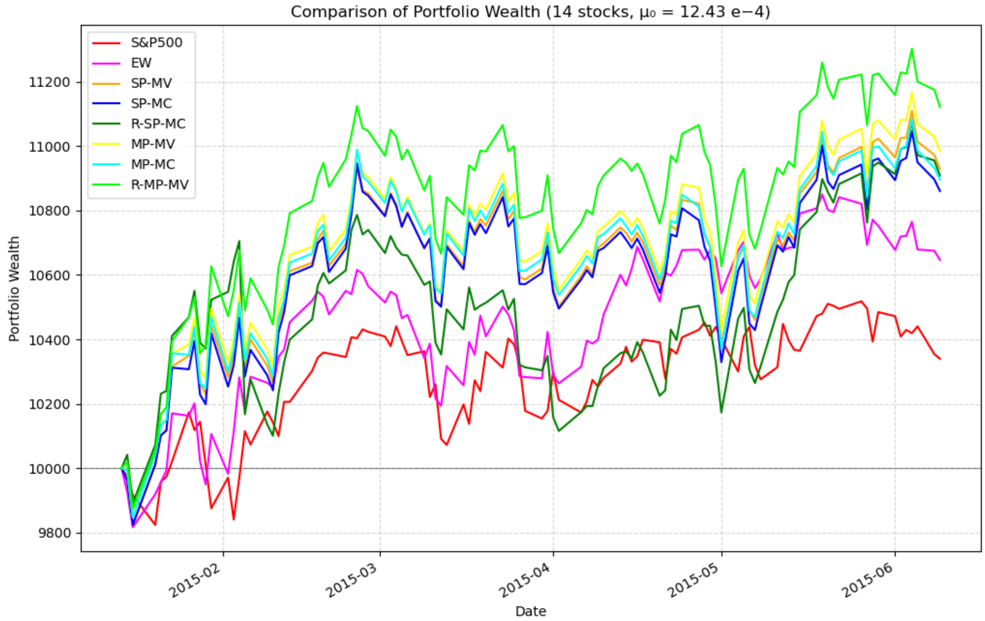


Figure 21: Portfolio wealth over 100-day testing period for experiment 5.

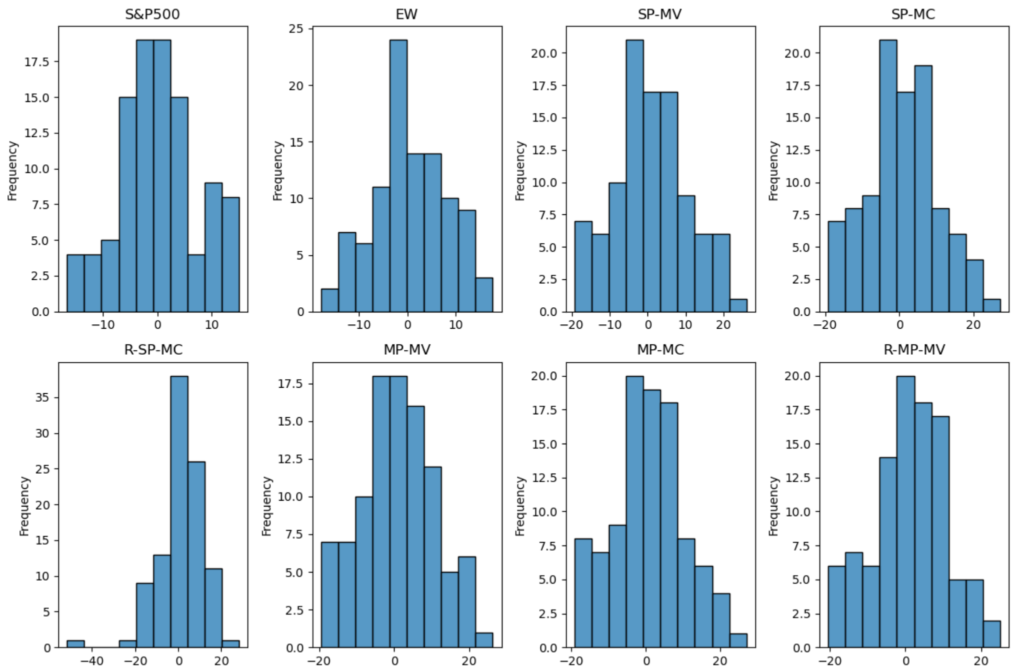


Figure 22: Distribution of portfolio wealth returns for experiment 5.

Experiment 6

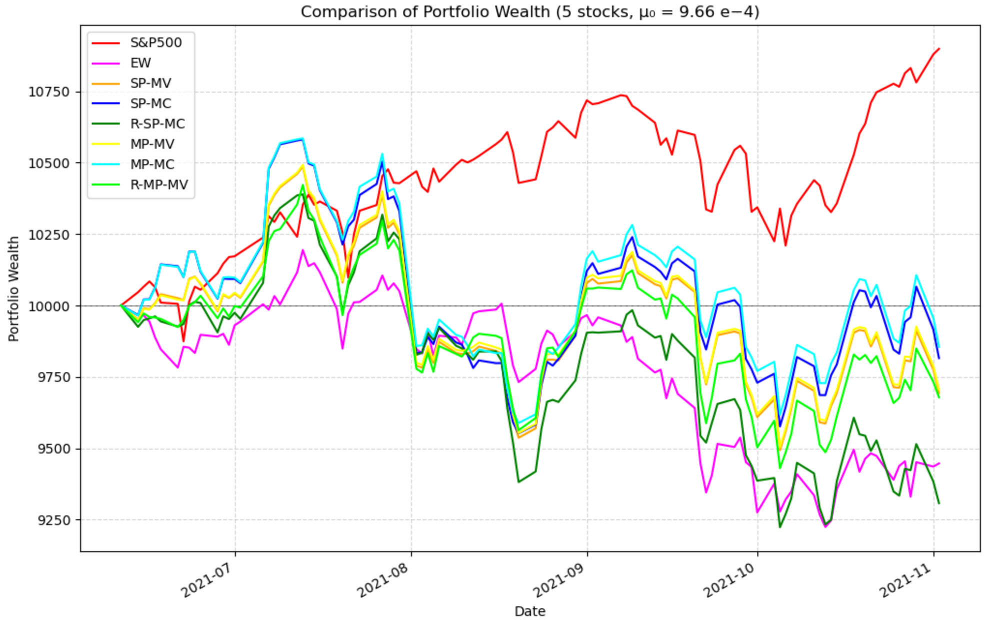


Figure 23: Portfolio wealth over 100-day testing period for experiment 6.

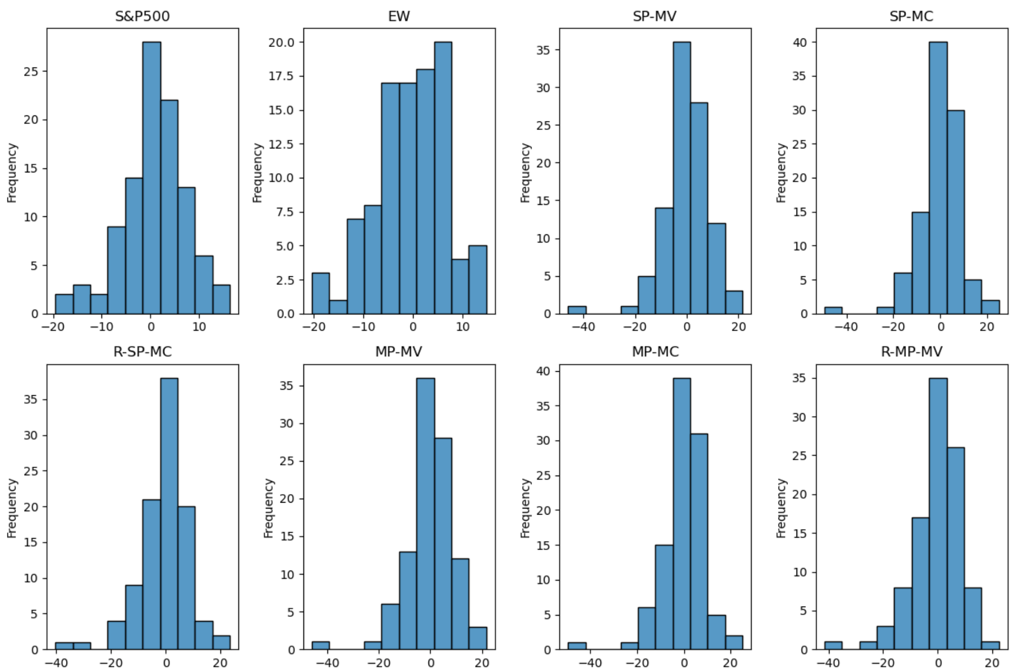


Figure 24: Distribution of portfolio wealth returns for experiment 6.

Experiment 7

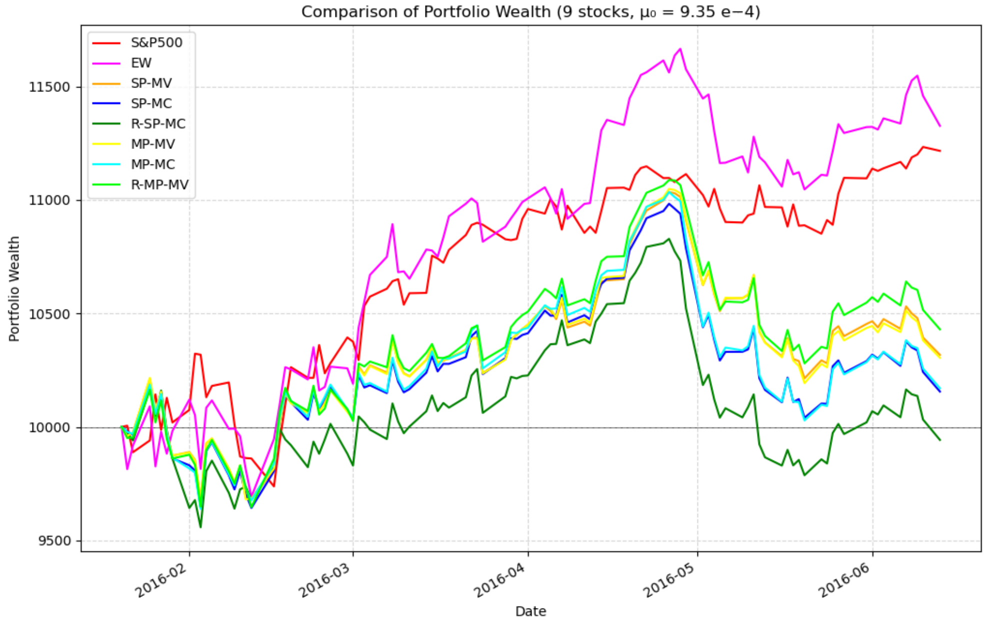


Figure 25: Portfolio wealth over 100-day testing period for experiment 7.

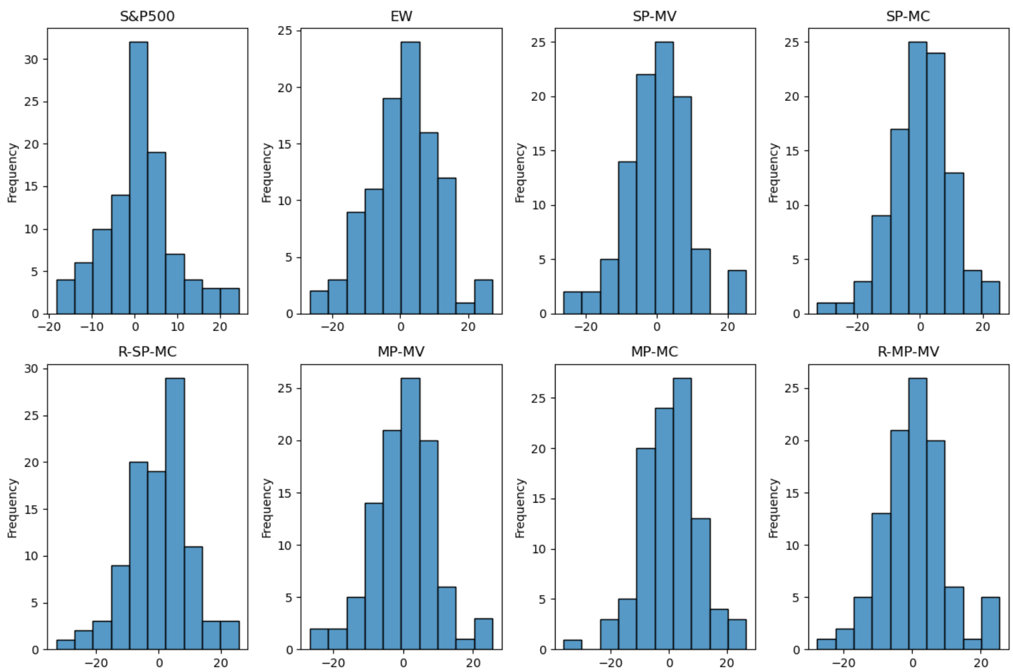


Figure 26: Distribution of portfolio wealth returns for experiment 7.

Experiment 8

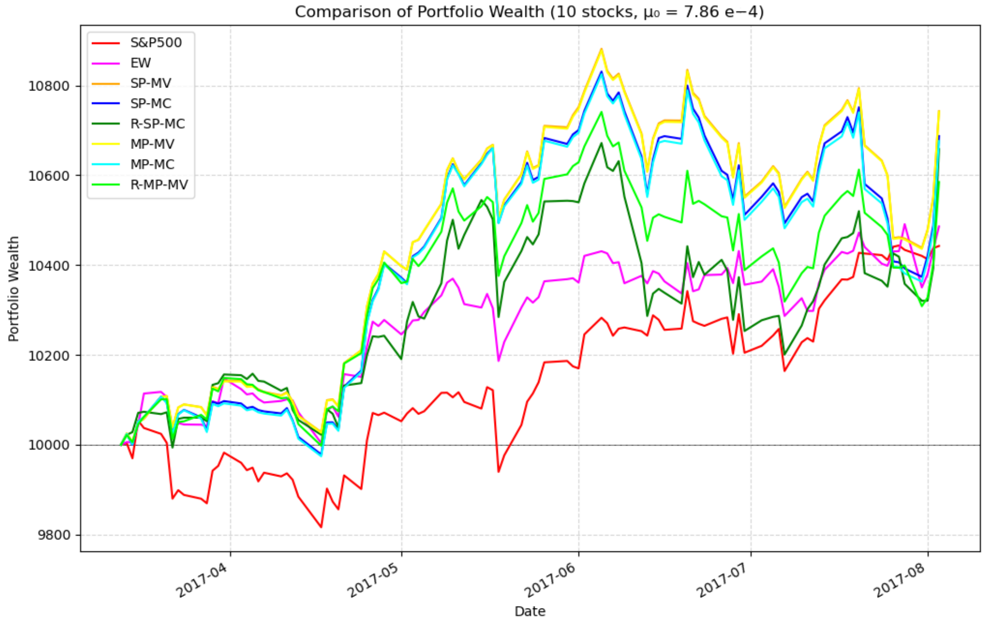


Figure 27: Portfolio wealth over 100-day testing period for experiment 8.

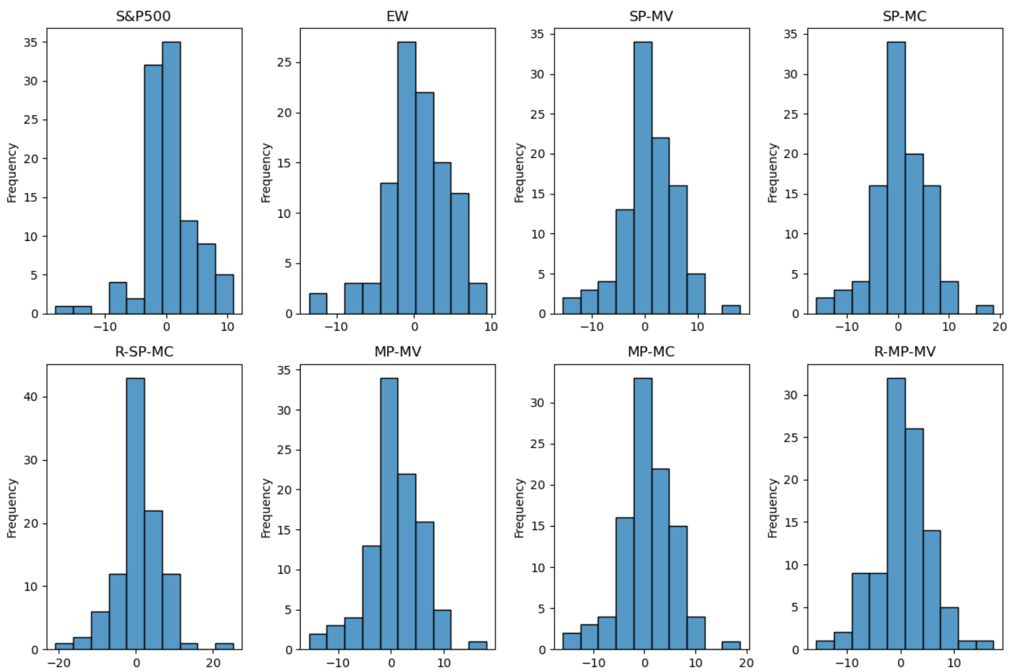


Figure 28: Distribution of portfolio wealth returns for experiment 8.

Experiment 9

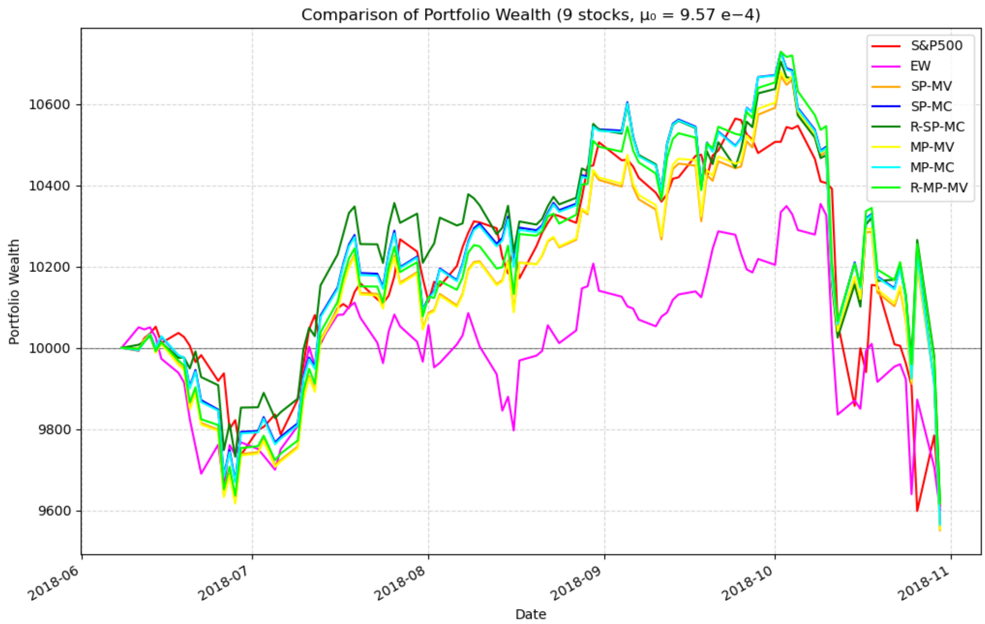


Figure 29: Portfolio wealth over 100-day testing period for experiment 9.

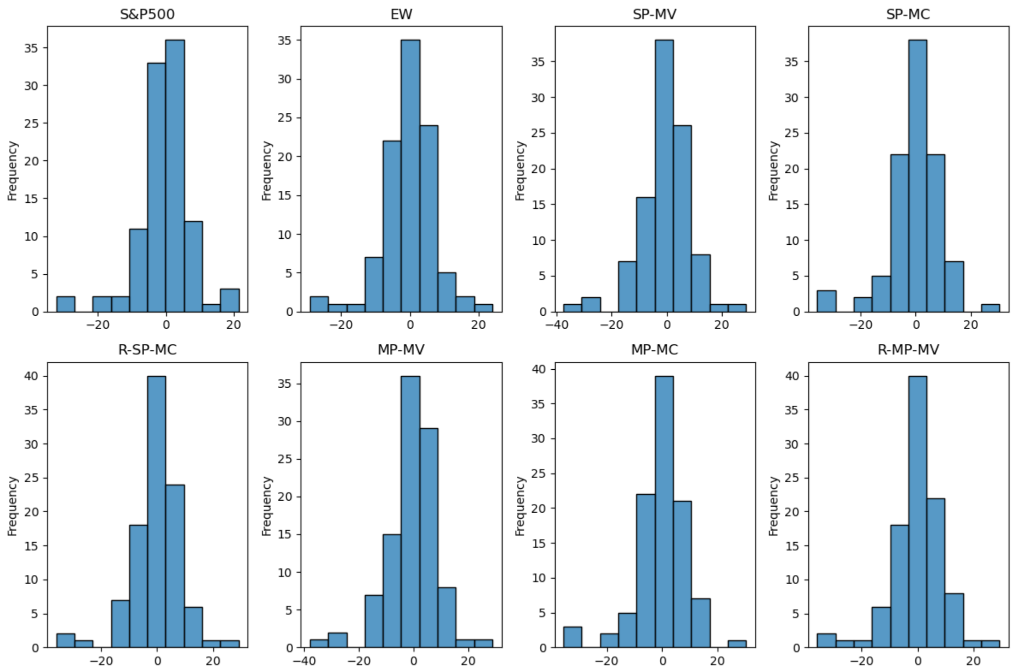


Figure 30: Distribution of portfolio wealth returns for experiment 9.

Experiment 10

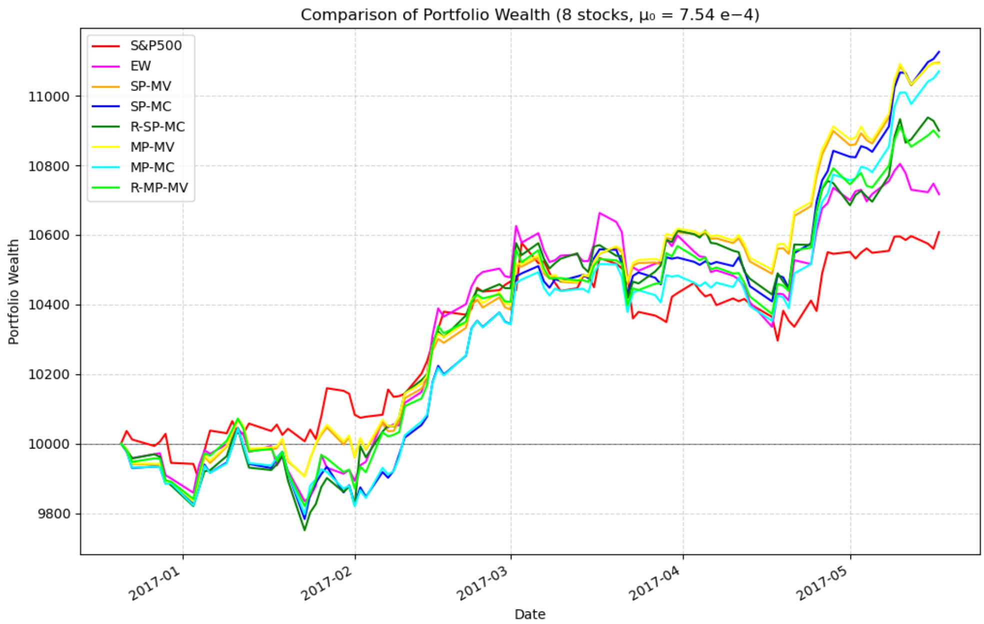


Figure 31: Portfolio wealth over 100-day testing period for experiment 10.

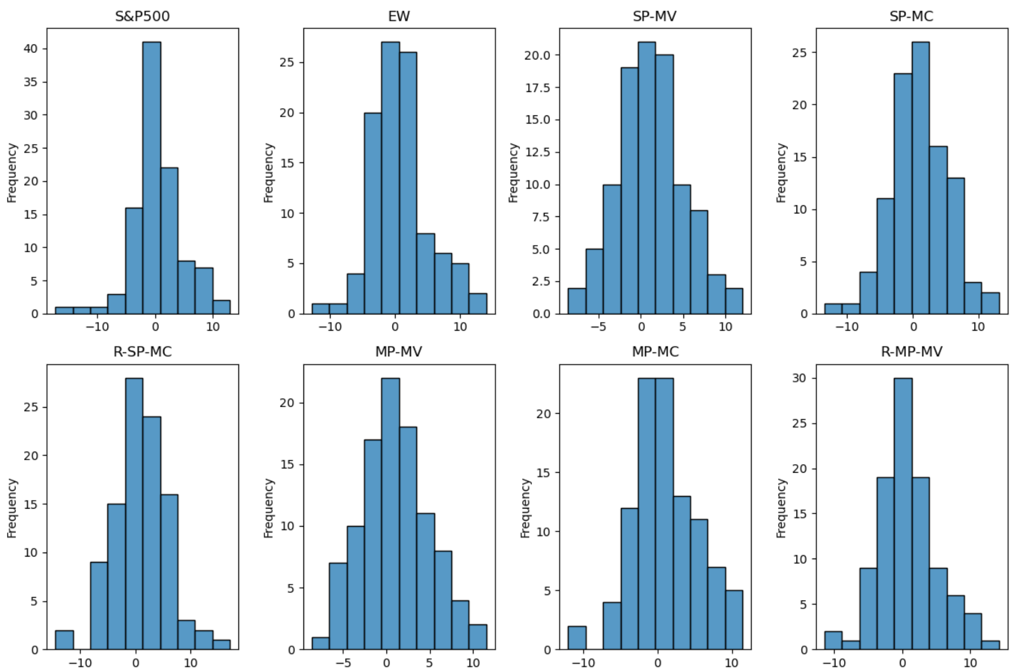


Figure 32: Distribution of portfolio wealth returns for experiment 10.

Experiment 11

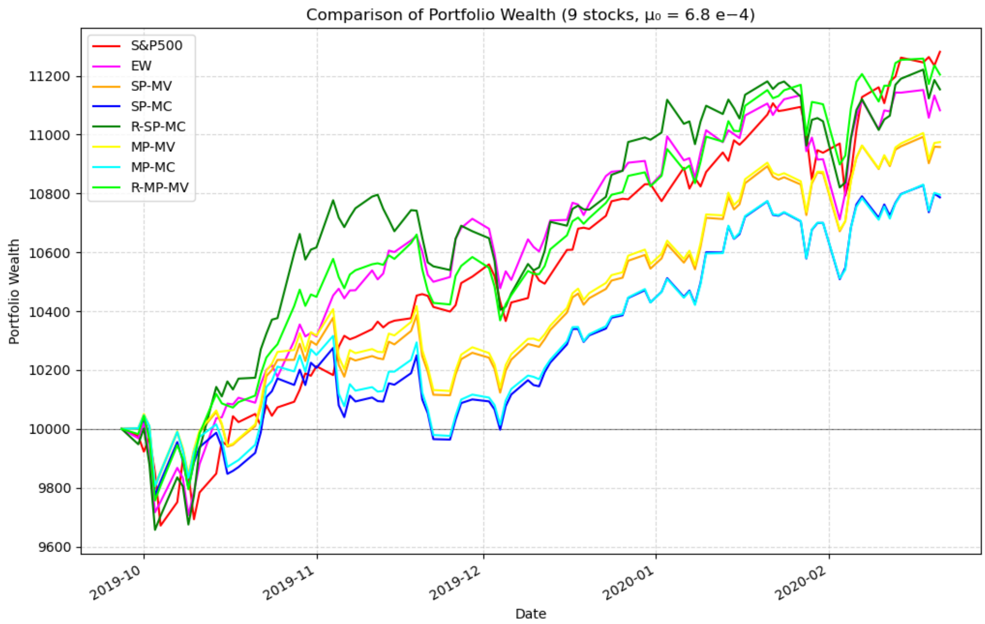


Figure 33: Portfolio wealth over 100-day testing period for experiment 11.

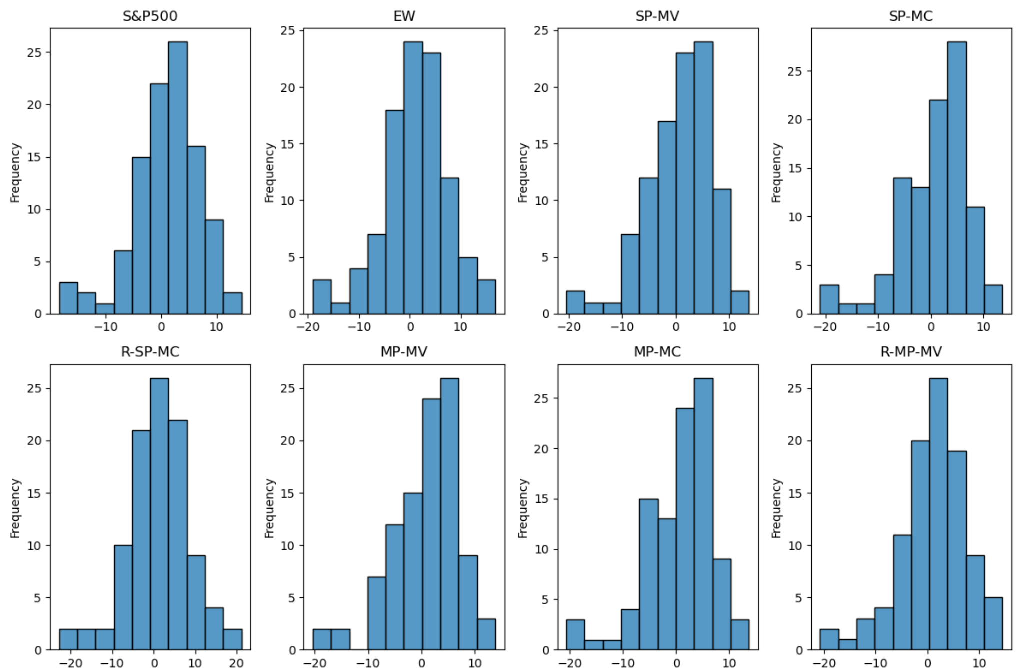


Figure 34: Distribution of portfolio wealth returns for experiment 11.

Experiment 12

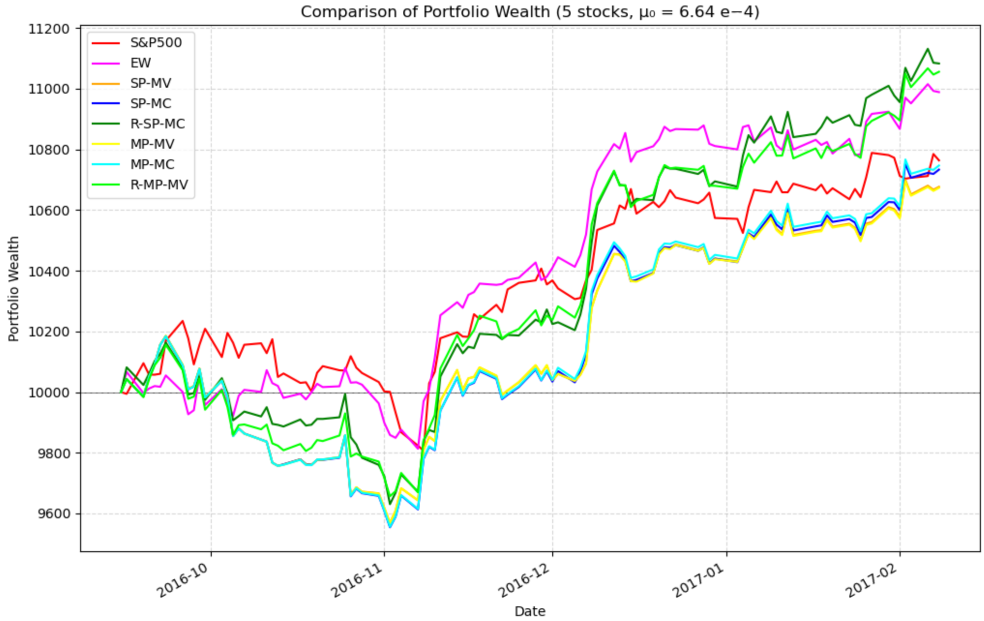


Figure 35: Portfolio wealth over 100-day testing period for experiment 12.

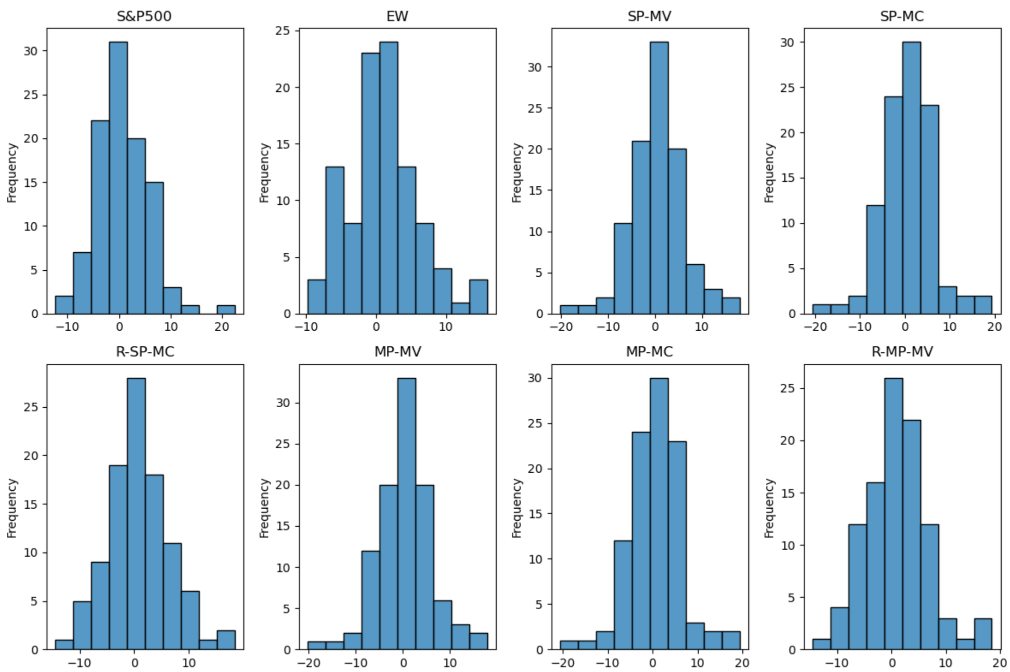


Figure 36: Distribution of portfolio wealth returns for experiment 12.

Experiment 13

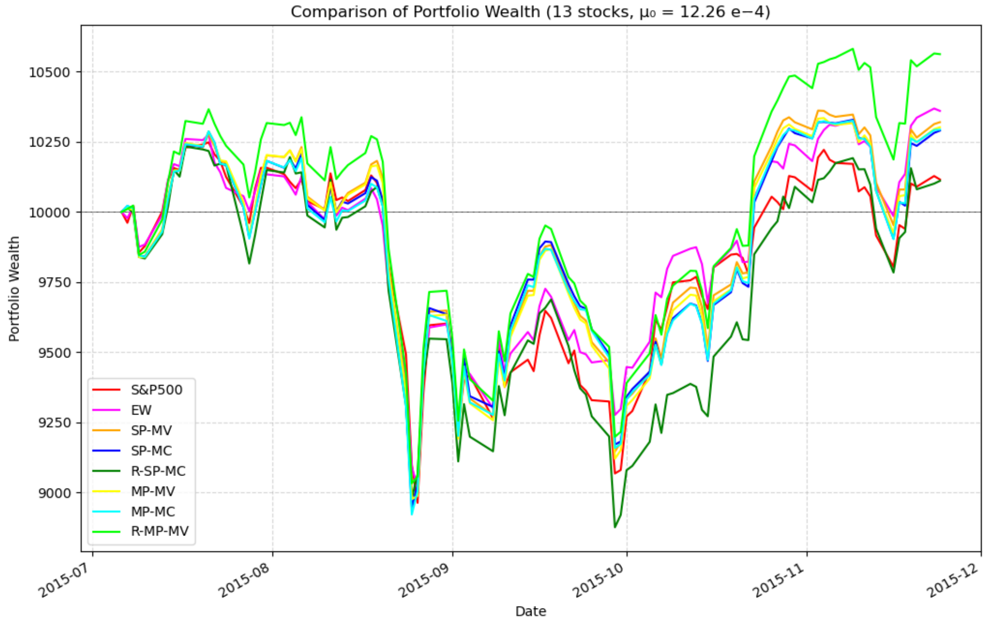


Figure 37: Portfolio wealth over 100-day testing period for experiment 13.

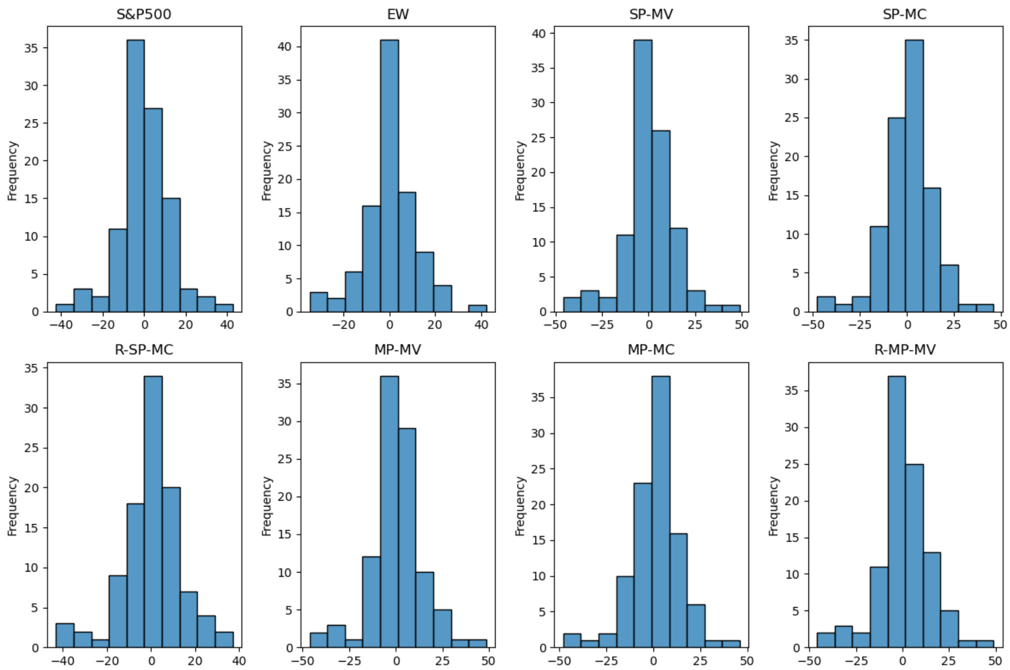


Figure 38: Distribution of portfolio wealth returns for experiment 13.

Experiment 14

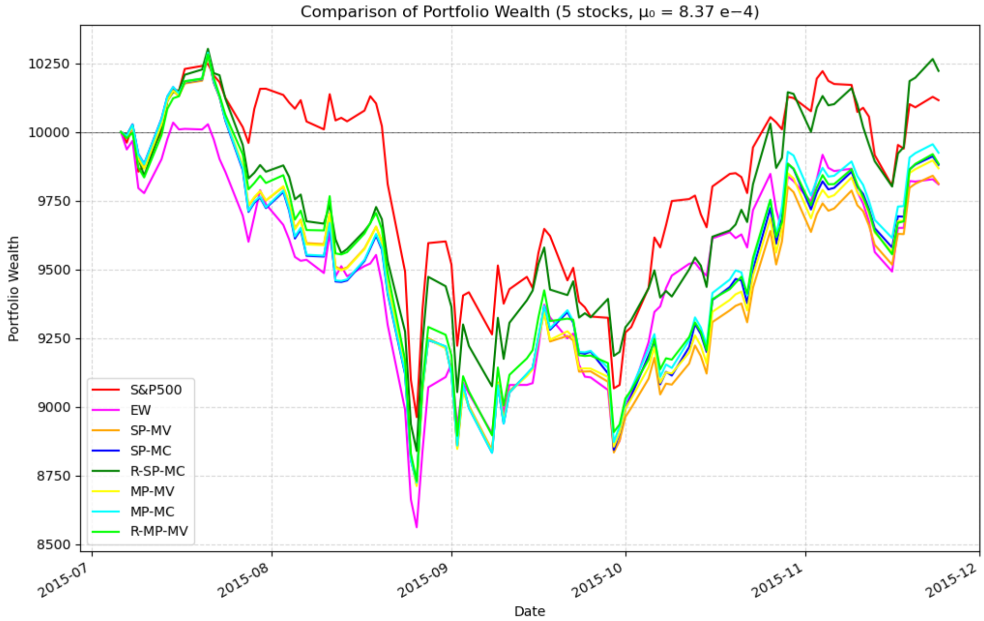


Figure 39: Portfolio wealth over 100-day testing period for experiment 14.

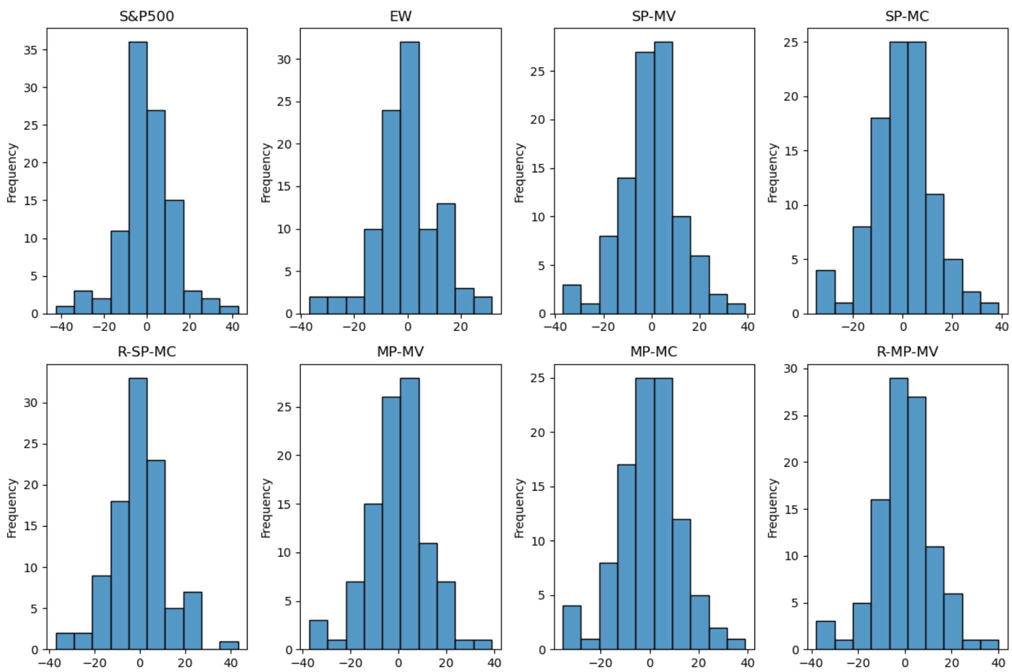


Figure 40: Distribution of portfolio wealth returns for experiment 14.

Experiment 15

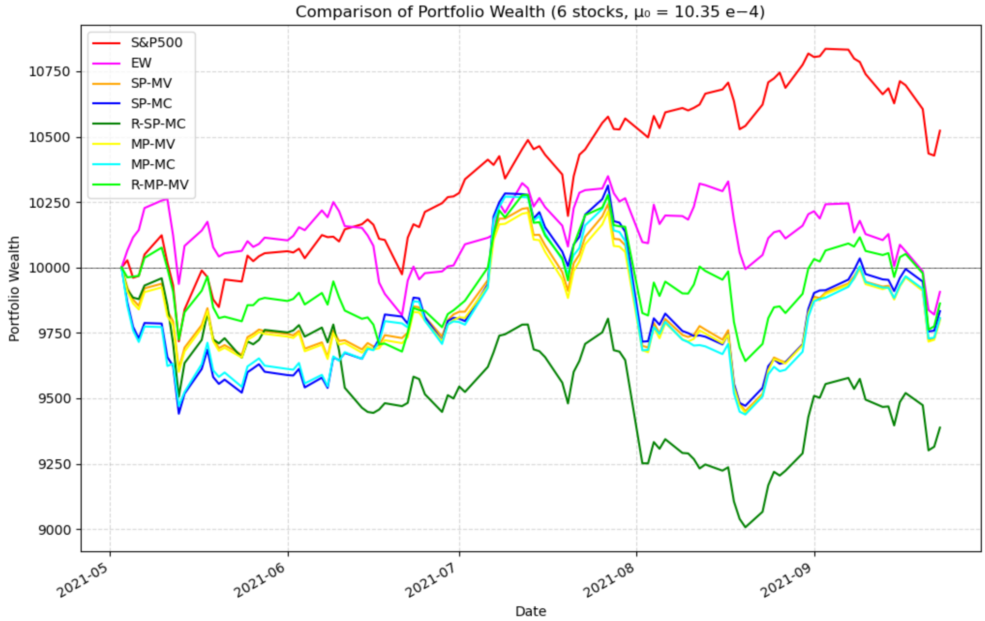


Figure 41: Portfolio wealth over 100-day testing period for experiment 15.

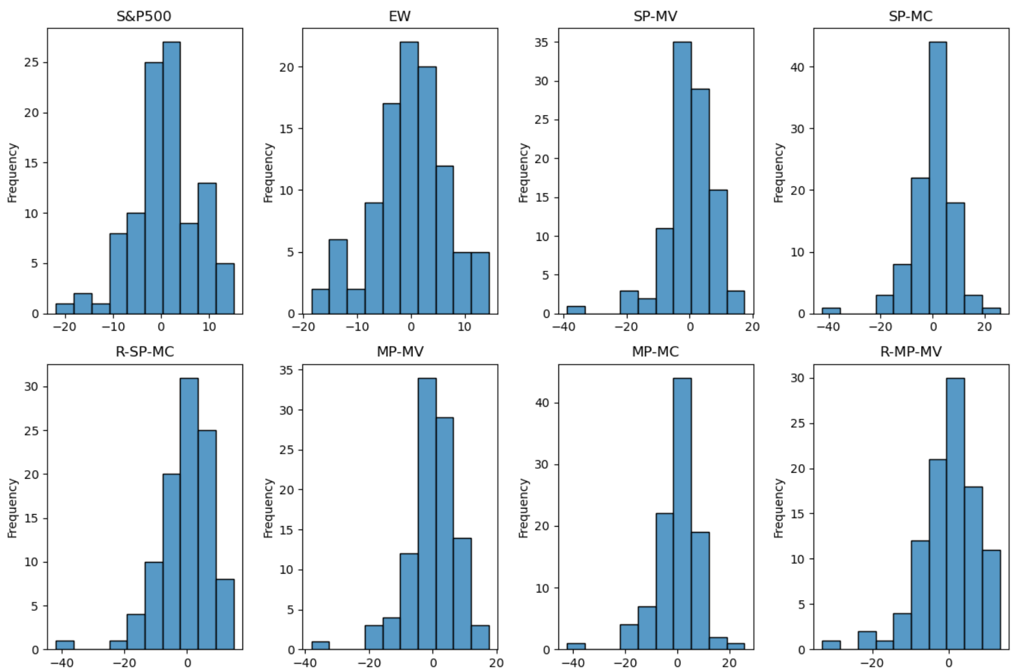


Figure 42: Distribution of portfolio wealth returns for experiment 15.

Experiment 16

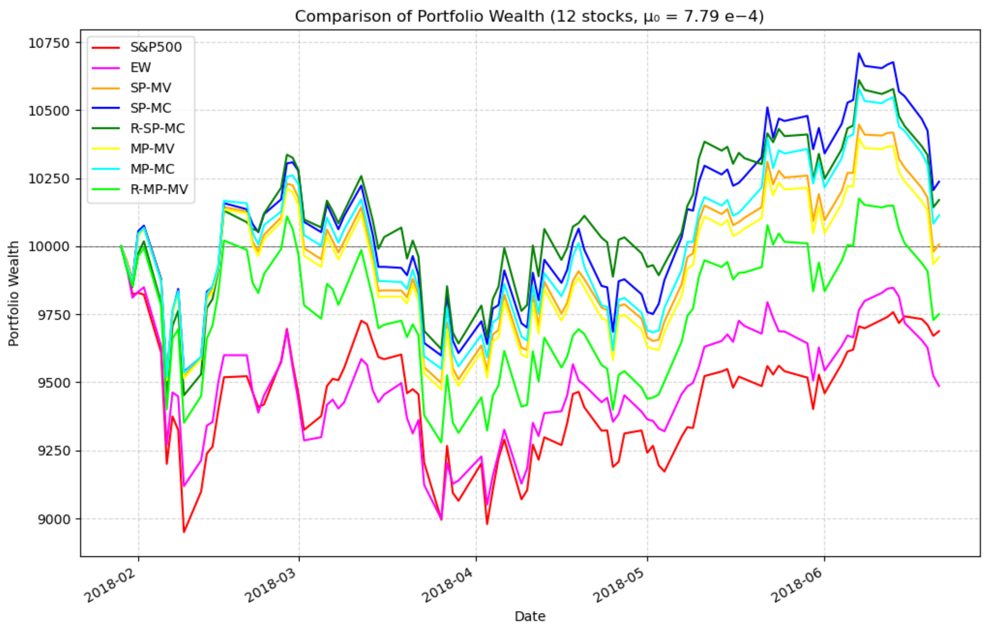


Figure 43: Portfolio wealth over 100-day testing period for experiment 16.

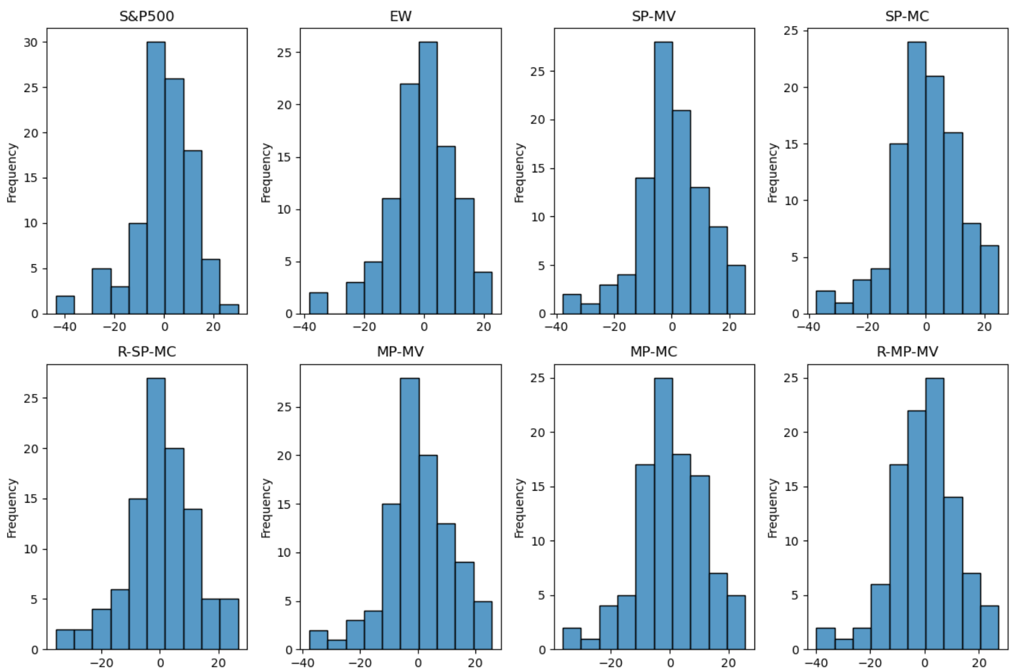


Figure 44: Distribution of portfolio wealth returns for experiment 16.

Experiment 17

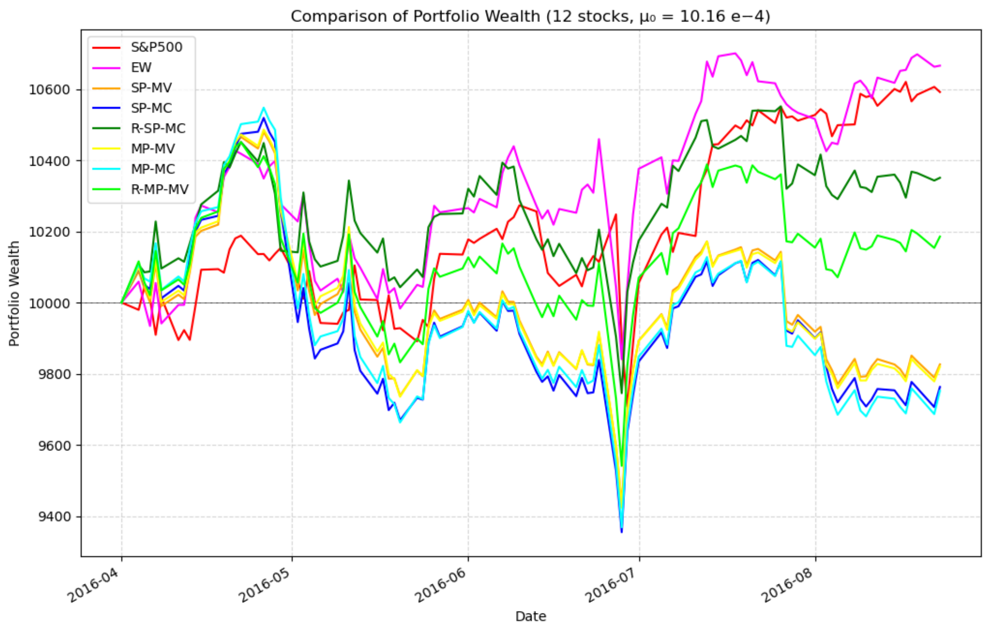


Figure 45: Portfolio wealth over 100-day testing period for experiment 17.

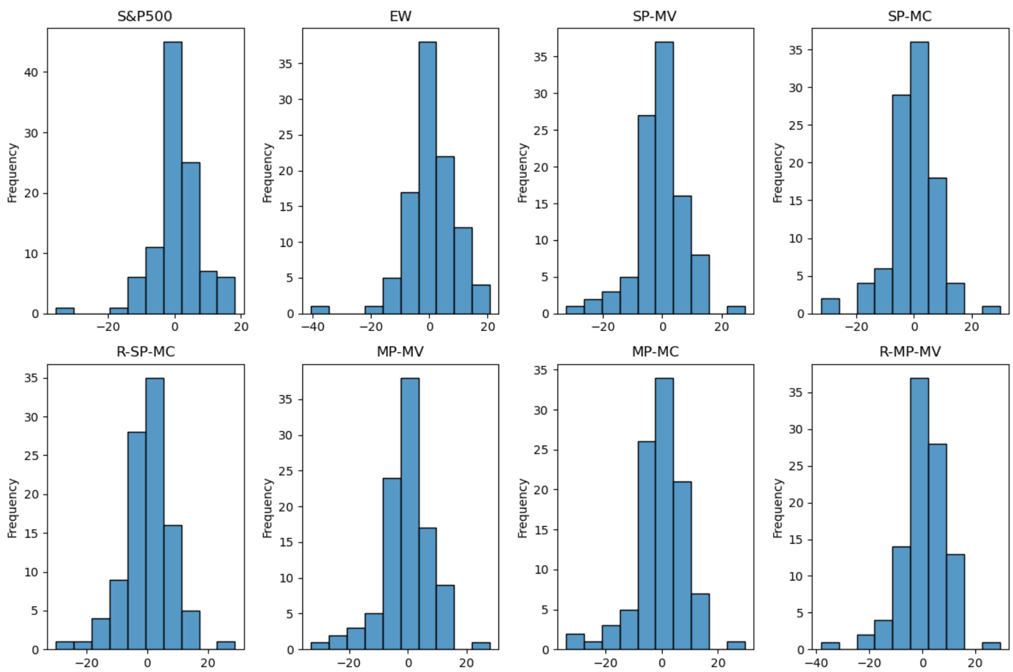


Figure 46: Distribution of portfolio wealth returns for experiment 17.

Experiment 18

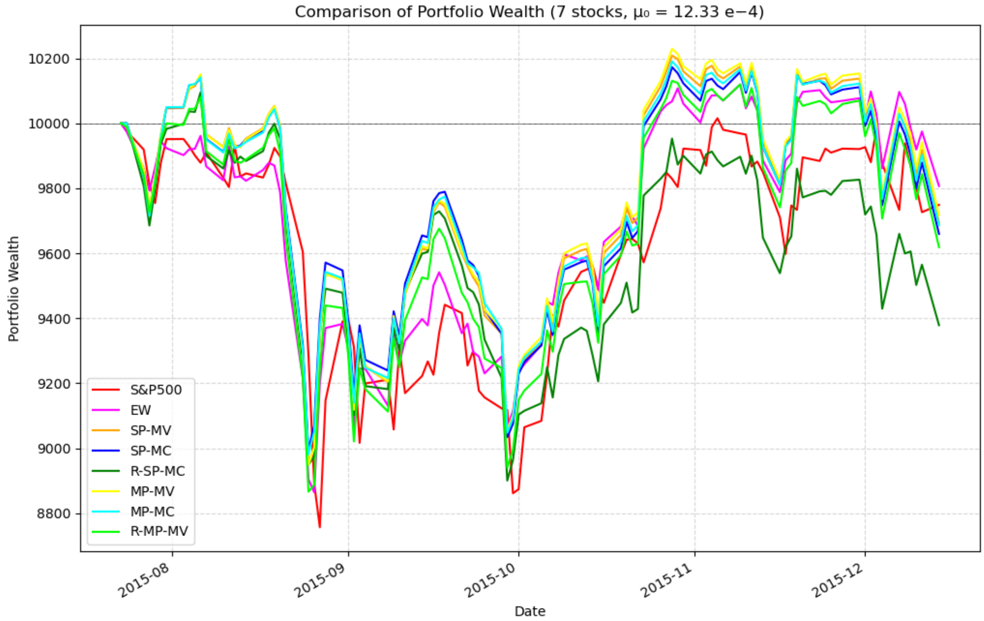


Figure 47: Portfolio wealth over 100-day testing period for experiment 18.

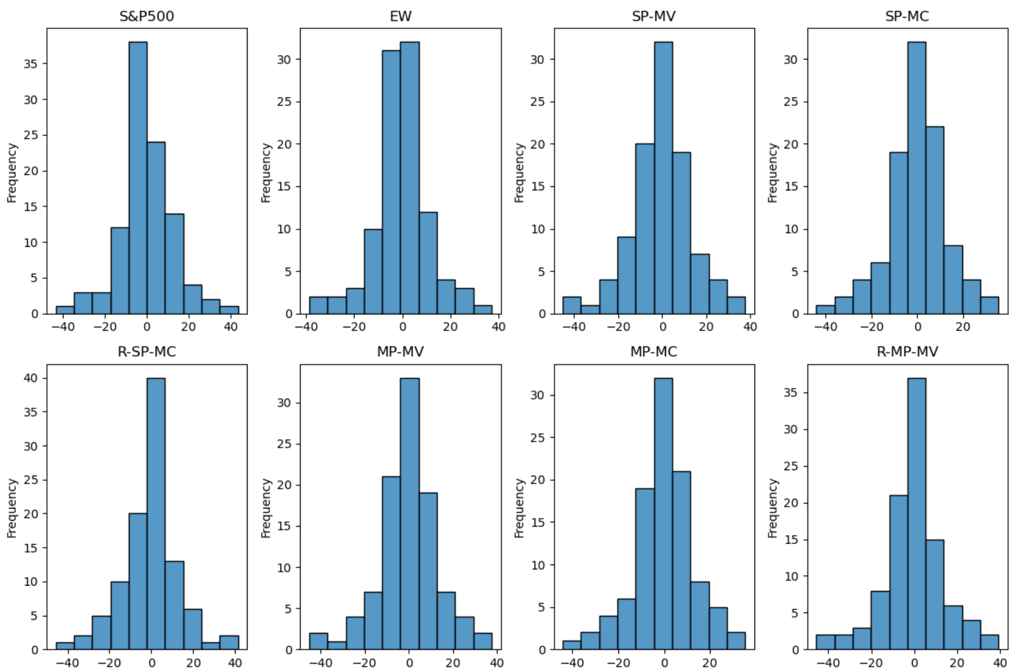


Figure 48: Distribution of portfolio wealth returns for experiment 18.

Experiment 19

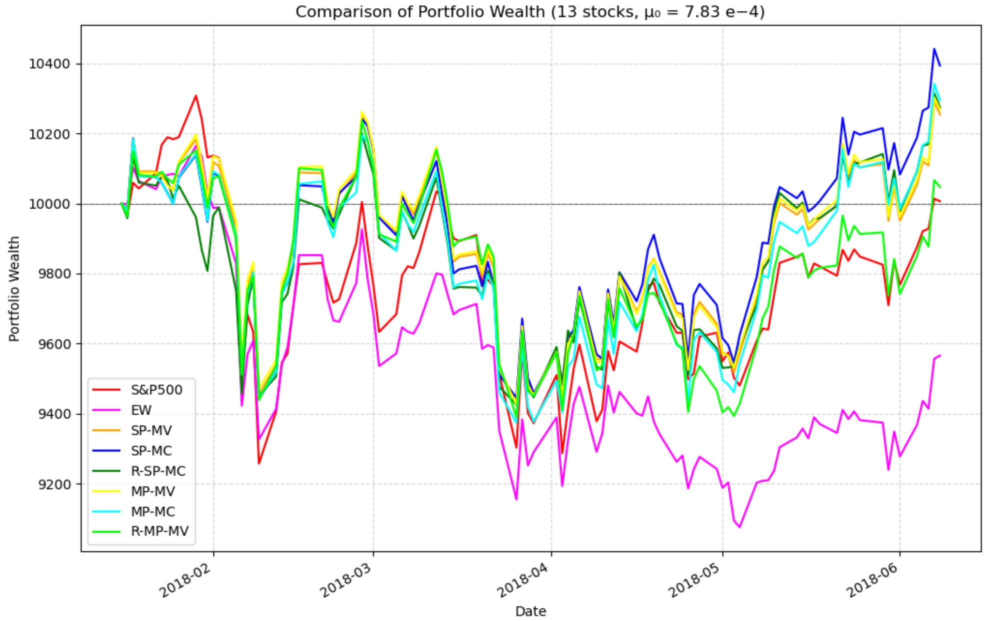


Figure 49: Portfolio wealth over 100-day testing period for experiment 19.

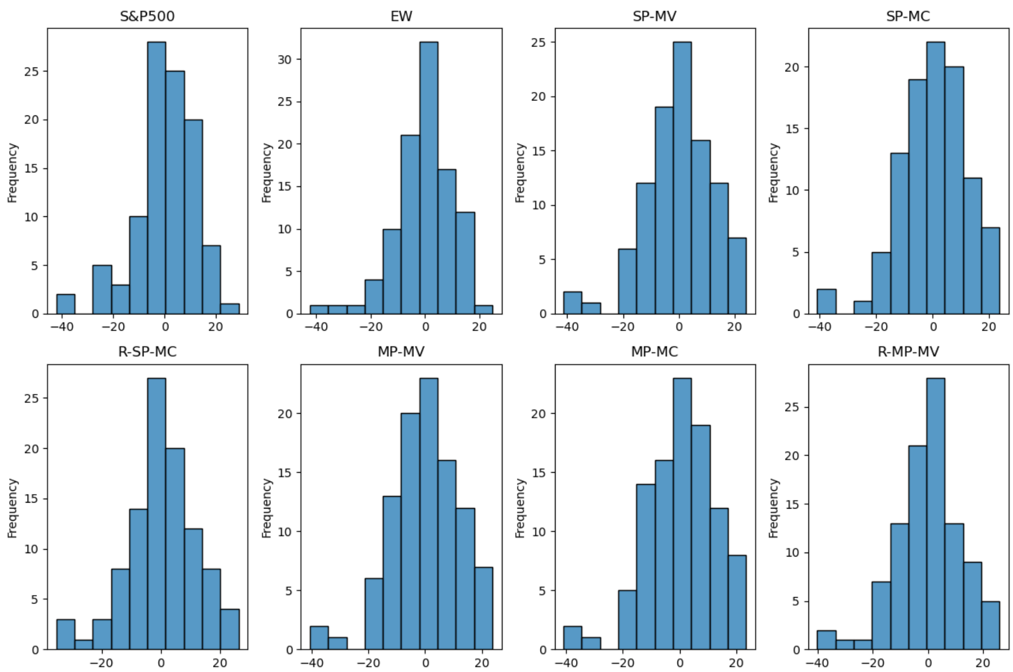


Figure 50: Distribution of portfolio wealth returns for experiment 19.

Experiment 20

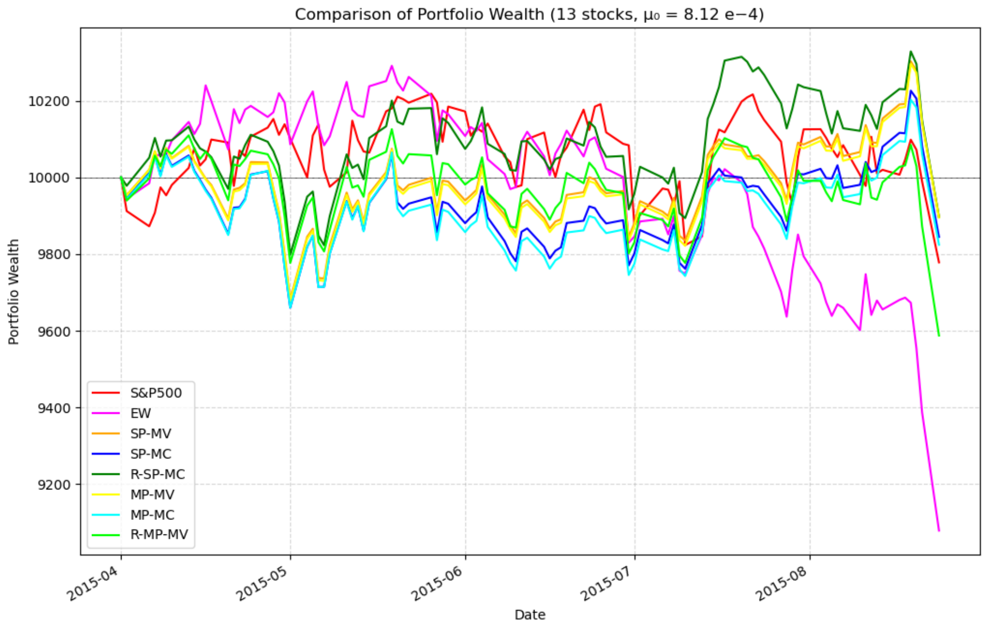


Figure 51: Portfolio wealth over 100-day testing period for experiment 20.

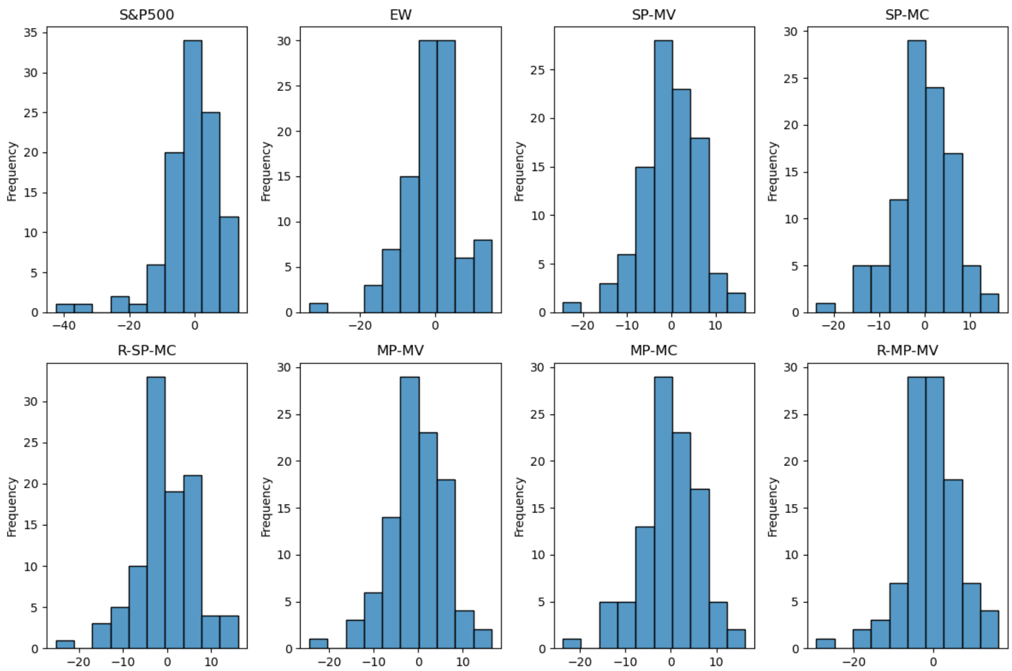


Figure 52: Distribution of portfolio wealth returns for experiment 20.

BIBLIOGRAPHY

Acerbi, C. and D. Tasche (2002). "On the coherence of expected shortfall." In: *Journal of Banking & Finance* 26.7, 1487–1503.

Agrawal, A., R. Verschueren, S. Diamond, and S. Boyd (2018). "A rewriting system for convex optimization problems." In: *Journal of Control and Decision* 5.1, pp. 42–60.

Andersen, M., J. Dahl, and L. Vandenberghe (2022). *CVXOPT*. http://cvxopt.org/index.html. Last updated on Mar 08, 2022. Copyright 2004–2022, Martin S. Andersen, Joachim Dahl, and Lieven Vandenberghe.

Arnott, R. D. and W. H. Wagner (1990). "The Measurement and Control of Trading Costs." In: *Financial Analysts Journal* 46.6, pp. 73–80.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). "Coherent measures of risk." In: *Mathematical Finance* 9.3, 203–228.

Basel Committee, o. B. S. (1996). *Supervisory Framework for the Use of "New" Basel Capital Accord.*

Ben-Tal, A. and A. Nemirovski (1998). "Robust convex optimization." In: *Mathematics of operations research* 23.4, pp. 769–805.

Bernardo, A. E. and O. Ledoit (2000). "Gain, loss, and asset pricing." In: *Journal of political economy* 108.1, pp. 144–172.

Black, F. and R. Litterman (1992). "Global Portfolio Optimization." In: *Financial Analysts Journal* 48.5, pp. 28–43.

Boda, K., J. A. Filar, et al. (2006). "Time consistent dynamic risk measures." In: *Mathematical Methods of Operations Research* 63.1, pp. 169–186.

Bodnar, T., D. Ivasiuk, N. Parolya, and W. Schmid (2020). "Mean-variance efficiency of optimal power and logarithmic utility portfolios." In: *Mathematics and Financial Economics* 14, pp. 675–698.

— (2023). "Multi-period power utility optimization under stock return predictability." In: *Computational Management Science* 20.1, p. 4.

Bodnar, T., N. Parolya, and W. Schmid (2013). "On the equivalence of quadratic optimization problems commonly used in portfolio theory." In: *European Journal of Operational Research* 229.3, pp. 637–644.

— (2015a). "A closed-form solution of the multi-period portfolio choice problem for a quadratic utility function." In: *Annals of Operations Research* 229, pp. 121–158.

— (2015b). “On the exact solution of the multi-period portfolio choice problem for an exponential utility under return predictability.” In: *European Journal of Operational Research* 246.2, pp. 528–542.

Boyd, S., E. Busseti, S. Diamond, R. N. Kahn, K. Koh, P. Nystrup, and J. Speth (2017). *Multi-Period Trading via Convex Optimization*.

Broadie, M. (1993). “Computing efficient frontiers using estimated parameters.” In: *Ann. Oper. Res.* 45.1, pp. 21–58.

Chen, Z.-p., G. Li, and J.-e. Guo (2013). “Optimal investment policy in the time consistent mean–variance formulation.” In: *Insurance: Mathematics and Economics* 52.2, pp. 145–156.

Chopra, V. K. and W. T. Ziemba (2013). “The effect of errors in means, variances, and covariances on optimal portfolio choice.” In: *Handbook of the fundamentals of financial decision making: Part I*. World Scientific, pp. 365–373.

Connor, G. (2000). “Transactions Costs, Turnover, and Trading.” In: *Active Portfolio Management: A Quantitative Approach to Providing Superior Returns and Controlling Risk*. Vol. 13. 4. Oxford University Press. Chap. 16.

Cui, X.-Y., J.-J. Gao, X. Li, and Y. Shi (2022). “Survey on multi-period mean–variance portfolio selection model.” In: *Journal of the Operations Research Society of China* 10.3, pp. 599–622.

Cui, X., J. Gao, X. Li, and D. Li (2014). “Optimal multi-period mean–variance policy under no-shorting constraint.” In: *European Journal of Operational Research* 234.2, pp. 459–468.

Cui, X., D. Li, S. Wang, and S. Zhu (2012). “Better than dynamic mean-variance: Time inconsistency and free cash flow stream.” In: *Mathematical Finance* 22 (2).

Dantzig, G. B. and G. Infanger (1993). “Multi-stage stochastic linear programs for portfolio optimization.” In: *Annals of Operations Research* 45.1, pp. 59–76.

Delage, E. and Y. Ye (2010). “Distributionally robust optimization under moment uncertainty with application to data-driven problems.” In: *Operations research* 58.3, pp. 595–612.

Diamond, S. and S. Boyd (2016). “CVXPY: A Python-embedded modeling language for convex optimization.” In: *Journal of Machine Learning Research* 17.83, pp. 1–5.

Fabozzi, F. J., S. M. Focardi, P. N. Kolm, and D. A. Pachamanova (2007a). *Robust portfolio optimization and management*. John Wiley & Sons.

Fabozzi, F., P. Kolm, D. Pachamanova, and S. Focardi (2007b). “Robust portfolio optimization - Recent trends and future directions.” In: *The Journal of Portfolio Management* 33, pp. 40–+.

Focardi, S. M., F. J. Fabozzi, et al. (2004). *The mathematics of financial modeling and investment management*. Vol. 138. John Wiley & Sons.

Ghaoui, L. E., M. Oks, and F. Oustry (2003). “Worst-case value-at-risk and robust portfolio optimization: A conic programming approach.” In: *Operations research* 51.4, pp. 543–556.

Hu, Z. and L. J. Hong (2013). “Kullback-Leibler divergence constrained distributionally robust optimization.” In: *Available at Optimization Online* 1.2, p. 9.

Hyndman, R. J. and Y. Fan (1996). “Sample Quantiles in Statistical Packages.” In: *The American Statistician* 50.4, pp. 361–365.

Ingersoll, J. E. (1987). *Theory of financial decision making*. Vol. 3. Rowman & Littlefield.

Kakouris, I. and B. Rustem (2014). “Robust portfolio optimization with copulas.” In: *European Journal of Operational Research* 235.1, pp. 28–37.

Kang, Z., X. Li, Z. Li, and S. Zhu (2019). “Data-driven robust mean-CVaR portfolio selection under distribution ambiguity.” In: *Quantitative Finance* 19.1, pp. 105–121.

Konno, H. and H. Yamazaki (1991). “Mean-Absolute Deviation Portfolio Optimization Model and Its Applications to Tokyo Stock Market.” In: *Management Science* 37.5, pp. 519–531.

Kuhn, D., P. M. Esfahani, V. A. Nguyen, and S. Shafieezadeh-Abadeh (2019). “Wasserstein distributionally robust optimization: Theory and applications in machine learning.” In: *Operations research & management science in the age of analytics*, pp. 130–166.

Leippold, M., F. Trojani, and P. Vanini (2004). “A geometric approach to multiperiod mean variance optimization of assets and liabilities.” In: *Journal of Economic Dynamics and Control* 28.6, pp. 1079–1113.

Li, D. and W.-L. Ng (2000). “Optimal dynamic portfolio selection: Multiperiod mean-variance formulation.” In: *Mathematical finance* 10.3, pp. 387–406.

Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. Yale University Press.

Markowitz, H. (1952). “Portfolio Selection.” In: *The Journal of Finance* 7.1, pp. 77–91.

McLeod, W. and G. van Vuuren (2004). "Interpreting the Sharpe ratio when excess returns are negative." In: *Investment Analysts Journal* 33.59, pp. 15–20.

Merton, R. C. (1972). "An analytic derivation of the efficient portfolio frontier." In: *Journal of financial and quantitative analysis* 7.4, pp. 1851–1872.

Michaud, R. O. (1989). "The Markowitz optimization enigma: Is 'optimized' optimal?" In: *Financial analysts journal* 45.1, pp. 31–42.

Mohajerin Esfahani, P. and D. Kuhn (2018). "Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations." In: *Mathematical Programming* 171.1-2, pp. 115–166.

Morgan, J., N. J.P. Morgan Chase & Co. (New York, J. Longerstaey, R. Limited, R. Ltd, M. Spencer, and M. G. T. C. of New York (1996). *RiskMetrics: Technical Document*. J. P. Morgan.

Pituch, K. A. and J. P. Stevens (2015). In: *Applied Multivariate Statistics for the Social Sciences: Analyses with SAS and IBM's SPSS*. 6th ed. Routledge.

Rockafellar, R. T. and S. Uryasev (2002). "Conditional value-at-risk for general loss distributions." In: *Journal of banking & finance* 26.7, pp. 1443–1471.

Rockafellar, R. T., S. Uryasev, et al. (2000). "Optimization of conditional value-at-risk." In: *Journal of risk* 2, pp. 21–42.

Sharpe, W. F. (1966). "Mutual Fund Performance." In: *The Journal of Business* 39.1, pp. 119–138.

Sharpe, W. F. (1994). *Journal of Portfolio Management*.

Sra, S., S. Nowozin, and S. J. Wright (2011). "Interior-Point Methods for Large-Scale Cone Programming." In: *Optimization for Machine Learning*. The MIT Press, pp. 55–84.

Wu, Z. and K. Sun (2023). "Distributionally robust optimization with Wasserstein metric for multi-period portfolio selection under uncertainty." In: *Applied Mathematical Modelling* 117, pp. 513–528.

Zhu, S. and M. Fukushima (2009). "Worst-case conditional value-at-risk with application to robust portfolio management." In: *Operations research* 57.5, pp. 1155–1168.

DECLARATION

A thesis submitted in fulfillment of the requirements for the degree of Master of Science.

Delft, September 2023

Vangelis Nakos