LUISS

DEPARTMENT OF IMPRESA E MANAGEMENT

Improving the Performance of the Mean-Variance Portfolio: GMV Portfolio and Short-Sales Constraints

SUPERVISOR
Prof. Giacomo Morelli

CANDIDATE
Alessandro Greco

CO-SUPERVISOR
Prof. Daniele Penza

STUDENT ID
770541

ACADEMIC YEAR 2023-2024

Table of Contents

Introduction

Quantitative portfolio management has emerged as a central framework in modern finance, providing investors with mathematical tools to construct optimal portfolios that maximize returns while managing risk. Rooted in Harry Markowitz's seminal Modern Portfolio Theory (MPT), the foundational principle of this approach is diversification, which suggests that spreading investments across multiple assets can reduce risk without sacrificing returns. Despite the theoretical appeal, the practical application of MPT, particularly through mean-variance optimization, is often hindered by significant estimation errors in two key parameters: expected returns and the covariance matrix of asset returns. These estimation errors can lead to highly unstable portfolios, with extreme weights on certain assets, and poor out-of-sample performance.

This thesis seeks to address these practical challenges by exploring advanced strategies that mitigate the impact of estimation errors on portfolio performance. The primary objective is to compare the effectiveness of four distinct portfolio optimization strategies: (1) the traditional Mean-Variance Portfolio (MV), (2) the Global Minimum Variance Portfolio (GMV), (3) the Mean-Variance Portfolio with Short-Sale Constraints (MV-C), and (4) the Global Minimum Variance Portfolio with Short-Sale Constraints (GMV-C). While the MV strategy seeks to maximize returns for a given level of risk, the GMV strategy focuses purely on minimizing portfolio variance without relying on estimates of expected returns. The MV-C and GMV-C strategies introduce short-sale constraints, limiting the portfolio to long positions, which helps to curb extreme allocations that typically arise due to estimation errors in the traditional MV approach.

The thesis adopts a rolling-window approach, re-estimating portfolio weights monthly over a 20-year period, using daily data from 13 major indices and sector-specific ETFs. These indices include benchmarks such as the S&P 500, Euro Stoxx 50, and S&P Asia 50, as well as sectoral ETFs covering industries like technology, healthcare, and energy. Performance is evaluated based on three key metrics: the Sharpe Ratio, which measures risk-adjusted returns; the Sortino Ratio, which focuses on downside risk; and the Turnover Rate, which captures the frequency of rebalancing and thus the stability and transaction costs of the portfolio. The ultimate goal is to provide empirical evidence on how these different strategies perform under real-world market conditions and to determine whether the GMV and constrained strategies offer a more reliable and stable alternative to the traditional mean-variance approach.

Chapter 1. Evolution of Portfolio Theory: from Efficient Frontier to Capital Market Line

1.1 Modern Porfolio Theory

Harry Markowitz (1927 – 2023) is widely recognized as the progenitor of Modern Portfolio Theory (MPT), a seminal framework in finance that revolutionized asset management practices. His pioneering research on efficient portfolios laid the groundwork for subsequent advancements in the field. Markowitz's contributions fundamentally altered the way investors perceive risk and return, providing a quantitative framework for portfolio optimization.

At the core of Markowitz's theory lies the principle of maximizing discounted expected returns, which serves as a guiding principle for portfolio allocation. Initially, it was posited that investors should allocate all funds to the security with the highest discounted value. However, Markowitz astutely recognized the inadequacy of this simplistic approach. In his seminal work, he asserted that any rule of behavior lacking a foundation in diversification must be dismissed. Diversification, he argued, is paramount in mitigating risk and enhancing portfolio performance. This foundational insight paved the way for the development of MPT, highlighting the benefits of spreading investments across a diverse array of assets.

Diversification hinges on the observation that the prices of individual securities in a portfolio tend to move independently of one another. This lack of perfect correlation between asset prices means that when one asset's price fluctuates, it does not necessarily entail a commensurate movement in another asset's price. Consequently, the overall variance of returns for a diversified portfolio is typically lower than the average variance of its constituent assets. While diversification can mitigate unsystematic or idiosyncratic risk associated with individual assets, it cannot eliminate systematic or market risk inherent in the broader market.

To underscore the importance of diversification, it is imperative to elucidate the distinction between variance and covariance. While variance measures the dispersion of returns for a single asset, covariance quantifies the degree to which the returns of two assets move in tandem. In the context of portfolio risk, covariance assumes greater significance than variance, as it encapsulates the interplay between asset returns. By diversifying across assets with low covariance, investors can effectively reduce portfolio risk without sacrificing potential returns.

1.1.1 Benefits of diversification

Let's demonstrate why diversification is so important. In particular, we want to prove that portfolio risk is mainly driven by covariance rather than variance.

Consider an equally weighted portfolio of N assets. Thus:

1. The weight on each asset is $w_n = \frac{1}{N}$
2. A typical variance term is $\left(\frac{1}{N}\right)^2 \sigma_{nn}$
    - Total number of variance terms is N
3. A typical covariance term is $\left(\frac{1}{N}\right)^2 \sigma_{nm}\ (n \neq m)$
    - Total number of covariance terms is $N^2 - N$
    - Total number of unique covariance terms is $(N^2 - N)/2$

Add all the variance and covariance terms:

$$\sigma_p^2 = \mathbb{V}[R_p] = \sum_{n=1}^{N} \sum_{m=1}^{N} w_n\, w_m \sigma_{nm}$$

$$= \sum_{n=1}^{N} \left(\frac{1}{N}\right)^2 \sigma_{nn} + \sum_{n=1}^{N} \sum_{m \neq n}^{N} \left(\frac{1}{N}\right)^2 \sigma_{nm}$$

$$= \left(\frac{1}{N}\right)\left(\frac{1}{N}\sum_{n=1}^{N}\sigma_n^2\right) + \left(\frac{N^2-N}{N^2}\right)\left(\frac{1}{N^2-N}\sum_{n=1}^{N}\sum_{m\neq n}^{N}\sigma_{nm}\right)$$

$$= \left(\frac{1}{N}\right)(average\ variance) + \left(1-\frac{1}{N}\right)(average\ covariance)$$

$$\lim_{N\to\infty}\sigma_p^2 = (0)(average\ variance) + (1-0)(average\ covariance)$$

$$= average\ covariance$$

Therefore, as N becomes very large, the contribution of variance terms goes to zero, while the contribution of covariance terms goes to "average covariance". For a "well-diversified" portfolio, covariances among assets determine portfolio risk. Variance of each asset contributes little to portfolio risk. Therefore, when considering which asset to add to a portfolio, it is important to look at covariances (or correlations) and not variances.

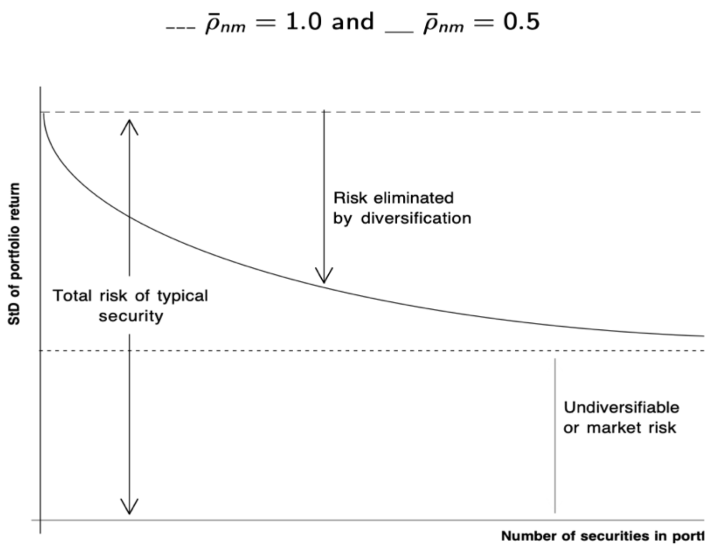

Figure 1. Benefits and limits of diversification

Once Markowitz rejected the foregoing rule, he based all his future work on the following maxim: "the investor does (or should) diversify his funds among all those securities which give maximum expected return". However, the assumption that a single portfolio can simultaneously offer both maximum expected return and minimum variance is flawed. The law of large numbers, while ensuring that the actual yield of the portfolio closely approximates the expected yield, cannot guarantee the existence of such an ideal portfolio. The intercorrelation among security returns precludes the application of the law of large numbers to portfolio returns, making it impossible to eliminate all variance through diversification. Consequently, the portfolio with maximum expected return may not necessarily have the minimum variance, leading to a trade-off between these two objectives.

The challenge then becomes finding the optimal balance between maximizing expected return and minimizing variance. While these goals are often in conflict, Markowitz recognized that an efficient portfolio strikes the optimal balance between risk and return. Consider two portfolios with identical expected returns but differing variances; intuitively, the portfolio with lower variance is preferred as it offers lower risk for the same expected return. This fundamental insight underpins Markowitz's definition of an efficient portfolio, where risk is minimized without sacrificing expected return.
Markowitz's seminal work extends beyond the simple expected returns – variance of returns (E-V) rule, which serves as the foundation for the derivation of the efficient frontier.

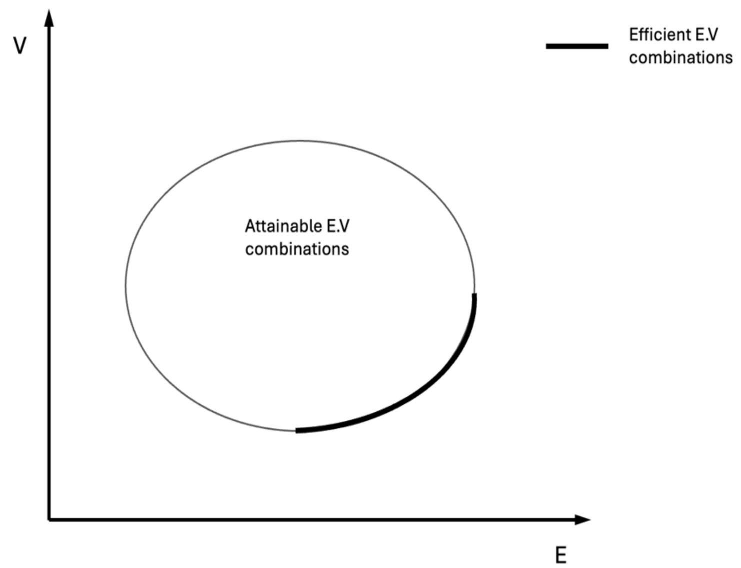

Figure 2. Attainable combinations of expected return and variance

This frontier delineates the set of all obtainable (E,V) combinations, representing the feasible combinations of expected return and variance, as shown in Figure1. According to the E-V rule, investors seek portfolios that generate efficient combinations of (E,V), minimizing variance for a given level of expected return or maximizing expected return for a given level of variance.

A portfolio is considered variance-efficient if, for a fixed level of expected return, no other portfolio exists with a smaller variance. Conversely, a portfolio is expected return-efficient if, for a fixed level of risk (variance), no other portfolio offers a higher expected return.

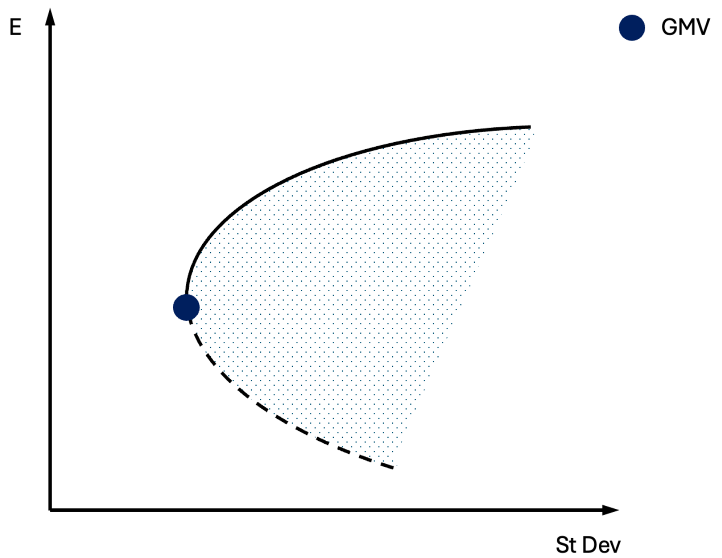

Figure 3. Global Minimum Variance Portfolio

The red dot in Figure2 is called Global Minimum Variance portfolio or simply minimum variance portfolio. As the name suggests, it’s the portfolio with the lowest variance (and consequently the lowest expected return). Points on the efficient frontier below the minimum variance point correspond to portfolios which are not efficient and therefore only the top half of the efficient frontier is used. Why are they inefficient? Consider any point on the lower part (dashed in Figure2), by simply moving up on a straight line we can find numerous portfolios with the same level of risk (standard deviation) but higher expected returns. Similarly, any point within the grey-dotted area is inefficient. Moving up and in a vertical direction would produce a portfolio with higher expected return at the same level of variance; moving to the left in a horizontal manner would produce a portfolio with decreased risk at the same level of expected return. Thus, all portfolios strictly within the efficient frontier are inefficient. The set of all efficient portfolios corresponds precisely to points on the top half of the efficient frontier.

1.2 Capital Market Line

In our exploration of Markowitz's Modern Portfolio Theory (MPT), we gained insights into the efficient frontier—a conceptual landscape showcasing optimal portfolios that either maximize expected returns for a given risk level or minimize risk for a given level of returns. Central to this theory is the mean-variance quadratic optimization problem, which forms the backbone of MPT strategies.

Transitioning further into portfolio theory, we encounter the Capital Market Line (CML) and the Two Fund Separation Theorem, pivotal concepts that expand our understanding of efficient portfolio construction.

The Capital Market Line, conceived by William F. Sharpe, stands as an extension of Markowitz's pioneering work. Its introduction was motivated by the need to offer investors a clear framework for navigating the risk-return trade-off in the presence of a risk-free asset. By delineating the optimal combinations of assets, the CML guides investors in constructing portfolios that maximize returns while prudently managing risk.

Mathematically, the CML is expressed as follows:

$$\mathbb{E}[R_p] = R_f + \frac{\mathbb{E}[R_m] - R_f}{\sigma_m} \sigma_p$$

Where $\mathbb{E}[R_p]$ represents the expected return of the portfolio, R_f is the risk-free rate, $\mathbb{E}[R_m]$ denotes the expected return of the market portfolio, σ_m is the standard deviation of the market portfolio, and σ_p signifies the standard deviation of the portfolio.
At the heart of the CML lies the concept of the tangent portfolio, which marks the point where the efficient frontier intersects with the CML. This portfolio, by virtue of its positioning, is deemed the most efficient due to its ability to yield the highest Sharpe Ratio – a metric gauging the excess return per unit of risk.

Mathematically, the Sharpe Ratio is defined as:

$$SR = \frac{\mathbb{E}[R_p] - R_f}{\sigma_p}$$

Graphically, the tangent portfolio manifests as the optimal blend of a risk-free asset and a risky portfolio, offering investors the highest Sharpe ratio and, consequently, maximizing utility.
By grasping the intricacies of the Capital Market Line and the tangent portfolio, investors gain a robust framework for tailoring their portfolios to align with their unique risk-return preferences. This comprehensive understanding empowers investors to navigate the complex landscape of investment decisions with confidence and precision.

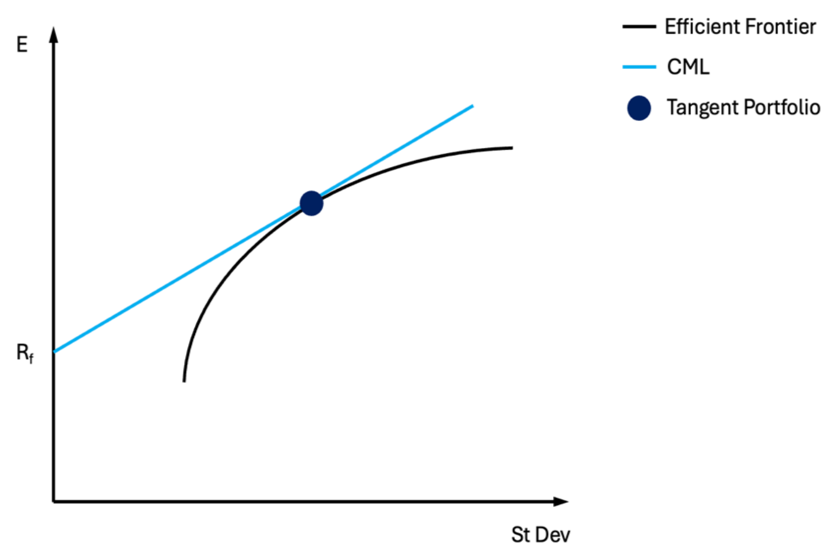

Figure 4. Efficient frontier and capital market line

Now that we understand the intuition behind this theory, we can move on to mathematical derivation.

1.3 Optimization problems in different scenarios

1.3.1 Useful notations

Consider a portfolio of N assets in which the proportion of total value invested in asset n is denoted by w_n.

1. The weights sum to one:

$$\sum_{n=1}^{N} w_n = w^T 1_N = 1$$

2. The return on the portfolio is:

$$R_p = w_1 R_1 + w_2 R_2 + \cdots + w_N R_N = \sum_{n=1}^{N} w_n R_n = w^T R$$

3. The expected return on the portfolio is:

$$\mu_p = \mathbb{E}[R_p] = w_1 \mathbb{E}[R_1] + w_2 \mathbb{E}[R_2] + \cdots + w_N \mathbb{E}[R_N] = \sum_{n=1}^{N} w_n \mathbb{E}[R_n] = w^T \mu$$

4. The variance of the portfolio is:

$$\sigma_p^2 = \mathbb{V}[R_p] = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n\, w_m \sigma_{nm}, \qquad where\ \sigma_{nn} = \sigma_n^2$$

5. The volatility (StD) of portfolio return is:

$$\sigma_p = \sqrt{\mathbb{V}[R_p]} = \sqrt{\sigma_p^2}$$

1.3.2 Mean-variance portfolio frontier without a risk-free asset

As explained before, choosing a mean-variance efficient portfolio is equal to minimize risk for a given expected return or maximize expected return for a given risk. In order to obtain the frontier portfolios, we need to solve the following problem:

$$Minimize_{\{w_1,\dots,w_N\}}\ \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n\, w_m \sigma_{nm} = w^T V w$$

$subject\ to$:

$$(1)\ \sum_{n=1}^{N} w_n = w^T 1_N = 1;$$

$$(2)\ \sum_{n=1}^{N} w_n \mathbb{E}[R_n] = w^T \mu = \mu_{target}$$

In plain language, we need to find the weights $\{w_1, \dots, w_N\}$ that minimize the portfolio variance σ_p^2 while satisfying two constraints:

1. the weights must add up to 1

2. The portfolio p must have an expected target return μ_{target}

This is a non-linear optimization problem. To solve this constrained optimization problem, we construct a Lagrangian function, $\mathcal{L}$:

$$\mathcal{L} = w^T V w + \lambda_w (1 - w^T 1_N) + \lambda_R (\mu_{target} - w^T \mu)$$

Where $\lambda_w$and λ_R are the Lagrange multipliers that correspond to the two constraints. By calculating the partial derivative with respect to the vector of weights, we obtain the first-order condition of the optimization problem:

$$\frac{\partial \mathcal{L}}{\partial w} = 2Vw - \lambda_w 1_N - \lambda_R \mu = 0$$

Assume that V is not a singular matrix, thus its inverse exists. Then:

$$w = V^{-1} \left(\frac{\lambda_w}{2} 1_N + \frac{\lambda_R}{2} \mu \right)$$

To get an explicit solution for the optimal weights, we need to find λ_w and λ_R, and substitute these values into the above expression for w. We must differentiate the Lagrangian function with respect to the Lagrange multipliers, which give us:

$$(1)\ 1 = w^T 1_N$$

$$(2)\ \mu_{target} = w^T \mu$$

Now, substitute the expression for *w* into equations (2) and (3):

$$1 = w^T 1_N = \left[V^{-1}\left(\frac{\lambda_w}{2}1_N + \frac{\lambda_R}{2}\mu\right)\right]^T 1_N = a_3\frac{\lambda_w}{2} + a_2\frac{\lambda_R}{2}$$

$$\mu_{target} = w^T\mu = \left[V^{-1}\left(\frac{\lambda_w}{2}1_N + \frac{\lambda_R}{2}\mu\right)\right]^T \mu = a_2\frac{\lambda_w}{2} + a_1\frac{\lambda_R}{2}$$

Where

$$a_1 = \mu^T V^{-1}\mu, \qquad a_2 = \mu^T V^{-1} 1_N, \qquad a_3 = 1_N^T V^{-1} 1_N$$

The above system of two equations can be written in matrix notation as follows, where the two unknowns are λ_w and λ_R:

$$\begin{bmatrix} 1 \\ \mu_{target} \end{bmatrix} = \begin{bmatrix} a_3 & a_2 \\ a_2 & a_1 \end{bmatrix} \cdot \begin{bmatrix} \frac{\lambda_w}{2} \\ \frac{\lambda_R}{2} \end{bmatrix}$$

Solving the two equations for $\frac{\lambda_w}{2}$ and $\frac{\lambda_R}{2}$, we have

$$\frac{\lambda_w}{2} = \frac{1}{D}(a_1 - a_2\mu_{target})$$

$$\frac{\lambda_R}{2} = \frac{1}{D}(-a_2 + a_3\mu_{target})$$

Where we denote by D the determinant of the 2 x 2 matrix:

$$D = \det\left(\begin{bmatrix} a_3 & a_2 \\ a_2 & a_1 \end{bmatrix}\right) = a_1 a_3 - a_2^2$$

Substituting this solution into the equation for w, we find the frontier portfolio weights for expected return μ_{target}:

$$w = \frac{1}{D}(a_1 V^{-1} 1_N - a_2 V^{-1}\mu) + \frac{1}{D}(a_3 V^{-1}\mu - a_2 V^{-1} 1_N)\, \mu_{target}$$

Note that this explicit solution is possible only when there are no other constraints on the portfolio weights. For example, if we want all weights to be positive or if we want all weights to be less than some maximum weight, then we will not be able to get an explicit solution. Instead, we will have to solve for the weights numerically.

Define by w_0 and w_1 the frontier portfolios with expected return $\mu_{target} = 0$ and $\mu_{target} = 1$, respectively:

$$w_0 = \frac{1}{D}(a_1 V^{-1} 1_N - a_2 V^{-1}\mu)$$

$$w_1 = \frac{1}{D}(a_1 V^{-1} 1_N - a_2 V^{-1}\mu) + \frac{1}{D}(a_3 V^{-1}\mu - a_2 V^{-1} 1_N)$$

Then any frontier portfolio p with expected return μ_{target} is equal to:

$$w_p = w_0 + (w_1 - w_0)\,\mu_{target} = w_0\left(1 - \mu_{target}\right) + w_1\,\mu_{target}$$

The last expression suggests that w_0 and w_1 generate the entire frontier. Given the above equation for any frontier portfolio with expected target return μ_{target}, the variance of frontier portfolio p is:

$$\sigma_p = w_p^T\, V\, w_p$$

$$= \left(w_0\left(1 - \mu_{target}\right) + w_1\,\mu_{target}\right)^T V \left(w_0\left(1 - \mu_{target}\right) + w_1\,\mu_{target}\right)$$

$$= (1 - \mu_{target})^2\, w_0^T\, V\, w_0 + 2\,\mu_{target}\left(1 - \mu_{target}\right) w_0^T\, V\, w_1 + \mu_{target}^2\, w_1^T\, V\, w_1$$

By choosing different levels of μ_{target} and then finding the corresponding w and σ_p, one can identify the entire mean-volatility portfolio frontier.

1.3.3 Mean-variance portfolio frontier with a risk-free asset

When investors can also invest in a risk-free asset, the optimal portfolio consists of the risk-free asset and risky assets. In the presence of a risk-free asset, frontier portfolios are combinations of the risk-free asset and the tangent portfolio (of risky assets).

Consider a portfolio *p* with:

- a invested in a risky portfolio q
- *(1-a)* invested in the risk-free asset

Then, we have:

1. The mean of the portfolio return is: $\mu_{target} = (1 - a)\, R_f + a\, \mathbb{E}[R_q]$

2. The variance of the portfolio return is: $\sigma_p^2 = a^2\, \sigma_q^2$

3. The standard deviation of the portfolio return is: $\sigma_p = a\, \sigma_q$

4. Let the total weight invested in risky assets be a, then: $a = \sum_{n=1}^{N} w_n$

5. The weight invested in the risk-free asset is: $w_f = 1 - a = 1 - \sum_{n=1}^{N} w_n$

In the presence of a risk-free asset, the optimal portfolio problem becomes the following:

$$Minimize_{\{w_1,\dots,w_N\}}\ \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n\, w_m \sigma_{nm} = w^T V w$$

$subject\ to$:

$$(1)\ \sum_{n=1}^{N} w_n + w_f = w^T 1_N + w_f = 1;$$

$$(2)\ \sum_{n=1}^{N} w_n \mathbb{E}[R_n] + w_f R_f = w^T \mu + w_f R_f = \mu_{target}$$

However, constraint (1) implies that: $w_f = 1 - w^T 1_N$

Substituting this expression for w_f into constraint (2), we get that the problem to be solved is:

$$Minimize_{\{w_1,\ldots,w_N\}}\ \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n\, w_m \sigma_{nm} = w^T V w$$

$subject\ to$:

$$w^T \mu + (1 - w^T 1_N)\, R_f = \mu_{target}$$

Let's apply the Lagrangian method to solve the above optimization problem.

$$\mathcal{L} = w^T V w + \lambda\left[\mu_{target} - w^T \mu - (1 - w^T 1_N) R_f\right]$$

To find the optimal vector of weights, we apply the first order condition. Thus, we differentiate the Lagrangian function with respect to *w* and set it equal to zero:

$$\frac{\partial \mathcal{L}}{\partial w} = 2Vw - \lambda\mu + \lambda R_f 1_N = 0$$

$$(1)\quad w = \frac{V^{-1}}{2}\left[\lambda\mu + \lambda R_f 1_N\right]$$

Now, let's differentiate the Lagrangian function with respect to λ:

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \mu_{target} - w^T \mu + (1 - w^T 1_N) R_f = 0$$

Substituting the expression (1) into this equation, we have:

$$\mu_{target} - \left(\frac{1}{2}\lambda\mu^T V^{-1}\mu - \frac{1}{2}\lambda R_f \mu^T V^{-1} 1_N\right) +$$
$$+\left(1 - \left(\frac{1}{2}\lambda\mu^T V^{-1} 1_N - \frac{1}{2}\lambda R_f 1_N^T V^{-1} 1_N\right)\right) R_f = 0$$

For simplicity, let's define:

$$a_1 = \mu^T V^{-1} \mu$$
$$a_2 = \mu^T V^{-1} 1_N$$
$$a_3 = R_f 1_N^T V^{-1} 1_N$$

Then, we can easily compute the expression of the optimal Lagrangian multiplier:

$$\lambda = \frac{\mu_{target} - R_f}{a_1 - a_2 + a_3}$$

Substituting the above expression into (1), we have:

$$w = \frac{\mu_{target} - R_f}{a_1 - a_2 + a_3} V^{-1} \left(\mu - R_f 1_N\right)$$

To conclude, the final solution for the optimal weights vector is:

$$w = \frac{\mu_{target} - R_f}{(\mu - R_f 1_N)^T \, V^{-1} \, (\mu - R_f 1_N)} \, V^{-1} \left(\mu - R_f 1_N\right)$$

Once we know the optimal weights for the risky assets, we can easily compute the weight in the risk-free asset:

$$w_f = 1 - w^T 1_N$$

We can also find out the weights in the tangency portfolio. The tangency portfolio is the portfolio with investment in only risky assets (and nothing in the risk-free asset). We can calculate the tangency portfolio weights by dividing the weight of each risky asset by the sum of the weights in the N risky assets:

$$w_{tang} = \frac{w}{w^T 1_N}$$

1.3.4 Mean-Variance utility

Markowitz shows that finding the mean-variance efficient portfolio in the presence of the risk-free asset, thus solving the previous optimization problem, is equivalent to maximizing mean-variance utility. Maximizing mean-variance utility implies maximizing the difference between the expected return on the portfolio and the variance of the portfolio, adjusted for the risk-aversion (γ) of the investor.
Investors risk-aversion is the tendency of investors to prefer lower levels of risk when making investment decisions, often resulting in a preference for assets with lower volatility or higher certainty of returns.

Let's translate it into an optimization problem.

$$Maximize_{\{w_1,\ldots,w_N\}}\ \mu_p - \frac{\gamma}{2}\sigma_p^2$$

We can view the portfolio of risk-free and risky assets as a portfolio of two sub-portfolios: (i) a sub-portfolio of only the risk-free asset, and (ii) a sub-portfolio of only risky assets. Thus, the portfolio expected return is equal to the sum of the expected return of the sub-portfolio of only the risk-free asset, and the expected return of the sub-portfolio of only risky assets. The sum of the weights in the two sub-portfolios must be 1.

$$Maximize_{\{w_1,\ldots,w_N\}}\ \left[(1 - w^T 1_N)R_f + w^T\mu\right] - \frac{\gamma}{2}w^T V w$$

Collecting the terms in *w*:

$$Maximize_{\{w_1,\ldots,w_N\}}\ \left[R_f +\ w^T\left(\mu - R_f 1_N\right)\right] - \frac{\gamma}{2}w^T V w$$

Again, we apply the Lagrange method and start by using the first-order condition.

$$\mathcal{L} = \left[R_f +\ w^T\left(\mu - R_f 1_N\right)\right] - \frac{\gamma}{2}w^T V w$$

$$\frac{\partial \mathcal{L}}{\partial w} = \left(\mu - R_f 1_N\right) - \gamma V w = 0_N$$

Solving the above equation, we find the optimal vector of portfolio weights:

$$w = \frac{1}{\gamma} V^{-1} \left(\mu - R_f 1_N\right)$$

We said that minimizing portfolio variance in the presence of a risk-free asset and maximizing mean-variance utility are equivalent problem. Still), the solutions for the optimal vector of portfolio weights are different. However, we can solve this dilemma by setting the following equation:

$$\frac{1}{\gamma} = \frac{\mu_{target} - R_f}{(\mu - R_f 1_N)^T V^{-1} (\mu - R_f 1_N)}$$

Are they the same economic quantity? Yes, both sides of the equation represent the same economic concept: the trade-off between risk and return. The left-hand side represents this trade-off in terms of the investor's risk aversion, while the right-hand side represents it in terms of the expected return of the target asset relative to its risk and the risk-free rate.

In essence, both sides of the equation provide different perspectives on how investors evaluate risk and return. The left-hand side quantifies this evaluation in terms of the investor's personal risk aversion, while the right-hand side quantifies it in terms of the asset's expected return adjusted for its risk relative to a risk-free investment.

Chapter 2. Estimation Error in Portfolio Optimization

In the first chapter, we explored models of optimal portfolio choice, culminating in the derivation of the optimal portfolio weight vector:

$$w = \frac{1}{\gamma}\, \mathbb{V}[R]^{-1} \left(\mathbb{E}[R] - R_f\right)$$

Where:

- γ is the risk aversion coefficient
- $\mathbb{V}[R]$ is the variance-covariance matrix of returns
- $\mathbb{E}[R]$ is the vector of expected returns
- R_f is the risk-free rate

To implement these optimal portfolios, one must estimate expected returns and the variances and covariances of returns. However, these estimates are marred by error. This chapter delves into the repercussions of estimation error on portfolio performance and examines models designed to mitigate its impact.

Estimation error is an unavoidable reality in financial modelling. The two primary components that need estimation in our portfolio optimization framework are the expected returns ($\mathbb{E}[R]$) and the variance-covariance matrix of returns ($\mathbb{V}[R]$). Both these elements are derived from historical data, and their precision is crucial for effective portfolio management. However, the errors in these estimations can significantly affect the portfolio's out-of-sample performance.

To comprehend how estimation errors influence portfolio performance, we need to examine the sensitivity of portfolio weights to these errors. The optimization formula shows that even slight miscalculations in expected returns or variances and covariances can lead to substantial changes in portfolio composition. This sensitivity can result in portfolios that perform poorly when applied to new data outside the estimation sample.

2.1 Estimation Error in Expected Returns ($\mathbb{E}[\boldsymbol{R}]$)

Errors in estimating expected returns are particularly problematic. Robert Merton (1980) elucidated the challenges in improving the precision of these estimates, highlighting that frequent sampling of returns (e.g., daily instead of monthly) does not enhance accuracy. This phenomenon occurs because the estimates of expected returns are essentially telescopic series; only the first and last price observations matter, while intermediate prices drop out. This characteristic renders the estimates of expected returns highly imprecise, regardless of the frequency of data sampling.

In financial theory, the Capital Asset Pricing Model (CAPM) provides a basic framework for understanding the equilibrium expected returns. According to CAPM, the expected return on a security is linearly related to its systematic risk, represented by beta (β_i). The relationship is given by:

$$R_i - R_f = \beta_i \, (R_M - R_f)$$

Where:

- R_i is the expected return on security ι
- R_f is the risk-free rate
- β_i is the ratio of the covariance of the return on security ι with the market return to the variance of the market return
- R_M is the expected return on the market portfolio

Empirical tests of the CAPM have shown significant deviations from this predicted relationship, leading to the development of more sophisticated models that account for other types of risks in addition to market risk. However, market risk remains a dominant factor in determining the equilibrium expected return for most common stocks.

2.1.1 The Challenges of Estimating Expected Returns

Despite the theoretical importance of expected return estimates, practical efforts to estimate these parameters from time-series data face substantial hurdles. One significant issue is that estimates of expected returns based on historical data are often subject to substantial error. This imprecision stems from the fact that these estimates are heavily influenced by the extreme values in the data series — primarily the initial and final observations. As Merton (1980) notes, this reliance on the end points of the data series means that even with a long history of returns, the estimates remain highly uncertain.

Moreover, under the Efficient Market Hypothesis (EMH), the unanticipated part of the market return (i.e., the difference between realized and expected return) should be unpredictable based on historical data. This unpredictability further complicates the task of estimating expected returns, as it implies that much of the variance in market returns is driven by factors that are not forecastable from past returns.

Merton (1980) claimed that the variance of returns could be estimated far more accurately from the available time series of realized returns that can the expected returns. We show now that this claim is correct provided that market returns can be described by a difusion-type stochastic process and that the mean and variance of these returns are slowly-varying functions of time. Under the hypothesis that the mean and variance are slowly-varying functions of time, the true process for market returns can be approximated by assuming that μ and σ^2 are constants over (non-overlapping) time intervals of length h, where μ is the expected logarithmic rate of return on the market per unit time and σ^2 is the variance per unit time.

Suppose that the realized return on the market can be observed over time intervals of length Δ, where $\Delta \ll h$. Then $n = h/\Delta$ is the number of observations of realized returns over a time interval of length h. So, for example, is h is equals 1 month and Δ equals 1 day, then n equals 30 (neglecting weekend and holidays). Let X_k denote the logarithmic return on the market over the k^{th} observation interval of length Δ during a typical period of length h for $k = 1,2, \dots, n$. X_k can be written as:

$$(1) \qquad X_k = \mu\Delta + \sigma\sqrt{\Delta}\,\varepsilon_k \quad , \qquad k = 1,2,\dots,n$$

Where $\{\varepsilon_k\},\ k = 1,2,\dots,n$ are independent and identically distributed standard normal random variables. From (1), the estimator for the expected logarithmic return, $\mu\ hat = \frac{[\sum_1^n X_k]}{h}$ will have the properties that

1. $\mathbb{E}[\mu\ hat] = \mu$
2. $Var[\mu\ hat] = \frac{\sigma^2}{h}$

Note that the accuracy of the estimator as measured by $Var[\mu\ hat]$ depends only upon the total length of the observation period h and not upon the number of observations *n*. That is, nothing is gained in ters of accuracy of the expected return estimate by choosing finer observation intervals for the returns and thereby, increasing the number of observations *n* for a fixed value of *h*.

Now, consider the followin estimator for the variance rate:

$$\sigma^2\ hat = \frac{[\sum_1^n X_k^2]}{h}$$

The estimator will have the following properties:

1. $\mathbb{E}[\sigma^2\ hat] = \sigma^2 + \mu^2\Delta = \sigma^2 + \mu^2\frac{h}{n}$
2. $Var\ [\sigma^2\ hat] = 2\frac{\sigma^4}{n} + 4\frac{\mu^2 h}{n^2}$

As inspection of the above reveals, $Var\ [\sigma^2\ hat]$ does depend uon the number of observations *n* for a fixed *h*, and indeed, to order *1/n*, it depends only upon the number of observations. Thus, unlike the accuracy of the expected return estimator, by choosing finer observation intervals Δ, the accuracy of the variance estimator can be improved for a fixed value of *h*. The practical advantage of the variance estimator's accuracy depending upon *n* rather than *h* is that a reasonably accurate estimate of the variance rate can be obtained using daily data while the estimates for expected return taken directly from the sample will be subject to so much error as to be almost useless.

2.1.2 Mathematical Example

Denote tha natural logarithm of prices by $p_{t,n} = \ln P_{t,n}$ so that

$$\ln R_{t,n} = \ln \frac{P_{t,n}}{P_{t-1,n}} = p_{t,n} - p_{t-1,n}$$

And for the multiperiod continuously compounded return:

$$\begin{aligned}
&\ln R_{t,n}(h) = \ln\,(R_{t,n}, R_{t-1,n}, \dots, R_{t-h+1,n}) \\
&= \ln R_{t,n} + \ln R_{t-1,n} + \cdots + \ln R_{t-h+1,n} \\
&= (p_{t,n} - p_{t-1,n}) + (p_{t-1,n} - p_{t-2,n}) + \cdots + (p_{t-h+1,n} - p_{t-h,n})
\end{aligned}$$

However, the last equation represents a telescopic series, meaning that the middle components are a set of couples of equivalent and opposite values that can be eliminated. Thus, the final result is:

$$\ln R_{t,n}(h) = p_{t,n} - p_{t-h,n}$$

This equation shows that for continuously compounded returns, only the first and last price observations matter, the rest drop out. That is, if you are computing returns over 100 years, the return that you get is the same whether you use just the data for the first price and the last price, annual data on prices for all 100 years, monthly data on prices for all the months in the last 100 years, or daily data on prices for all the days in the last 100 years. That is, no additional data is useful for estimating the return. Consequently, estimates of expected return are very imprecise. However, as we have seen, this is not true for the variance. Estimates of the variance are much more precise as their accuracy increases with the number of observations.

2.1.3 How Precise is the Estimate of the Mean E[R]

How precise is the estimate of the mean $\mathbb{E}[R_n]$? to answer this question, we need to find the standard error (standard deviation) of the estimator of $\mathbb{E}[R_n]$. If the annual variance of a return is σ_n^2, then the annual variance of the average return is $\frac{\sigma_n^2}{T}$ and, thus, the standard deviation of the expected return is $\frac{\sigma_n}{\sqrt{T}}$.

The quantity $\frac{\sigma_n}{\sqrt{T}}$ is quite large because:

- For individual stocks, σ_n is about 0.30 p.a.
- Thus, for nine years of data (meaning T=9) $\frac{\sigma_n}{\sqrt{T}} = \frac{0.30}{\sqrt{9}} = 10\%$
- Thus, a 95% confidence interval (± 2 standard deviations) will be a band of about $\pm 2 * 10\% = 4 * 10\% = 40\%$
-

Given that expected returns are of the order of 10%, the 95% confidence interval is

$$0.10 \pm (2 * 10\%) = 0.10 \pm 0.20 = -10\% \; to \; 30\%$$

Such a large confidence interval implies an estimate that is so imprecise that it is useless for a practical purpose.

2.2 Estimation Error in the Covariance Matrix ($\mathbb{V}[\boldsymbol{R}]$)

The variance-covariance matrix of returns, $\mathbb{V}[R]$, poses its own set of estimation challenges. This matrix contains a large number of elements — $N(N+1)/2$, to be precise — each of which must be estimated from historical data. For a portfolio of N assets, the number of covariance terms increases rapidly, exacerbating the problem of estimation error.

The estimated covariance matrix is often ill-conditioned and close to being singular. An ill-conditioned matrix has a high condition number, meaning small changes in input can cause large changes in the output, making it unreliable for inversion, which is essential for calculating optimal portfolio weights. As N increases, the number of unique elements in the covariance matrix grows, requiring an impractically large amount of data to estimate accurately. For instance, for a portfolio of N=50 assets, the total number of covariance terms is 1275, which, if one were using monthly data, would require $T > \frac{1275}{12} = 107$ years of monthly data.

For a portfolio of N=100 assets, the total number of variance and covariance terms is 5050, which would require:

- $T > \frac{5050}{12} = 421$ years of monthly data, or
- $T > \frac{5050}{250} = 21$ years of daily data

Thus, the estimated covariance matrix is ill-conditioned because relative to the number of data points, the number of parameters to be estimated is very large.
The inaccuracies in estimating $\mathbb{E}[R]$ and $\mathbb{V}[R]$ have a pronounced effect on portfolio weights. Sample-based estimates of these parameters often lead to unreasonable portfolio weights—extremely large or small positions that fluctuate significantly over time. Such portfolios typically perform poorly out of sample, highlighting the detrimental effect of estimation error.

2.3 Performance of Mean-Variance Portfolio

De Miguel, Garlappi, and Uppal (2009) conducted a comprehensive evaluation of the out-of-sample performance of the sample-based mean-variance portfolio rule. They compared it to a naïve diversification strategy, where an equal fraction of wealth is allocated to each of the N available assets. This naïve rule, despite its simplicity, serves as a valuable benchmark because it does not rely on the estimation of moments or optimization and is easy to implement.

Their findings indicate that unconstrained policies incorporating estimation error often perform worse than strategies that constrain short sales and even worse than the 1/N strategy. Imposing constraints on the sample-based mean-variance strategy results in only modest improvements in Sharpe Ratios and certainty-equivalent (CEQ) returns, though it does significantly reduce turnover.

2.3.1 Critical Estimation Window

One critical insight from their research is the length of the estimation window required for the sample-based mean-variance strategy to outperform the 1/N strategy. The critical estimation-window length depends on the number of assets, the ex-ante Sharpe ratio of the mean-variance portfolio, and the Sharpe ratio of the 1/N policy. For US stock-market data, they found that the critical length is about 3000 months for a portfolio with 25 assets and over 6000 months for a portfolio with 50 assets. These findings starkly contrast with the much shorter estimation windows typically used in practice, often only 60 or 120 months of data.

2.3.2 Results

Their simulation results suggest that portfolio strategies from optimizing models are expected to outperform the 1/N benchmark only if the estimation window is sufficiently long, the ex-ante Sharpe ratio of the mean-variance efficient portfolio is significantly higher than that of the 1/N portfolio, and the number of assets is relatively small. The latter condition is intuitive, as fewer assets imply fewer parameters to estimate and thus less room for estimation error. Additionally, a smaller number of assets makes naive diversification less effective relative to optimal diversification.

The practical implications of these findings are profound. The allocation errors introduced by using 1/N weights can be smaller than those caused by using weights from an optimizing model with inputs estimated with significant error. De Miguel, Garlappi, and Uppal's study emphasizes that the severity of estimation error can entirely erode the benefits of optimal diversification.

They use three performance criteria in their analysis: the out-of-sample Sharpe Ratio, the CEQ return for the expected utility of mean-variance investors, and portfolio turnover. Their results reveal a substantial difference between the Sharpe Ratio of the in-sample mean-variance strategy (which has no estimation error) and the 1/N strategy. This difference is significant, underscoring the impact of estimation error.

Strategy	S&P sectors $N = 11$	Industry portfolios $N = 11$	Inter'l portfolios $N = 9$	Mkt/ SMB/HML $N = 3$	FF 1-factor $N = 21$	FF 4-factor $N = 24$
$1/N$	0.1876	0.1353	0.1277	0.2240	0.1623	0.1753
mv (in sample)	0.3848	0.2124	0.2090	0.2851	0.5098	0.5364
mv	0.0794	0.0679	−0.0332	0.2186	0.0128	0.1841

Figure 5. Sharpe Ratio Comparison. Data from De Miguel, Garlappi, Uppal (2009)

Strategy	S&P sectors $N = 11$	Industry portfolios $N = 11$	Inter'l portfolios $N = 9$	Mkt/ SMB/HML $N = 3$	FF 1-factor $N = 21$	FF 4-factor $N = 24$
$1/N$	0.0069	0.0050	0.0046	0.0039	0.0073	0.0072
mv (in sample)	0.0478	0.0106	0.0096	0.0047	0.0300	0.0304
mv	0.0031	−0.7816	−0.1365	0.0045	−2.7142	−0.0829

Figure 6. CEQ Comparison. Data from De Miguel, Garlappi, Uppal (2009)

Strategy	S&P sectors $N = 11$	Industry portfolios $N = 11$	Inter'l portfolios $N = 9$	Mkt/ SMB/HML $N = 3$	FF- 1-factor $N = 21$	FF- 4-factor $N = 24$
$1/N$	0.0305	0.0216	0.0293	0.0237	0.0162	0.0198
			Panel A: Relative turnover of each strategy			
mv (in sample)	–	–	–	–	–	–
mv	38.99	606594.36	4475.81	2.83	10466.10	3553.03

Figure 7. Turnover Comparison. Data from De Miguel, Garlappi, Uppal (2009)

Furthermore, the difference between the in-sample and out-of-sample Sharpe Ratios for the mean-variance strategy highlights the severity of estimation error. For all datasets examined, the sample-based mean-variance strategy has a substantially lower Sharpe Ratio out-of-sample than in-sample. Moreover, the out-of-sample Sharpe Ratio for the sample-based mean-variance strategy is often less than that of the 1/N strategy, indicating that estimation errors can completely negate the advantages of optimal diversification.

The journey through estimation error in portfolio optimization reveals a sobering reality: classical mean-variance optimization, while theoretically robust, suffers significantly in practice due to the inaccuracies in estimating expected returns and the variance-covariance matrix. Merton's insights underscore the inherent difficulty in improving the precision of expected return estimates, and the findings of De Miguel, Garlappi, and Uppal illustrate the substantial impact of these estimation errors on portfolio performance. This chapter has highlighted the critical need for robust methods to mitigate estimation errors, suggesting that simple strategies like equal-weighted portfolios often outperform more sophisticated models out-of-sample. The next chapter will explore alternative approaches and advanced techniques aimed at improving the robustness of portfolio optimization in the face of estimation errors, providing practical solutions for achieving better investment outcomes.

Chapter 3. Proposed Solution to Estimation Errors

3.1 Bayesian and Non-Bayesian Models

The poor performance of the Markowitz modek was recognized as early as the 1970s. Considerable research effort has been devoted to dealing with estimation error. Below, some of the approaches proposed to deal with estimation error, divided into "Bayesian" and "Non-Bayesian" approaches. Bayesian models interpret probability as a degree of belief, incorporating prior knowledge and updating it with new data using Bayes' Theorem, resulting in posterior distributions of parameters. Non-Bayesian models, or frequentist models, treat probability as long-run frequencies and estimate parameters purely from observed data without using priors, typically providing point estimates and confidence intervals. Bayesian models offer a dynamic approach to uncertainty quantification but are computationally intensive, while non-Bayesian models are simpler, faster, and focus on objective inference.

Non-Bayesian models for dealing with estimation error:

1. Shortselling constraints:
   - Frost and Savarino (1988), Chopra (1993), and Jagannathan and Ma (2003) "Robust" portfolio optimization:
   - Goldfarb and Iyengar (2003), Uppal and Wang (2003), Garlappi, Uppal, and Wang (2007), Boyle, Garlappi, Uppal, and Wang (2012)
2. Optimally diversity across market and estimation risk:
   - Kan and Zhou (2005)
3. Resampling methods:
   - Michaud (1998); discussed in Scherer (2002) and Harvey, Liechty, Liechty, and Muller (2003)

Bayesian models for dealing with estimation error:

1. Purely statistical approaches relying on diffuse-priors:

- Barry (1974) and Bawa, Brown, and Klein (1979)

2. Shrinkage estimators:
   - Jobson, Korkie, and Ratti (1979), Jobson and Korkie (1980), and Jorion (1985)

3. Methods to reduce error in estimating the covariance matrix:
   - Best and Grauer (1992), Chan, Karceski, and Lakonishok (1999), Ledoit and Wolf (2004a, 2004b)

4. Black-Litterman (1990, 1992)

The central idea for dealing with estimation error is shrinkage. All these methods can be interpreted in terms of shrinking the estimated $\mathbb{E}[R]$ and $\mathbb{V}[R]$ towards reasonable quantity.

For example, if you are not sure of your estimate of the expected return on a particular risky asset, then one way to reduce estimation error is to shrink the estimated expected return of this asset toward a sensible value, such as the average of all expected returns or the expected return on the market portfolio.

If you are not sure of your estimate of the return covariance matrix, then one way to reduce estimation error is to shrink the estimated covariance matrix toward a sensible matrix, such as the covariance matrix in a one-factor (market) model or the covariance matrix where all the covariances are equal.

3.2 The Importance of Short-Sale Contraints

Incorporating short-sale constraints in the mean-variance optimization framework addresses critical issues arising from estimation errors in expected returns and the covariance matrix. These estimation errors can significantly impact portfolio performance, leading to suboptimal asset allocations that deviate from the intended risk-return profile. Imposing a short-sale constraint effectively "shrinks" the expected returns towards the overall average, mitigating the undue influence of extreme values that often result from sampling variability.

Empirical evidence consistently demonstrates that mean-variance portfolios with short-sale constraints outperform their unconstrained counterparts. This improvement in performance is attributed to the constraints' ability to curb exaggerated positions driven by estimation noise, thereby enhancing the stability and robustness of the portfolio's returns. Moreover, short-sale constraints contribute to significantly lower portfolio turnover, as they prevent the over-reliance on aggressive trading strategies that could otherwise amplify transaction costs and erode returns.

Additionally, the implementation of short-sale constraints is straightforward, making it accessible to a broader range of investors, including retail participants who may lack the sophisticated tools or capital requirements necessary for short-selling. This accessibility, combined with the performance benefits and turnover reduction, underscores the practical value of short-sale constraints as a strategic enhancement to the traditional mean-variance optimization approach.

3.2.1 Demonstration of Short-Sale Constraints Leading to Shrinkage

Let's show mathematically that short-sale constraints lead to shrinkage. We've already seen that an investor can obtain the optimal portfolio weights by maximizing mean-variance utility:

$$max_w \; MVU = \left[R_f + w^T\left(\mu - R_f 1_N\right)\right] - \frac{\gamma}{2} w^T V w$$

When we implement short-sale constraints, the above problem becomes:

$$max_w \; MVU = \left[R_f + w^T\left(\mu - R_f 1_N\right)\right] - \frac{\gamma}{2} w^T V w$$

$$s.t. \qquad w_n \geq 0, \forall \; n \; \in [1, N]$$

Imposing short-sale constraints implies that all the weights must be positive, meaning the investor can go only long (buy) on the assets.
Setting short-sale constraints in the mean-variance optimization problem yields the following Lagrangian:

$$\mathcal{L} = \left[R_f + w^T\left(\mu - R_f 1_N\right)\right] - \frac{\gamma}{2} w^T V w + (\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_N w_N) =$$
$$= \left[R_f + w^T\left(\mu - R_f 1_N\right)\right] - \frac{\gamma}{2} w^T V w + w^T \lambda_{ss}$$

Where λ_{ss} referes to the vector of Lagrange multipliers for the N constraints on short-selling (ss).

$$\mathcal{L} = \left[R_f + w^T\left(\mu + \lambda_{ss} - R_f 1_N\right)\right] - \frac{\gamma}{2} w^T V w$$

Thus, the constrained mean-variance weights are equivalent to the unconstrained weights but with the adjusted mean vector $\mu + \lambda_{ss}$
To understand why the above expression implies a shrinkage of expected returns, note that the short-sale constraint on asset *n* is likedly to be binding when its expected return μ_n is low. When the constraint for a particular asset *n* binds, $\lambda_n > 0$ and the expected return is increased from μ_n to $\mu_n + \lambda_n$. Hence, imposing a short-sale constraint on the sample-based mean-variance problem is equivalent to "shrinking" the expected return toward the average.

3.3 Global Minimum Variance Portfolio

In the next section, we explore the Global Minimum Variance (GMV) portfolio as an alternative approach to addressing the limitations of mean-variance optimization. Extensive academic research has shown that portfolios relying on sample estimates of expected returns often perform poorly, frequently failing to consistently outperform the much simpler equally-weighted portfolio, even when adjustments for estimation errors are made through Bayesian and non-Bayesian methods. By focusing on the GMV portfolio, we avoid the challenge of precisely estimating expected returns (E[R]), thereby refining the optimization process. This approach aligns with findings from notable studies, including DeMiguel, Garlappi, and Uppal (2009), and Jacobs, Müller, and Weber (2014).

Thus, instead of solving the following optimization problem:

$$Minimize_{\{w_1,\ldots,w_N\}}\ \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n\, w_m \sigma_{nm} = w^T V w$$

$subject\ to$:

$$(1)\ \sum_{n=1}^{N} w_n = w^T 1_N = 1;$$

$$(2)\ \sum_{n=1}^{N} w_n \mathbb{E}[R_n] = w^T \mu = \mu_{target}$$

One could assume that expected returns for all N assets are equal. Therefore, in the optimization problem, $\mathbb{E}[R]$ can be ignored. Thus, constraint (2) can be removed from the optimization problem. This reduced the above optimization problem to the following:

$$Minimize_{\{w_1,\ldots,w_N\}}\ \sigma_p^2 = \sum_{n=1}^{N}\sum_{m=1}^{N} w_n\, w_m \sigma_{nm} = w^T V w$$

$subject\ to$:

$$(1)\ \sum_{n=1}^{N} w_n = w^T 1_N = 1;$$

Let’s solve the new optimization problem. As always, let’s apply the Lagrangian method:

$$\mathcal{L} = w^T V w + \lambda(1 - w^T 1_N)$$

Applying the FOC (First Order Condition):

$$\frac{\partial \mathcal{L}}{\partial w} = 2Vw - \lambda 1_N = 0$$

$$2Vw = \lambda 1_N$$

$$w = \frac{\lambda}{2} V^{-1} 1_N$$

Now, let's substitute the above result into the constraint to find λ:

$$w^T 1_N = 1$$

$$\left(\frac{\lambda}{2} V^{-1} 1_N\right)^T 1_N = 1$$

$$\frac{\lambda}{2} 1_N^T V^{-1} 1_N = 1$$

$$\lambda = \frac{2}{1_N^T V^{-1} 1_N}$$

Finally, substituting λ back into w:

$$w = \frac{\frac{2}{1_N^T V^{-1} 1_N}}{2} V^{-1} w$$

Thus, the vector of optimal weights of the Global Minimum Variance portfolio is:

$$w = \frac{V^{-1} 1_N}{1_N^T V^{-1} 1_N}$$

The GMV portfolio does not depend on expected returns, which we know cannot be estimated precisely. At the same time, the weights of the GMV portfolio are optimized to have the lowest risk, which can be estimated precisely. So, possibly, these portfolios will perform well out of sample.

3.3.1 Interpreting the Weights on the GMV Portfolio

To understand the minimum-variance portfolio weights, we study the special case with zero correlation between the returns of all assets. That is, instead of the general covariance matrix

$$V = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \dots & \sigma_N^2 \end{bmatrix}$$

We consider the special case of a diagonal matrix

$$V = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_N^2 \end{bmatrix}$$

For the special case where V is a diagonal matrix, its inverse becomes simply:

$$V^{-1} = \begin{bmatrix} 1/\sigma_1^2 & 0 & \dots & 0 \\ 0 & 1/\sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/\sigma_N^2 \end{bmatrix}$$

By simply substituting the above result into the vector of optimal weights, we find that:

$$w_n = \frac{\frac{1}{\sigma_n^2}}{\sum_{n=1}^{N} \frac{1}{\sigma_n^2}}$$

This formula shows that there is an inverse relation between the variance of an asset and its weight in the portfolio.

The GMV portfolio weights will be reasonable if the elements of the covariance matrix V can be estimated precisely. We already stated that the elements of the covariance matrix can be estimated more precisely than expected excess returns.

Having concluded the discussion on models that address errors in estimating expected returns, we now shift our focus to those that aim to reduce estimation errors in the return covariance matrix—a critical factor influencing the risk profile and diversification potential of a portfolio. In scenarios involving a large number of assets, even precise estimates of individual elements of the variance-covariance matrix may not prevent the matrix inversion from becoming ill-conditioned, leading to unstable and unreliable optimization outcomes. Imposing short-sale constraints can effectively mitigate this issue by enhancing the conditioning of the inverse matrix, thereby improving the overall stability of the portfolio construction process.

3.4 Short-Sale Constraints to Manage Ill-Conditioned Variance-Covariance Matrix

Now we need to address the second problem of an ill-conditioned variance-covariance matrix. To solve this issue, Jagannathan and Ma (2003) study the effect of imposing short-sale constraints on the global minimum-variance portfolio. Shortsales-constrained GMV portfolio is:

$$min_w \quad w^T \, V \, w$$
$$s.t. \quad w^T 1_N = 1$$
$$w_n \geq 0, \quad \forall \, n \, \epsilon \, [1, \dots, N]$$

The solution coincides with unconstrained GMV portfolio

$$min_w \quad w^T \, V_{JM} \, w$$
$$s.t. \quad w^T 1_N = 1$$

where the sample covariance matrix has been replaced by V_{JM} (JM stands for Jagannathan and Ma), the variance-covariance matrix after shrinkage:

$$V_{JM} = V - \lambda_{ss}\, 1_N^T - 1_N\, \lambda_{ss}^T$$

Note that the elements of the vector of Lagrange multipliers can take only non-negative values, meaning $\lambda_{ss} \geq 0_N$. Therefore, the matrix V_{JM} can be interpreted as the sample covariance matrix after shrinkage, because if the short-sale constraint for the n^{th} asset is binding $(\lambda_{ss,n} > 0)$, then the sample covariance of this asset with any other asset is reduced by $\lambda_{ss,n}$.

Chapter 4. Performance Analysis

In this chapter, I aim to evaluate the performance of four portfolio optimization strategies: the Mean-Variance Portfolio (MV), the Global Minimum Variance Portfolio (GMV), the Mean-Variance Portfolio with Short-Sale Constraints (MV-C), and the Global Minimum Variance Portfolio with Short-Sale Constraints (GMV-C). These strategies, grounded in the framework introduced by Harry Markowitz, have been widely utilized in portfolio theory but present practical challenges, particularly due to estimation errors in expected returns and the covariance matrix.

4.1 Methodology Used

To address these challenges, I apply a rolling-window approach, which enables a dynamic assessment of each strategy's performance across various market conditions over time. The rolling-window method ensures that the results are not influenced by static market conditions, but rather evolve in response to changing data.

The core of this methodology lies in the monthly re-estimation of portfolio weights based on historical returns. I selected a 60-month (5-year) rolling window to estimate the required parameters—expected returns and the covariance matrix for MV portfolios, and only the covariance matrix for GMV portfolios. The 60-month window strikes a balance between using sufficient historical data to make informed estimates and maintaining the adaptability to evolving market conditions. After each window, I calculate the portfolio weights and apply them to out-of-sample data for the next month to assess performance. This process continues by rolling forward the window by one month and repeating the estimation and portfolio construction steps until the end of the dataset is reached.

4.1.1 Step-by-Step Methodology Algorithm

1. Choose the Rolling Window:

A rolling window of 60 months was selected. This length ensures that the parameter estimation is based on a significant amount of data, while remaining sensitive to market changes.

2. Estimate Parameters:

For each rolling window, I estimate the expected returns (for MV and MV-C portfolios) and the covariance matrix (for all strategies). In the case of GMV and GMV-C portfolios, I can ignore the expected return estimation to focus solely on minimizing portfolio variance. The expected return estimates for the MV strategies are computed as the historical average return of each asset within the 60-month window.

3. Impose Short-Sale Constraints (if applicable):

For the MV-C and GMV-C strategies, short-sale constraints are introduced. These constraints prevent negative portfolio weights, ensuring that no asset is shorted. This step aims to reduce the potential for extreme portfolio allocations, which are often the result of estimation errors, especially in expected returns.

4. Compute Portfolio Weights:

Using the estimated parameters, I calculate the portfolio weights that would maximize the Sharpe Ratio for the MV portfolios, and minimize variance for the GMV portfolios. For constrained portfolios (MV-C and GMV-C), the optimization problem includes the short-sale constraint, which limits portfolio weights to non-negative values.

5. Out-of-Sample Performance Calculation:

Once the portfolio weights are determined, they are applied to the out-of-sample returns of the next month (i.e., the month immediately following the rolling window). The out-of-sample returns are then recorded for each strategy.

6. Roll the Window Forward:

The window is shifted forward by one month, and the oldest month's data is dropped, while the next month's data is added to the window. This process continues until the end of the dataset is reached, generating a series of out-of-sample returns for each strategy.

7. Performance Metrics:

Using the series of out-of-sample returns generated for each strategy, I compute three key performance metrics:

- Sharpe Ratio: To evaluate risk-adjusted returns, using both the excess return over the risk-free rate and the standard deviation of returns.
- Sortino Ratio: A variation of the Sharpe Ratio, focusing on downside risk by penalizing only negative returns.
- Turnover Rate: To measure portfolio stability and transaction frequency, reflecting the extent of portfolio rebalancing across different strategies.

4.2 Datasets Selected

The datasets used in this analysis consist of 13 different indices and sector-specific exchange-traded funds (ETFs) spanning global markets and major economic sectors. This diverse selection allows for a comprehensive analysis of portfolio performance across varying market environments, geographic regions, and industries. By using daily data from January 1, 2000, to December 31, 2019, I ensured that the dataset covers a wide range of market conditions, including periods of both bull and bear markets.

The indices and ETFs were chosen with the goal of capturing broad market movements while also analyzing performance in specific sectors. Each dataset reflects a different facet of global economic activity, which is essential for demonstrating how the portfolio optimization strategies perform under different conditions.

Here is a detailed explanation of each index and ETF used:

1. S&P 500 (SPX) – The S&P 500 is one of the most widely recognized equity indices in the world, representing 500 of the largest publicly traded companies in the United States. It spans multiple sectors, making it a comprehensive benchmark for U.S. equity markets. The S&P 500 provides exposure to large-cap U.S. companies, making it an ideal benchmark for evaluating the performance of portfolio strategies in one of the most important global markets. Its breadth allows for the analysis of sectoral diversification and performance within the U.S. economy.

2. Euro Stoxx 50 (SX5E) – This index represents 50 of the largest blue-chip companies from 11 Eurozone countries. It is the leading benchmark for European equity markets. By including Euro Stoxx 50, I capture the dynamics of the Eurozone economy, providing geographic diversification in the portfolio analysis. Europe has experienced significant economic shifts during the period of analysis, which adds valuable insights into the performance of the strategies under varying market conditions.

3. S&P Asia 50 (SPAS50) – The S&P Asia 50 Index tracks 50 leading companies from Hong Kong, South Korea, Singapore, and Taiwan. It is a key indicator of the performance of Asia's largest economies. The inclusion of this index introduces exposure to the Asia-Pacific region, which has been one of the fastest-growing economic areas during the study period. This geographic diversification is critical for analyzing the effectiveness of the portfolio strategies in rapidly developing markets.

4. Financial Select Sector SPDR Fund (XLF) – XLF tracks the financial sector of the S&P 500, which includes banks, insurance companies, and real estate firms. This sector is heavily influenced by interest rates and financial regulations. Given the importance of the financial sector in global markets, XLF offers targeted exposure to the U.S. financial system, making it a key sector for understanding how financial institutions' stock prices react to economic changes.

5. Technology Select Sector SPDR Fund (XLK) – XLK focuses on technology companies within the S&P 500, encompassing hardware, software, IT services, and semiconductor companies. The technology sector has been a major driver of global equity returns over the past two decades. Analyzing this sector provides insight into how portfolio strategies perform in one of the most dynamic and high-growth industries.

6. Consumer Discretionary Select Sector SPDR Fund (XLY) – XLY provides exposure to companies in the consumer discretionary sector, which includes non-essential goods and services such as retail, automobiles, and travel. This sector is highly cyclical, with its performance closely tied to consumer spending patterns. Including XLY allows for the analysis of how different portfolio strategies fare in a consumer-driven market.

7. Industrial Select Sector SPDR Fund (XLI) – XLI tracks the industrial sector, covering companies involved in manufacturing, transportation, and infrastructure. The industrial sector is often correlated with broader economic growth, making it a valuable component for analyzing how portfolio strategies respond to macroeconomic factors such as GDP growth and industrial production.

8. Health Care Select Sector SPDR Fund (XLV) – XLV provides exposure to the healthcare sector, including pharmaceutical companies, biotechnology firms, and healthcare providers. Healthcare is a defensive sector, often showing resilience during economic downturns. By including XLV, the analysis can capture how portfolio strategies perform in a sector that is less sensitive to economic cycles.

9. Energy Select Sector SPDR Fund (XLE) – XLE tracks the energy sector, which includes companies involved in oil, gas, and renewable energy.
- Why Chosen: Energy is one of the most volatile sectors, often impacted by geopolitical events and commodity prices. Including XLE helps assess how portfolio strategies handle high volatility and sector-specific risk.

10. Consumer Staples Select Sector SPDR Fund (XLP) – XLP focuses on companies that produce essential goods like food, beverages, and household products. The consumer staples sector tends to perform well during periods of economic uncertainty, as demand for these products is relatively inelastic. This makes it a critical sector for understanding how portfolio strategies perform in more stable, less volatile sectors.

11. Utilities Select Sector SPDR Fund (XLU) – XLU provides exposure to utilities companies that offer essential services such as electricity, gas, and water. Utilities are known for their stability and dividend yield, making them less susceptible to market fluctuations. XLU offers insight into how portfolio strategies perform in sectors that prioritize income and stability.

12. Real Estate Select Sector SPDR Fund (VNQ) – VNQ focuses on real estate investment trusts (REITs) and other real estate-related businesses. Real estate serves as a hedge against inflation and provides diversification within a portfolio. VNQ allows for analysis of how the strategies perform in this asset class, which has different risk-return characteristics compared to equities.

13. Materials Select Sector SPDR Fund (XLB) – XLB tracks companies in the materials sector, including chemicals, metals, and forestry. The materials sector is heavily influenced by global demand for raw materials and commodity prices, providing exposure to industries sensitive to economic and industrial growth.

4.3 Performance Metrics

To thoroughly evaluate the performance of the four portfolio strategies, I utilized three key metrics: the Sharpe Ratio, Sortino Ratio, and Turnover Rate. Each of these metrics provides unique insights into different aspects of portfolio performance, from risk-adjusted returns to stability and trading activity.

4.3.1 Sharpe Ratio

The Sharpe Ratio is one of the most widely used performance metrics in finance. It measures the risk-adjusted return of a portfolio by comparing its excess return (over the risk-free rate) to its volatility, represented by the standard deviation of portfolio returns. The Sharpe Ratio is computed as follows:

$$Sharpe\ Ratio = \frac{\mathbb{E}[R_p - R_f]}{\sigma_p}$$

Where:
- $\mathbb{E}[R_p]$ is the expected return of the portfolio.
- R_f is the risk-free rate of return.
- σ_p is the standard deviation of the portfolio's excess returns.

The higher the Sharpe Ratio, the better the portfolio's return relative to the risk taken. This makes the Sharpe Ratio an ideal metric for comparing portfolios with differing levels of risk. In this thesis, the Sharpe Ratio is used to assess the risk-adjusted returns of each of the four strategies across different indices and ETFs. The results highlight the performance differences between the traditional Mean-Variance approach, the Global Minimum Variance approach, and the impact of imposing short-sale constraints.

4.3.2 Sortino Ratio

The Sortino Ratio is a variation of the Sharpe Ratio that focuses specifically on downside risk, which is often more relevant for investors concerned with avoiding losses rather than just minimizing volatility. Instead of using the standard deviation of all returns as a measure of risk, the Sortino Ratio only considers negative returns, using downside deviation as the risk measure.

It is calculated as follows:

$$Sortino\ Ratio = \frac{\mathbb{E}[R_p - R_f]}{Downside\ Deviation}$$

Where:
- The downside deviation measures the volatility of returns that fall below a certain threshold, often the risk-free rate or 0.

By focusing on negative returns, the Sortino Ratio provides a more accurate measure of risk-adjusted performance for portfolios that seek to minimize downside risk. It is especially useful when evaluating strategies like the Mean-Variance portfolio, which can be overly sensitive to estimation errors and may result in extreme allocations that increase downside risk. In contrast, portfolios constrained by short-sales, such as the MV-C and GMV-C strategies, often improve downside risk-adjusted returns, as shown by higher Sortino Ratios.

4.3.3 Turnover Rate

The Turnover Rate is an important metric that captures the level of trading activity in a portfolio, defined as the frequency with which the portfolio's assets are bought and sold. A high Turnover Rate indicates frequent portfolio rebalancing, while a low Turnover Rate suggests a more stable portfolio composition. The Turnover Rate is computed as:

$$Turnover\ Rate = \frac{Total\ value\ of\ trader\ over\ a\ period\ of\ time}{Total\ portfolio\ value}$$

High turnover can result in higher transaction costs, which can erode portfolio returns over time. It also reflects portfolio instability, as frequent rebalancing suggests that the portfolio is highly reactive to changes in the estimated parameters. For this reason, turnover is a key consideration when evaluating the practical implementation of portfolio strategies. In this thesis, I analyze the Turnover Rate to compare the stability of each strategy.

- MV Portfolio: This strategy typically has a higher Turnover Rate due to its sensitivity to changes in the expected return and covariance matrix estimates. Frequent rebalancing may introduce significant transaction costs.
- GMV Portfolio: Since this strategy minimizes portfolio variance without estimating expected returns, it generally results in lower Turnover Rates, reflecting greater portfolio stability.
- MV-C and GMV-C Portfolios: By imposing short-sale constraints, both of these strategies limit extreme portfolio allocations, which not only improves risk-adjusted performance but also reduces trading frequency. The lower Turnover Rate observed in the constrained strategies underscores the practical advantages of short-sale constraints in managing transaction costs and portfolio stability.

4.4 Performance Analysis

In this section, I analyze the performance of the four portfolio strategies—Mean-Variance (MV), Global Minimum Variance (GMV), Mean-Variance with Short-Sale Constraints (MV-C), and Global Minimum Variance with Short-Sale Constraints (GMV-C)—based on the performance metrics outlined earlier: Sharpe Ratio, Sortino Ratio, and Turnover Rate. The analysis focuses on identifying trends in the performance of each strategy across different indices and sectors, as well as understanding how the introduction of short-sale constraints impacts portfolio efficiency and risk-adjusted returns.

Sharpe Ratio	MV	GMV	MV-C	GMV-C
S&P500	0,4624	0,7217	0,9683	0,7543
Euro Stoxx 50	0,2749	0,4391	0,4998	0,4707
S&P Asia 50	0,1548	0,2932	0,6422	0,2903
XLF	0,2887	0,2918	0,3563	0,4362
XLK	-0,0589	0,4826	0,8545	0,5343
XLY	0,3841	0,7267	0,6599	0,8171
XLI	0,2283	0,4705	0,6372	0,5433
XLV	0,1694	0,5887	0,6833	0,7197
XLE	-0,1508	0,1513	0,1312	0,2829
XLP	0,2904	0,6289	0,5313	0,7232
XLU	0,4444	0,3355	0,6421	0,5894
VNQ	0,3657	0,3991	0,4403	0,4745
XLB	0,4813	0,6339	0,4829	0,6126

Figure 8. Results table for Sharpe Ratio

4.4.1 Sharpe Ratio Analysis

The Sharpe Ratio measures the risk-adjusted return of each portfolio and is the primary metric for comparing the performance of the four strategies.

1. Mean-Variance Portfolio (MV):
Across all indices and sectors, the MV portfolio consistently underperforms relative to the other strategies. For example, the Sharpe Ratio for the MV strategy is only 0.4624 for the S&P 500, which is significantly lower than that of the other strategies. Similar underperformance is observed across other datasets, such as the Euro Stoxx 50 (0.2749) and the S&P Asia 50 (0.1548). The poor performance of the MV strategy can be attributed to its reliance on expected return estimates, which are often subject to substantial estimation errors, leading to suboptimal portfolio allocations and lower risk-adjusted returns.

2. Global Minimum Variance Portfolio (GMV):
The GMV portfolio, which focuses solely on minimizing portfolio variance without estimating expected returns, outperforms the MV portfolio in most cases. For example,

the GMV strategy achieves a Sharpe Ratio of 0.7217 for the S&P 500, a marked improvement over the MV strategy. This trend is consistent across 12 out of 13 datasets, with notable examples including the Euro Stoxx 50 (0.4391) and the S&P Asia 50 (0.2932). The GMV portfolio's superior performance highlights the advantages of minimizing variance, particularly in cases where return estimates are unreliable or inaccurate.

3. Mean-Variance Portfolio with Short-Sale Constraints (MV-C):
The introduction of short-sale constraints significantly improves the performance of the MV portfolio. The MV-C strategy consistently outperforms the MV strategy in all 13 datasets. For example, in the S&P 500, the Sharpe Ratio for MV-C is 0.9683, which is a substantial improvement over the unconstrained MV portfolio. Similarly, the Euro Stoxx 50 achieves a Sharpe Ratio of 0.4998 for the MV-C strategy, compared to 0.2749 for the MV strategy. This improvement is due to the short-sale constraints limiting extreme portfolio weights, thereby mitigating estimation errors and improving risk-adjusted returns.

4. Global Minimum Variance Portfolio with Short-Sale Constraints (GMV-C):
The GMV-C strategy shows mixed results when compared to the GMV strategy. While the GMV-C strategy outperforms GMV in some datasets, such as the S&P 500 (0.7543) and XLY (0.8171), it underperforms in others, such as XLK (0.5343) and XLU (0.5894). The mixed performance suggests that while short-sale constraints are effective in managing estimation errors, the benefits of these constraints may vary depending on the dataset or sector. However, the GMV-C strategy still provides a viable alternative, especially in cases where downside risk or extreme portfolio allocations are a concern.

Sortino Ratio	MV	GMV	MV-C	GMV-C
S&P500	0,1689	0,2873	0,4426	0,3033
Euro Stoxx 50	0,1012	0,1684	0,2047	0,1835
S&P Asia 50	0,0483	0,1567	0,3049	0,1495
XLF	0,1179	0,0931	0,1351	0,1481
XLK	-0,0117	0,2067	0,3248	0,2273
XLY	0,1351	0,2893	0,2811	0,3269
XLI	0,1338	0,1811	0,2353	0,1962
XLV	0,0555	0,2535	0,3148	0,2875
XLE	-0,0312	0,0623	0,0359	0,1076
XLP	0,1098	0,2171	0,1941	0,2529
XLU	0,1829	0,1173	0,2377	0,2258
VNQ	0,1776	0,1528	0,1768	0,1621
XLB	0,2026	0,2551	0,1867	0,2197

Figure 9. Results table for Sortino Ratio

4.4.2 Sortino Ratio Analysis

The Sortino Ratio offers additional insights by focusing on downside risk-adjusted returns.

1. Mean-Variance Portfolio (MV):
The MV strategy once again shows the weakest performance in terms of downside risk-adjusted returns. For example, the Sortino Ratio for the S&P 500 is only 0.1689, reflecting the MV portfolio's vulnerability to downside risk. The underperformance of the MV strategy in terms of the Sortino Ratio mirrors its poor performance in the Sharpe Ratio, confirming that it is not well-suited to handle downside risk, particularly when expected returns are prone to estimation errors.

2. Global Minimum Variance Portfolio (GMV):
The GMV portfolio outperforms the MV strategy in 10 out of 13 datasets based on the Sortino Ratio. For instance, the Sortino Ratio for the S&P 500 is 0.2873, compared to 0.1689 for the MV strategy. Similar improvements are seen across other datasets, including the Euro Stoxx 50 (0.1684) and the S&P Asia 50 (0.1567). These results demonstrate the effectiveness of variance minimization in controlling downside risk.

3. Mean-Variance Portfolio with Short-Sale Constraints (MV-C):

The MV-C strategy significantly improves downside risk-adjusted returns compared to the unconstrained MV portfolio. The Sortino Ratio for the S&P 500 rises to 0.4426 for the MV-C strategy, compared to only 0.1689 for the MV strategy. Similar improvements are observed across all datasets, confirming that short-sale constraints help mitigate downside risk by preventing extreme negative allocations.

4. Global Minimum Variance Portfolio with Short-Sale Constraints (GMV-C):

The GMV-C strategy performs well in terms of downside risk-adjusted returns, though its superiority over the GMV strategy is less consistent. For example, the Sortino Ratio for the S&P 500 is 0.3033 for GMV-C, slightly better than 0.2873 for GMV. However, in other cases, such as XLK, GMV-C underperforms GMV (0.2273 vs. 0.2067). While short-sale constraints generally improve downside protection, the benefits may vary by sector.

Turnover Rate	MV	GMV	MV-C	GMV-C
S&P500	0,4776	0,0569	0,0925	0,0362
Euro Stoxx 50	7,3749	0,0267	0,1111	0,0269
S&P Asia 50	12,5469	0,0241	0,0532	0,0239
XLF	0,5028	0,1458	0,1192	0,0523
XLK	5,6034	0,1121	0,0934	0,0669
XLY	0,2919	0,0937	0,0986	0,0649
XLI	1,8112	0,1541	0,1074	0,0699
XLV	0,5266	0,1116	0,1067	0,0689
XLE	7,9184	0,1239	0,1465	0,0625
XLP	0,5026	0,1163	0,0997	0,0549
XLU	1,0889	0,2739	0,1191	0,0816
VNQ	0,8222	0,1926	0,1549	0,0931
XLB	0,3684	0,1176	0,1109	0,0542

Figure 10. Results table for Turnover Rate

4.4.3 Turnover Rate Analysis

Turnover Rate provides insight into the trading activity and stability of each portfolio strategy.

1. Mean-Variance Portfolio (MV):
The MV portfolio exhibits the highest Turnover Rate across all datasets, reflecting its sensitivity to parameter estimates and frequent rebalancing. For example, the Turnover Rate for the S&P 500 is 0.4776, indicating that the portfolio is frequently adjusted in response to changes in expected returns and the covariance matrix. This high turnover introduces potential transaction costs, which could further erode the portfolio's net returns.

2. Global Minimum Variance Portfolio (GMV):
In contrast to the MV strategy, the GMV portfolio shows significantly lower Turnover Rates, highlighting its stability. For the S&P 500, the GMV strategy has a Turnover Rate of only 0.0569, suggesting much less frequent rebalancing. This stability reduces transaction costs and contributes to the GMV portfolio's overall efficiency.

3. Mean-Variance Portfolio with Short-Sale Constraints (MV-C):
The MV-C strategy shows a noticeable reduction in turnover compared to the unconstrained MV portfolio. For example, the Turnover Rate for the S&P 500 is 0.0925 for MV-C, compared to 0.4776 for MV. The introduction of short-sale constraints prevents extreme portfolio weight adjustments, which reduces trading frequency and helps control transaction costs.

4. Global Minimum Variance Portfolio with Short-Sale Constraints (GMV-C):
The GMV-C strategy exhibits the lowest Turnover Rates across most datasets. For the S&P 500, the Turnover Rate is 0.0362, reflecting the portfolio's high stability. Similar trends are observed across other datasets, confirming that short-sale constraints, when combined with the GMV strategy, provide the most stable portfolio with minimal trading activity.

4.4.4 Summary of Findings

1. MV Portfolio: The Mean-Variance strategy consistently underperforms due to its reliance on expected return estimates, which introduces significant estimation errors. It also exhibits high turnover, leading to greater transaction costs.
2. GMV Portfolio: The Global Minimum Variance portfolio generally outperforms the MV strategy by focusing on minimizing variance, resulting in higher risk-adjusted returns and lower turnover.
3. MV-C Portfolio: The introduction of short-sale constraints improves the performance of the MV strategy, both in terms of Sharpe and Sortino Ratios, while reducing turnover.
4. GMV-C Portfolio: While the GMV-C strategy shows mixed results relative to GMV, it provides excellent stability, with the lowest turnover rates across all strategies.

Conclusion

The findings of this thesis reveal significant insights into the practical application of portfolio optimization strategies, particularly in dealing with estimation errors that often plague the traditional mean-variance approach. The results show that the Mean-Variance Portfolio (MV), while theoretically sound, performs poorly in practice due to its reliance on highly uncertain estimates of expected returns. The MV strategy consistently exhibits lower Sharpe and Sortino Ratios, reflecting inferior risk-adjusted returns and vulnerability to downside risk. Additionally, its high Turnover Rate suggests excessive rebalancing, leading to increased transaction costs and portfolio instability.

In contrast, the Global Minimum Variance Portfolio (GMV), which focuses on minimizing variance without estimating expected returns, performs substantially better across most metrics. The GMV strategy consistently achieves higher Sharpe and Sortino Ratios compared to the MV portfolio, particularly in markets where estimation errors in returns are more pronounced. This improvement is largely due to its avoidance of extreme portfolio weights driven by inaccurate return forecasts. However, the GMV strategy still faces challenges in highly volatile markets, where variance alone may not fully capture the portfolio's risk exposure.

The introduction of short-sale constraints, particularly in the Mean-Variance with Short-Sale Constraints (MV-C) and Global Minimum Variance with Short-Sale Constraints (GMV-C) strategies, significantly enhances portfolio performance. The MV-C strategy shows a marked improvement in both risk-adjusted returns and downside risk protection, as evidenced by its superior Sharpe and Sortino Ratios across nearly all datasets. By restricting the portfolio to long-only positions, the short-sale constraints help mitigate the extreme asset weightings often induced by estimation errors in expected returns. This reduction in portfolio volatility, coupled with lower Turnover Rates, makes the constrained strategies more practical for real-world application, as they avoid excessive trading and the associated costs.

Among all the strategies, the Global Minimum Variance with Short-Sale Constraints (GMV-C) portfolio stands out for its stability and robustness. It consistently delivers the lowest Turnover Rates, indicating minimal rebalancing and reduced transaction costs, while also achieving competitive risk-adjusted returns. Although the GMV-C strategy does not always outperform the unconstrained GMV portfolio in terms of the Sharpe and Sortino Ratios, its enhanced stability and lower exposure to estimation errors make it a valuable alternative, particularly for investors seeking long-term, low-volatility investment strategies.

In conclusion, this thesis demonstrates that while traditional mean-variance optimization can be severely impacted by estimation errors, the Global Minimum Variance and short-sale constrained portfolios provide more robust and reliable alternatives. The GMV and GMV-C strategies, in particular, offer compelling solutions for investors looking to minimize risk and avoid the pitfalls of inaccurate return estimates. These findings highlight the importance of adopting strategies that are less reliant on precise estimates of expected returns, while also emphasizing the practical advantages of imposing constraints to ensure portfolio stability and reduce turnover. Overall, this research provides a strong foundation for improving portfolio construction in the face of uncertainty, offering insights that are both theoretically sound and practically applicable.

References

A. Craig MacKinlay, Ľuboš Pástor, Asset Pricing Models: Implications for Expected Returns and Portfolio Selection, *The Review of Financial Studies*, Volume 13, Issue 4, October 2000, Pages 883–916, https://doi.org/10.1093/rfs/13.4.883

Benartzi, S., & Thaler, R. H. (2001). Naive Diversification Strategies in Defined Contribution Saving Plans. *The American Economic Review*, *91*(1), 79–98. http://www.jstor.org/stable/2677899

Best, Michael J. (2010). *Portfolio Optimization*. Chapman & Hall/CRC Finance Series

Doron Avramov, Stock Return Predictability and Asset Pricing Models, *The Review of Financial Studies*, Volume 17, Issue 3, July 2004, Pages 699–738, https://doi.org/10.1093/rfs/hhg059

Frost, P. A., & Savarino, J. E. (1988). For better performance: Constrain portfolio weights. *Journal of portfolio management*, *15*(1), 29.Dangl, T., and M. Kashofer. 2013. *Minimum-variance stock picking – a shift in preferences for minimum-variance constituents*.

Goldfarb, D., & Iyengar, G. (2003). Robust portfolio selection problems. *Mathematics of operations research*, *28*(1), 1-38.

Green, R.C. and Hollifield, B. (1992) When Will Mean-Variance Efficient Portfolios Be Well Diversified? Journal of Finance, 47, 1785-1809.

Herzing, M. *Using the Lagrangian Method to Solve Optimization Problems*. Stockholm University, Department of Economics https://doi.org/10.1111/j.1540-6261.1992.tb04683.x

Jacobs, H., Müller, S., & Weber, M. (2014). How should individual investors diversify? An empirical evaluation of alternative asset allocation policies. *Journal of Financial Markets*, *19*, 62-85.

Jagannathan, R. and Ma, T. (2003), Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. The Journal of Finance, 58: 1651-1683. https://doi.org/10.1111/1540-6261.00580

Ledoit, O., & Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of empirical finance*, *10*(5), 603-621.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91. https://doi.org/10.2307/2975974

Mengmeng Ao, Li Yingying, Xinghua Zheng, Approaching Mean-Variance Efficiency for Large Portfolios, *The Review of Financial Studies*, Volume 32, Issue 7, July 2019, Pages 2890–2919, https://doi.org/10.1093/rfs/hhy105

Merton, R. C. (1972). An Analytic Derivation of the Efficient Portfolio Frontier. *The Journal of Financial and Quantitative Analysis*, *7*(4), 1851–1872. https://doi.org/10.2307/2329621

Merton, Robert C., *On Estimating the Expected Return on the Market: An Exploratory Investigation* (February 1980). NBER Working Paper No. w0444, Available at SSRN: https://ssrn.com/abstract=262067

Michael W. Brandt, Amit Goyal, Pedro Santa-Clara, Jonathan R. Stroud, A Simulation Approach to Dynamic Portfolio Choice with an Application to Learning About Return Predictability, *The Review of Financial Studies*, Volume 18, Issue 3, Fall 2005, Pages 831–873, https://doi.org/10.1093/rfs/hhi019

Olivier Ledoit, Michael Wolf, Nonlinear Shrinkage of the Covariance Matrix for Portfolio Selection: Markowitz Meets Goldilocks, *The Review of Financial Studies*, Volume 30, Issue 12, December 2017, Pages 4349–4388, https://doi.org/10.1093/rfs/hhx052

Parrino, R., & Bates, T.W., & Gillan, S.L., & Kidwell, D.S. (2021). *Fundamentals of Corporate Finance* (5th ed). Wiley

Pástor, Ľ. (2000). Portfolio selection and asset pricing models. *The Journal of Finance*, *55*(1), 179-223.

Tobin, J. (1958). Liquidity Preference as Behavior Towards Risk. *The Review of Economic Studies*, *25*(2), 65–86. https://doi.org/10.2307/2296205

Victor DeMiguel, Lorenzo Garlappi, Raman Uppal, Optimal Versus Naive Diversification: How Inefficient is the 1/*N* Portfolio Strategy?, *The Review of Financial Studies*, Volume 22, Issue 5, May 2009, Pages 1915 1953, https://doi.org/10.1093/rfs/hhm075