ROBUST PORTFOLIO OPTIMIZATION: A STUDY OF BSE30 AND BSE100

A Project Report Submitted

for the Course

MA498 Project II

by

Mohammed Bilal Girach

(Roll No. 150123024)

Shashank Oberoi

(Roll No. 150123047)

to the

DEPARTMENT OF MATHEMATICS

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI - 781039, INDIA

April 2019

CERTIFICATE

This is to certify that the work contained in this project report entitled "**Robust Portfolio Optimization: A Study of BSE30 and BSE100**" submitted by **Mohammed Bilal Girach (Roll No. 150123024)** and **Shashank Oberoi (Roll No. 150123047)** to the Department of Mathematics, Indian Institute of Technology Guwahati towards partial requirement of **Bachelor of Technology** in Mathematics and Computing has been carried out by them under my supervision.

It is also certified that, along with literature survey, a few new results are established using computational implementations carried out by the students under the project.

Turnitin Similarity: 10 %

Guwahati - 781 039 (Dr. Siddhartha Pratim Chakrabarty)

November 2018 Project Supervisor

ABSTRACT

We begin with a discussion on the classical Markowitz portfolio optimization, its drawbacks and consequent motivation of the alternate approach of robust portfolio optimization. This is followed by presenting several robust optimization models. Using uncertainty sets, we then present computational results with BSE30 and BSE100 followed by a simulation study using true mean and covariance of asset returns. We undertake a comparison of performance of the robust optimization approaches as compared to Markowitz optimization. We finally discuss the advantages of the robust optimization from the standpoint of number of stocks, number of samples and types of data.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Investment in an individual security always has an associated risk, which can be minimized through diversification, a process involving investment in a portfolio consisting of several securities. For optimal allocation of weights in a diversified portfolio, one of the well-established methods is the classical mean-variance portfolio optimization introduced by Markowitz [31, 32]. Mean and covariance matrix of returns of securities are used as the measures for giving a quantitative sense to the return and the risk, respectively, of the portfolio. Despite being considered as the most basic theoretical framework in the field of portfolio optimization, there are several drawbacks associated with incorporating the Markowitz model in a practical setup.

Theoretically, Markowitz based portfolio optimization can result in assigning extreme weights to the securities comprising the portfolio. However, investment in securities can not be made in such extreme positions like large short positions if one takes active trading into account. Such kind of scenarios can be avoided by introducing appropriate constraints on the weights. Black and Litterman [12] argued that there is an added disadvantage since there are high chances of the optimal portfolio lying in the neighborhood of the imposed constraints. Thus, imposition of constraints leads to strong dependence of the constructed portfolio upon the constraints. For example, disallowing short sales often results in assigning zero weights to many securities and largely positive weights to the securities having small market capitalization.

One of the most major limitations of the mean-variance model is the sensitivity of the optimal portfolios to the errors in the estimation of return and risk parameters. These parameters are estimated using sample mean and sample covariance matrix, which are maximum likelihood estimates (MLEs) (calculated using historical data) under the assumption that the asset returns are normally distributed. According to DeMiguel and Nogales [18], since the efficiency of MLEs is extremely sensitive to deviations of the distribution of asset returns from the assumed normal distribution, it results in the optimal portfolios being vulnerable to the errors in estimation of input parameters. Additionally, the historical data neglects various other market factors and is not an accurate representation for estimates of future returns. Taking into account the above reasons, Michaud [33] argued that the mean-variance analysis tends to maximize the impact of estimation errors associated with the return and the risk parameters for the securities. As a result, Markowitz portfolio optimization often overweighs (underweighs) the securities having higher (lower) expected return, lower (higher) variance of returns and negative (positive) correlation between their returns. Labelling the model as "estimation-error maximizers", he stated that it often leads to financially counter-intuitive portfolios, which, in some cases, perform worse than the equal-weighted portfolio. Broadie [13] investigated the error maximization property of mean-variance analysis. Accordingly, he conducted a simulation based study to compare the estimated efficient frontier with the actual frontier computed using true parameter values. He observed that points on the estimated efficient frontier show superior performance as compared to the corresponding points on the actual frontier. He supported his argument of over-estimation of expected returns of optimal portfolios through his simulated results of obtaining the estimated frontier lying above the actual frontier. Additionally, he pointed out that non-stationarity in the data of returns can further increase the errors in computing the efficient frontier. Chopra and Ziemba [16] performed the sensitivity analysis of performance of optimal portfolios by studying the relative effect of estimation errors in means, variances and covariances of security returns, taking the investors' risk tolerance into consideration as well. They observed that at a high risk tolerance (to be defined in later chapters) of around fifty, cash equivalent loss for estimation errors in means is about eleven times greater than that for errors in variances or covariances. Accordingly, they pointed out that if the investors have superior estimates for means of security returns, they should prefer using them over the sample means calculated from historical data. Best and Grauer [9, 10] also arrived at similar conclusions by studying the sensitivity of weights of optimal portfolios with respect to changes in estimated means of returns on individual securities. Further, on imposition of no short selling constraint on the securities, they observed that a small change in estimated mean return of an individual security can assign zero weights to almost half the securities comprising the portfolio, which is counter-intuitive.

The discussed literature arrives at a common conclusion that the optimal portfolios are extremely sensitive towards the estimated values of input parameters, particularly expected returns of individual securities. In order to address this issue, there has been significant progress in recent years in the area of robust portfolio optimization. Several methods have been proposed in this area. **We are particularly interested in the approaches falling in the category related to enhancing robustness by optimizing the portfolio performance in worst-case scenarios.**

Significant efforts have been made towards formulating these kinds of approaches from Markowitz based mean-variance analysis. The robust optimization approach incorporates uncertainty in the input parameters directly into the optimization problem. Tütüncü and Koenig [40] described uncertainty, using an uncertainty set that includes almost all possible realizations of the uncertain input parameters. Accordingly, they formulated the problem of robust portfolio optimization by optimizing the portfolio performance under the worst possible realizations of the uncertain input parameters. They conducted numerous experiments applying the robust allocation methods to the market data and concluded that robust optimization can be considered as a viable asset allocation alternative for conservative investors. According to Ceria and Stubbs [15], the standard approach of robust optimization is too conservative. They argued that it is too pessimistic to adjust the return estimate of each asset downwards. Accordingly, they introduced new variants of robust optimization, taking into account the estimation errors in input parameters while formulating the optimization problem. They observed that the constructed robust portfolios perform superior in comparison to those constructed using mean-variance analysis in most of the cases but not in each month with certainty. Utilizing the standard framework of robust optimization, Scherer [38] showed that robust methods are equivalent to Bayesian shrinkage estimators and do not lead to significant change in the efficient set. Constructing an example, he showed that robust portfolio underperforms out of the sample in comparison to Markowitz portfolio, especially in the case of low risk aversion and high uncertainty aversion. He also argued that performance of robust portfolio is dependent upon the consistency between uncertainty aversion and risk aversion which is quite complicated. Santos [37] performed similar experiments to compare two types of robust approaches, namely, the standard robust optimization discussed in Scherer's work [38] and zero net alpha-adjusted robust optimization proposed by Ceria and Stubbs [15], with the traditional optimization methods. The empirical results indicated better performance of robust approaches in comparison to the portfolios constructed using mean-variance analysis in the case of simulated data unlike in the case of real market data.

Chapter 2

Robust Portfolio Optimization Models using Uncertainty Sets

All the real world optimizing problems inevitably have uncertain parameters embedded in them. In order to tackle such problems, a framework called "Stochastic Programming" [11] is used, which can model such problems having uncertain parameters. These models take the probability distributions of the underlying data into consideration. To improve the stability of the solutions, robust methods such as re-sampling techniques, robust estimators and Bayesian approaches were developed. One of the approaches is **robust optimization**, which is used when the parameters are known to lie in a certain range. In this chapter, we discuss some robust models with worst-case optimization approaches for a given objective function within a predefined "uncertainty" set.

The concept of uncertainty sets was introduced by Soyster [39], where he uses a different definition for defining a feasible region of a convex programming problem. In this definition, the convex inequalities are replaced by convex sets with a condition that the finite sum of convex sets again should be within another convex set. In another way, he defines a new linear programming problem (LPP) with uncertain truth value, but it is bound to lie within a defined convex set. Later, El Ghaoui and Lebret [20] extend these uncertainty sets to define a robust formulation while tackling the least-squares problem having uncertain parameters, but they are bounded matrices. In their work, they describe the problem of finding a worst-case residual and refer the solution as a robust least-squares solution. Furthermore, they show that it can be computed via semi-definite or second order cone programming. El Ghaoui, Oustry, and Lebret [21] further study how to integrate bounded uncertain parameters in semidefinite programming. They introduce robust-formulations for semidefinite programming and provide sufficient conditions to guarantee the existence of such robust solutions. Ben-Tal and Nemirovski [6] mainly focus on the uncertainty related with *hard* constraints and which are *ought* to be satisfied, irrespective of the representation of the data. They suggest a methodology where they replace an actual uncertain linear programming problem by its robust counterpart. They show that the robust counterpart of an LPP with the ellipsoidal uncertainty set is computationally attractive, as it reduces to a polynomial time solvable conic quadratic program. Additionally, they use interior points methods [7] to compute the solutions efficiently. Along the same lines, Goldfarb and Iyengar [23], focus on the robust convex quadratically constrained programs which are a subclass of the robust convex programs of Ben-Tal and Nemirovski [6]. They mainly work on finding uncertainty sets which structures this subclass of programs as second-order cone programs.

In its early phases, the major directions of research were to introduce robust formulations and to build uncertainty sets for robust counterparts of the LPP as they are computationally attractive. Once the basic framework of robust optimization was established, it is now applied across various domains such as learning, statistics, finance and numerous areas of engineering.

2.1 Uncertainty Sets

The determination of the structure of the uncertainty sets, so as to obtain computationally tractable solutions has been a key step in robust optimization. In the real world, even the distribution of asset returns has an uncertainty associated. In order to address this issue, a most frequently used technique is to find an estimate of the uncertain parameter and to define a geometrical bound around it. Empirically, historical data is used to compute the estimates of these uncertain parameters. For a given optimization problem, determining the geometry of the uncertainty set is a difficult task. In this section, we discuss a couple of popular types of uncertainty sets which are used in portfolio optimization. Accordingly, we first introduce the common notations.

1. N: Number of assets.

2. $\mathbf{x}$: Weight vector for a portfolio.

3. $\boldsymbol{\mu}$: Vector for expected return.

4. Σ: Covariance matrix for asset returns.

5. λ: Risk aversion.

6. $\mathcal{U}_{\mu,\Sigma}$: General uncertainty set with μ and Σ as uncertain parameters.

7. $\mathbf{1}$: Unity vector of length N.

The classical Markowitz model formulation with no short selling constraint is given by the following problem:

$$\max_{\mathbf{x}} \left\{\boldsymbol{\mu}^\top \mathbf{x} - \lambda \mathbf{x}^\top \Sigma \mathbf{x}\right\} \text{ such that } \mathbf{x}^\top \mathbf{1} = 1 \text{ and } \mathbf{x} \geq 0. \tag{2.1}$$

Most of the robust models deal with optimizing a given objective function with a predefined "uncertainty set" for obtaining computationally tractable solutions. For any general uncertainty set $\mathcal{U}_{\mu,\Sigma}$, the worst case classical Markowitz model formulation [27, 25] with no short selling constraint is given as:

$$\max_{\mathbf{x}} \left\{ \min_{(\boldsymbol{\mu},\boldsymbol{\Sigma}) \,\in\, \mathcal{U}_{\mu,\Sigma}} \boldsymbol{\mu}^\top \mathbf{x} - \lambda \mathbf{x}^\top \Sigma \mathbf{x} \right\} \text{ such that } \mathbf{x}^\top \mathbf{1} = 1 \text{ and } \mathbf{x} \ \geq 0, \tag{2.2}$$

In literature, there are many extensions of uncertainty sets varying from simple polytopes to statistically derived conic-representable sets. A *polytopic* [19] uncertainty set which resembles a "box", it is defined as

$$U_\delta(\hat{\mathbf{a}}) = \{\mathbf{a} : |a_i - \hat{a_i}| \leq \delta_i, i = 1, 2, 3, ..., N\}, \tag{2.3}$$

where $\mathbf{a} = (a_i, a_2, ..., a_N)$ is a vector of values of uncertain parameters of dimension N and $\hat{\mathbf{a}} = (\hat{a_1}, \hat{a_2}, ..., \hat{a_N})$ is generally the estimate for $\mathbf{a}$.

In order to capture more information from the data, the consideration of the second moment gives rise to another class of uncertainty sets, namely, ellipsoidal uncertainty sets. One of the most popular way of defining them [19] is

$$U(\hat{\mathbf{a}}) = \left\{\mathbf{a} : \mathbf{a} = \hat{\mathbf{a}} + \mathbf{P}^{1/2}\mathbf{u}, ||\mathbf{u}|| \leq 1\right\}, \tag{2.4}$$

where the choice of $\mathbf{P}$ is driven by the optimization problem. The main motivation behind the use of such kind of sets is that they come up naturally when one tries to estimate uncertain parameters using techniques like regression. Additionally, these sets take probabilistic properties into account. We further discuss how to model the uncertainties for some of the financial indicators.

2.1.1 Uncertainty in Expected Returns

Recently, many attempts have been made to model the uncertainty in the expected returns because of several reasons. When compared with variances and covariances, it is known that the effect on the performance of portfolio due to the estimation error is more in case of expected returns. Though it is unlikely that the future returns of the assets are equal to the estimated value of expected return, one can foresee that they will be within a certain range of the estimated return. Accordingly, one can define uncertainty sets in such a way so that expected values lie inside the geometric bound around the estimated value, say $\hat{\boldsymbol{\mu}}$.

In a simple scenario, one can define possible intervals for the expected returns of each individual asset by using box uncertainty set. Mathematically, it can be expressed as [27]:

$$U_{\boldsymbol{\delta}}(\hat{\boldsymbol{\mu}}) = \left\{ \boldsymbol{\mu} : |\mu_i - \hat{\mu}_i| \leq \delta_i, i = 1, 2, 3, ..., N \right\}, \tag{2.5}$$

where N represents the number of stocks and δ_i represents the value which determines the confidence interval region for individual assets. Clearly from the above expression, for the asset i, the estimated error has an upper bound limit of δ_i. On incorporating the box uncertainty set (2.5), the max-min robust formulation (2.2) reduces to the following maximization problem:

$$\max_{\mathbf{x}} \quad \hat{\boldsymbol{\mu}}^\top \mathbf{x} - \lambda \mathbf{x}^\top \Sigma \, \mathbf{x} - \boldsymbol{\delta}^\top |\mathbf{x}| \quad \text{such that } \mathbf{x}^\top \mathbf{1} = 1 \text{ and } \mathbf{x} \geq 0, \tag{2.6}$$

The most popular choice is to use ellipsoidal uncertainty set, as it takes the second moments into account. Uncertainty in expected return using ellipsoidal uncertainty set is expressed as [27] :

$$U_\delta(\hat{\boldsymbol{\mu}}) = \left\{ \boldsymbol{\mu} : (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})^\top \Sigma_{\boldsymbol{\mu}}^{-1} (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}) \leq \delta^2 \right\}, \tag{2.7}$$

where Σ_μ is a variance-covariance matrix of the estimation error of expected returns of the assets. The max-min robust formulation (2.2) in conjunction with the ellipsoidal uncertainty set (2.7) results in the following maximization problem:

$$\max_{\mathbf{x}} \left\{ \hat{\boldsymbol{\mu}}^\top \mathbf{x} - \lambda \mathbf{x}^\top \Sigma \, \mathbf{x} - \delta \sqrt{\mathbf{x}^\top \Sigma_{\boldsymbol{\mu}} \, \mathbf{x}} \right\} \text{such that } \mathbf{x}^\top \mathbf{1} = 1 \text{ and } \mathbf{x} \geq 0. \tag{2.8}$$

While dealing with the box uncertainty, it is assumed that the returns follow normal distribution as it eases the task of computing the desired confidence intervals for each individual asset. We define δ_i for $100(1-\alpha)\%$ confidence level as follows:

$$\delta_i = \sigma_i \, z_{\frac{\alpha}{2}} \, n^{-\frac{1}{2}} \tag{2.9}$$

where $z_{\frac{\alpha}{2}}$ represents the inverse of standard normal distribution, σ_i is the standard deviation of returns of asset i and n is the number of observations of returns for asset i.

For the same reason, if the uncertainty set follows ellipsoid model, the underlying distribution is assumed to be tracing a χ^2 distribution with the number of assets being the degrees of freedom (df). Accordingly, for $100(1-\alpha)\%$ confidence level, δ is defined in the following manner

$$\delta^2 = \chi^2_N(\alpha) \tag{2.10}$$

where $\chi^2_N(\alpha)$ is the inverse of a chi square distribution with N degrees of freedom.

2.1.2 Separable Uncertainty Set

As mentioned earlier, portfolio performance is more sensitive towards estimation error in mean returns of assets in comparison to variances and covariances of asset returns. This is one of the major reasons behind research works laying less emphasis upon the uncertainty set for covariance matrix of asset returns. The robust approaches discussed in previous section model only the expected returns using uncertainty sets. Hence, in order to also encapsulate the uncertainty in the covariances, box uncertainty set for the covariance matrix of returns is defined on similar lines as that for expected returns. Lower bound $\underline{\Sigma}_{ij}$ and upper bound $\overline{\Sigma}_{ij}$ can be specified for each entry Σ_{ij} of the covariance matrix. Using this methodology, the constructed box uncertainty set for covariance matrix is expressed in the following form [40] :

$$U_\Sigma = \{\Sigma : \underline{\Sigma} \leq \Sigma \leq \overline{\Sigma},\ \Sigma \succeq 0\}. \tag{2.11}$$

In the above equation, the condition $\Sigma \succeq 0$ implies that Σ is a symmetric positive semidefinite matrix. This condition is necessary in most of the robust optimization approaches, particularly those involving Markowitz model as the basic theoretical framework.

Tütüncü and Koenig [40] discuss a method to solve the robust formulation of Markowitz optimization problem having non-negativity constraints upon the weights of assets. They define the uncertainty set for covariance matrix as per above equation and uncertainty set for expected returns as $U_{\boldsymbol{\mu}} = \{\boldsymbol{\mu} : \underline{\boldsymbol{\mu}} \leq \boldsymbol{\mu} \leq \overline{\boldsymbol{\mu}}\}$, where $\underline{\boldsymbol{\mu}}$ and $\overline{\boldsymbol{\mu}}$ represent lower and upper bounds on mean return vector $\boldsymbol{\mu}$ respectively. Consequently, the max-min robust formulation (2.2) can be formulated as the following maximization problem:

$$\max_{\mathbf{x}} \left\{\underline{\boldsymbol{\mu}}^\top \mathbf{x} - \lambda \mathbf{x}^\top \overline{\Sigma} \mathbf{x}\right\} \text{ such that } \mathbf{x}^\top \mathbf{1} = 1 \text{ and } \mathbf{x} \geq 0. \tag{2.12}$$

Above approach involves the use of "separable" uncertainty sets, which implies, the uncertainty sets for mean returns and covariance matrix are defined independent of each other.

2.1.3 Joint Uncertainty Set

There are certain drawbacks associated with separable uncertainty sets. Lu [30] argues that such kind of uncertainty sets don't take the knowledge of actual confidence level into consideration. Secondly, separable uncertainty sets don't incorporate the joint behavior of mean returns and covariance matrix. As a result, these uncertainty sets are completely or partially similar to box uncertainty sets. This is one of the major reasons behind robust portfolios being conservative or highly non-diversified as observed in numerous computations. In order to address these drawbacks, Lu proposes a "joint uncertainty set". This uncertainty set is constructed as per desired confidence level using a statistical procedure that takes the factor model [24] for asset returns into consideration.

Delage and Ye [17] define a joint uncertainty set that takes into consideration the uncertainty in distribution of asset returns as well as moments (mean returns and covariance matrix of returns). The proposed uncertainty set having confidence parameters, $\gamma_1 \geq 0$ and $\gamma_2 \geq 1$, is given by:

$$\begin{aligned} &(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})^\top \ \hat{\Sigma}^{-1} \ (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}) \leq \gamma_1, \\ &E[(\mathbf{r} - \hat{\boldsymbol{\mu}})(\mathbf{r} - \hat{\boldsymbol{\mu}})^\top] \leq \gamma_2 \hat{\Sigma}. \end{aligned} \tag{2.13}$$

In the above equation, $\hat{\boldsymbol{\mu}}$ and $\hat{\Sigma}$ represent the estimates of mean return vector and covariance matrix of asset returns respectively, and $\mathbf{r}$ is the random return vector. Using this uncertainty set, they formulate the portfolio optimization problem as a Distributionally Robust Stochastic Program (DRSR). Accordingly, they demonstrate that the problem is computationally tractable by solving it as a semidefinite program.

Chapter 3

Computational Results

We analyze the performance of robust portfolio optimization methods discussed in the preceding chapter vis-à-vis the Markowitz model in a practical setup involving domestic market data and simulated data. The analysis is performed under two scenarios, namely, the number of stocks N being 31 and the number of stocks N being 98. This is done in order to observe the effect of increase in number of stocks on the performance of robust methods with respect to the Markowitz model. These numbers were chosen since they represent the number of stocks in S&P BSE 30 and S&P BSE 100 indices, respectively.

For the first scenario, we use the daily log-returns based on daily adjusted close price of the 31 stocks comprising BSE 30 (data source: Yahoo Finance [3]). Accordingly, we have considered the period from December 18, 2017 to September 30, 2018 (both inclusive) comprising of a total of 194 active trading days. Corresponding to this market data, we prepare two sets of simulated data for the 31 assets by sampling returns from a multivariate normal distribution with mean and covariance matrix set equal to those obtained from the S&P BSE30 data. The first set of sample returns comprises of number of samples same as the daily log-return observations in S&P BSE30 market data, namely, 193, whereas the second set comprises of 1000 samples. The two sets of simulated sample returns of different sizes were used to facilitate the study of the impact of the number of samples in simulated data on the performance of the robust portfolio optimization approaches. We make a comparative study of robust portfolio optimization approaches, in case of the historical S&P BSE 30 data, as well as the two sets of simulated data, in order to analyze whether the worst case robust portfolio optimization approaches are useful in a real market setup.

For the second scenario, we use the log-returns based on daily adjusted close price data of the 98 stocks comprising S&P BSE 100 (data source: Yahoo Finance [3]) with the period spanning from December 18, 2016 to September 30, 2018. Two sets of simulated data are constructed using multivariate normal distribution, on the similar lines as the first scenario.Similar kind of comparative study is performed for the second scenario.

The robust portfolio optimization approaches that we have taken into consideration for analyzing their performance with respect to Markowitz model without short-selling (**Mark**) are as follows:

1. Robust Model involving box uncertainty set in expected return without short-selling (**Box**).

2. Robust Model involving ellipsoidal uncertainty set in expected return without short-selling (**Ellip**).

3. Robust Model involving separable uncertainty set without short-selling (**Sep**).

For Box and Ellip model, we construct uncertainty sets in expected mean return with $100(1-\alpha)\%$ confidence level by considering $\alpha = 0.05$. Separable uncertainty set in Sep model is constructed as a $100(1-\alpha)\%$ confidence interval for both μ and Σ using Non-parametric Boostrap Algorithm with same α as in other robust models and assuming β, *i.e* the number of simulations, equal to 8000.

The performance analysis for these robust portfolio models vis-à-vis the Mark model is performed by taking into consideration the "Sharpe Ratio" of the portfolios constructed having λ representing risk-aversion in the ideal range [19] *i.e.*, $\lambda \in [2, 4]$. Since the yield for Treasury Bill in India from 2016 to 2018 has been found to oscillating around 6% [2], so, we have assumed the annualized riskfree rate to be equal to 6%. In the following sections, we present the computational results observed in case of two scenarios as discussed above.

3.1 Performance with $N = 31$ assets

We begin with the analysis for $N = 31$ assets, in the case of the simulated data with 1000 samples. In Figure 3.1 and Table 3.1, we present the efficient frontier and performance of portfolios constructed by applying the Mark model and three robust models to the simulated returns. We observe that efficient frontiers for the Ellip and Sep models lie below the one for the Mark model. This supports the argument made by Broadie [13] regarding over-estimation of efficient frontier in case of the Mark model. Overlap of efficient frontiers for the Mark and Box models indicates that utilizing box uncertainty set for robust optimization does not prove to be of much use in this case. This claim is supported quantitatively from Table 3.1 as well as since the average Sharpe ratio for portfolios constructed in the ideal range of risk-aversion is same in case of both the models. From Table 3.1, we infer that

Sep model performs at par with Mark model if we take into consideration the average Sharpe Ratio. This is evident from Figure 3.1 as well, since Mark model starts outperforming Sep model in terms of Sharpe ratio after the risk-aversion crosses 3. We also note that the Ellip model outperforms all the models including the Mark model in the entire ideal range of risk-aversion.

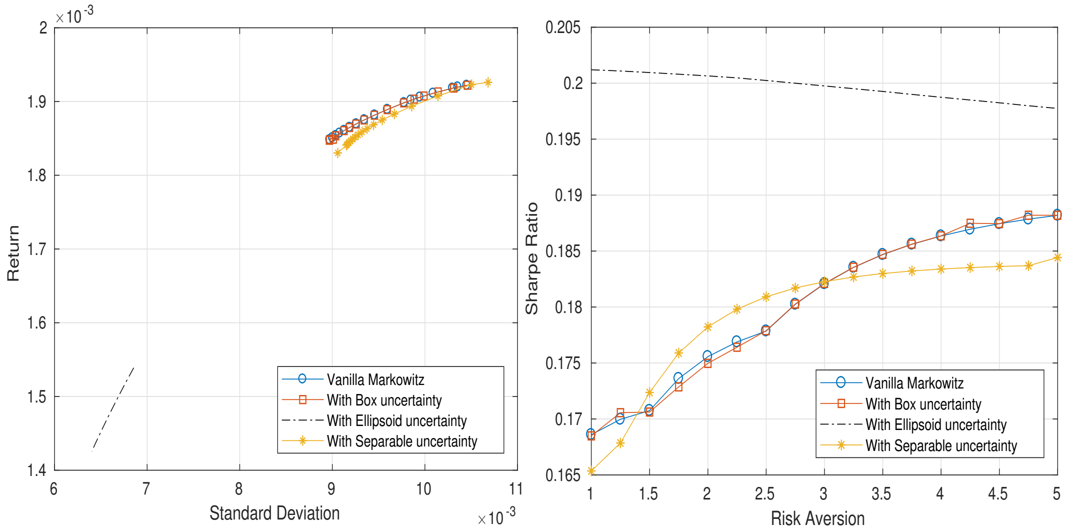

Figure 3.1: Efficient Frontier plot and Sharpe ratio plot for different portfolio optimization models in case of Simulated Data with 1000 samples (31 assets)

λ	SR_{Mark}	SR_{Box}	SR_{Ellip}	SR_{Sep}
2	0.176	0.175	0.201	0.178
2.5	0.178	0.178	0.2	0.181
3	0.182	0.182	0.2	0.182
3.5	0.185	0.185	0.199	0.183
4	0.186	0.186	0.199	0.183
Avg	0.181	0.181	0.2	0.182

Table 3.1: Comparison of different portfolio optimization models in case of Simulated Data with 1000 samples (31 assets)

On performing the simulation study with same number of samples as in case of market data (Figure 3.2 and Table 3.2), we observe similar results on comparing Box model with the Mark model. However, we observe slight inconsistency in performance of the Box model as evident from the plot of the Sharpe Ratio in Figure 3.2. The efficient frontiers for the Sep and Ellip model lie below that for the Mark model. We also infer that the Sep model and the Ellip model outperform the Mark model in terms of Sharpe Ratio in the ideal range of risk-aversion. However, it is difficult to compare the performance of the Sep model with that of the Ellip model in this case since the average Sharpe Ratio for both of them is almost the same.

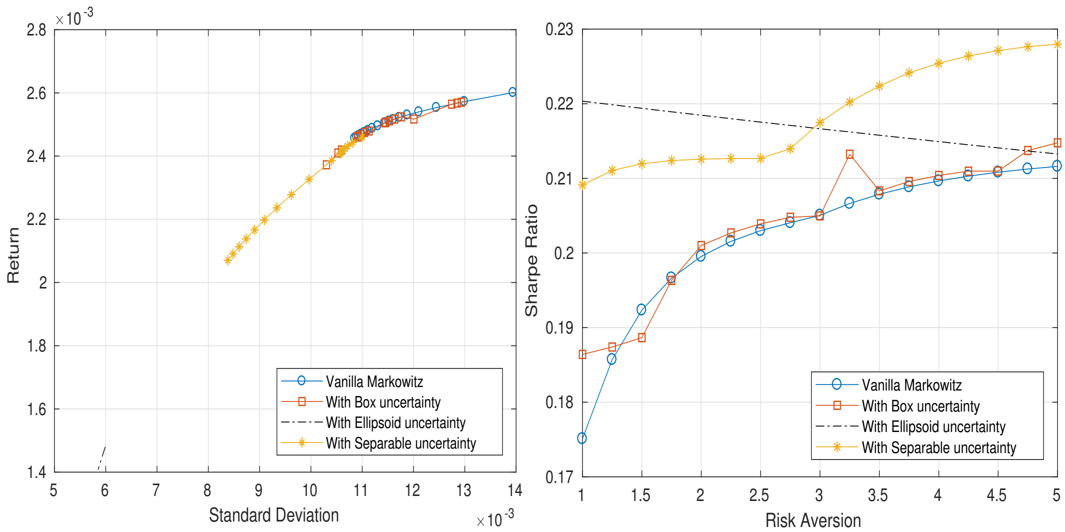

Figure 3.2: Efficient Frontier plot and Sharpe ratio plot for different portfolio optimization models in case of Simulated Data with same number of samples as market data (31 assets)

In a "real" market setup involving stocks comprising S&P BSE 30, we observe that efficient frontier for the Box model almost overlaps with that for the Mark model. However, the performance of the Box model in terms of Sharpe ratio is quite inconsistent as evident from the plot in Figure 3.3. Efficient frontier for the Sep model lies below that of the Mark model and

λ	SR_{Mark}	SR_{Box}	SR_{Ellip}	SR_{Sep}
2	0.2	0.198	0.218	0.213
2.5	0.203	0.204	0.218	0.213
3	0.205	0.207	0.217	0.217
3.5	0.208	0.209	0.216	0.222
4	0.21	0.21	0.215	0.225
Avg	0.205	0.206	0.217	0.218

Table 3.2: Comparison of different portfolio optimization models in case of Simulated Data with same number of samples as market data (31 assets)

the gap between the plots widens to a great extent, incase of the Ellip model. We also observe that the Sep model outperforms the Mark model in the ideal range of risk-aversion on taking the Sharpe Ratio into consideration as the performance measure. This is not true in case of the Ellip Model as evident from the Sharpe ratio plot in Figure 3.3. Even from Table 3.3, we observe that average Sharpe ratio for Ellip model is only slightly greater than that for the Mark model. Thus, unlike the simulated data, the Sep model performs superior in comparison to the Ellip model when applied to market data.

λ	SR_{Mark}	SR_{Box}	SR_{Ellip}	SR_{Sep}
2	0.181	0.181	0.193	0.186
2.5	0.181	0.181	0.192	0.193
3	0.186	0.191	0.192	0.202
3.5	0.194	0.195	0.191	0.209
4	0.201	0.202	0.19	0.213
Avg	0.189	0.19	0.192	0.2

Table 3.3: Comparison of different portfolio optimization models in case of Market Data (31 assets)

A common observation that could be inferred from three cases considered in the scenario involving less number of assets ($N = 31$) is that the Sep and Ellip models perform superior or equivalent in comparison to the Mark model in the ideal range of risk-aversion.

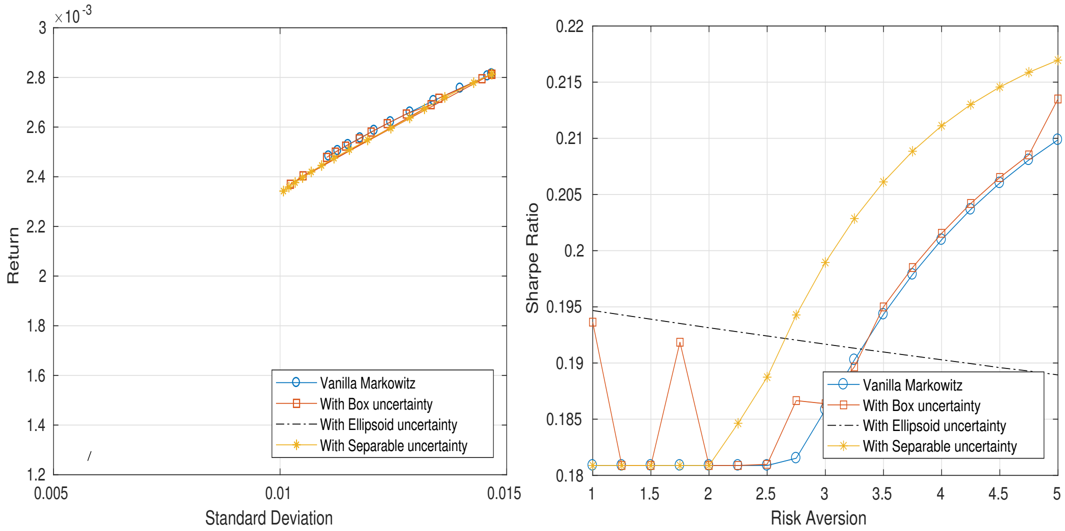

Figure 3.3: Efficient Frontier plot and Sharpe ratio plot for different portfolio optimization models in case of Market Data (31 assets)

3.2 Performance with $N = 98$ assets

We now analyze the scenario involving $N = 98$ assets. On applying robust models along with the Mark model on a simulated data having 1000 samples, we observe results similar to the corresponding case for the previous scenario when we compared the Box model with the Mark model. This is evident from the coinciding plots of the efficient frontier and the Sharpe ratio for both the models in Figure 3.4. Contrary to the similar case for the scenario involving less number of assets ($N = 31$), we observe that not only does the Ellip model but also the Sep model outperforms the Mark model taking into consideration the portfolios constructed in the ideal range of risk-aversion. Additionally, from Table 3.4, we infer that the Ellip model performs superior in comparison to the Sep model in terms of greater average value of the Sharpe ratio.

Figure 3.5 and Table 3.5 present the results of simulation study with

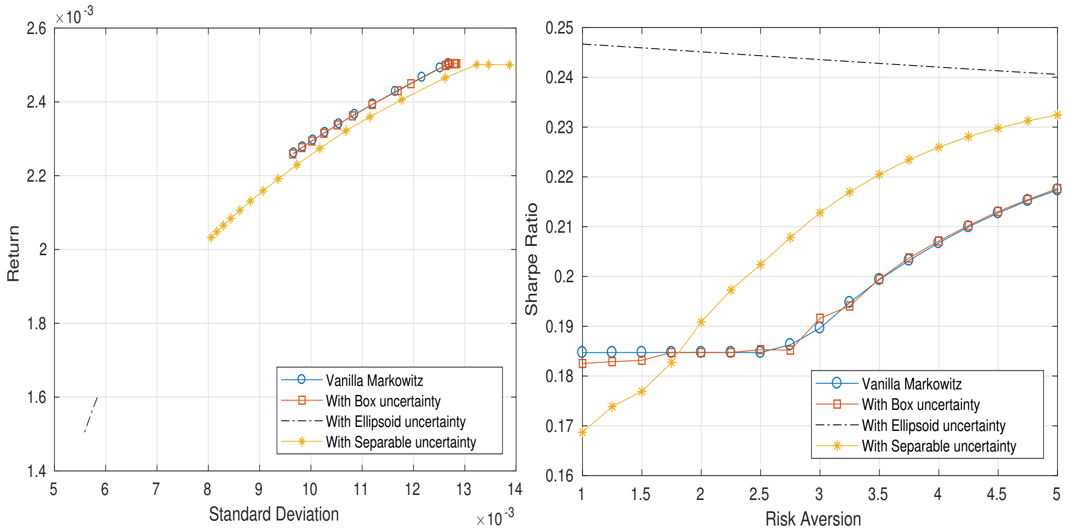

Figure 3.4: Efficient Frontier plot and Sharpe ratio plot for different portfolio optimization models in case of Simulated Data with 1000 samples (98 assets)

λ	SR_{Mark}	SR_{Box}	SR_{Ellip}	SR_{Sep}
2	0.185	0.185	0.245	0.191
2.5	0.185	0.185	0.244	0.202
3	0.19	0.192	0.244	0.213
3.5	0.199	0.199	0.243	0.221
4	0.207	0.207	0.242	0.226
Avg	0.193	0.194	0.244	0.21

Table 3.4: Comparison of different portfolio optimization models in case of Simulated Data with 1000 samples (98 assets)

same number of samples as that of log-returns of S&P BSE 100 data. Results observed on comparing the Box model with the Mark model are similar to the previous case of 1000 simulated samples. In the ideal range of risk aversion, we observe that efficient frontier for the Ellip as well as the Sep model lie below the Mark model. Additionally, both the models perform better than the Mark model in terms of the Sharpe Ratio. From the Sharpe ratio plot in Figure 3.5, it is difficult to compare the Sep model and the Ellip model since each outperforms the other in a different sub-interval of risk-aversion. The similar values of the average Sharpe ratio in Table 3.5 supports the claim of equivalent performance of these two models in this case.

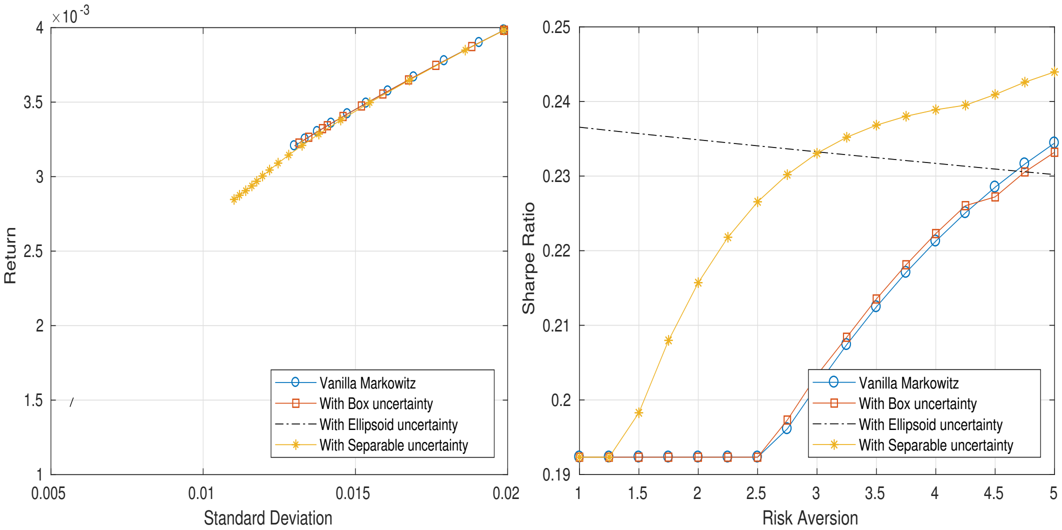

Figure 3.5: Efficient Frontier plot and Sharpe ratio plot for different portfolio optimization models in case of Simulated Data with same number of samples as market data (98 assets)

λ	SR_{Mark}	SR_{Box}	SR_{Ellip}	SR_{Sep}
2	0.192	0.192	0.235	0.216
2.5	0.192	0.192	0.234	0.227
3	0.202	0.203	0.233	0.233
3.5	0.212	0.213	0.232	0.237
4	0.221	0.222	0.232	0.239
Avg	0.204	0.205	0.233	0.23

Table 3.5: Comparison of different portfolio optimization models in case of Simulated Data with same number of samples as market data (98 assets)

The results for the case involving the market data (that contains log-returns of stocks comprising BSE 100) are presented in Figure 3.6 and Table 3.6. The efficient frontier plot leads to observations similar to the previous case. However, there is a slight inconsistency in the performance of the Box model as observed from the plot of the Sharpe Ratio in Figure 3.6. The robust portfolios constructed using the Sep and the Ellip model outperform the ones constructed using the Mark model in the ideal range of risk-aversion. Additionally, the Ellip model performs slightly better than the Sep model as evident from the Sharpe Ratio plot. Marginal difference in average Sharpe ratio between these two models supports this inference.

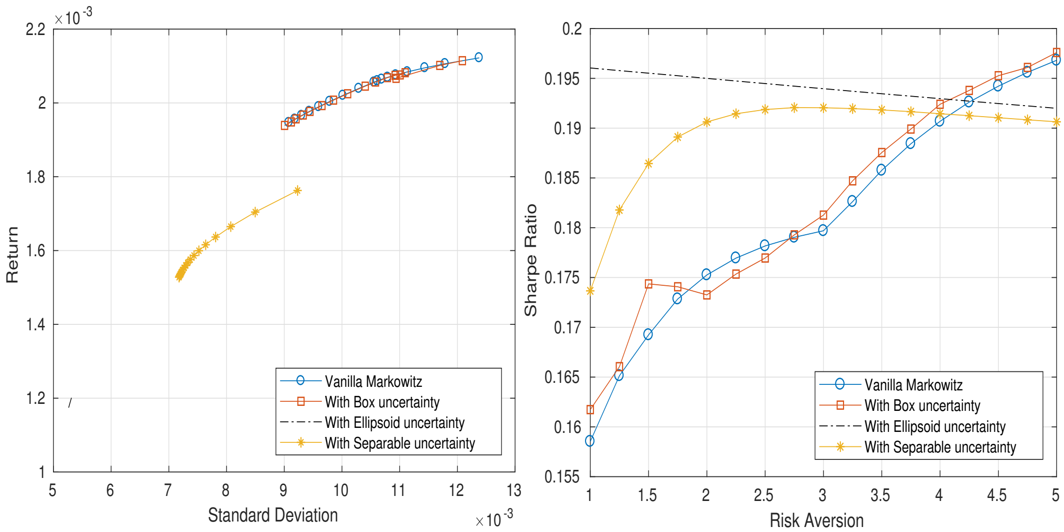

Figure 3.6: Efficient Frontier plot and Sharpe ratio plot for different portfolio optimization models in case of Market Data (98 assets)

λ	SR_{Mark}	SR_{Box}	SR_{Ellip}	SR_{Sep}
2	0.175	0.173	0.195	0.193
2.5	0.178	0.177	0.194	0.193
3	0.18	0.181	0.194	0.193
3.5	0.186	0.188	0.193	0.192
4	0.191	0.192	0.193	0.192
Avg	0.182	0.182	0.194	0.192

Table 3.6: Comparison of different portfolio optimization models in case of Market Data (98 assets)

We draw a common inference from the three cases considered in the scenario involving greater number of assets, *i.e.*, Sep and Ellip model outperform the Markowitz model in the ideal range of risk aversion.

Chapter 4

VaR and Its Robust Formulation

In the preceding chapters, we observed several limitations of the mean-variance framework such as sensitivity to errors in data as well as in the estimation of mean and variance of the underlying distribution. Furthermore, another criticism usually associated with the mean-variance setup is the use of standard deviation as a measure of risk. On the other hand, from a practitioners' point of view, the upside and downside risk can't be perceived in the same way because most of the times, the upside risk can improve the overall performance of the portfolio whereas the downside fluctuation usually brings impactful losses. Variance is not a reliable/appropriate risk measure, if the underlying distribution is leptokurtic. In order to address these issues, models involving other measures of risk have been developed and accordingly their corresponding robust models have been studied as well. In the forthcoming chapters, we will elaborately discuss some of the most widely measures of risk like Value at Risk (VaR) and Conditional Value at Risk (CVaR) and also study their corresponding robust worst case formulations.

4.1 Introduction

In this chapter, we start with the definition of VaR and formulate the optimization problems associated with VaR. Furthermore, we incorporate separable uncertainty set to model the worst case formulation. Next, we analyze and compare the performance of both VaR and Worst case VaR (WVaR) with respect to S&P BSE 100 and S&P BSE 30. Finally, we conclude with relevant insightful comments and conclusions.

4.2 Value at Risk (VaR)

Unlike the mean-variance setup, VaR framework takes into account the probability of losses[29]. Ghaoui et al. [22] defined VaR as the minimum value of γ such that the probability of loss exceeding γ is less than ϵ. Accordingly,

$$V(\mathbf{x}) = \min \left\{\gamma : \; P\left\{\gamma \leq -r(\mathbf{x}, \boldsymbol{\mu})\right\} \leq \epsilon\right\}, \tag{4.1}$$

where $\epsilon \in (0, 1]$. When we deal with mean-variance setup, only the mean and the variance *i.e., the first and the second moments* of the asset returns are required but in VaR framework, the knowledge of the entire distribution is necessary for the computation part. If the underlying distribution is Gaussian with the moments' pair as $(\hat{\boldsymbol{\mu}}, \, \Sigma)$, then VaR can be computed via the following analytical form:

$$V(x) = \kappa(\epsilon)\sqrt{\mathbf{x}^\top \Sigma \mathbf{x}} - \hat{\boldsymbol{\mu}}^\top \mathbf{x}, \tag{4.2}$$

where $\kappa(\epsilon) = -\Phi^{-1}(\epsilon)$ and $\Phi(z)$ represents the cumulative distribution for standard normal random variable. The value of the κ can be determined only if the cumulative distribution function of the underlying distribution is known apriori. However, if the distribution is unknown, then we have to rely on the Chebyshev's inequality. The bound (upper) obtained due to Chebyshev's only requires the knowledge of the first two moments' pair. We call a bound to be exact if the upper bound is computationally tractable. If not, we use the bound given by Bertsimas and Popescu [8] *i.e.,* $\kappa(\epsilon) = \sqrt{\dfrac{1-\epsilon}{\epsilon}}$. Finally, we formulate the generalised VaR as follows:

$$\min \kappa\sqrt{\mathbf{x}^\top \Sigma \mathbf{x}} - \hat{\boldsymbol{\mu}}^\top \mathbf{x} \quad \text{subject to} \quad \mathbf{x} \in \mathcal{X}, \tag{4.3}$$

where κ is an appropriate factor of risk chosen in accordance with the underlying distribution of asset returns and $\mathcal{X} = \left\{\mathbf{x} : \mathbf{x}^\top \mathbf{1} = 1 \text{ and } \mathbf{x} \geq 0\right\}$. The function $V(x)$ is convex and the global optimum can be obtained using techniques like interior-point methods and second order cone programming (SOCP).

Though VaR takes probability of losses into account, it has its own limitations. For the computation part, it requires the knowledge of the whole distribution. The computation also involves high dimensional numerical integration which may not be tractable at times and also there hasn't been an extensive research on methods using Monte Carlo simulations [29] for the design of the portfolio. Black and Litterman [12], Pearson and Ju [26] discussed the issues regarding the computational difference between the true VaR and the calculated VaR and determined that the error in the computation of the VaR can be attributed errors in the estimation of the first and second moments of the asset returns.

4.3 Worst Case VaR

The concept of worst-case VaR not only allows to approach the solution in a more tractable way but also relaxes the assumptions on the information known to us apriori. Here, we assume that only partial information about the underlying distribution is known. We also assume that the distribution of the asset returns belong to a family of allowable probability distributions $\mathcal{P}$. For example, given component wise bounds of ($\hat{\boldsymbol{\mu}}$, Σ), $\mathcal{P}$ could comprise of Normally distributed random variables with $\hat{\boldsymbol{\mu}}$ and Σ as the moments' pairs.

Given a probability (confidence) level ϵ, the worst-case VaR can be formulated as

$$V_{\mathcal{P}}(\mathbf{x}) = \min \left\{ \gamma : \sup_{P \in \mathcal{P}} P\left\{ \gamma \leq -r(\mathbf{x}, \boldsymbol{\mu}) \right\} \leq \epsilon \right\}, \tag{4.4}$$

and accordingly, the robust formulation can be written as

$$V_{\mathcal{P}}^{\text{opt}}(\mathbf{x}) = \min V_{\mathcal{P}}(\mathbf{x}) \quad \text{subject to} \quad \mathbf{x} \in \mathcal{X}, \tag{4.5}$$

The above optimization problems can be computed by a semi-definite programming (SDP) problem which again uses the above mentioned interior-point methods. We deal with the high dimensional problems by using bundle methods which are mainly used for large-scale (sparse) problems.

4.3.1 Polytopic Uncertainty

Motivated by the work of separable uncertainty sets by Tütüncü and Koenig [40], one can view the robust formulation given by Ghaoui et al. [22] in case of Polytopic uncertainty, as a robust formulation of WVaR involving separable uncertainty (section: 2.1.2). The formulation the optimization problem for worst case VaR:

$$\min\ \kappa(\epsilon)\sqrt{\mathbf{x}^\top\overline{\Sigma}\mathbf{x}} - \underline{\hat{\boldsymbol{\mu}}}^\top\mathbf{x} \quad \text{subject to} \quad \mathbf{x} \in \mathcal{X}, \tag{4.6}$$

where $\overline{\Sigma}$ and $\underline{\hat{\boldsymbol{\mu}}}$ are higher bound for covariance matrix and lower bound for the estimated mean of asset returns respectively. We obtain these values by using the Non-Parametric Bootstrap algorithm where the type of distribution is unknown. This can also be viewed as a robust formulation involving polytopic uncertainty because "Separable uncertainty" is a special case of "Polytopic uncertainty". In case of models involving ellipsoidal uncertainty sets, the robust worst-case VaR formulation is not that trivial and such models mainly revolve around the assumption of factor models.

4.4 Computational Results

In this section, we carry out the comparative analysis on the performance of the optimal portfolios obtained when VaR (referred as "Base VaR" from hereon) and WVaR are incorporated as measures of risk in the robust portfolio optimization problem. We carry out the analysis in a similar fashion as we have completed for mean-variance setup. On the same lines, we use "Sharpe Ratio (SR)" as a metric to compare the performance of the portfolios where the annualised risk-free rate is taken as 6% [2]. We perform the empirical analysis for the available historical data for S&P BSE 30 ($N = 31$) and S&P BSE 100 ($N = 98$). The reason behind this setup is to observe the trends/patterns in the performance of the portfolio when the number of stocks taken into consideration for constructing the optimal portfolio are changed. Furthermore, we also analyze the performance of the portfolio when in place of real market data, simulated data with true moments' pair is fed to the robust optimization problem. We dive one step further deep in above setup and vary the number of simulations to observe the changes in the performance of the optimal portfolio with increasing the number of simulations. For the above cause, we simulate two data-sets with true moments' pair, one with exact number of samples as in the available real market data, say ζ ($\leq$ 1000) and another with a large number of samples say 1000. We also use the same setup in order to analyze the performance of the portfolio when different type of data is taken into consideration *i.e.,* real market and simulated data. For the computational part of the robust formulation, we choose $\mathcal{P}$ to be family of distribution with the first and the second moments as the true mean and variance of the historical data. As the knowledge of the distribution is unknown, we use κ from equation (4.2), where we only consider the values of ϵ to be in (0,0.1] *i.e.,* the confidence level is greater than or equals 90% In the following subsections, we analyze each scenario in detail with appropriate figures and tables. We allow a slight notation abuse and refer the model which uses Base VaR/ WVaR as a measure of risk as VaR and WVaR respectively. We start our discussion with smaller number of stocks for which we obtain data from S&P BSE 30 and accordingly compute the log returns from the asset values.

4.4.1 Performance with $N = 31$ stocks

We proceed with the analysis of the performance of the portfolio when real market data is used to construct the optimal portfolio. From figure 4.1 one can observe that the performance of the portfolio is better when basic definition of VaR is used in the optimization problem. The Base VaR model outperforms the WVaR model over the complete range of ϵ. For a better

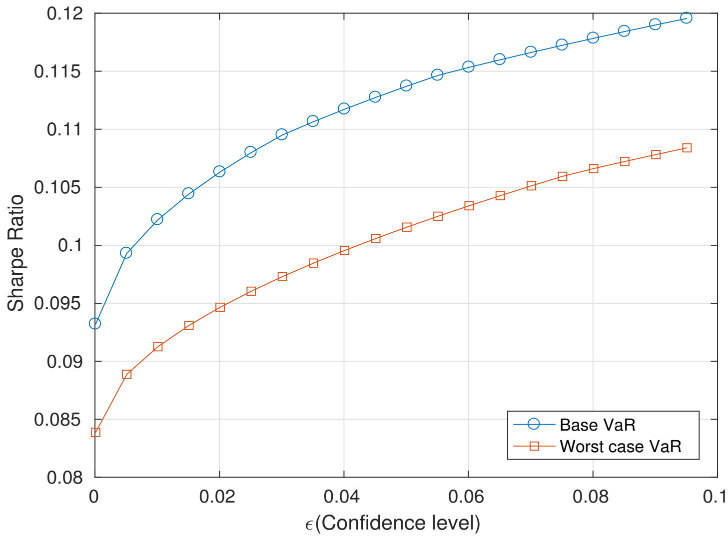

Figure 4.1: Sharpe ratio plot for Base VaR and WVaR models in case of Market data (31 assets)

understanding of the trends, we mainly tabulated the values of Sharpe ratio for the optimal portfolios obtained with different values of ϵ in table 4.1. We can observe that the average value of Sharpe ratio is more for Base VaR model when compared to WVaR model and is obvious from the figure 4.1.

ϵ	μ_{VaR}	σ_{VaR}	μ_{WVaR}	σ_{WVaR}	SR_{VaR}	SR_{WVaR}
0.0001	0.000646	0.00522	0.000603	0.00528	0.0932	0.0839
0.0201	0.000715	0.00522	0.00066	0.00529	0.106	0.0947
0.0401	0.000744	0.00523	0.000687	0.0053	0.112	0.0995
0.0601	0.000763	0.00523	0.000708	0.0053	0.115	0.103
0.0801	0.000777	0.00524	0.000726	0.00531	0.118	0.107
				Avg	0.111	0.0998

Table 4.1: Empirical Analysis of Base VaR and WVaR models in case of Market Data (31 assets)

Now, we move to the domain of the simulated data. Firstly, we discuss the trends in the performance of the optimal portfolio when the number of the samples (ζ) generated are same as there are in the real market data. We observe from the figure 4.2 that there is hardly any difference in trends of the performance when we used ζ number of simulations when compared to the real market data. Also from table 4.2, we can observe that the Base

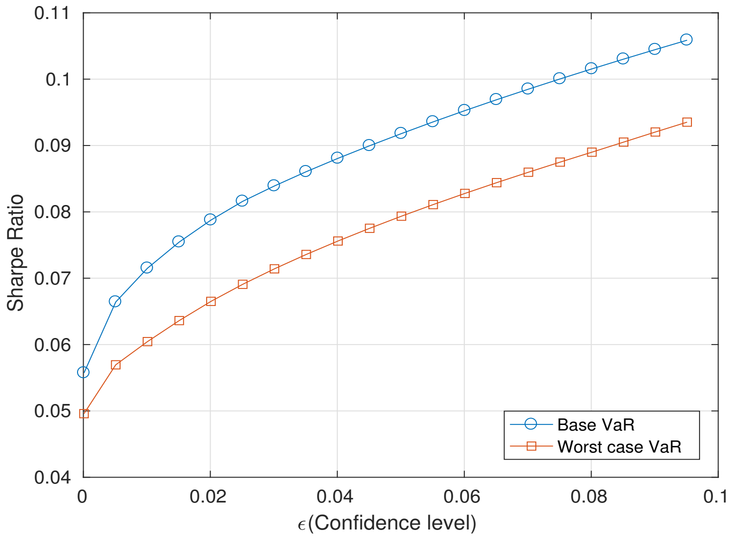

Figure 4.2: Sharpe ratio plot for Base VaR and WVaR models in case of simulated data with ζ number of samples (31 assets)

VaR model exhibits superior performance over WVaR model on the complete range of ϵ.

ϵ	μ_{VaR}	σ_{VaR}	μ_{WVaR}	σ_{WVaR}	SR_{VaR}	SR_{WVaR}
0.0001	0.000444	0.00509	0.000417	0.00519	0.0557	0.0496
0.0201	0.000562	0.0051	0.000506	0.0052	0.0788	0.0665
0.0401	0.00061	0.00511	0.000554	0.00521	0.0881	0.0756
0.0601	0.000647	0.00512	0.000592	0.00522	0.0953	0.0828
0.0801	0.000681	0.00513	0.000625	0.00523	0.102	0.089
				Avg	0.0884	0.0765

Table 4.2: Empirical Analysis of Base VaR and WVaR models in case of simulated data with ζ number of samples (31 assets)

As the model with ζ number of samples outputs the same trends as in the real market data. We now consider the case where we simulated larger number of samples, say 1000 with a justification, that larger the number of samples, less is the difference between true moments' pair with their computed counterparts.

In this case, from figure 4.3, we observe that the WVaR model delivers better performance than the Base VaR model in almost 40% of the range of ϵ. So, for more conservative investors, WVaR model will benefit more than the corresponding VaR model. In the table 4.3, one can cross-check the observation made above, the WVaR model performs better than Base VaR model in the interval (0,0.04] of ϵ. However, the average values remains almost equal because after the interval (0,0.04], the Base VaR model outperforms the WVaR model. So, we conclude that in this case both Base VaR and WVaR models are indifferent. We now extend the same type of analysis for a larger set of stocks for which we obtain data from S&P BSE 100 ($N = 98$).

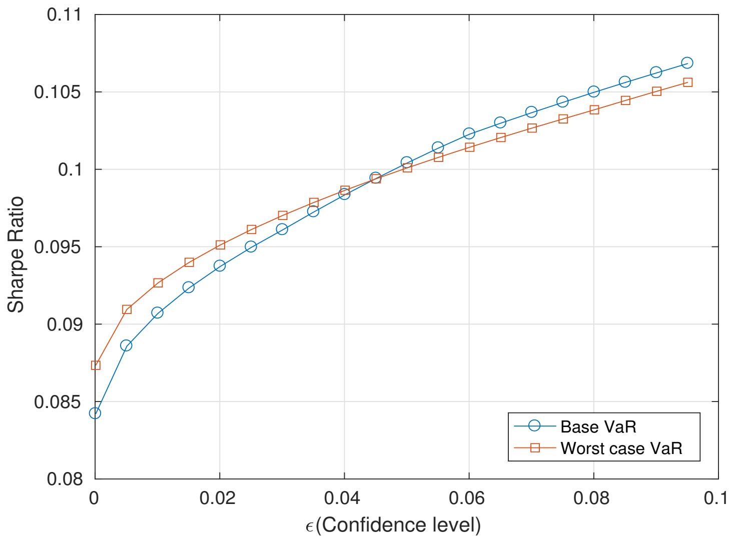

Figure 4.3: Sharpe ratio plot for Base VaR and WVaR models in case of simulated data with 1000 samples (31 assets)

ϵ	μ_{VaR}	σ_{VaR}	μ_{WVaR}	σ_{WVaR}	SR_{VaR}	SR_{WVaR}
0.0001	0.000587	0.00507	0.000605	0.0051	0.0842	0.0873
0.0201	0.000636	0.00508	0.000645	0.0051	0.0937	0.0951
0.0401	0.00066	0.00508	0.000663	0.00511	0.0984	0.0986
0.0601	0.00068	0.00509	0.000678	0.00511	0.102	0.101
0.0801	0.000694	0.00509	0.000691	0.00512	0.105	0.104
				Avg	0.0987	0.0989

Table 4.3: Empirical Analysis of Base VaR and WVaR models in case of simulated data with 1000 samples (31 assets)

4.4.2 Performance with $N = 98$ stocks

In the same lines, we begin our discussion of analyzing the performance of the Base VaR and WVaR models when real market data from S&P BSE 100 is taken into consideration. We present the empirical results in figure 4.4 and tabulate the observations in table 4.4, one can observe from the plot that the WVaR model exhibits superior performance in comparison to the corresponding Base VaR model on the entire range of ϵ. This observation can be quantitatively justified by the table 4.4 where the Sharpe ratios of the portfolios obtained from WVaR model is greater than that of Base VaR model.

ϵ	μ_{VaR}	σ_{VaR}	μ_{WVaR}	σ_{WVaR}	SR_{VaR}	SR_{WVaR}
0.0001	0.000667	0.00484	0.00071	0.00496	0.105	0.111
0.0201	0.000713	0.00485	0.000755	0.00497	0.114	0.12
0.0401	0.000735	0.00485	0.000776	0.00498	0.119	0.124
0.0601	0.000751	0.00486	0.000793	0.00499	0.122	0.127
0.0801	0.000765	0.00486	0.000807	0.005	0.124	0.13
				Avg	0.119	0.124

Table 4.4: Empirical Analysis of Base VaR and WVaR models in case of market data (98 assets)

The same type of trend is observed for $N = 98$ stocks in the case of simulated data as well. When the number of simulations equals ζ *i.e.,* the

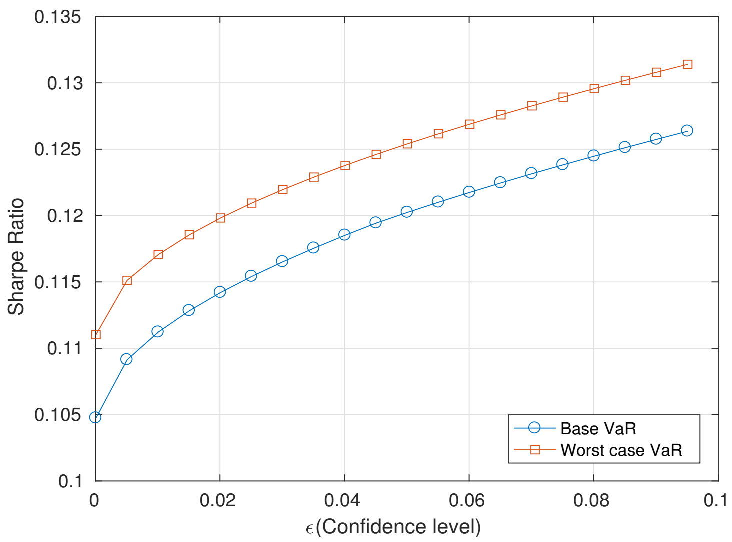

Figure 4.4: Sharpe ratio plot for Base VaR and WVaR models in case of market data (98 assets)

number of instances available in the real market data of S&P BSE 100. From figure 4.5 and the table 4.5, we infer that the WVaR model performs better than the Base VaR model when Sharpe ratio is used as performance measure for every $\epsilon \in (0,0.1]$.

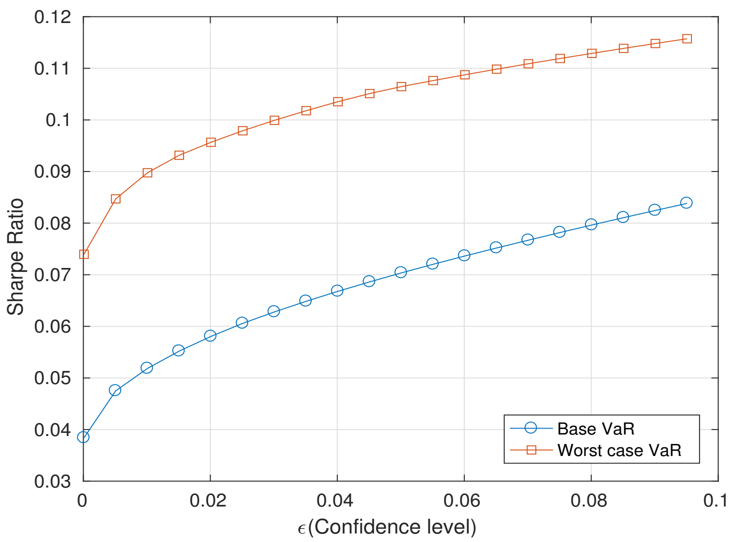

Figure 4.5: Sharpe ratio plot for Base VaR and WVaR models in case of simulated data with ζ number of samples (98 assets)

ϵ	μ_{VaR}	σ_{VaR}	μ_{WVaR}	σ_{WVaR}	SR_{VaR}	SR_{WVaR}
0.0001	0.000341	0.00471	0.00052	0.00487	0.0384	0.074
0.0201	0.000434	0.00472	0.000629	0.00491	0.0581	0.0957
0.0401	0.000475	0.00473	0.00067	0.00493	0.0668	0.104
0.0601	0.000508	0.00473	0.000697	0.00494	0.0736	0.109
0.0801	0.000537	0.00474	0.000719	0.00495	0.0796	0.113
				Avg	0.0674	0.103

Table 4.5: Empirical Analysis of Base VaR and WVaR models in case of simulated data with ζ number of samples(98 assets)

Lastly, we consider the case where we simulated larger number of samples in order to distinguish the fluctuations in the performance of the portfolio when number of simulations are varied. Similar kind of inferences can be drawn from figure 4.6 and table 4.6. The optimal portfolios obtained from the WVaR model have greater values of Sharpe ratio when compared to those obtained from Base VaR model irrespective of the value of ϵ.

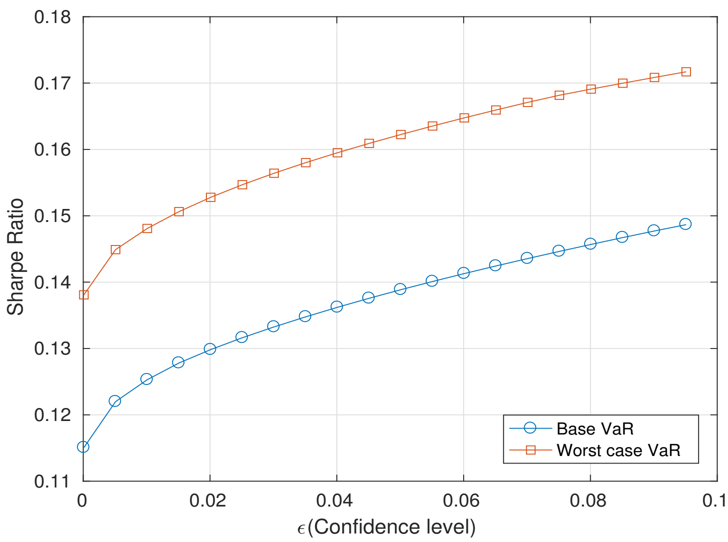

Figure 4.6: Sharpe ratio plot for Base VaR and WVaR models in case of simulated data with 1000 samples (98 assets)

ϵ	μ_{VaR}	σ_{VaR}	μ_{WVaR}	σ_{WVaR}	SR_{VaR}	SR_{WVaR}
0.0001	0.000702	0.00471	0.000822	0.0048	0.115	0.138
0.0201	0.000772	0.00472	0.000897	0.00483	0.13	0.153
0.0401	0.000803	0.00472	0.000931	0.00484	0.136	0.16
0.0601	0.000828	0.00473	0.000959	0.00485	0.141	0.165
0.0801	0.00085	0.00474	0.000982	0.00486	0.146	0.169
				Avg	0.137	0.16

Table 4.6: Empirical Analysis of Base VaR and WVaR models in case of simulated data with 1000 samples (98 assets)

Chapter 5

CVaR and Its Robust Formulation

Despite its popularity as a measure of downside risk, VaR has received fair amount of criticism [14, 41, 28]. VaR for a diversified portfolio may exceed that for an investment in a single asset. Thus, one of the major limitations of VaR is the lack of sub-additivity in the case of general distributions. In accordance with this observation, VaR is not a coherent measure of risk as per the definition laid out by Artzner et al [5]. Secondly, VaR does not provide any information on the size of losses in adverse scenarios *i.e,* those beyond the confidence level $1 - \epsilon$. Additionally, VaR is non-convex and non-differentiable when incorporated into portfolio optimization problem. As a result, it becomes difficult to optimize VaR since global minimum may not exist.

5.1 Introduction

Conditional Value-at-Risk (CVaR), introduced by Rockafellar and Uryasev [35, 36], has addressed various concerns centred around VaR. For continuous distributions, CVaR is the expected loss conditioned on the loss outcomes exceeding VaR. For $\epsilon \in (0,1)$, CVaR, at confidence level $1-\epsilon$, is defined as:

$$CVaR_\epsilon(\mathbf{x}) \triangleq \frac{1}{\epsilon} \int\limits_{-\mathbf{r}^\top \mathbf{x} \geq VaR_\epsilon(\mathbf{x})} p(\mathbf{r}) d\mathbf{r} \tag{5.1}$$

where $\mathbf{r}$ is the random return vector and $\mathbf{x}$ is the weight vector for a portfolio. In recent years, CVaR has emerged as a viable risk measure in portfolio optimization problems over its use in reducing downside risk. Since CVaR is a coherent risk measure, it follows the property of sub-additivity. Therefore, CVaR can be minimized through investment in a diversified portfolio. Unlike VaR, CVaR takes into consideration the impact of losses beyond the threshold value [14]. Also, the minimization of CVaR is a convex optimization problem [28].

Similar to Markowitz optimization and VaR minimization, there is an issue of lack of robustness in the classical framework of CVaR minimization. Computing CVaR requires the complete knowledge of the return distribution as argued by Zhu and Fukushima [41]. They have addressed this shortcoming of the classical CVaR (henceforth, referred to as base-case CVaR) by proposing a new risk measure, namely, **Worst-Case CVaR (WCVaR)**. For a fixed weight vector $\mathbf{x}$, at confidence level $1-\epsilon$, WCVaR is defined as:

$$WCVaR_\epsilon(\mathbf{x}) \triangleq \sup_{p(\mathbf{r}) \in \mathcal{P}} CVaR_\epsilon(\mathbf{x}) \tag{5.2}$$

where the density function $p(\mathbf{r})$ of returns is known to belong to a set $\mathcal{P}$ of probability distributions. The next section discusses coherent measures of risk that include both CVaR and WCVaR.

5.2 Coherent measures of risk

A risk measure ρ is a mapping of random gain X to a real value to represent the risk associated with X quantitatively. In their seminal work on Coherent Measures of Risk, Artzner et al [5] presented and justified following consistency rules for a risk measure ρ to be coherent:

1. Monotonicity: $X \leq Y \implies \rho(X) \geq \rho(Y)$
2. Translation Invariance: For any constant $m \in \mathcal{R}$, $\rho(X+m) = \rho(X)+m$
3. Positive Homogeneity: For any positive constant $\lambda \geq 0$, $\rho(\lambda X) = \lambda\rho(X)$
4. Sub-additivity: For any X and Y, $\rho(X+Y) \leq \rho(X) + \rho(Y)$

It has been proved by Pflug [34] and Acerbi and Tasche [4] that CVaR is a coherent measure of risk. Zhu and Fukushima [41] have proved the coherence of WCVaR as a risk measure by analyzing it in terms of worst-case risk measure ρ_w:

$$\rho_w(X) \triangleq \sup_{p(\mathbf{r})\in\mathcal{P}} \rho(X) \tag{5.3}$$

5.3 Mathematical Formulations

In this section, we discuss the mathematical formulations of the problems of optimizing CVaR along with WCVaR using mixture distribution uncertainty. We skip discussion on WCVaR formulations using box uncertainty set and ellipsoidal uncertainty set since they require a set of possible return distributions to be assumed so as to obtain bounds and scaling matrix respectively. However, our problem setup doesn't involve a set of return distributions since we make use of only market data (where return distribution is not known) and simulated data (where we generate a known return distribution).

5.3.1 Minimizing Base-Case CVaR

As proved by Rockafeller and Uryasev [35] , $CVaR_\epsilon(\mathbf{x})$, defined in equation (5.1), can be transformed into:

$$CVaR_\epsilon(\mathbf{x}) = \min_{\gamma \in \mathcal{R}^n} F_\epsilon(\mathbf{x}, \gamma) \tag{5.4}$$

where n is number of assets in the portfolio and $F_\epsilon(\mathbf{x}, \gamma)$ is defined as:

$$F_\epsilon(\mathbf{x}, \gamma) \triangleq \gamma + \frac{1}{\epsilon} \int\limits_{\mathbf{r} \in \mathcal{R}^n} [-\mathbf{r}^\top \mathbf{x} - \gamma]^+ p(\mathbf{r}) d\mathbf{r} \tag{5.5}$$

In the above equation, $[t]^+ = \max\{t, 0\}$. The problem of approximation of the integral involved in the equation (5.5) can be dealt by sampling the probability distribution of $\mathbf{r}$ as per its density $p(\mathbf{r})$. Assuming there are S samples, $\{\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_i, .., \mathbf{r}_S\}$ for the return vector $\mathbf{r}$, $F_\epsilon(\mathbf{x}, \gamma)$ can be approximated as [35]:

$$F_\epsilon(\mathbf{x}, \gamma) \approx \gamma + \frac{1}{S\,\epsilon} \sum_{i=1}^{S} [-\mathbf{r}_i^\top \mathbf{x} - \gamma]^+ \tag{5.6}$$

In accordance with above approximation, the problem of minimization of classical CVaR, assuming no short-selling constraints, can be formulated as the following Linear Programming Problem (LLP) [35, 41]:

$$\begin{aligned}
&\min_{(\mathbf{x},\mathbf{u},\gamma,\theta)} \theta \ s.t. \\
&\mathbf{x}^\top \mathbf{1} = 1, \mathbf{x} \geq \mathbf{0}, \\
&\gamma + \frac{1}{S\,\epsilon}\mathbf{1}^\top \mathbf{u} \leq \theta, \\
&u_i \geq -\mathbf{r}_i^\top \mathbf{x} - \gamma,\ u_i \geq 0,\ i = 1, 2, .., S.
\end{aligned} \tag{5.7}$$

where the auxiliary vector $\mathbf{u} \in \mathcal{R}^S$.

5.3.2 Minimizing Worst-Case CVaR using Mixture Distribution Uncertainty

In this section, we describe the formulation of the optimization problem involving WCVaR by assuming that the return distribution belongs to a set of distributions comprising of all possible mixtures of some prior likelihood distributions [41]. Mathematically, it is assumed that:

$$p(\mathbf{r}) \in \mathcal{P}_M = \triangleq \left\{ \sum_{j=1}^{l} \lambda_j p^j(\mathbf{r}) \ : \ \sum_{j=1}^{l} \lambda_j = 1, \lambda_j \geq 0, j = 1, 2, .., l \right\} \tag{5.8}$$

In the above equation, $p^j(\mathbf{r})$ denotes the j^{th} likelihood distribution and l is the number of the likelihood distributions. In accordance with above assumption of mixture distribution uncertainty (that involves $\mathcal{P}_M$ as a compact convex set) , $WCVaR_\epsilon(\mathbf{x})$, defined in equation (5.2), can be rewritten as the following min-max problem [41]:

$$\begin{aligned}
&WCVaR_\epsilon(\mathbf{x}) = \min_{\alpha \in \mathcal{R}} \max_{j \in \mathcal{L}} F_\epsilon^j(\mathbf{x}, \gamma)\,,\ where \\
&\mathcal{L} \triangleq \{1, 2, .., l\} \\
&F_\epsilon^j(\mathbf{x}, \gamma) \triangleq \gamma + \frac{1}{\epsilon} \int\limits_{\mathbf{r} \in \mathcal{R}^n} [-\mathbf{r}^\top \mathbf{x} - \gamma]^+ p^j(\mathbf{r}) d\mathbf{r}
\end{aligned} \tag{5.9}$$

Similar to the case involving classical CVaR in previous subsection, $F_\epsilon^j(\mathbf{x}, \gamma)$ can be approximated via discrete sampling as:

$$F_\epsilon^j(\mathbf{x}, \gamma) \approx \gamma + \frac{1}{S_j\, \epsilon} \sum_{i=1}^{S_j} [-\mathbf{r}_{i,j}^\top \mathbf{x} - \gamma]^+ \tag{5.10}$$

In the above equation, $\mathbf{r}_{i,j}$ is the i^{th} sample of the return with respect to j^{th} likelihood distribution and S_j is the number of samples corresponding to j^{th} likelihood distribution. Accordingly, assuming non-negative weights, the problem of minimization of WCVaR over a feasible set of portfolios can be formulated as the following LLP [41]:

$$\begin{aligned}
&\min_{(\mathbf{x}, \mathbf{u}, \gamma, \theta)} \ \theta\ s.t. \\
&\mathbf{x}^\top \mathbf{1} = 1, \mathbf{x} \geq \mathbf{0}, \\
&\gamma + \frac{1}{S_j\, \epsilon} \mathbf{1}^\top \mathbf{u}^j \leq \theta,\ j = 1, 2, .., l, \\
&u_i^j \geq -\mathbf{r}_{i,j}^\top \mathbf{x} - \gamma,\ u_i^j \geq 0,\ i = 1, 2, .., S_j,\ j = 1, 2, .., l.
\end{aligned} \tag{5.11}$$

In the above equation, the auxiliary vector $\mathbf{u} = (\mathbf{u}^1; \mathbf{u}^2; ..; \mathbf{u}^l) \in \mathcal{R}^S$ where $S = \sum_{j=1}^{l} S_j$.

5.4 Computational Results

Similar to the computational analysis of robust methods in Markowitz optimization and VaR minimization, we perform empirical analysis of the Worst-Case CVaR model vis-à-vis the Base-Case CVaR model, making use of the historical market data and simulated data. Taking into consideration the number of stocks N, the comparative study is conducted under two scenarios, namely, $N = 31$ and $N = 98$. These numbers represent the number of stocks in S&P BSE 30 and S&P BSE 100 indices, respectively.

For each scenario, we use the same three sets of data as used in the performance analysis of robust approaches in chapters 3 and 5 *i.e.* a set of historical market data comprising daily log-returns and two sets of simulated data. The first set of simulated data comprises of the number of sample returns matching the number of return instances of the historical market data, say $\zeta(< 1000)$. On the other hand, the second set has a larger number of samples, namely, 1000.

During computation of the Base-Case CVaR model, S in equation (5.7) is set equal to the number of return samples, *i.e,* either ζ or 1000, depending upon the set of data used. We perform computation of the Worst-Case CVaR model, as formulated in equation (5.11), for values of l in $\{2, 3, 4, 5\}$ by setting $S_j = \dfrac{S}{l}$ where $j \in \{1, 2, .., l\}$.

We perform comparative analysis of the Worst-Case CVaR model with respect to the Base-Case CVaR model using the Sharpe Ratio of the constructed portfolios having confidence level greater than 90%, *i.e,* $\epsilon \in (0, 0.1)$. Similar to preceding chapters, for computing the Sharpe Ratio, we assume the annualized risk free rate equal to 6%. In the following subsections, we present the computational results for the two scenarios.

5.4.1 Performance with $N = 31$ assets

l	$Avg.\ SR_{CVaR}$	$Avg.\ SR_{WCVaR}$	$Diff.\ in\ Avg.\ SR$
2	0.0856	0.0611	-0.0245
3	0.0856	0.0345	-0.0511
4	0.0856	0.0439	-0.0417
5	0.0856	0.031	-0.0546

Table 5.1: Comparison of CVaR and WCVaR in case of Market Data (31 assets) for different values of l

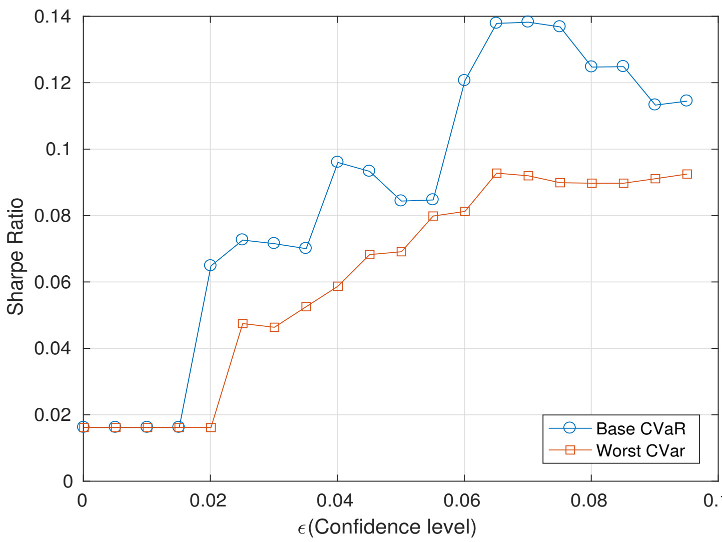

Figure 5.1: Sharpe ratio plot for CVaR and WCVaR in case of Market Data (31 assets)

ϵ	μ_{CVaR}	σ_{CVaR}	μ_{WCVaR}	σ_{WCVaR}	SR_{CVaR}	SR_{WCVaR}
0.0001	0.000266	0.00661	0.000266	0.00661	0.0162	0.0162
0.0201	0.000545	0.00595	0.000266	0.00661	0.0648	0.0162
0.0401	0.000706	0.00569	0.000514	0.00603	0.096	0.0587
0.0601	0.000832	0.00557	0.000645	0.00598	0.121	0.0812
0.0801	0.000877	0.00576	0.000696	0.00597	0.125	0.0897

Table 5.2: Empirical Analysis of CVaR and WCVaR in case of Market Data (31 assets)

We begin with the analysis for $N = 31$ assets, in the case of the historical market data having log-returns of the stocks comprising S&P BSE 30. In Table 5.1, we present the average Sharpe Ratio of the portfolios constructed for the Base-Case CVaR and Worst-Case CVaR models (denoted by $Avg.\ SR_{CVaR}$ and $Avg.\ SR_{WCVaR}$, respectively) for different values of l. For this case as well as other cases, we choose the value of l having maximum difference in the average Sharpe Ratio between the WCVaR and CVaR model and accordingly, perform comparative study based on the Sharpe Ratio plot and tabulation of results for some selected values of ϵ. The motivation behind following this kind of methodology for empirical analysis is to assess the practical viability of the robust technique over the classical method in CVaR minimization. In accordance with our methodology, the empirical results are presented for $l = 2$ in Figure 5.1 and Table 5.2. From Figure 5.1, we observe that the Base-Case CVaR model performs better than its robust counterpart in terms of the Sharpe Ratio of the constructed portfolios having $\epsilon \in (0, 0.1)$. Table 5.2 supports the above observation quantitatively.

l	$Avg.\ SR_{CVaR}$	$Avg.\ SR_{WCVaR}$	$Diff.\ in\ Avg.\ SR$
2	0.103	0.102	-0.000411
3	0.103	0.103	0.000195
4	0.103	0.104	0.00124
5	0.103	0.106	0.00358

Table 5.3: Comparison of CVaR and WCVaR in case of Simulated Data with ζ samples (31 assets) for different values of l

On performing the simulation study with ζ samples, we draw an inference from Table 5.3 that the performance of the WCVaR model vis-à-vis the CVaR model is maximal for $l = 5$. Similar to the methodology followed in the previous case, we present the relevant results for $l = 5$ in Figure 5.2 and Table 5.4. From the plot of the Sharpe Ratio in Figure 5.2, we observe

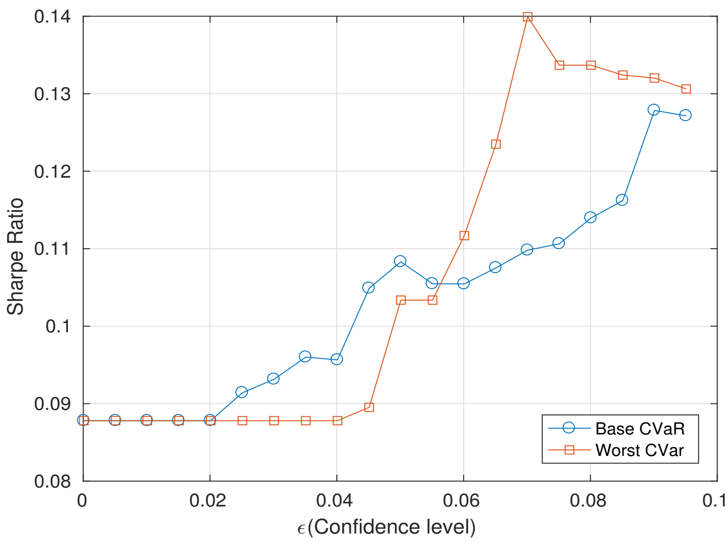

Figure 5.2: Sharpe ratio plot for CVaR and WCVaR in case of Simulated Data with ζ samples (31 assets)

ϵ	μ_{CVaR}	σ_{CVaR}	μ_{WCVaR}	σ_{WCVaR}	SR_{CVaR}	SR_{WCVaR}
0.0001	0.000675	0.00587	0.000675	0.00587	0.0878	0.0878
0.0201	0.000675	0.00587	0.000675	0.00587	0.0878	0.0878
0.0401	0.000694	0.00559	0.000675	0.00587	0.0957	0.0878
0.0601	0.000731	0.00542	0.000838	0.00607	0.105	0.112
0.0801	0.000776	0.00541	0.000889	0.00545	0.114	0.134

Table 5.4: Empirical Analysis of CVaR and WCVaR in case of Simulated Data with ζ samples (31 assets)

that the Worst-Case CVaR model starts outperforming the Base-Case CVaR model after ϵ crosses 0.05 *i.e*, incorporating robust optimization in CVaR minimization is advantageous in this case only for less conservative investors. We infer that the performance of the two models is almost equivalent, which is evident from Table 5.4 as well.

The comparative analysis in terms of the average Sharpe Ratio with the number of simulated samples being 1000 is presented in Table 5.5. Accordingly, we select the value of l equal to 4 and perform empirical study of the Base-Case CVaR and Worst-Case CVaR models, as presented in Figure 5.3

l	$Avg.\ SR_{CVaR}$	$Avg.\ SR_{WCVaR}$	$Diff.\ in\ Avg.\ SR$
2	0.0929	0.0967	0.00374
3	0.0929	0.0945	0.00161
4	0.0929	0.0969	0.00402
5	0.0929	0.0954	0.00249

Table 5.5: Comparison of CVaR and WCVaR in case of Simulated Data with 1000 samples (31 assets) for different values of l

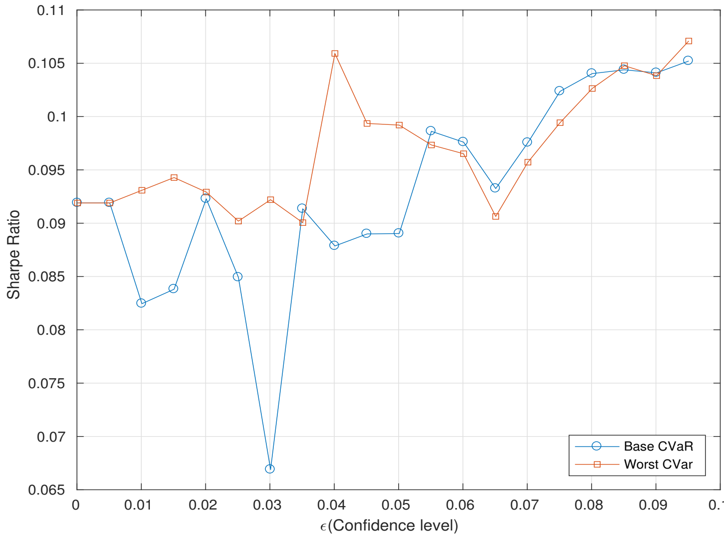

Figure 5.3: Sharpe ratio plot for CVaR and WCVaR in case of Simulated Data with 1000 samples (31 assets)

ϵ	μ_{CVaR}	σ_{CVaR}	μ_{WCVaR}	σ_{WCVaR}	SR_{CVaR}	SR_{WCVaR}
0.0001	0.000677	0.00563	0.000677	0.00563	0.0919	0.0919
0.0201	0.000658	0.0054	0.000661	0.0054	0.0923	0.0929
0.0401	0.000629	0.00534	0.000722	0.00531	0.0879	0.106
0.0601	0.00067	0.00523	0.000667	0.00526	0.0976	0.0965
0.0801	0.000701	0.0052	0.000692	0.00519	0.104	0.103

Table 5.6: Empirical Analysis of CVaR and WCVaR in case of Simulated Data with 1000 samples (31 assets)

and Table 5.6. It is difficult to draw any comparative inference from the Sharpe Ratio plot of the two models in Figure 5.3 since each outperforms the other in a different sub-interval of the range of ϵ. The similar values of the

Sharpe Ratio in Table 5.6 as well as the marginal difference in the average Sharpe Ratio for $l = 4$ in Table 5.5 supports the claim of almost equivalent performance of the CVaR and WCVaR models in this case.

5.4.2 Performance with $N = 98$ assets

l	$Avg.\ SR_{CVaR}$	$Avg.\ SR_{WCVaR}$	$Diff.\ in\ Avg.\ SR$
2	0.121	0.102	-0.0189
3	0.121	0.105	-0.0165
4	0.121	0.0991	-0.0219
5	0.121	0.0927	-0.0283

Table 5.7: Comparison of CVaR and WCVaR in case of Market Data (98 assets) for different values of l

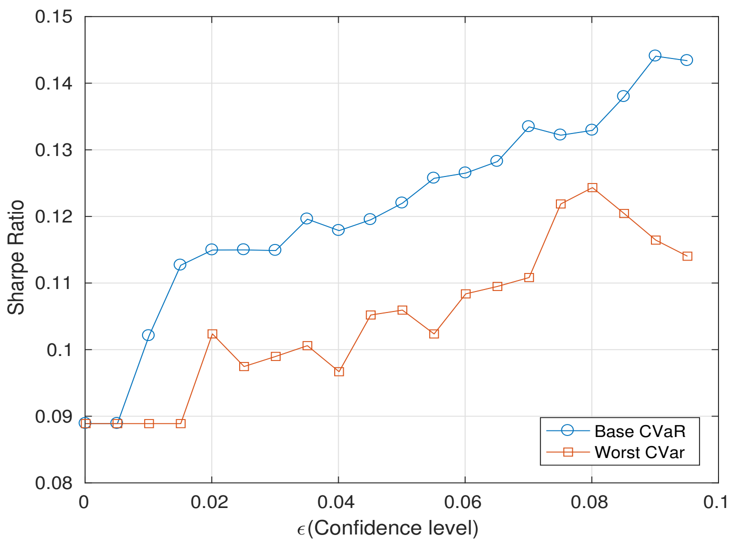

Figure 5.4: Sharpe ratio plot for CVaR and WCVaR in case of Market Data (98 assets)

In this subsection, we consider the scenario involving $N = 98$ assets. Table 5.7 summarizes the comparative results observed using the CVaR and WCVaR models on the historical market data (involving stocks comprising S&P BSE 100). Since the difference between the average Sharpe Ratio of the

ϵ	μ_{CVaR}	σ_{CVaR}	μ_{WCVaR}	σ_{WCVaR}	SR_{CVaR}	SR_{WCVaR}
0.0001	0.000687	0.00593	0.000687	0.00593	0.0889	0.0889
0.0201	0.000786	0.00544	0.000755	0.00582	0.115	0.102
0.0401	0.00079	0.00535	0.000692	0.0055	0.118	0.0967
0.0601	0.000839	0.00537	0.000738	0.00533	0.127	0.108
0.0801	0.000847	0.00517	0.00082	0.00531	0.133	0.124

Table 5.8: Empirical Analysis of CVaR and WCVaR in case of Market Data (98 assets)

WCVaR and CVaR models is maximum for $l = 3$, so, we plot and tabulate the relevant results summarizing the empirical analysis of the two models for the same value of l (Figure 5.4 and Table 5.8). Similar to the corresponding case for the previous scenario, Figure 5.4 and Table 5.8 lead to an observation that the Base-Case CVaR model exhibits superior performance in comparison to the Worst-Case CVaR model, taking into consideration the Sharpe Ratio of the constructed portfolios as the performance measure.

l	$Avg.\ SR_{CVaR}$	$Avg.\ SR_{WCVaR}$	$Diff.\ in\ Avg.\ SR$
2	0.0963	0.0978	0.00156
3	0.0963	0.0968	0.000541
4	0.0963	0.102	0.00581
5	0.0963	0.0915	-0.00479

Table 5.9: Comparison of CVaR and WCVaR in case of Simulated Data with ζ samples (98 assets) for different values of l

ϵ	μ_{CVaR}	σ_{CVaR}	μ_{WCVaR}	σ_{WCVaR}	SR_{CVaR}	SR_{WCVaR}
0.0001	0.000554	0.00579	0.000554	0.00579	0.0681	0.0681
0.0201	0.000703	0.0055	0.000755	0.00547	0.0988	0.109
0.0401	0.000727	0.00546	0.000723	0.00536	0.104	0.105
0.0601	0.000648	0.00522	0.000698	0.00527	0.0935	0.102
0.0801	0.000706	0.00514	0.000768	0.00525	0.106	0.116

Table 5.10: Empirical Analysis of CVaR and WCVaR in case of Simulated Data with ζ samples (98 assets)

Table 5.7 presents the comparison of the average Sharpe Ratio for the

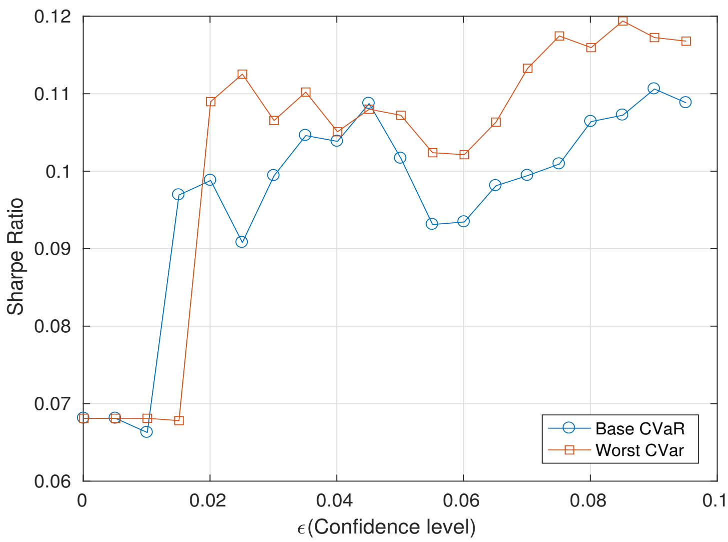

Figure 5.5: Sharpe ratio plot for CVaR and WCVaR in case of Simulated Data with ζ samples (98 assets)

simulation study with ζ samples *i.e,* same number of samples as that of log-returns of S&P BSE 100 data. In accordance with the discussed methodology, we choose the value of l equal to 4 and present the empirical results in Figure 5.5 and Table 5.10. From Figure 5.5, we observe that the Worst-Case CVaR model almost outperforms the Base-Case CVaR model in terms of the Sharpe Ratio of the constructed portfolios having $\epsilon \in (0, 0.1)$. From Table 5.10, we can support this inference quantitatively by observing the greater Sharpe Ratio of the portfolios for the WCVaR model vis-à-vis the CVaR model.

l	$Avg.\ SR_{CVaR}$	$Avg.\ SR_{WCVaR}$	$Diff.\ in\ Avg.\ SR$
2	0.156	0.156	-0.000382
3	0.156	0.16	0.0037
4	0.156	0.163	0.00667
5	0.156	0.165	0.00868

Table 5.11: Comparison of CVaR and WCVaR in case of Simulated Data with 1000 samples (98 assets) for different values of l

Finally, the comparative analysis on the basis of the average Sharpe Ratio

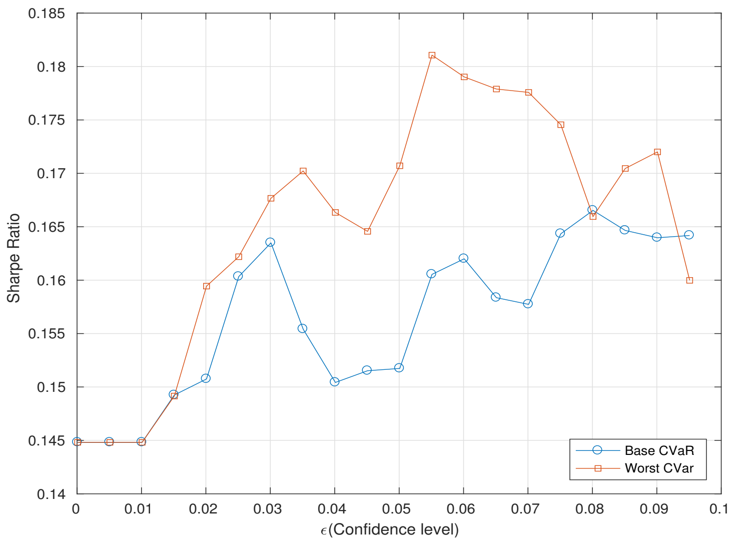

Figure 5.6: Sharpe ratio plot for CVaR and WCVaR in case of Simulated Data with 1000 samples (98 assets)

ϵ	μ_{CVaR}	σ_{CVaR}	μ_{WCVaR}	σ_{WCVaR}	SR_{CVaR}	SR_{WCVaR}
0.0001	0.000935	0.00536	0.000935	0.00536	0.145	0.145
0.0201	0.00097	0.00537	0.00101	0.00536	0.151	0.159
0.0401	0.000927	0.0051	0.00104	0.00531	0.15	0.166
0.0601	0.000963	0.00496	0.00106	0.00504	0.162	0.179
0.0801	0.000974	0.00489	0.000981	0.00495	0.167	0.166

Table 5.12: Empirical Analysis of CVaR and WCVaR in case of Simulated Data with 1000 samples (98 assets)

for the simulated data with 1000 samples is presented in Table 5.11. Accordingly, the results corresponding to the empirical study for $l = 5$ are presented in Figure 5.6 and Table 5.12. As observed from the plot of the Sharpe Ratio in Figure 5.6, the Worst-Case CVaR model performs superior with respect to the Base-Case CVaR model almost entirely for $\epsilon \in (0, 0.1)$. This observation is supported by the results presented in Table 5.12.

Chapter 6

Relevant Insights

6.1 Robust Optimization in Mean-Variance Analysis

In this section, we analyze the different kinds of scenarios in the context of trends of the Sharpe Ratio. Recall that, we have considered the "adjusted closing prices" data of S&P BSE 30 and S&P BSE 100 to illustrate our analysis. Further, we have also generated simulated samples using the true mean and covariance matrix of log-returns obtained from the aforesaid actual market data of "adjusted closing prices". Since the number of instances in market data for the assets comprising the two indices, was very less, we simulated two sets of samples, one where the number of simulated samples matches the number of instances of real market data available, say $\zeta(< 1000)$ and another where the number of simulated samples is large (a constant, which in our case was taken to be 1000), irrespective of the number of stocks. The motivation behind this setup was to understand if the market data we obtained (which was limited) is able to capture the trends and results in better portfolio performance.

6.1.1 From the Standpoint of Number of Stocks

	#stocks = 31	#stocks = 98
#generated_simulations = 1000	0.2	0.244
#generated_simulations = ζ	0.218	0.233
Market data	0.2	0.194

Table 6.1: The maximum average Sharpe ratio compared by varying the number of stocks in different kinds of scenarios.

We begin with a description of the results summarized in Table 6.1, wherein for a particular row and a particular column, we presented the maximum possible Sharpe Ratio that was obtained for that particular scenario. For example, in case of the tabular entry for the case when $N = 98$ where we simulated ζ samples using true mean vector and the true covariance matrix of S&P BSE 100, we refer to Table 3.5 (which explains the simulation corresponding to S&P BSE 100 with ζ simulated samples) and take the maximum of its last row *i.e.*, maximum of average Sharpe ratios that was attained using the available robust and Mark models.

More the number of stocks, the better is the performance of the portfolios constructed using robust optimization. This claim can be supported via both qualitative and quantitative approaches. Qualitatively, the number of stocks in a portfolio represent its diversification. According to Modern Portfolio Theory (MPT), investors get the benefit of better performance from diversifying their portfolios as it reduces the risk of relying on only one security to generate returns. Value Research Online [1] provides us with the information that on an average basis, large-cap funds hold around 38 shares, mid-cap funds around 50-52 assets for balanced funds in which around 65-

70% of their assets are in equity. This is because of great stability of returns in case of companies with large market capitalization, whereas this is not the case with mid-cap companies. Hence diversification requirements drives greater percentages in equities in case of mid-cap funds. From the above table, we can quantitatively justify by observing that the Sharpe ratio was more for portfolios with larger number of stocks when compared to portfolios with smaller number of stocks. However, we observe opposite behavior for the market data which can be attributed to the following two reasons:

1. The insufficient availability of market data, when it comes to larger number of stocks.

2. The error in the estimation of return and covariance matrix accumulating as the number of stocks increases, impacting the performance of the model [33].

6.1.2 From the Standpoint of Number of Samples Generated

In this section, we solely focus on the performance when different number of samples were generated. We tabulated Table 6.2 in the same way as we described in the preceding section. Here, we notice some interesting performance trends. One can observe that in the case of smaller number of stocks, the performance when same number of instances (ζ) were simulated is better than the scenario when large (1000) number of simulations were generated. On the contrary, exactly opposite trend can be observed when higher number of stocks are taken into consideration. We explain this type of behaviour as follows (as mentioned above): In the available real market data, the number of instances available for larger number of stocks is relatively low.

So, when more number of samples were generated, we observe higher Sharpe ratios when compared to ζ number of simulations. However, the reason behind such a pattern of opposite behavior when smaller number of stocks are considered is not obvious.

	#samples = 1000	#samples = ζ
#stocks = 31	0.2	0.218
#stocks = 98	0.244	0.233

Table 6.2: The maximum average Sharpe ratio compared by varying the number of stocks in different kinds of scenarios.

6.1.3 From the Standpoint of Type of the Data

Finally, we discuss about the performance of the portfolio from the standpoint of kind of data that we have used in this work. Accordingly, the relevant results are tabulated in Table 6.3, from where the behavior is observed to be fairly consistent. For both the cases, the performance in case of the simulated data is better than in case of the real market data. This is clear from the fact that the real market data is difficult to model as it hardly follows any distribution, whereas the simulated data is generated from multivariate normal distribution with mean and covariances as the true values obtained from the data.

	Simulated data	Real Market data
#stocks = 31	0.218	0.2
#stocks = 98	0.244	0.194

Table 6.3: The maximum average Sharpe ratio compared by varying the type of the data in different kinds of scenarios.

6.2 Robust Optimization in VaR minimization

	$N = 31$			$N = 98$		
	Market data	Sim. data ζ samples	Sim. data 1000 samples	Market data	Sim. data ζ samples	Sim. data 1000 samples
VaR	0.111	0.0884	0.0987	0.119	0.0674	0.137
WVaR	0.0998	0.0765	0.0989	0.124	0.103	0.16

Table 6.4: Comparison of the average Sharpe ratio for the VaR and WVaR models in various scenarios.

In this section, we discuss regarding the practical viability of incorporating robust optimization in VaR minimization from the point of view of number of stocks, sample size and types of data. For the sake of convenience, we tabulate all the results obtained in preceding chapters in the table 6.4, where for a particular scenario, we tabulated the average Sharpe ratio obtained over the range of ϵ.

6.2.1 From the Standpoint of Number of stocks

From the table 6.4 we point out a common inference that irrespective of the type of the data, the WVaR model exhibits superior performance than the Base VaR model in case of larger number of stocks ($N = 98$) and vice-versa in case of smaller number of stocks ($N = 31$). The qualitative argument for this kind of behaviour can't be attributed to the diversification of the portfolio because at times, VaR may not be sub-additive (coherent). We justify this behaviour in the lines of Michaud and Ghaoui. As Michaud [33] points out that the errors in the estimation of mean and covariances of the asset returns accumulates as the number of stocks increases. This implies that the data uncertainty in case of larger number of stocks is more in comparison to the smaller number of stocks. The Worst-case robust model can handle the data uncertainty in a better way than the Base VaR model [22]. Therefore, we observe an increment in the Sharpe ratio in both the models but the increment is more in case of WVaR model such that it outperforms the Base VaR in all types of data environments. In the scenario where the simulated data is taken into consideration, the same argument justifies the behaviour as the estimated moments' pair are used as true moments' pair for the generation of the data. So the error in the estimation of the mean and the variance of the asset returns effects the behaviour in the same way as in the market data.

6.2.2 From the Standpoint of Number of Simulations

When we observe the results from the perspective of number of simulations, we can draw some insightful comments. In $N = 98$ case, the better performance of WVaR model is attributed to the reason above. But when $N = 31$ one can observe that the performance of the optimal portfolio when 1000 samples were simulated is more than that of the portfolio obtained when ζ number of samples were simulated. The reason for this sort of behaviour lies in the subroutine of the Non parametric Bootstrap Algorithm where we use sampling with replacement. Therefore more the number of samples, better the bounds one can obtain from the algorithm. So when we generate more number of samples (1000) the WVaR model which uses these bounds performs better than the case where ζ number of samples are used for computing the bounds. In this setup, the increment in the number of simulation transits the robust portfolio from under-performing (ζ case) to be par with the portfolio obtained from Base VaR model.

6.2.3 From the Standpoint of Type of data

As explained in the previous section, when $N = 31$ the equivalent performance of the WVaR model with Base VaR model in case of simulated data can be attributed to the following reason: The real market data is difficult to model and may not follow any distribution whereas the simulated data follows the multivariate normal distribution with their true moments' pair as the tuple of estimated mean and covariance matrix of the asset returns from the real market data. Therefore, the Base VaR model exhibits superior performance in case of real market data. The reason for the out-performance of WVaR model over the Base VaR model when $N = 98$ stocks are considered is discussed in the above sections.

6.3 Robust Optimization in CVaR minimization

Similar to VaR minimization, we conduct the performance analysis of the Worst-Case CVar with respect to its classical counterpart from different standpoints. Table 6.5 presents the average Sharpe Ratio of the Base-Case CVaR and the Worst-Case CVaR models for each scenario based on the chosen value of l, as per the methodology discussed in the preceding chapter.

	$N = 31$			$N = 98$		
	Market data	Sim. data ζ samples	Sim. data 1000 samples	Market data	Sim. data ζ samples	Sim. data 1000 samples
CVaR	0.0856	0.103	0.0929	0.121	0.0963	0.156
WCVaR	0.0611	0.106	0.0969	0.105	0.102	0.165

Table 6.5: Comparison of the average Sharpe ratio for the CVaR and WCVaR models in various scenarios.

6.3.1 From the Standpoint of Number of stocks

We begin with a discussion of the results presented in Table 6.5 from the standpoint of number of stocks. For the case involving market data, we observe that the CVaR models performs better than the WCVaR model in the scenario of less number of stocks ($N = 31$). Even after increasing N to 98, we draw the same inference. The reason behind this trend can be attributed to the lack of knowledge regarding the distribution of the returns in the real market data. Computation of CVaR assumes that the probability distribution is perfectly known and optimization of WCVaR is based on the assumption of the return distribution belonging to a mixture of some prior likelihood distributions. Since the market data hardly follows any distribution, so, there is a sense of ambiguity associated with optimizing CVaR and WCVaR using the discrete sampling technique (discussed in the previous chapter) for the market data.

On the other hand, we observe an expected trend for the case of simulated data with 1000 samples. For $N = 31$, the WCVaR model performs at par with the VaR model. As N increases to 98, the WCVaR model exhibits superior performance vis-à-vis the CVaR model. Since CVaR and WCVaR are coherent risk measures, so, increase in the number of stocks enhances diversification. As a result, an uptrend is observed in the performance of these two models for $N = 98$. But, WCVaR, being a robust risk measure, diversifies over worst-case scenarios as well through mixture distribution uncertainty. As a result, the WCVaR model performs better than the CVaR model when larger number of stocks are taken into consideration.

However, in the case of simulated data with ζ samples, an unusual trend is observed despite involving comparative inference similar to the previous case. On increasing N to 98, we observe a decline in the performance of the CVaR and WCVaR models. The reason behind such observation is not obvious.

6.3.2 From the Standpoint of Number of Simulations

We now compare the performance of the two models based on the number of simulated samples. From Table 6.5, for the case involving ζ simulations, we infer that the WCVaR model performs at par with the CVaR model irrespective of the number of stocks. Same inference can be drawn with 1000 simulated samples as well. Hence, we note equivalent performance of the two models in each simulation study.

6.3.3 From the Standpoint of Type of data

Due to similar reasons as in VaR minimization, an opposite trend is observed in the case of real market data (as discussed above)when the number of stocks is less ($N = 31$). Similar observation is inferred for the market data on taking into account the larger number of stocks ($N = 98$).

Chapter 7

Conclusion

7.1 Concluding Remarks for Robust Optimization in Mean-Variance Analysis

Robust optimization is an emerging area of portfolio optimization. Various questions have been raised on the advantages of robust methods over the Markowitz model. Through computational analysis of various robust optimization approaches followed by a discussion from different standpoints, we try to address this skepticism. We observe that robust optimization with ellipsoidal uncertainty set performs superior or equivalent as compared to the Markowitz model, in the case of simulated data, similar to the results reported by Santos [37]. In addition, we observe favorable results in the case of market data as well. Better performance of the robust formulation having separable uncertainty set in comparison to the Markowitz model is in line with the previous study on the same robust model by Tütüncü and Koenig [40]. Empirical results presented in this work advocate enhanced practical use of the robust models involving ellipsoidal uncertainty set and separable uncertainty set and accordingly, these models can be regarded as possible alternatives to the classical mean-variance analysis in a practical setup.

7.2 Concluding Remarks for Robust Optimization in VaR and CVaR minimization

Akin to mean variance analysis, there is a problem of lack of robustness in the classical formulations of VaR and CVaR minimization. We discuss and assess the performance of the robust counterparts for these optimization problems that have been formulated to address this concern. Motivated by the results by Ghaoui, we formulate the worst case robust version of the VaR model using separable uncertainty set. Regardless of the type of the data, be it from real market or from a simulated environment, we observe favourable results for the worst case VaR model with Sharpe ratio as the performance measure when the portfolio comprises higher number of stocks.

In contrast to the results reported by Zhu, we observe that the base case CVaR performs better than the robust counterpart (formulated by incorporating mixture distribution uncertainty) in the case of Market data irrespective of the number of stocks comprising in the optimal portfolio. The reason could be attributed to the following two reasons:

- Incorporation of different weight constraints in our optimization problem.
- Unlike Zhu, our work uses Sharpe ratio as a performance measure.

Whereas in the case of simulated data, we draw a favourable inference by noting superior or equivalent performance of the worst case CVaR vis-à-vis the base case CVaR. In accordance with these results, we advocate for consideration of worst case models as a viable alternative to their classical counterparts mainly in the case of higher number of stocks and in a simulated environment.

Bibliography

[1] Portfolios: do number of stocks matter? https://www.valueresearchonline.com/story/h2_storyView.asp?str=200396&fbclid=IwAR3YkUaNLCY3keZGXfooehh0WU90zkZ5s94h0Uazj1-rzOFSy5bw6R1E-MY.

[2] Reserve bank of india - weekly statistical supplement. https://www.rbi.org.in/scripts/WSSViewDetail.aspx?TYPE=Section&PARAM1=4%0A.

[3] Yahoo finance. https://in.finance.yahoo.com/.

[4] C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503, 2002.

[5] P. Artzner, F. Delbaen, J. Eber, and D. Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.

[6] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.

[7] A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Oper. Res. Lett.*, 25(1):1–13, August 1999.

[8] Dimitris Bertsimas and Ioana Popescu. Optimal inequalities in probability theory: A convex optimization approach. *SIAM Journal on Optimization*, 15(3):780–804, 2005.

[9] M. J. Best and R. R. Grauer. On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results. *The Review of Financial Studies*, 4(2):315–342, 1991.

[10] M. J. Best and R. R. Grauer. Sensitivity analysis for mean-variance portfolio problems. *Management Science*, 37(8):980–989, 1991.

[11] J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming (Springer Series in Operations Research and Financial Engineering)*. Springer, 2011.

[12] F. Black and R. Litterman. Global portfolio optimization. *Financial Analysts Journal*, 48(5):28–43, 1992.

[13] M. Broadie. Computing efficient frontiers using estimated parameters. *Annals of Operations Research*, 45(1):21–58, Dec 1993.

[14] M. J. Capiński and E. Kopp. *Portfolio theory and risk management*, volume 5. Cambridge University Press, 2014.

[15] S. Ceria and R. A. Stubbs. Incorporating estimation errors into portfolio selection: Robust portfolio construction. *Journal of Asset Management*, 7(2):109–127, Jul 2006.

[16] V. K. Chopra and W. T. Ziemba. The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management*, 19(2):6–11, 1993.

[17] E. Delage and Y. Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations research*, 58(3):595–612, 2010.

[18] V.r DeMiguel and F. J. Nogales. Portfolio selection with robust estimation. *Operations Research*, 57(3):560–577, 2009.

[19] F. J. Fabozzi, P. N. Kolm, D. Pachamanova, and S. M. Focardi. *Robust Portfolio Optimization and Management*. Wiley, 2007.

[20] L. E. Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18(4):1035–1064, 1997.

[21] L. E. Ghaoui, F. Oustry, and H. Lebret. Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9(1):33–52, 1998.

[22] Laurent El Ghaoui, Maksim Oks, and Francois Oustry. Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations research*, 51(4):543–556, 2003.

[23] D. Goldfarb and G. Iyengar. Robust convex quadratically constrained programs. *Mathematical Programming*, 97(3):495–515, Aug 2003.

[24] D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics of operations research*, 28(1):1–38, 2003.

[25] Bjarni V Halldórsson and Reha H Tütüncü. An interior-point method for a class of saddle-point problems. *Journal of Optimization Theory and Applications*, 116(3):559–590, 2003.

[26] Xiongwei Ju and Neil D Pearson. Using value-at-risk to control risk taking: how wrong can you be? *OFOR Working Paper Series, no. 98-08*, 1998.

[27] J. H. Kim, W. C. Kim, and F. J. Fabozzi. Recent developments in robust portfolios with a worst-case approach. *Journal of Optimization Theory and Applications*, 161(1):103–121, 2014.

[28] C. Lim, H. D. Sherali, and S. Uryasev. Portfolio optimization by minimizing conditional value-at-risk via nondifferentiable optimization. *Computational Optimization and Applications*, 46(3):391–415, 2010.

[29] Thomas J Linsmeier and Neil D Pearson. Risk measurement: An introduction to value at risk. Technical report, 1996.

[30] Z. Lu. A new cone programming approach for robust portfolio selection. *Optim. Methods Softw.*, 26(1):89–104, 2006.

[31] H. M. Markowitz. Portfolio selection. *The journal of finance*, 7(1):77–91, 1952.

[32] H. M. Markowitz. *Portfolio Selection: Efficient Diversification of Investments*. Yale University Press, 1959.

[33] R. O. Michaud. The markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal*, 45(1):31–42, 1989.

[34] G. Pflug. Some remarks on the value-at-risk and the conditional value-at-risk, probabilistic constrained optimization: Methodology and applications. *Kluwer Academic Publishers*, 38:272–281, 2000.

[35] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of risk*, 2:21–42, 2000.

[36] R. T. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of banking & finance*, 26(7):1443–1471, 2002.

[37] A. Santos. The out-of-sample performance of robust portfolio optimization. *Brazilian Review of Finance*, 8(2):141–166, 2010.

[38] B. Scherer. Can robust portfolio optimisation help to build better portfolios? *Journal of Asset Management*, 7(6):374–387, Mar 2007.

[39] A. L. Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21(5):1154–1157, 1973.

[40] R. H. Tütüncü and M. Koenig. Robust asset allocation. *Annals of Operations Research*, 132(1-4):157–187, 2004.

[41] S. Zhu and M. Fukushima. Worst-case conditional value-at-risk with application to robust portfolio management. *Operations research*, 57(5):1155–1168, 2009.