Portfolio Optimization Models

An Insight into Five Methods of Portfolio Optimization: MVO, Robust MVO, Resampling MVO, Most-Diverse MVO, and Conditional Value at Risk

MIE377 Project 2

April 11th, 2018

Presented to Giorgio Costa, Alexia Yeo, and Hassan Anis

Prepared by Eric Huang 1002236653, Jun Han 1002214819, Andre D'Souza 1002268397

Introduction

The goal of this project was to form portfolios using five different variations of MVO, as well as CVaR with Monte Carlo. This report will discuss and quantitatively compare the different investment strategies on 20 stocks U.S. S&P500 between 2012 and 2015. A practical method of applying these portfolio optimization strategies is presented along with portfolio results using optimization solvers and follow by an analysis and discussion.

Background and Purpose

Portfolio optimization is the process of choosing specific proportions of a group of given assets to be held as a collective group (aka a portfolio) that satisfy a group of requirements better than any other. Important characteristics include the expected rate of return, variance, deviations between estimates and actual values, the Sharpe Ratio, and finally the portfolio's actual performance in terms of return and variance. Each of the optimization techniques has a different set of goals and constraints to satisfy, and seeks to find the best set of assets based on those constraints. In our case, the five selected portfolio optimization techniques are: Mean-Variance Optimization (MVO), Robust MVO, Resampling MVO, Most-Diverse MVO and Conditional Value-at-Risk (CVaR) Optimization.

By applying five different optimization techniques, we can visualize varying weights of each portfolio over each rebalancing period. We can also analyze the benefits and shortcomings of each investment strategy and determine which illustrates the best balance between risk and performance, as well as discuss reasons behind the differentiation.

Mean Variance Optimization (MVO) is an optimization technique that builds of the Modern Portfolio Theory developed by Harry Markowitz. It seeks to either minimize variance given a target portfolio return, or maximize return based on a selected risk level. MVO achieves this by forming portfolios that lie on the efficient frontier from Markowitz's theory. Its formulation can include other constraints such as prevented short selling, however, there are several other variations to increase or limit asset diversity, the units of the weighting or brace against uncertainty. The added complexity aims to solve a few of the drawbacks of MVO such as irregular quantities, or high transaction costs, and is explained more in methodology.

Most diverse MVO is an optimization method that utilizes the MVO framework, however it addresses the issue of portfolio diversity and over-concentration. It limits the number assets; however, it ensures that the assets selected ae a subset of the original that minimize correlation. Finding these set of the most uncorrelated assets is done through a process called 'Index tracking' and as a result, more representative of the overall market. Although the overall target return is still an average of the overall market, it serves as a better benchmark than VO as stocks in similar sectors/related variances are not over-invested in. This allows the portfolio to be less subjective to risk imposed when a subset of assets suffer in the market.

Robust MVO accounts for the uncertainty in the parameters, especially in the case of expected return. Due to the nature of the markets and reliance on external factors, historical data is not a reliable indicator of future asset performance. By introducing an ellipsoidal uncertainty set, the model is more likely to be inclusive of the actual asset return. Another more practical method of Robust MVO is Resampling MVO, which averages several different portfolios to form a single one. This brace against the estimations and assumptions that would be better suited to each individual portfolio.

Conditional Value at Risk (CVaR), otherwise known as expected shortfall, estimates the risk of an investment not by its volatility, but rather by focusing on a confidence interval surrounding losses. The metric itself calculates the weighted average of the Value at Risk (VaR) and the losses past the VaR. and can be combined with scenario based methods to optimize portfolios using linear programming.

Methodology

Transaction Costs

Although transaction costs are not used in our return calculations, they are kept as a metric to observe how drastic changes are within portfolios during rebalancing periods. In other words, transaction costs can be seen as a measure of volatility. The transaction costs will be calculated at 0.5% of the traded volume during each rebalancing session, calculated as:

$$tCost(t) = 0.005(Change\ in\ Shares\ of\ each\ asset)\cdot(Price\ of\ each\ Asset)$$

Rate and Variance Estimations

Factor models are financial models that are used to explain asset prices and market phenomena. It compares two or more ‘factors’, and analyzes relationships between different attributes and the resulting returns.

The Fama French Factor Model is a three factor model is an asset pricing model that improves upon the CAPM pricing model by adding factors that describe the size and value factors in addition to market risk factors (which is traditionally used for CAPM). A common trend in the market is that both value and small-cap stocks outperform markets regularly (due to their excess risk), hence they need to be differentiated using additional factors.

$$r_i - R_f = \alpha_i + \sum_i \beta_{ik} f_i + \epsilon_i = \alpha_i + \beta_{im}(f_m - r_f) + \beta_{iv} HML + \beta_{is} SMB + \epsilon_i$$

where

- α_i, β_{ik}: coefficients of the factor model
- ϵ_i : error between estimated and real return
- HML: High Minus Low Factor, premium for spread between value and growth stocks
- SMB: Small Minus Big, premium for spread between small and large firms

Using the Fama French Factor Model, we can find the excess returns given the factor data for the previous period to predict the expected returns. OLS regression will be used to find the coefficients for the factor model, producing the matrices α (intercepts) and V (the slopes),

which we can use to find our returns and variances.
The original rates can be represented as:

$$r = \alpha + V^T f + \epsilon$$

Our estimated rates, μ can be estimated using the geometric means of our factors:

$$\mu = \alpha + V^T \bar{f}$$

where $\bar{f}$ is the matrix representing the geometric mean of the factors.
Finally, the covariance of the excess returns can be found using:

$$Q = V^T F V + D \; where \; D = diagonals \; of \; \sigma^2(\epsilon_i)$$

Using this variance of excess returns, Q, we can express any portfolio's variance as:

$$\sigma_p = x^T Q x$$

where x is the weights allocated to each asset.

Data Estimation

One year of data is always used to estimate the performance in the next six months. Once the Fama –French factor model is used, the geometric mean of these returns are projected into the future for 6 months, and will be used as the inputs to our optimization functions. The geometric mean is represented as:

$$Return = \prod_i (1 + r_i)^{\frac{1}{N}} - 1$$

Once the expected return is found for each of the aassets, the correlation is found using the correlation function in MaLAB. Finally, for each of the optimization problems the arithmetic mean of the returns is used as the target return for the period.

Mean Variance Optimization Formulation

Mean Variance Optimization is an optimization program where we seek to minimizing σ_p, or the variance of the portfolio, while meeting a specific rate of return. The target rate of return (R) will be found using the mean of the predicted excess returns for the period. Furthermore, there will be no constraints imposed on short selling. The problem is formulated as follows:

$$\begin{gathered} \min_x x^T Q x \\ s.t. \mu^T x \geq R \\ 1^T x = 1 \end{gathered}$$

The function will be solved in MatLab using the QuadProg function, which is designed to solve MQPs when given a quadratic equation and its set of constraints.

Robust Mean Variance Optimization Formulation

Robust Mean Variance Optimization is an MVO optimization program that incorporates the uncertainty of its parameters explicitly into the model as deterministic variability. The prime drive to account for this uncertainty is because a small error in the estimated parameters could result in a drastic difference in the optimal solution for the problem. Also, there is a problem of feasibility, Uncertainty of parameters can be incorporated into the model as an uncertainty set in various ways. Uncertainty set essentially contains all the possible variability of the parameters that one wants to account for. This project uses ellipsoidal uncertainty set. Ellipsoidal uncertainty set $U(\mu)$ introduces an uncertainty set around the estimated expected returns, $U(\mu) = (\mu_{true} - \mu)^T \Theta^{-1}(\mu_{true} - \mu) \leq \epsilon_2^2$. $\theta^{\frac{1}{2}}$ is the standard error of the estimated μ, where $\theta = \frac{diag(Q)}{N}$. It essentially introduces a confidence level around the estimate μ. The confidence interval used in this project will be 95%. Thus the measure of distance($\epsilon_2 \left\|\Theta^{\frac{1}{2}}\right\| = \epsilon_2\sqrt{x^T Q x}$) of the estimate from the true μ can be added to the target return constraint to allow for the variability. The problem is then formulated as follows:

$$\min_x x^T Q x$$
$$s.t\ \mu^T x - \epsilon_2 y \geq R$$
$$\mu^2 = x^T Q x$$
$$1^t x = 1$$
$$y \geq 0$$
$$x_i \geq 0, i = 1, \dots\dots, n$$

The function will be solved using Gurobi. Because of this reason, the measure of distance was reformulated by adding an auxiliary variable y as Gurobi is not able to handle $\epsilon_2\|.\|_2$ term.

Resampling Mean Variance Optimization Formulation

Resampling Mean Variance Optimization is an alternative robust MVO. This model was driven by the fact that even a small variability in constraints results in drastic difference in the optimal solution. The method of incorporating the parameter uncertainty is where the resampling MVO diverge from a typical MVO. Instead of deterministically including the uncertainty into the model, this model generates different scenarios for the estimated parameters (in this project μ and Q). Then it finds the optimal solution for each scenario and find an average of all the generated solutions. One way to do this is to assume normal distribution with mean μ and standard variance Q and randomly generate observation from this normal distribution. Then the generated parameters are used to estimate new μ and Q. Using the new estimated parameters, portfolio weights are calculated using the nominal MVO model. These steps are repeated then the weights are averaged to arrive at the final weights for the portfolio. This project will generate 100 scenarios 50 times.

Most-Diverse Mean Variance Optimization Formulation

Most-Diverse Mean Variance Optimization generated a portfolio containing assets that are the most uncorrelated to each other as the name would suggest. Since it is very difficult to hold all of the assets listed in a certain index for tracking purposes, it is economical to find a smaller set of assets that represents other assets listed in the index. This is called bucketing as each asset representing a bucket that contain that are the most similar to each other. Similarity measure that will be used is the correlation coefficient $\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$. This project will use k = 12 assets to represent the 20 assets. The problem is formulated as follows:

$$\max_{z,y} \sum\sum \rho_{ij} z_{ij}$$
$$\sum y_j = k$$
$$\sum z_{ij} = 1 \; for \; i = 1, \dots \dots ., n$$
$$z_{ij} \leq y_j \; for \; i = 1, \dots \dots , n; j = 1, \dots \dots ., n$$
$$z_{ij} \in \{0,1\}, \quad y_j \in \{0,1\}$$

The y_j variable represents an asset that will be the most representative in the j[th] bucket. The y_{ij} variable represents the assets that will belong to the j[th] bucket. Since each bucket contains assets that are the most similar to each other, conversely this means that all the buckets will be the least similar to each other. Hence, using the y variable, it is possible to find the most diverse set of assets out of the pool. Gurobi was used to calculate this model. Using the 12 assets acquired from this model will be fed into the MVO to generate the portfolio weights.

Conditional Value at Risk (CVaR)

Monte Carlo Scenarios

Monte Carlo Sampling is utilized to generate the scenarios for CVaR optimization. Monte Carlo methods are a set of stochastic simulation algorithms that utilize random sampling to obtain hypothetical scenarios, or in the scope of this project, stock price movements. Ideally, large experiments and amounts of scenarios will converge to the deterministic (true) solution.

It is assumed that stock returns follow a normal distribution, such that $r_t = \mu + \sigma\epsilon_{t-1}$. Thus, individual price steps can be formulated as:

$$S_{t+1} = S_t \exp\left[\left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma\sqrt{dt}\epsilon_t\right]$$

To simulate the price paths, the same equation can be repeated over the number of steps in a certain time period. For the purposes of the assignment, only **one** individual step is taken to estimate the stock price change in 6 months. However, when there is correlation between the matrices (which there is from our 20 assets), an element ζ is defined to take into account correlation ρ, that follows the following constraints:

$$\zeta = L\epsilon, where \; L = chol(\rho, 'lower')$$

so that the price path can be computed as:

$$S_{t+1}^i = S_t^i \exp\left[\left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma\sqrt{dt}\zeta_t^i\right]$$

where i represents the assets.

Conditional Value at Risk

The key difference between CVaR and MVO is that MVO penalizes portfolios for any type of excessive movements, whereas CVaR only minimizes losses/downside risk. The problem for the scope of this project can be formulated as follows:

$$\min_{x} CVaR_{beta}(x)$$
$$s.t. x \epsilon \chi$$

where χ represents the constraints relating to budget, short selling, and target returns. CVaR can be defined as:

$$F_{beta}(x,\gamma) = \gamma + \frac{1}{1-\beta}\int_{f(x,y)\geq VaR_{\beta}} f(x,y)^{+} p(y)dy$$

where γ is the Value at Risk with respect to β. Since our cases were generated using a finite amount of scenarios, the problem will be redefined as follows:

$$\min_{x} \gamma + \frac{1}{(1-\beta)S}\Sigma z_s$$
$$s.t. z_s \geq 0, z_s \geq \ (x,y) - \gamma, x\epsilon\chi$$

By generating a sufficient amount of scenarios that project all possible stock paths, we can construct our linear program accordingly. In this project, shorting selling was restricted in order to prevent the linear programming from 'blowing up'.

Sharpe Ratio

Sharpe ratio represents a measure of risk-adjusted return on investment. There are two types of Sharpe ratios, Ex Ante and Ex Post.

Ex Ante: This method uses estimated parameters to calculate the Sharpe ratio. This represents expected Sharpe ratio of the portfolio.

$$SR_p = \frac{\mu^T x - R_f}{\sqrt{x^T Q x}}$$

Ex Post: This method uses realized values to compute the Sharpe ratio. This represents realized Sharpe ratio of the portfolio.

$$SR_p = \frac{\mu_p - R_f}{\sigma_p}$$

Hence, the portfolio's Sharpe ratio is a good indicator to measure and compare the performance. In this project, both types of Sharpe ratios will be used as a performance metric to compare the effectiveness and efficiency of different optimization's model.

Performance Metrics

The following table summarizes the metrics that will be analyzed across the different portfolio strategies, followed by their respective equations.

Metric	Purpose
Return	Compares the gain on the portfolio over the time period
Portfolio Weights	The percentage of each asset in the portfolio
Portfolio Variance	Measures the volatility of the portfolio with its respective weights
Average Return	Represents the return of the portfolios across all periods
Return Accuracy	Compares estimated returns compared to the actual returns
Sharpe Ratio	Average excess return earned of the risk free rate relative to portfolio volatility. This can be calculated ex-ante or ex-post.
Value at Risk	Measures the financial risk using the potential for loss based on β
Conditional VaR	Represents the expected loss with respect to a probability level β

Metric Equations

$$Return\ (any\ time\ frame) = \frac{Final\ Value - Initial\ Value}{Initial\ Value}$$

$$Portfolio\ Variance = x^T Q x$$

$$Average\ Return = \prod_i (1 + r_i)^{\frac{1}{N}} - 1$$

$$Return\ Accuracy = \left| \frac{\mu - Actual\ Return}{\mu} \right|$$

$$Sharpe\ Ratio = \frac{\mu_x - r_f}{\sigma_p}$$

$$Conditional\ Value\ at\ Risk\ = \frac{1}{1-\beta} \int_{f(x,y) \geq VaR_\beta} f(x,y) p(y) dy$$

$$SR_p = \frac{\mu^T x - R_f}{\sqrt{x^T Q x}}$$

$$SR_p = \frac{\mu_p - R_f}{\sigma_p}$$

Sensitivity Analysis

Effects of changes in the following metrics will be analyzed on our portfolio performance.

- Effect of number rebalance periods on portfolio performance
- Changes in confidence level and its effect on CVaR and Robust MVO performance
- Changes in the iteration of the resampling MVO
- Changes in the number of scenarios of the resampling MVO and the Monte Carlo simulations of CVaR

Standard Model Results

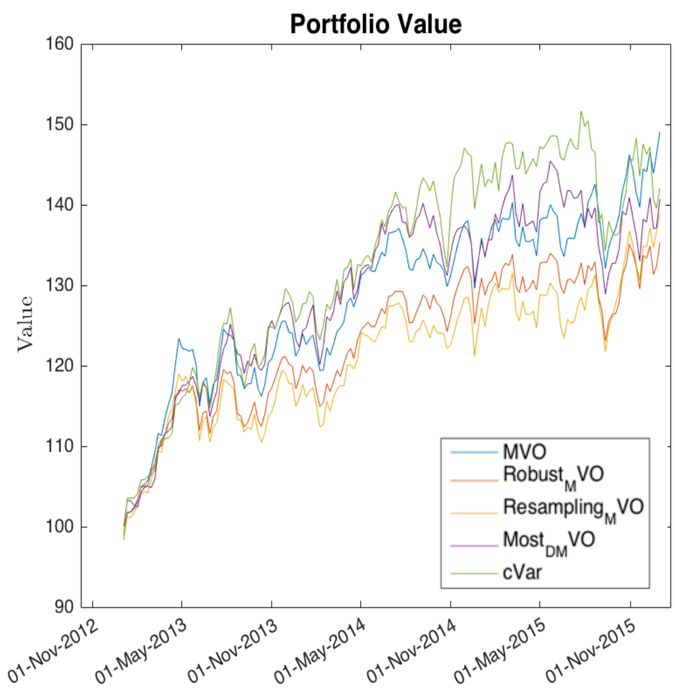
Table 1: Summary Statistics

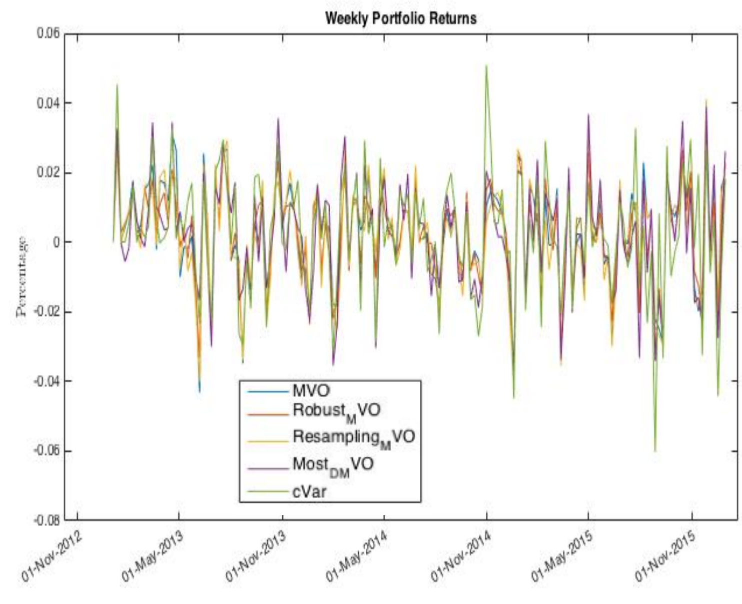
The following are calculated based on

PORTFOLIO TYPE	MVO	ROBUST MVO (BETA = 0.99)	RESAMPLING MVO	MOST DIVERSE MVO	CVAR (BETA = 0.9
FINAL PORTFOLIO VALUE	149.1	135.27	138.973	140.698	142.0702
OVERALL RETURN (4 YEARS)	50.8%	37.5%	37.7%	40.6%	43.5%
VALUE AT RISK (PERCENT) (BETA = 0.95)	-0.5091	-0.02058	-0.04730	-0.05312	0.03042
CONDITIONAL VAR (PERCENT) (BETA = 0.95)	-0.068331	-0.03369	-0.063289	-0.06897	-0.01735
AVERAGE PORTFOLIO VARIANCE	0.1171e-03	0.13733e-03	0.12545e-03	0.15437e-03	
AVERAGE PER PERIOD RETURN	6.02%	4.52%	4.39%	4.85%	5.11%
PERIOD RETURN VARIANCE	0.0052	0.00264	0.00412	0.00475	0.00494
AVERAGE WEEKLY RETURN	0.264%	0.204%	0.205%	0.219%	0.232%
WEEKLY RETURN VARIANCE	0.00002323	0.00001995	0.00002406	0.00002589	0.0000307
AVERAGE ESTIMATED WEEKLY RETURN	0.243%	0.238%	0.199%	0.332%	0.516%
AVERAGE EX-ANTE SHARPE RATIO	0.2617	0.2319	0.2098	0.2656	0.3421
AVERAGE EX-POST SHARPE RATIO	12.873	13.565	10.243	13.895	13.473
TOTAL TRANSACTION COSTS (0.5% OF TRADED VOLUME)	$3.42	$1.67	$2.87	$4.35	$2.91

Data

Figure 1. Portfolio Value and Weekly Portfolio Returns





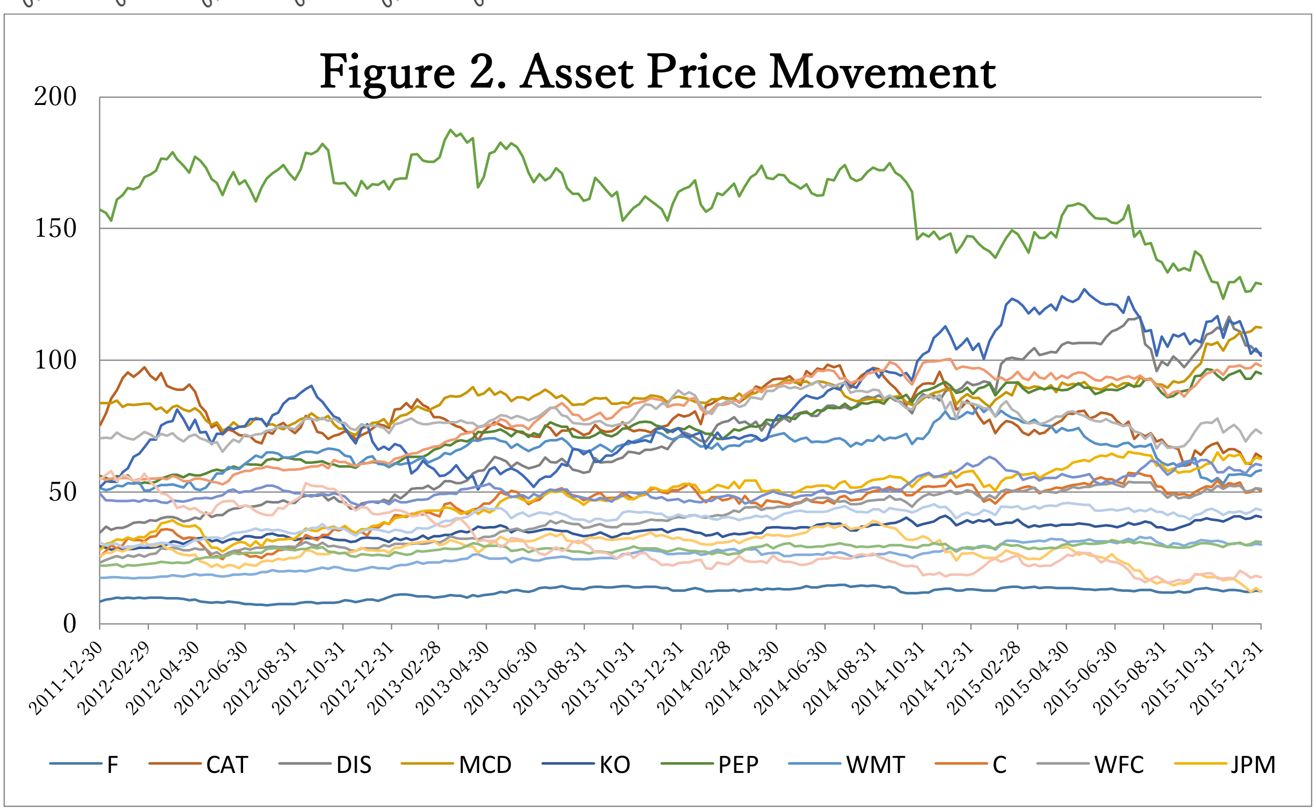


Figure 3: Nominal MVO Weights

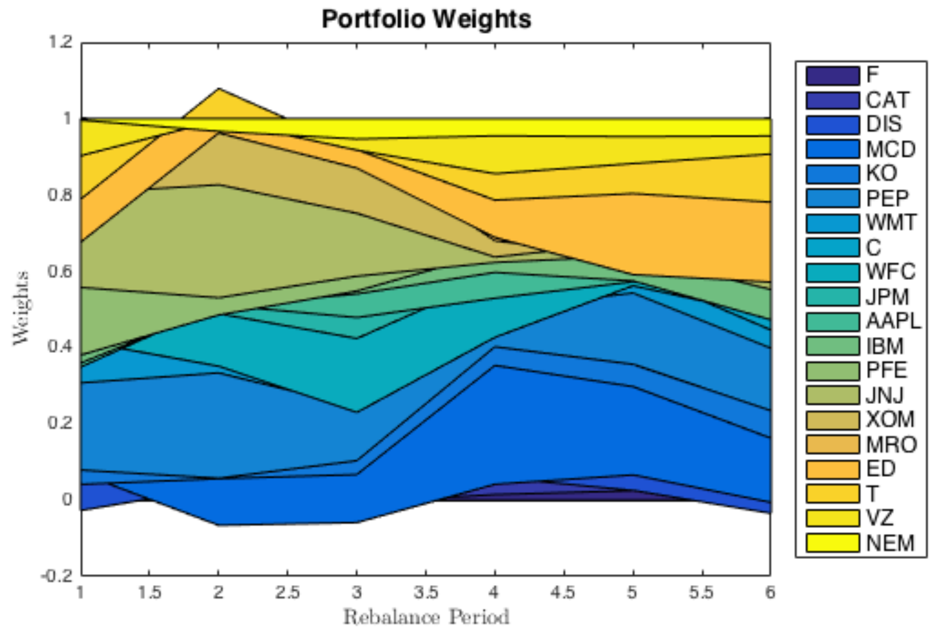


Figure 4. Robust MVO Portfolio Weights

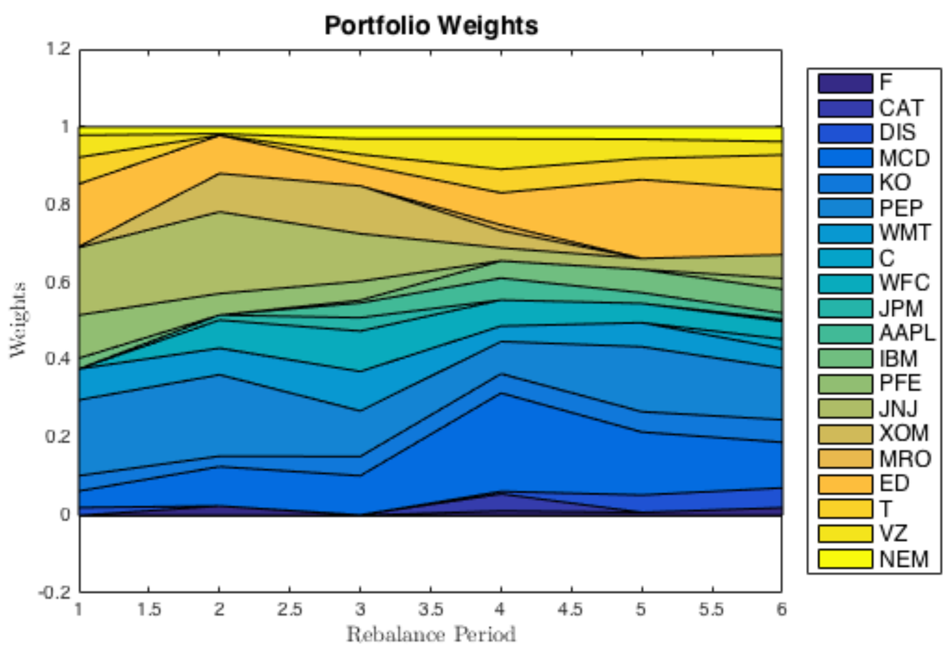


Figure 5. Resampling MVO Portfolio Weights

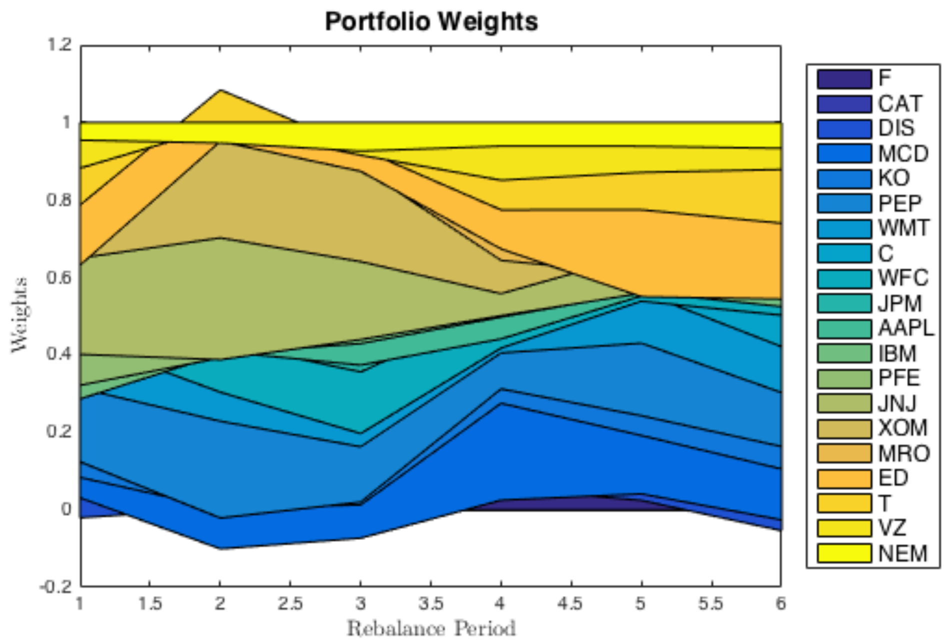


Figure 6. Most-Diverse Portfolio Weights

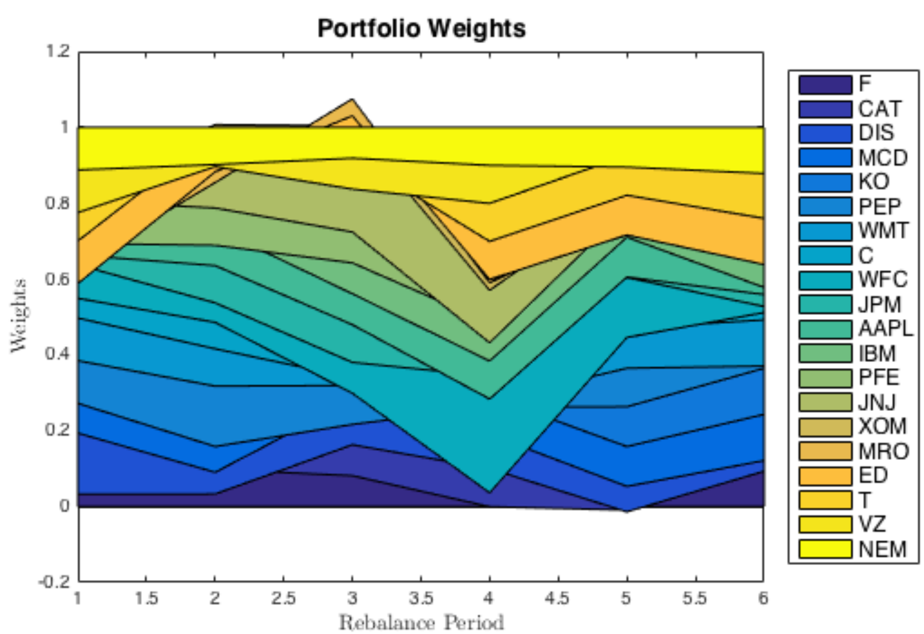


Figure 7. CVaR Portfolio Weights

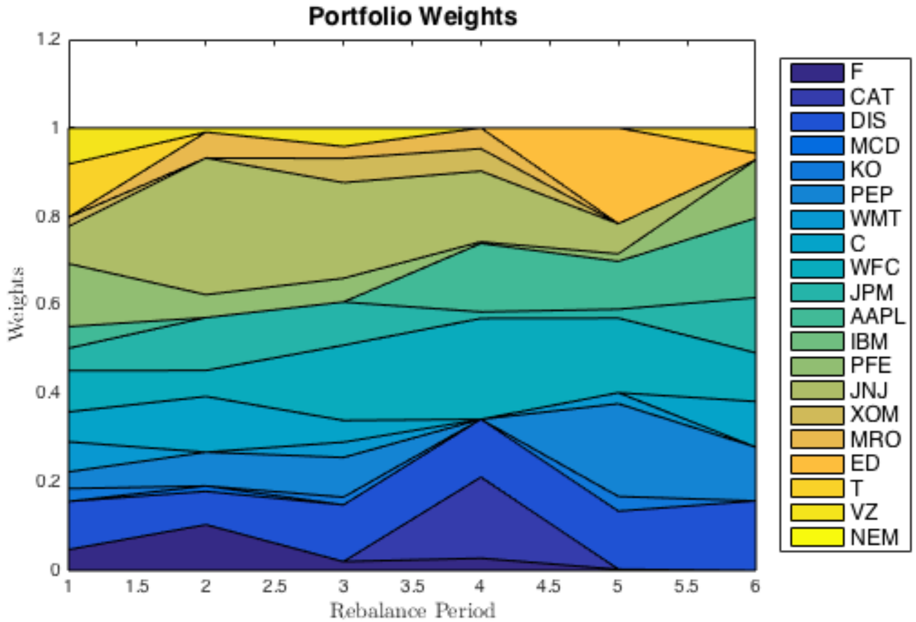


Figure 8. Ex-Post Sharpe Ratio and Ex-Ante Sharpe Ratio

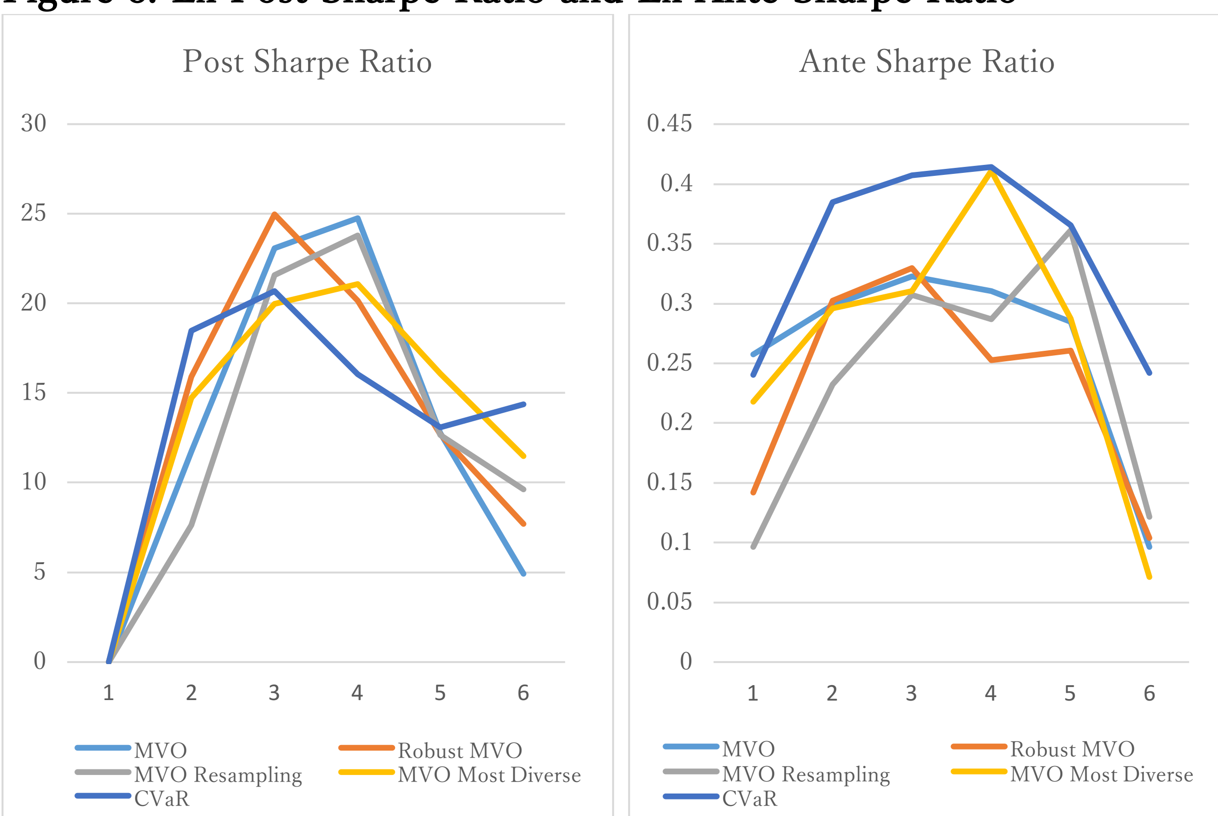


Figure 9. Transaction costs per rebalance period of various portfolios

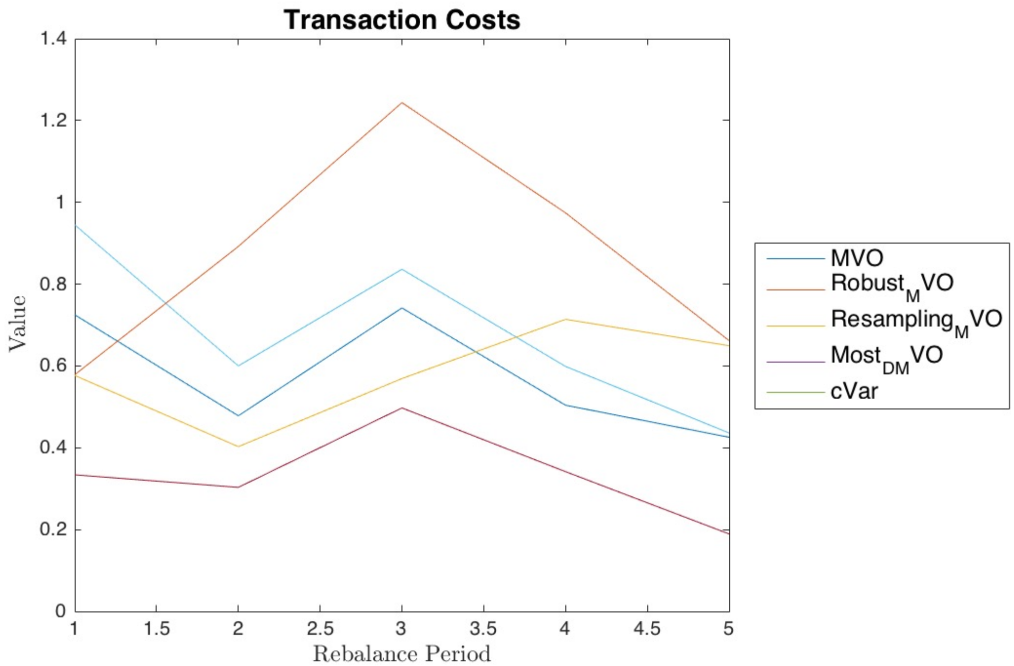
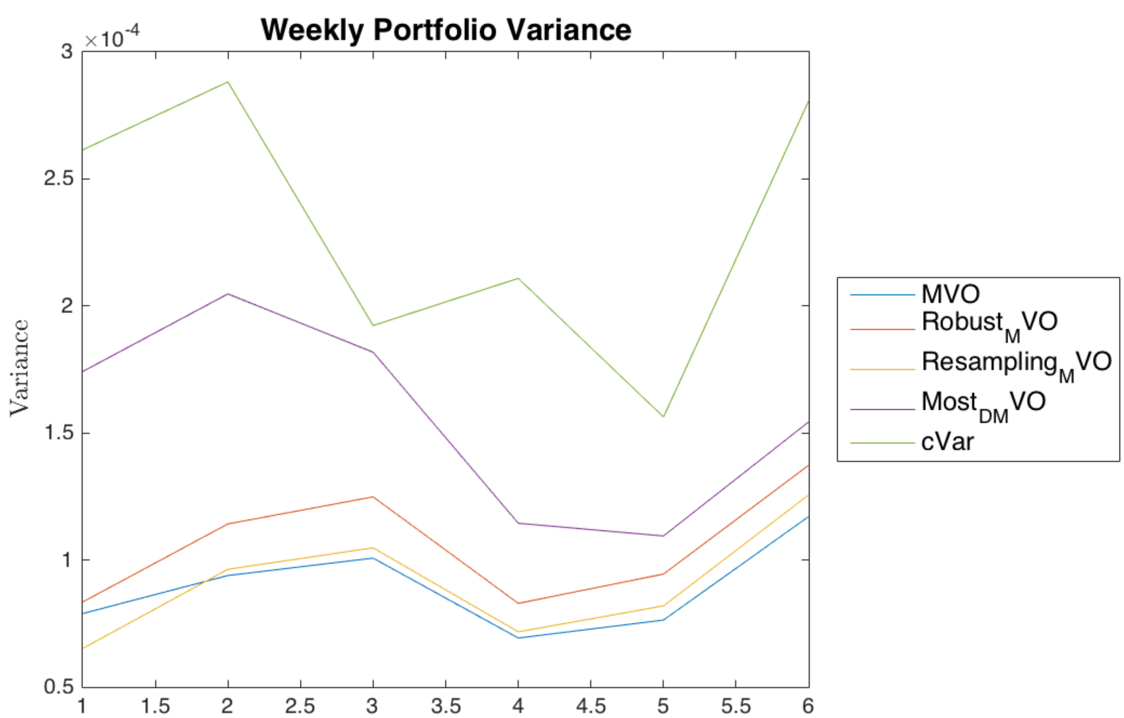


Figure 10: Portfolio Return Variance over Rebalance Periods



Model Variant Results

Note that when a certain variable is altered for sensitivity analysis, all the other variables are kept constant. Furthermore, the MATLAB function rng(1) is used to ensure that random variables remain consistent across all trials.

Figure 11 : VaR of Different Portfolios

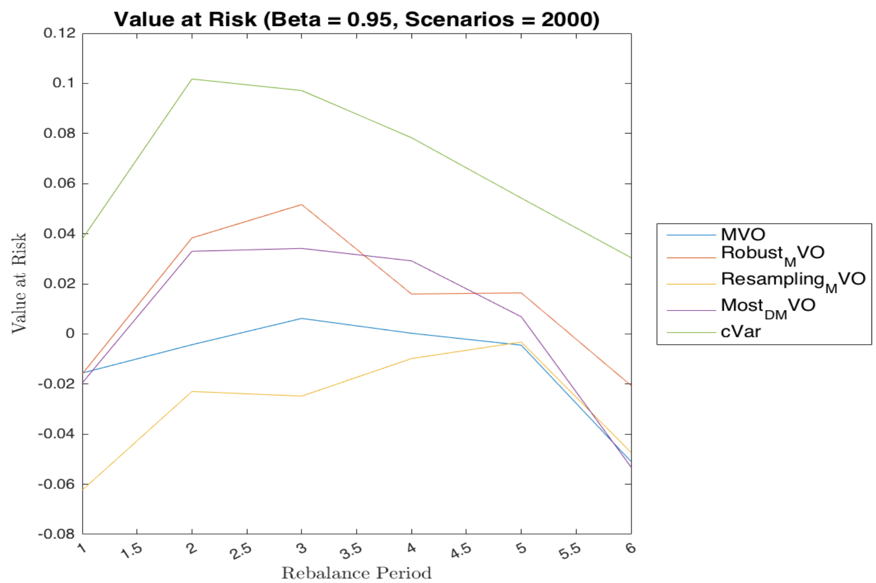


Figure 12 : CVaR of Different Portfolios

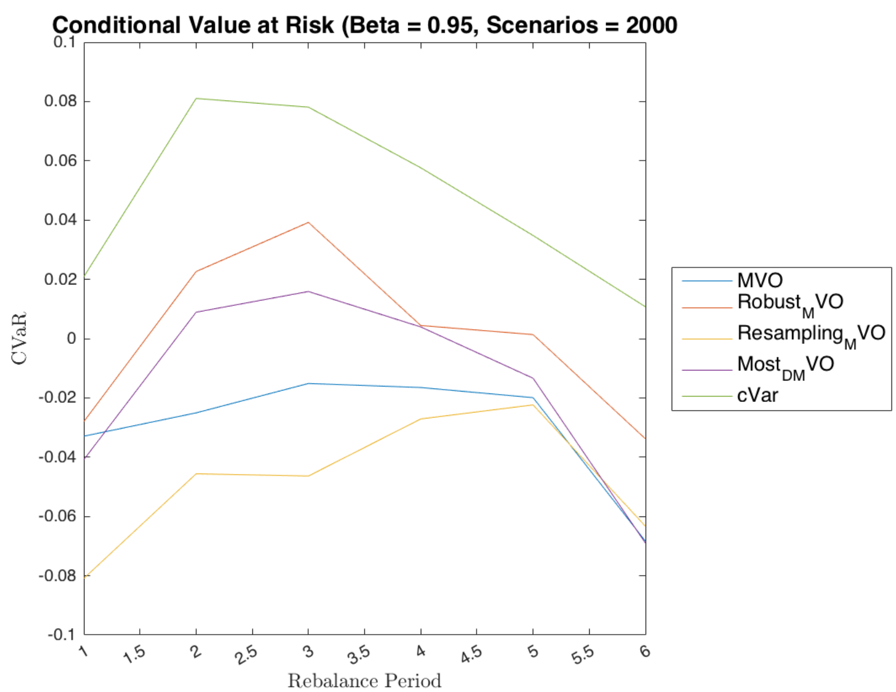


Figure 13: Effect of Confidence Level on CVaR Portfolio Value

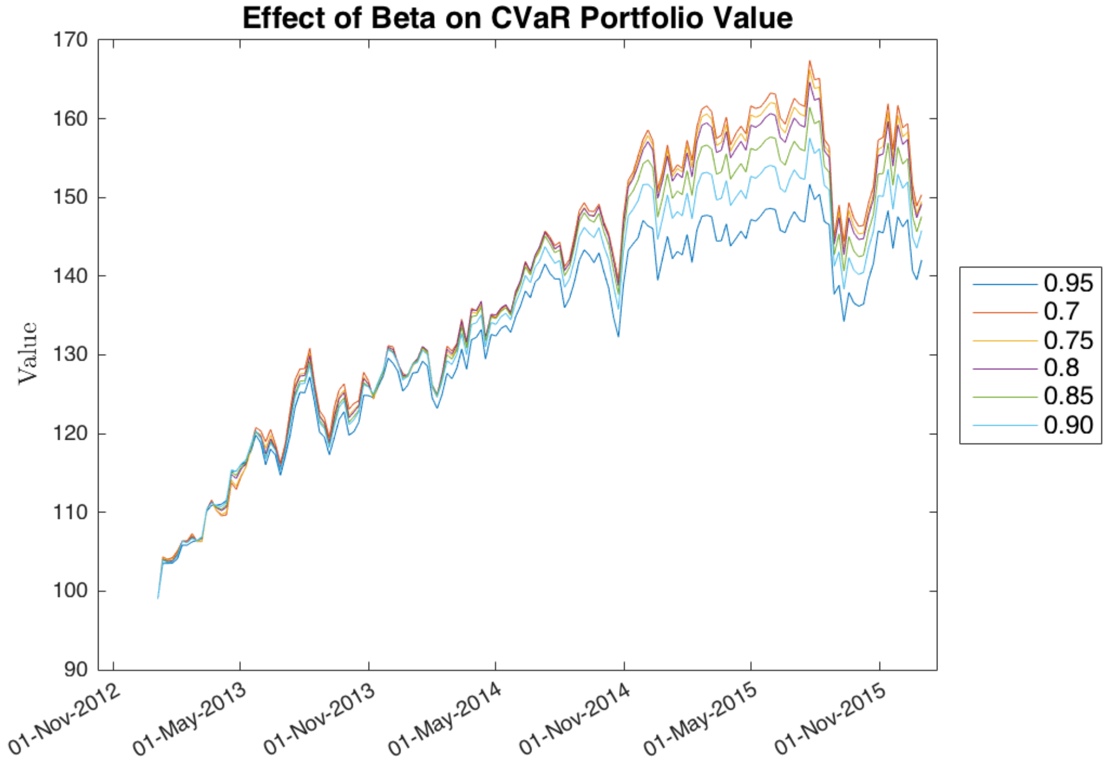


Figure 14: Effect of Number of Scenarios on CVaR Portfolio Value

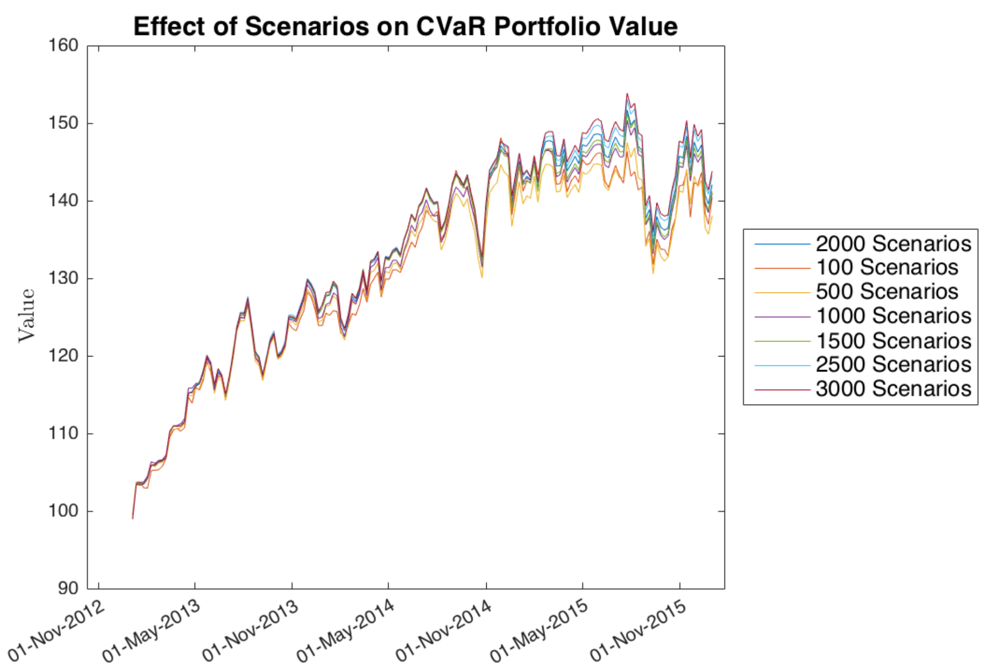


Figure 15: CVaR Runtime Dependence on Number of Scenarios

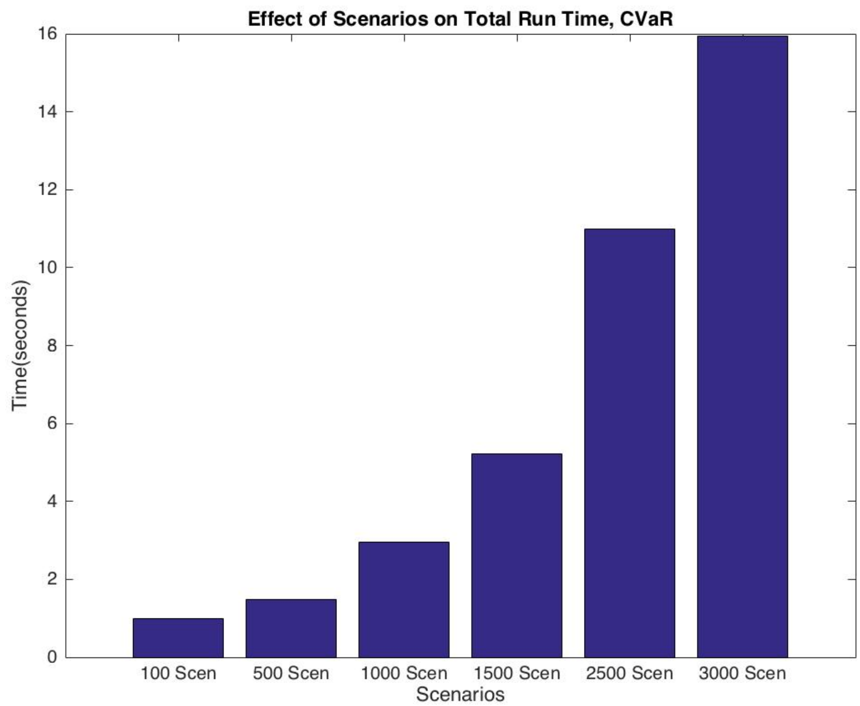


Figure 16: Effect of Number of Scenarios on Resampling Portfolio

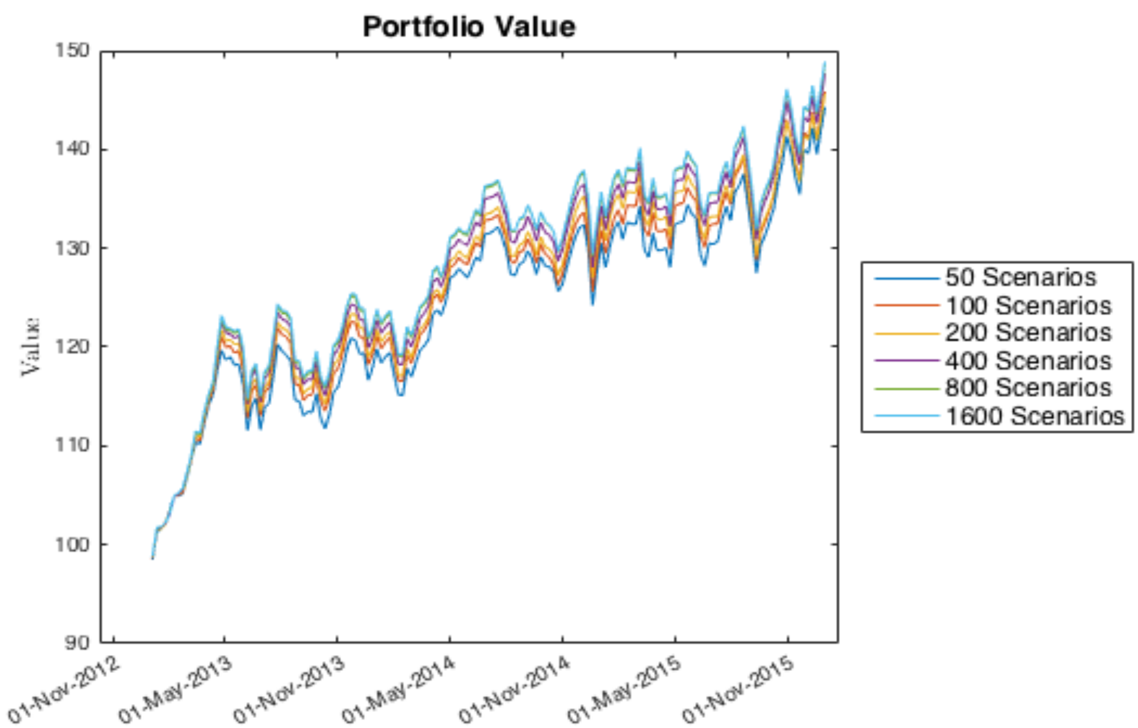


Figure 17: Resampling Run Time Dependence on Scenarios

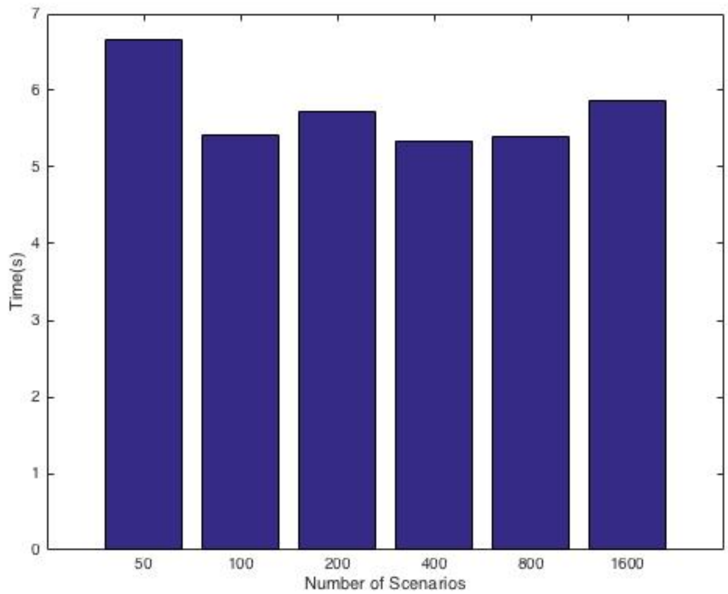


Figure 18: Effect of Number of Iterations on Resampling Portfolio

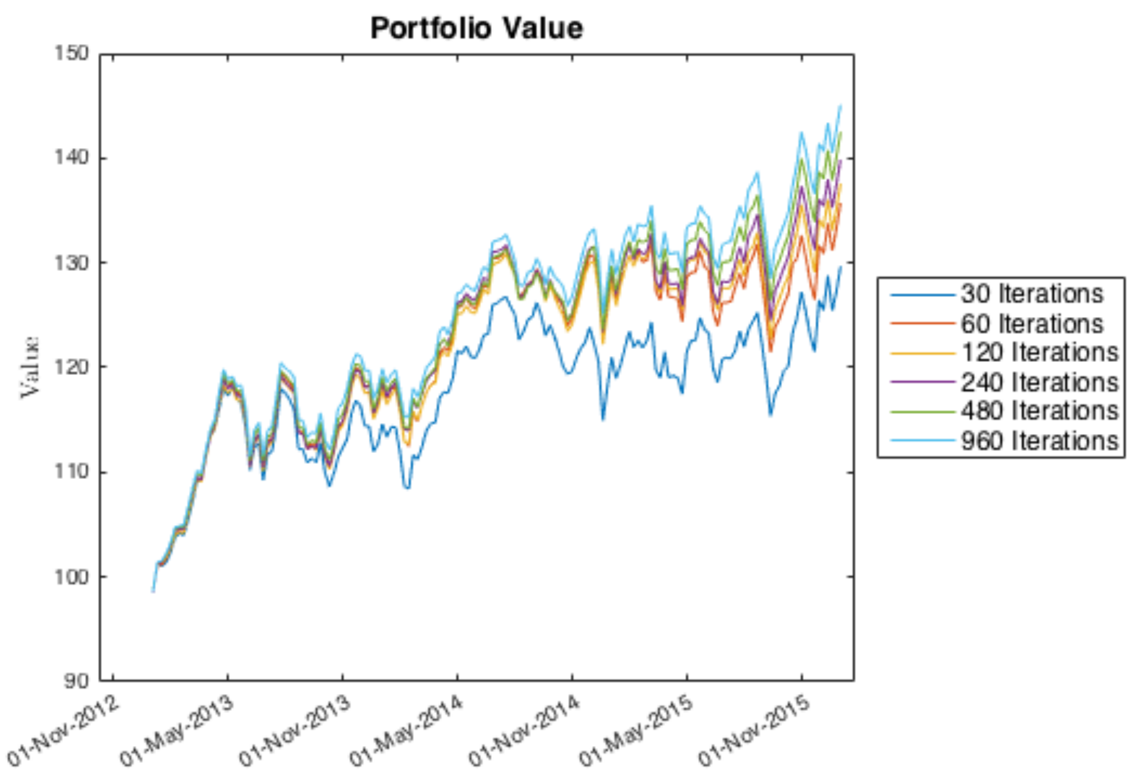


Figure 19: Resampling Run Time Dependence on Iterations

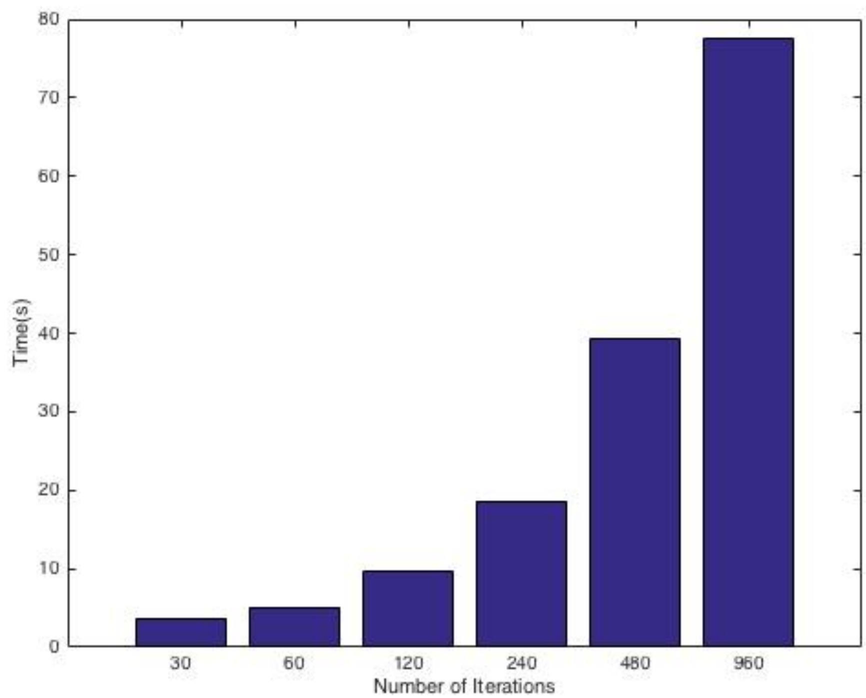


Figure 20: Effect of Confidence Level on Robust MVO Portfolio Value

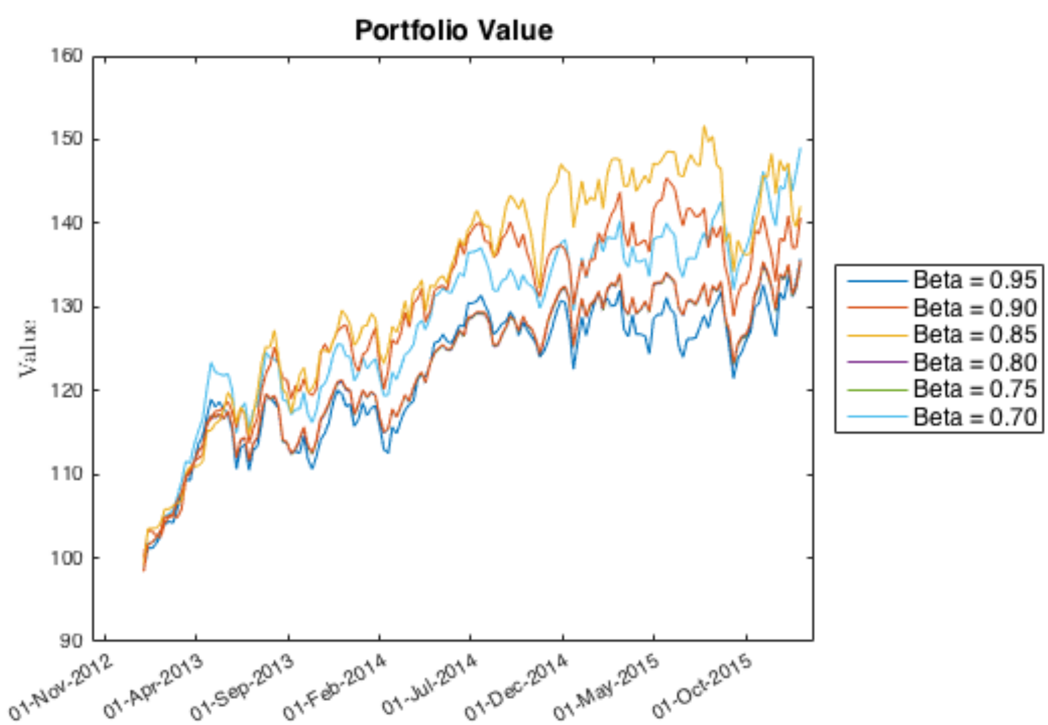
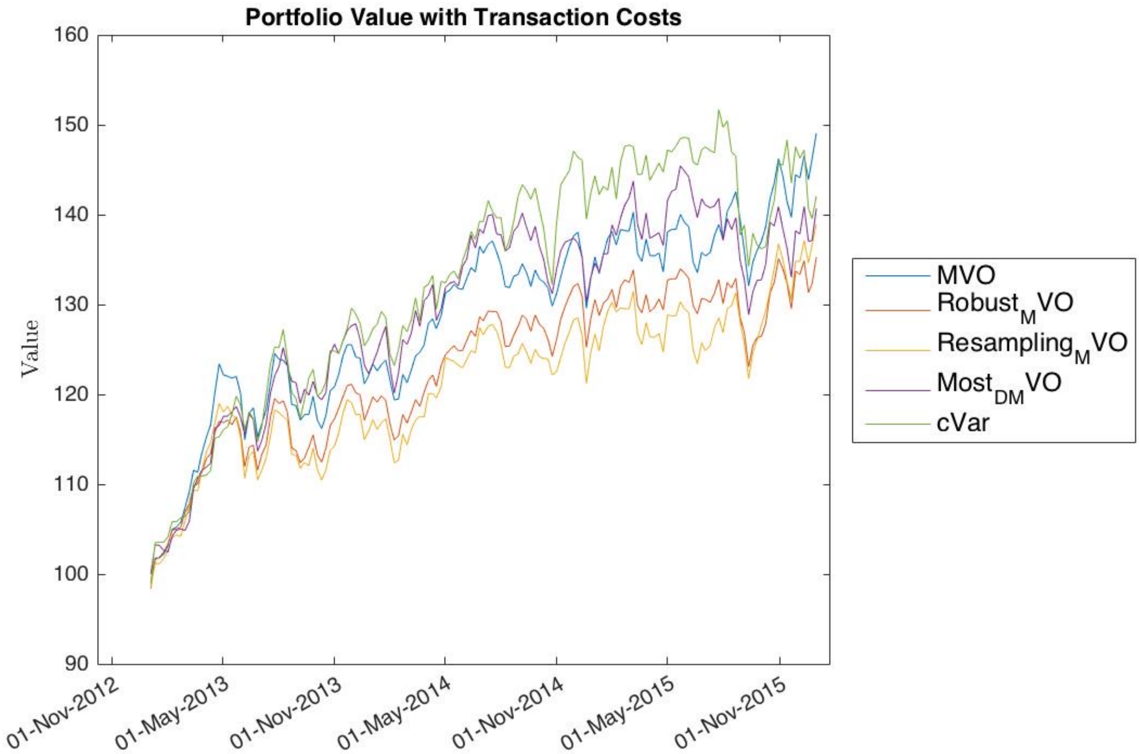


Figure 21: Portfolio Value with Transaction Costs



Analysis

Important Results

- CVaR optimization, as expected, had the lowest VaR and CVaR values
- CVaR values for CVaR exhibited positive returns for $\beta = 0.95$, meaning that even in the worst cases the portfolio expects to exhibit positive returns
- Transaction costs are all relatively even across all portfolios, however, most diverse MVO has the highest costs
- The MVO-optimized portfolios had a much lower variance than CVaR (CVaR had >200% higher variance)
- Higher Beta does not exhibit higher returns for CVaR: Beta=0.7 had the highest overall returns for CVaR portfolios
- The number of scenarios has a positive effect on the overall return of the portfolios for CVaR: 3000 scenarios had the best results, whereas 100-500 scenarios had the lowest returns

Portfolio Performance

The performance of each portfolio can be found in Figure 1, in terms of value over each rebalancing period as well as weekly return. At first glance the takeaways are CVaR performed the strongest throughout the three years, although Nominal MVO pulled ahead in value on the final date. Robust MVO finished the simulation with the lowest value, while the remaining three investment strategies were almost identical. The weekly return saw the portfolios overlap throughout the simulation. The best weekly return was seen by CVaR in November 2014.

Looking again at Figure 1, it is clear all figure methods yielded very similar patterns when plotting portfolio value over time. This observation is expected, since it is common for the majority of the market to move in the same direction on any given day, whether up or down. It is speculation to guess the movement of a stock on a single day. The common shape is highlighted by the severe correction around October 2015, as well as the other peaks and valleys that overlap. The portfolio value is practically identical over the first half of 2013. Therefore, it is interesting that the portfolios perform nearly identically over small time spans, however, the small differences add up over annual returns and even more over the full three-year period. The compounding effect is highlighted as the spread between portfolio values grows over time, with CVaR, Most Diverse MVO and Nominal MVO. On the other hand, Robust MVO and Resampling MVO trail, however, the spread becomes smaller towards the end of the simulation. This is further shown by the small distance between each line in 2013, before increasing in 2014.

Figure 1 yields the very important conclusions, that if investing over a small time span, there is practically no difference between the performance of each investment strategy, however, it can compound into significant value differences over longer time periods. This is confirmed by the overlapping weekly returns.

Now looking again at the correction that occurred in October 2015, all five investment methods experienced a large value drop. However, the correction in the CVaR portfolio was much more severe than the others. To explain this, we can compare Figure 7, which contains the weights for the CVaR portfolio, to Figure 2, Asset Prices over time. CVaR had the misfortune of investing in assets that experienced larger drops such as IBM. Over a short period of time, its substantial lead over the other investment strategies was diminished. For this reason, Nominal MVO finishing the simulation at as highest portfolio value at the end of 2015 does not make it the best strategy. It is very likely that if the data was expanded into 2016, the nominal version would once again underperform other strategies.

All portfolios were able to gain significant value, increasing between 30% to 50% over the three-year span. This again highlights that investment strategy is important, but is still reliant on the trends of the market. If an investor was more comfortable with any of the five strategies versus the others, their portfolio would still change value in the same direction of the others even if the returns are smaller.

Figure 10 shows the weekly portfolio variance over rebalance periods. CVaR had the highest variance throughout the simulation. Nominal MVO, Robust MVO, and Resampling MVO all had very similar variances.

Robust MVO Analysis

The confidence levels in Figure 20 illustrate the volatility and differences in portfolios when different confidence levels are used. Beta of 0.85 had the best performance overall, and Beta of 0.95 had the lowest portfolio value. Finally, Beta of 0.7 had the best final value, whereas the intermediate portfolios performed between these two. This is likely because the higher confidence levels do not give the algorithm as much 'freedom', thus safer assets are chosen instead. Thus, it is evident that the highest confidence level does not necessarily result in the highest return for Resampling MVO.

The robust MVO graphs had consistently low Sharpe ratios relative to other portfolios. This is likely due to the additional constraints that resulted in a lower portfolio return compared to the others. The Robust model consistently had conservative returns compared to the other models, however it exhibited very low variance across all periods.

Resampling MVO Analysis

Resampling MVO had the second lowest variance next to MVO (and was only slightly greater), and this is because of its investments in uncorrelated assets. However, as a result of this it could not capitalize on increases in the market for related assets, and thus had one of the lowest returns out of all the methods. Furthermore, it also had a very similar Sharpe ratio to the MVO method as well, but overall it exhibited much lower returns.

Varying the number of scenarios for resampling MVO had a positive impact on portfolio returns, as can be seen by Figure 18. However, unlike other scenario dependent methods, the run time does not increase significantly for resampling MVO. Because the resampling in these scenarios is done by the mvnrnd() MATLAB function, which is optimized by MATLAB, there is not a significant impact of these scenarios as they do not increase the difficulty of the problem. Also, the number of scenarios does not have a significant impact on the final portfolio returns as can be seen by Figure 16 (they all exhibit the same upwards and downwards movements), however there is still a benefit that arises from having more scenarios. Thus, there is no advantage to having less scenarios for resampling MVO.

However, a key result is the effect of iterations on the portfolio performance. The effect of number of iterations is much larger than scenarios, however, the iteration time increases exponentially as well. Thus, if more analysis is done, it is suggested that a convergence point between time, iterations and payoff is found to find the most optimal scenario/tradeoff for run time. Less than 100 iterations exhibited significantly worse performance, thus the tradeoff in those areas is obvious.

CVaR Analysis

CVaR was one of the best portfolios across different metrics. It consistently outperformed the portfolio values of the other stocks until the end (likely because an unexpected drop happened in the market that was not accounted for in the scenarios), and furthermore it has one of the highest Sharpe Ratios as well despite a much higher variance than the other portfolios.

Based on the figures related to CVaR, it is interesting to see the effect of different parameters on performance accuracy and returns. First of all, the CVaR portfolio consistently exhibited positive Value at Risk (when Beta = 0.95), as this means that even in the worst 5% of scenarios the portfolio expects to exhibit positive returns. This scenario is likely because the scenarios are not as 'strict' in terms of optimization constraints, thus the CVaR with lower Beta tend to invest in assets that exhibit higher returns rather than avoiding assets with higher downsides. Thus, Beta=0.75 actually exhibited the best returns across all the CVaR portfolios, while Beta = 0.95 exhibited the lowest. A good middle ground between the two is Beta = 0.85; despite having a decently high confidence interval, the portfolio still performed significantly better than the base case. This scenario is only likely, however, if the market moves as the scenarios predict; in the cases where sudden drops or crashes occur, there is not as much accuracy and higher confidence intervals will be better to prevent losses.

The effects on number of scenarios was not as drastic as expected. It is clear that the lower numbers of scenarios had lower returns than the higher amounts, with the best results exhibited by 3000 and 2000 scenario cases. In the initial year, the portfolios were nearly identical in terms of performance. However, the larger differences were exhibited in sharp drops and rises in the market; slightly differing weights in these scenarios (as can be seen between Nov 2014 – Nov 2015) results in large differences in portfolio value. These effects typically compound with the rebalancing periods, and result in lower returns overall for portfolios with 100-500 scenarios. This effect can also be seen in the difference between expected returns and actual returns; as lower scenarios are not able to accurately project future price movements, the resultant difference between the expected and actual returns is higher. Thus, having a higher number of scenarios is beneficial, however once the amount is representative ‘enough’ of possibilities, there is not much point in further scenario time. This is key especially when algorithms are run on short term data and for higher efficiency (i.e algorithmic trading), that a higher number of scenarios may delay run times significantly as was seen by the code when different scenarios were used. Thus, for the purposes of this project, 2000-3000 scenarios exhibited very similar returns, and anything much greater may not result in as effective results compared to run times when used on high frequency algorithms.

Further analyses into these phenomena can be examined by altering the number of time steps used in the Monte Carlo Algorithm. However, based on previous testing and sampling, greater number of time steps only results in higher computational power and not necessarily greater accuracy when used in the algorithms: perhaps extreme cases in scenario amount and data estimates may exhibit significant results.

Transaction Costs

Similar to the portfolio values, the transaction costs displayed similar patterns. This can be seen in figure 9. Out of the five investment strategies, transactions costs at rebalance 2 were lower than rebalance 1 for all portfolios except Robust MVO. The costs then all rose between periods 2 and 3, before dropping throughout the rest of the simulation. Resampling transaction costs continued to rise until rebalance 4. The transaction had a decreasing trend overall, which was expected. As portfolios exist over longer time spans, they move closer to the ideal portfolio, therefore less rebalancing should be needed. However, the need for rebalancing and magnitude of transaction costs will never approach zero due to the unexpected nature of the markets.

In Figure 21, the effects of transaction costs can be seen on the effects of the portfolios. The price of the transaction costs is incorporated as such:

$$Beginning\ Period\ Price_t = Final\ Period\ Price_{t-1} - Tcost_{t-1}$$

The only noticeable effect of these transaction costs is slightly decreased final portfolio values. However, they do not significantly impact the returns of portfolios relative to each other. The conclusion from this is that the group of optimization tactics selected do not result in severe rebalancing techniques, and the market illustrated relatively consistent returns across the studied time period.

Areas of Improvement

Parameter Uncertainty

Stock prices change significantly throughout the week. By only using weekly factor models and stock prices for estimation purposes, we are significantly decreasing the accuracy of our estimates as this only outlines 20% of stock closing data. Given that stock data is readily accessible for all time periods, an improved iteration of this model should incorporate daily stock changes (or even hourly) for an improved model. This will greatly decrease the variance in our estimates and allow us to make more specific and accurate models for the short term. This phenomena is known as mean blur. Increasing the data set by moving to daily returns would reduce the blur, however, the variance to man value ratio would increase.

The effects of weekly values have a great impact on the accuracy of our variance (Q matrix) estimates. Since the Adjusted Close stock data is retrieved based on weekly prices, the returns are found as weekly returns, thus having a direct impact on our Variance matrix. Furthermore, the accuracy of the variance estimate, however, can be improved if there are daily return samples rather than weekly (as there are 5 data samples versus a single sample).

This can be summarized in the equation below:

$$\sigma^{week}(\sigma^2) = \sqrt{\frac{2}{5-1}\sigma^2}$$

Thus, assuming we want to continue using weekly returns to find portfolio variance, it will be better to use daily data to estimate the weekly returns to minimize the deviations in variance by a significant amount.

Lack of Rebalance Periods

A major area of improvement for these portfolios would be to increase the frequency of rebalancing periods. Simply projecting the returns using past data is not sufficient as many activities can shift actual returns significantly. Although increasing the frequency can significantly increase transaction costs, the risk of elongated periods of unstable markets can be avoided by reinvesting into more stable stocks (or shorting the market). Transaction cost constraints can also be set and solved using other linear programming methods, in the case that they become too high. Building off the parameter uncertainty, rebalancing more often would also require more complete data sets, such as daily returns instead of weekly returns

Between rebalancing periods 4 and 5, all portfolios exhibit a significant drop-off due to an overall weakening in the 20 assets available. See Figure Price Movement. Hence, had there been a rebalancing period between these periods, losses across all portfolios could have been decreased, either through shorting the market or by investing in stable stocks. In the real world, a portfolio manager would adapt to the changes in the market, instead of just watching the value loss until the next rebalancing period. If the manager believed in their stock picks, they would increase their position to take advantage of the recovery, however, this is more of a qualitative strategy than quantitative portfolio optimization.

However, more rebalancing period means more transaction costs which will decrease the overall portfolio return. Also, practically it takes time to acquire and sell the assets thus too many rebalancing could result in a less profitable portfolio. With recent improvements in technology and the emergence of algorithmic trading, this becomes less of a concern. Although, the trade-off between the possible gain opportunity and transaction costs should be carefully compared.

Round Lots

In practice, it is not feasible to buy or sell fraction of a share. Instead, shares are traded in multiples of a unit size. Therefore, in order to use these portfolio weights in order to invest, integer programming that is able to incorporates the round lots should be used. This change with the variable type will affect the all models.

Multiple Objective Functions

Optimizing based on multiple objective functions would allow investors to optimize for multiple goals all at once. In theory, a portfolio could be formed while optimizing for CVaR as well as Robust versions of MVO all at once. The drawback here is the optimization program may be unable to solve the problem into a portfolio, and the problem would be difficult to formulate. However, if successful the portfolio variance would likely be lower than the five strategies illustrated in Figure 1.

Different Stochastic Programming Models

Another interesting point of comparison for our current studied algorithms would be to study other stochastic programming models. These models can be compared to our existing Monte Carlo simulations, and more in depth models to model scenarios can be used instead.

Extensions to the Model

For a more extensive analysis of the Optimization Models (outside the scope of this project), the following can be performed to further understanding of the models and better understanding of the parameters:

- More Assets: Instead of just 20 assets, a test on all assets in the S&P500 or in volatile 'penny' stocks would prove to be an interesting analysis, as not all optimization methods may yield positive returns
- Effect of Rebalance Frequency on Portfolio Returns: Increase the rebalancing periods to once a month or once a week; find a balance between portfolio returns (including transaction costs)
- Machine Learning Regression for Factor Models: Using machine learning regression models, generate more in-depth factor models for our training period to decrease measurement inaccuracy

Conclusion

The project yielded the expected results in terms of variant metrics, however, it was unexpected that MVO finished with the highest portfolio value. It is evident that portfolio value is not the most important aspect of the returns, and neither is minimizing variance. CVaR proves to be a very effective measure of downside risk and especially during stable markets the scenarios can be estimated relatively effectively. Problems come into play when random drops and volatility occur in the market however. Interestingly, likely due to the time period studied, the VaR values for CVaR remained positive across all rebalancing periods. The Robust MVO portfolio had the highest transaction costs, while Most Diverse MVO had the lowest. Across all portfolios, it is evident that large divergences multiply after significant time has passed. Further analyses can be made into the studied methods using increased metrics and portfolio optimization techniques. Overall, it can be concluded that CVaR and variance are both effective measures for performance, however imposing large degree of constraints can limit the constraints of these portfolios significantly.

References

"MIE377 Lecture Notes". Giorgio Costa. Winter 2018. Web.

Luenberger, David. "Investment Science." *Stanford University.* Oxford University Press. 1998.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1):3–56.

French, K. R. (2016). Data library. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. [Online; accessed 20-Sep-2017].

Quandl.com (2017). Wiki – various end-of-day stock prices. https://www.quandl.com/product/WIKIP/usage/export. [Online; accessed 07-Nov-2017].