Financial Optimization Models for Portfolio Asset Allocation

Matthew Reiter[1], Daniel Kecman[2] and Yiqing (Louis) Luo[3]

[1] matthew.reiter@mail.utoronto.ca
[2] daniel.kecman@mail.utoronto.ca
[3] louis.luo@mail.utoronto.ca

Abstract. The purpose of this report is to compare the performances of five financial optimization models: (1) Mean-Variance Optimization (MVO), (2) Robust MVO with an ellipsoidal uncertainty set, (3) Re-sampling MVO, (4) Most-diverse MVO and (5) Conditional Value-at-Risk (CVaR) Optimization with Monte Carlo Simulations. Techniques 1, 2, 3, and 4 were designed to use the Fama-French three factor model for parameter estimates, and technique 5 incorporated a Monte-Carlos simulation to generate future prices. The performance of each optimization model was compared by running an investment simulation based on the weekly adjusted closing prices of twenty stocks from S&P500, beginning in January 2013 and ending in December 2015. It was observed that the largest return was realized under MVO and re-sampling MVO, with both models returning over 60% (of time zero value) by the end of the simulation. The behavior of different re-sampling parameters was measured. The performances of other risk ratios of such as the Value-at-Risk (VaR), CVaR, and Sharpe Ratios were evaluated for the portfolio generated by each model.

1 Methodology

We motivate our work by formally stating our goal: we are interested in exploring the capabilities of financial models to construct an optimal portfolio of stocks. We turn to optimization models because we are interested in developing an efficient framework to justify our asset allocations to minimize risk of loss and ensure a confidence of gain.

The models presented in section 2 will be used to construct portfolios which will consist of up to twenty assets, all from the S&P500. The simulation is run from 2013 to 2015 based on the weekly adjusted prices of the twenty stocks. Since we are using historical data, we have the luxury of immediately comparing our results to realized returns.

The remainder of this section will be dedicated to explaining the methodology used in this project. All topics will be relevant to some application of a portfolio model presented in section 2.

Portfolio Mean and Variance: Central to an optimization model are the parameters we rely on. Suppose we are interesting in tracking up to n assets, each having the following properties

- a realized return of r_i,
- an expected return of μ_i,
- a return variance of σ_i^2, and
- a return covariance of σ_{ij} with a different asset j

We can setup our optimization problem by establishing the decision variables x_i which correspond to the proportions of our wealth to invest in each asset i. With that, we can state the expected return μ_p and variance of our portfolio σ_p^2

$$\mu_p := \sum_{i=1}^{n} \mu_i x_i \tag{1}$$

$$\sigma_p^2 := \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij} \tag{2}$$

It is convenient to develop our parameters in matrix form, so we define $\mu \in \mathbb{R}^{n \times 1}$ to be the vector of expected returns, $\mathbf{x} \in \mathbb{R}^{n \times 1}$ to be the vector of optimized asset allocations and $Q \in \mathbb{R}^{n \times n}$ to be a symmetric covariance matrix. We express these quantities explicitly

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad Q = \begin{pmatrix} \sigma_{11}^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_n^2 \end{pmatrix}$$

The equations above provide the parameters for the typical optimization model, but we need to justify how to arrive at parameter estimates. For such a task, we use a factor model.

Factor Model: A factor model attempts to explain the returns of asset by defining the returns as a function of several observed factors. Factor models can range in complexity being driven by different economic views which are considered important drivers for the return rate of an asset. The generic form of a k factor model is a regression model:

$$r_i := \alpha_i + \sum_{i=1}^{k} \beta_{ik} f_k + \epsilon_i \tag{3}$$

Where r_i is taken to be the rate of return of asset i. Note that we preserve randomness of the return r_i by noting that our factors may be noisy and ϵ_i represents idiosyncratic risk particular to the asset. Note that the following conditions are assumed when in an ideal environment

– The factors are independent from one another, $\sigma_{f_i,f_j} = 0 \quad \forall i \neq j$
– The factors are independent to idiosyncratic noise, $\sigma_{f_i,\epsilon_j} = 0 \quad \forall i, j$
– Idiosyncratic risk from one asset is independent to idiosyncratic risk from another, $\sigma_{\epsilon_i,\epsilon_j} = 0 \quad \forall i \neq j$
– Idiosyncratic risk is normally distributed with mean 0 and variance σ_{ϵ_i}, meaning $\epsilon_i \sim \mathcal{N}(0, \sigma_{\epsilon_i})$

For the sake of developing our parameter estimates, we will not be assuming ideal conditions in our project. We do this because, as we shall soon see, it often is the case that it is difficult to establish parameters which are independent of each other. With that, we proceed with our parameter estimates from the following

$$\mu_i := \mathbb{E}(r_i) \tag{4}$$
$$\sigma_i^2 := \mathbb{E}\big(r_i - \mathbb{E}(r_i)\big)^2 \tag{5}$$
$$\sigma_{ij} := \mathbb{E}\big(r_i - \mathbb{E}(r_i)\big)\big(r_j - \mathbb{E}(r_j)\big) \tag{6}$$

It is easier to work with factor models when in matrix form, so we express equations for $\mu \in \mathbb{R}^{n\times 1}$ and $S \in \mathbb{R}^{n\times n}$

$$\mu = \beta^T \mathbf{f} + \epsilon \tag{7}$$
$$S = \mathbf{x}^T Q \mathbf{x} + D \tag{8}$$

Where $\bar{f}_i$ is a geometric return of factor i. For clarity, we explicitly state the following:

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} \quad \mathbf{f} = \begin{pmatrix} \bar{f}_1 \\ \bar{f}_2 \\ \vdots \\ \bar{f}_n \end{pmatrix} \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix} \quad D = \begin{pmatrix} \sigma_{\epsilon_1}^2 & \sigma_{\epsilon_{12}}^2 & \dots & \sigma_{\epsilon_{1n}}^2 \\ \sigma_{\epsilon_{21}}^2 & \sigma_{\epsilon_2}^2 & \dots & \sigma_{\epsilon_{2n}}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{\epsilon_{n1}}^2 & \sigma_{\epsilon_{n2}}^2 & \dots & \sigma_{\epsilon_n}^2 \end{pmatrix}$$

Fama-French Three Factor Model: The Fama-French three-factor model[4] will be what is used in this project for MVO parameter estimation. The model is as follows

$$r_i - R_f := \alpha_i + \beta_{im}(f_m - R_f) + \beta_{is} SMB + \beta_{iv} HML \tag{9}$$

For which

– $r_i - R_f$ is the excess return of asset i with respect to the risk free rate R_f
– α_i is the intercept term, representing the absolute return
– β_{im} is the market risk factor loading, tracks how an asset moves with the market portfolio

[4] Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1):356.

- $f_m - R_f$ is a factor that tracks the excess return of the market portfolio
- β_{is} is the size risk factor loading, relates the performance of the stock's return to it's company's size
- SMB is a factor taken as the average return of a portfolio composed of small cap stocks minus the average return of large cap stocks
- β_{iv} is the value risk factor loading, relates the perceived value of the company's stock to performance
- HML is a factor computed as the average return of a portfolio composed of stocks with the highest Book-to-Market ratio minus the stocks with the lowest Book-to-Market ratio.

The Fama-French Three factor model does not satisfy the ideal conditions for a factor model, since the factors may very well be dependent. So, it is justified to use equations 7 and 8 for parameter estimates. The next part of this section will summarize how to determine the factor loadings through a linear regression.

Linear Regression with the Fama-French Three Factor Model: The general factor model can be represented as a linear regression model. For this, we use a dataset of size p containing known returns and factors for up to n assets. This project will perform the regression for $n = 20$ stocks based on $p = 52$ sets of returns and factors.

With the Fama-French Three Factor model, we will solve for the factor loading terms by writing equation 9 in the following form.

$$\mathbf{R} = X\beta + \epsilon \tag{10}$$

Here we have are

$$\mathbf{R} = \begin{pmatrix} r_{11} - R_f & r_{21} - R_f & \dots r_{n1} - R_f \\ r_{12} - R_f & r_{22} - R_f & \dots r_{n2} - R_f \\ \vdots & \vdots & \vdots \\ r_{1p} - R_f & r_{2p} - R_f & \dots r_{np} - R_f \end{pmatrix} \quad X = \begin{pmatrix} 1 & f_{m1} - R_F & SMB_1 & HML_1 \\ 1 & f_{m2} - R_F & SMB_2 & HML_2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & f_{mp} - R_F & SMB_p & HML_p \end{pmatrix}$$

$$\beta = \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \beta_{11} & \beta_{21} & \dots & \beta_{n1} \\ \beta_{12} & \beta_{22} & \dots & \beta_{n2} \\ \beta_{13} & \beta_{23} & \dots & \beta_{n3} \end{pmatrix}$$

In which the matrix $X \in \mathbb{R}^{p\times 4}$ is our design matrix containing all of the regression information about our factors, $R \in \mathbb{R}^{p\times n}$ is the realized returns, $\beta \in \mathbb{R}^{4\times n}$ is the regression coefficients we wish to solve for and. Completing the linear regression (requiring that X is nonsingular) we calculate for β using the following closed form solution

$$\mathbf{b} = (X^T X)^{-1} X^T \mathbf{R} \tag{11}$$

Based on equation 11, we can determine our model error terms by taking

$$\epsilon = \mathbf{R} - X\mathbf{b} \tag{12}$$

Which provides a ready means to determine D in equation 8.

Ellipsoidal Uncertainty Sets: Our factor model gets us so far so to arrive at parameter estimates, but we concede that these estimates may still be 'noisy'. We would like to develop a framework to account for this uncertainty and work this into the robust optimization model presented in 2.2. For this task, we define the following ellipsoidal uncertainty set $\mathcal{U}$ for expected return estimates of n assets.

$$\mu_{true} \in \mathcal{U}(\mu) = \{\mu_{true} \in \mathbb{R}^n : (\mu_{true} - \mu)^T \Theta^{-1} (\mu_{true} - \mu) \leq \varepsilon_2^2\}$$

Note that $\Theta \in \mathbb{R}^{n \times n}$ represents the standard variance derived from the covariance matrix Q and $\varepsilon_2^2 \in \mathbb{R}$ is the uncertainty tolerance of $\mathcal{U}$. These values are calculated as follows

$$\Theta := \frac{diag(Q)}{n} = \frac{1}{n} \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix} \tag{13}$$

$$\varepsilon_2^2 := \chi_n^2(1 - \alpha) \tag{14}$$

In which $\chi_n^2(1 - \alpha)$ is the inverse cumulative distribution function for a chi-squared distribution with n degrees of freedom. We base this percentile estimate at a $1 - \alpha$ confidence level.

When optimizing under a robust framework, we penalize our expected returns by the amount that we expect the returns to vary around their true value. Defined in terms of Θ and ε_2^2, the penalty terms looks like

$$\varepsilon_2 \sqrt{\mathbf{x}^T \Theta \mathbf{x}}$$

And as we will see in 2.2, the application of robust optimization with an ellipsoidal uncertainty set involves taking the difference between the portfolio's expected return and the penalty term shown above.

Re-sampling Process: Exercising re-sampling during the optimization process improves the robustness of the optimizer by introducing small perturbations to the expected returns and covariances.

The re-sampling process is a technique to generate samples of the assets' returns based on some historical distribution. An alternative to robust optimization, optimization based on re-sampling can achieve similar forms of robustness as with the aforementioned uncertainty sets.

Specifically, this process assumes a normal distribution of the asset returns with μ, which is the expected asset excess returns for n assets, and Q, which is the covariance of these returns. At each cycle of the re-sampling process, a vector $\mathbf{r_t}$ of the sampled asset return is drawn from this distribution:

$$\mathbf{r_t} \sim \mathcal{N}(\mu, Q) \quad \forall \text{t} = 1,\ 2,\ ...,\ \text{T and} \quad \mathbf{r_t} \in \mathbb{R}^n$$

Then, geometric mean and covariance of these T sampled returns are computed. The new mean and covariance will likely contain some bias from the initial sample mean and covariance. By iterating this process numerous times, the output of the optimizer would be more stable with respect to the perturbations introduced.

Index Tracking for Diverse Optimization: The metric used to measure how closely a portfolio of k assets tracks the market is with the total correlation of the portfolio, defined as the sum of the correlations that each asset in the market has with the one asset in the portfolio it is best represented by. To attempt to create a portfolio that will most accurately track the market, we seek to maximize the correlation, which is done with the following optimization problem:

$$\begin{aligned}
\max_{z,\mathbf{y}} \quad & \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{i,j} z_{i,j} \\
st \quad & \mathbf{1}^T \mathbf{y} = k \\
& z\mathbf{1} = \mathbf{1} \\
& z_{i,j} \leq y_j, \quad i = 1...n, \quad j = 1...n \\
& z_{i,j} \in \{0, 1\} \\
& y_j \in \{0, 1\}
\end{aligned}$$

with $z \in \mathbb{R}^{n \times n}$ representing whether asset j is representative of asset i in the resulting portfolio and $\mathbf{y} \in \mathbb{R}^n$ representing whether an asset was selected to be in the portfolio.

The first constraint in the problem ensures that only k assets are selected to be held at any time. The value of k is determined prior to running the optimization and given as an input argument.

The second constraint ensures that each asset in the market is represented by only one asset in the portfolio, which will correspond a single 1 in each row of z will all other values being 0. If the asset itself is in the portfolio, then $z_{i,i} = 1$, and $z_{i,j} = 0, i \neq j$ by definition.

The final constraint ensures that an asset can't be representative of another asset without also being in the portfolio, thereby linking the two decision variables of the problem. By combining this constraint with the second constraint, we obtain the relation that every asset is represented by exactly one asset in the portfolio.

This model uses the approach of separating the assets in the market into 'buckets', with one asset from each being chosen to represent the whole bucket. So, the value for each asset that is added to the objective function is the correlation between the asset and its representative. We see that if the number of assets in the portfolio k is allowed to equal the number of assets in the market, each asset would represent itself ($zi, i = 1, i = 1...n$, since $\rho_{i,i} = 1$), and the resulting objective function value would be n, the least upper bound of the problem.

Monte Carlo Simulation for Correlated Asset Prices: We assume that our asset prices follow a Geometric Brownian Motion. From this, we can readily derive a stochastic equation governing the asset prices, starting first from the asset returns.

We have previously seen that $\mathbf{r_t} \sim \mathcal{N}(\mu, Q)$ so that for asset i with price P at time t, we have

$$r_t = \frac{S_t^{(i)} - S_{t-1}^{(i)}}{S_{t-1}^{(i)}} = \mu_i + \xi_t^{(i)} \tag{15}$$

Where ξ_i is a normal random variable which accounts for correlation between your holding assets. We will address shortly how this is determined. Under a continuous pricing model we have the following stochastic differential equation (SDE)

$$dS_t^{(i)} = \mu_i S_t^{(i)} dt + \sigma_i \xi_t^{(i)} \sqrt{dt} \tag{16}$$

Solving this SDE leads to the following equation

$$S_{t+1}^{(i)} = S_t^{(i)} e^{(\mu_i - \frac{1}{2}\sigma_i^2)dt + \sigma_i \xi_t^{(i)} \sqrt{dt}} \tag{17}$$

Now we will address where ξ_i comes from. Given our covariance matrix Q we can readily define our correlation matrix $\rho \in \mathbb{R}^{n \times n}$ such that

$$\rho = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1n} \\ \rho_{21} & 1 & \dots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \dots & 1 \end{pmatrix} \tag{18}$$

where $\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$. With that, we define $L \in \mathbb{R}^{n \times n}$ to be the lower Cholesky factorization of ρ and take $\xi \in \mathbb{R}^{n \times 1}$ to be defined as follows

$$\xi := L\varepsilon \tag{19}$$

Where $\varepsilon \in \mathbb{R}^{n \times 1} \sim \mathcal{N}(\mathbf{0}, \mathit{1})$. It readily follows then that $\xi \sim \mathcal{N}(\mathbf{0}, LL^T)$. If we were considering assets in isolation, then it would suffice to use a standard normal variable in equation 17 instead of ξ_i. However we are considering a portfolio of correlated assets, so it follows that our development of ξ is required.

Given the above framework for determining the price of an asset, a Monte Carlo simulation can be run by repeatably calculating the price of an asset as it follows different paths. Upon each iteration, the asset's price will be simulated from a different random path and given a sufficient number of simulations, parameters can be estimated as done with any other Monte Carlo method.

Finally, we conclude this section by addressing risk measures which will be used in section 2.5 and throughout the analysis of each optimization model.

Value-at-Risk and Conditional Value-at-Risk: Value-at-Risk (VaR) is a downside risk measure which indicates at what value an expected loss would exceed over a specified period and at a certain confidence level. For instance, a twenty-day $\text{VaR}_{95\%}$ of \$17 million would specify that over a 17 day investment horizon, there is a 5% chance that the portfolio will have losses that exceed \$17 million. Where the VaR provides a threshold for which losses might exceed, the Conditional Value-at-Risk (CVar) specifies what the expected loss is over a specified period and at a certain confidence level. Meaning that if a twenty-day $\text{CVaR}_{95\%}$ is \$17 million, then the portfolio is expected to lose \$17 million in one of those twenty days.

To calculate VaR_β and CVar_β, we need access to the following:

- $\mathbf{x}$ which is a portfolio of assets
- $\mathbf{r}$ which is a random realization of asset returns
- $p(\mathbf{r})$ which is the probability density function of the random returns
- Π which is the cumulative distribution function of the random returns
- $f(\mathbf{x}, \mathbf{r})$ which is the loss/gain of the portfolio for a given return

Given the above, we calculate VaR_β and CVaR_β as follows:

$$\text{VaR} = \min\{\gamma : \mathbb{R} | \Pi(\gamma; \mathbf{x}) \geq \beta\} \tag{20}$$

$$\text{CVaR} = \frac{1}{1-\beta} \int_{f(\mathbf{x},\mathbf{r}) \geq \text{VaR}_\beta} f(\mathbf{x}, \mathbf{r}) p(\mathbf{r}) d\mathbf{r} \tag{21}$$

Ideally, CVaR_β should be minimized as this reduces a portfolio's expected loss. This is the task at hand in section 2.5 however minimizing CVaR_β is different from an optimization perspective because it is non-convex. Additionally, probability functions are seldom known analytically and so historical-based CVaR optimization or simulation-based CVaR optimization must be used. This project elects for a simulation-based CVaR optimization, using the methods summarized in the Monte Carlo simulation section as a guideline for generating returns scenarios in which CVaR is approximated empirically. More details for this will come in section 2.5.

2 Portfolio Optimization Models

We use three optimization techniques based on the MVO introduced above. In this section, each model will be presented and briefly explained. The models are as follows:

1. The standard MVO model.
2. Robust MVO with an ellipsoidal uncertainty set
3. Re-sampling MVO
4. Most-diverse MVO
5. Conditional Value-at-Risk (CVaR) Optimization with Monte Carlo Simulations

2.1 Standard MVO Model

The MVO model we take as standard will seek to minimize portfolio risk subjected to the following constraints.

Return Constraint: We demand a *minimum* portfolio mean of **R**, which will be taken as the average yearly return for each asset. This constraint ensures that we are adequately rewarded for the risk we assume.

Budget Constraint: We wish to use all of our available funds for asset allocation, so the sum of the proportions of our constituent assets should sum to unity.

Short-Selling Permitted: We do not prohibit taking on a short position in our portfolio. This means that the proportions of our assets are unbounded.

With these constraints, we present our standard MVO model as

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{x}^T Q \mathbf{x} \\ st \quad & \mu^T \mathbf{x} \geq \mathbf{R} \\ & \mathbf{1}^T \mathbf{x} = 1 \\ & x_i \in \mathbb{R} \quad \forall i = 1 \ldots n \end{aligned}$$

Parameter Estimates: Most of section 1 was dedicated to explaining how parameters are estimated in an MVO formulation. Here we use this the Fama-French Three Factor model to estimate our expected returns μ and covariances Q.

2.2 Robust MVO with an Ellipsoidal Uncertainty Set

The robust MVO aims at improving the stability of the solutions by including an uncertainty set, which in this case belongs exclusively to the constraint variables.

Most of the heavy-lifting for this robust optimization section came in the relevant portion of the Methodology section. All that is left to do here is present the optimization model and summarize our constraints. With that, we are optimizing the following

$$
\begin{aligned}
\min_{\mathbf{x},y} \quad & 50\mathbf{x}^T Q\mathbf{x} - (\mu^T\mathbf{x} - \varepsilon_2 y) \\
st \quad & \mathbf{1}^T\mathbf{x} = 1 \\
& y = \mathbf{x}^T \Theta \mathbf{x} \\
& x_i \in \mathbb{R} \quad \forall i = 1 \dots n \\
& y \geq 0
\end{aligned}
$$

This optimization seeks to minimize the variance while maximizing the robust expected return of the portfolio.

Confidence Level: Recall that $\varepsilon_2^2 = \chi_n^2(1-\alpha)$. For this model, we elect for a 90% confidence level so that $\varepsilon_2^2 = 28.4$.

Auxiliary Variable: The auxiliary variable y was introduced so that $y = \sqrt{\mathbf{x}^T\Theta\mathbf{x}}$. We do this for practical reasons given that our solver of choice, Gurobi, requires the objective function to follow a standard quadratic form. By making this variable definition we introduce a new quadratic constraint which the solver can handle.

Short-Selling Permitted: We permit short-selling, allowing our asset-holding values to take on any real number.

Budget Constraint: It should come as no surprise that we are again looking to fully-utilize our investment budget.

2.3 Re-sampling MVO

The re-sampling MVO is combines the re-sampling process with nominal MVO model, where the optimizer seeks to minimize the variance subjected to the target return constraint. The parameters μ, which is the expected asset excess returns for n assets, and Q, which is the covariance of these returns, are generated with the returns through a normal distribution described in section 1.

Then the geometric mean and covariance of these sampled returns are computed, which in turn are feed into the MVO model described in section 1. This sampling and feeding process is repeated for pre-specified number of times. The output weights for each cycle is recorded, and a final set of output weights are determined through averaging the weights generated through this cyclic sampling process.

2.4 Most-Diverse MVO

The most-diverse MVO combines the nominal MVO problem with the added constraint of only wanting to hold a pre-determined number of assets $k < n$, the number of assets in the market, so as to be able to track the performance of the market as closely as possible without holding every single asset.

In our example, with $n = 20$ assets chosen to represent the market, we decide to hold exactly $k = 12$ different assets during each investment period. After finding the assets that should be included in the final portfolio, their condensed return and covariance matrices are passed into the nominal MVO problem to yield the final portfolio of k assets. A few different things to note are:

1. The maximum possible value when calculating the total correlation of the portfolio is 20, which would occur if the 12 selected assets had perfect correlation with the other 8 assets that aren't selected.
2. Because the correlation optimization is done for every single investment period, the assets that are represented by each asset in the portfolio can be different each time.
3. As a direct consequence of the above result, assets that are held for one investment period may be liquidated for the next investment period, introducing the risk of high transaction costs for the portfolio over its lifespan.

2.5 CVaR Optimization with Monte Carlo Asset Price Simulation

Following the development of section 1, we look to formulate a convex equivalent to the following CVaR optimization

$$\begin{aligned} \min_{\mathbf{x}} \;& \mathrm{CVaR}_{\beta} \\ st \quad & \mathbf{1}^T \mathbf{x} = 1 \\ & x_i \geq 0 \quad \forall i = 1 \dots n \end{aligned}$$

We take a step towards simplifying the objective function by accepting the role that a Monte-Carlo simulation will play in generating unrealized portfolio losses and gains. The general approach will be to simulate many asset prices in the return to extract returns for which the loss of the portfolio can be approximated. We look to take a sufficiently large number of simulated asset prices so as to adequately cover the variation implied by the Geometric Brownian Motion Process.

Under the Geometric Brownian Motion framework for asset prices, we simulate $S = 2000$ prices for all $n = 20$ holding assets. For simplicity we take a single time-step and look to use the weekly return estimates that we get from our factor model to get our asset prices for six-months ahead. From the framework developed in section 1, we take $dt = 26$.

Shown in figure 1 is the simulated prices of our twenty holding assets. Note that these prices are generated at each scenario and are based on the current prices. Using a single time-step, the prices are estimated six months in the future.

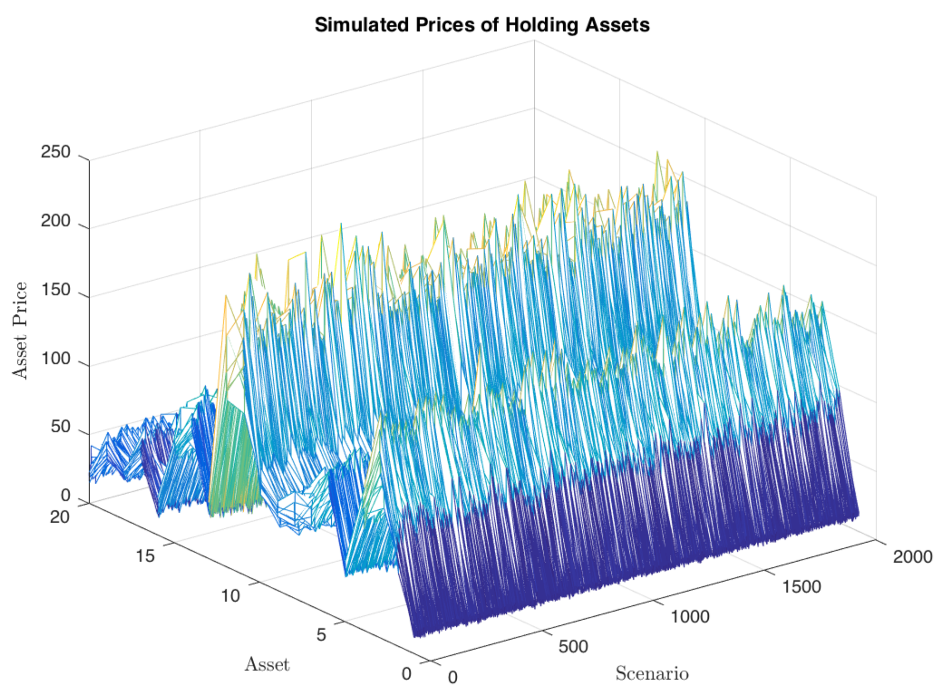

Fig. 1. Monte Carlo Simulation for Holding Asset Prices 6-months up to 01/01/2015 using 1 timestep.

With the simulate asset prices, at each scenario the returns are calculated $\mathbf{r}_s$ based on the current asset price. From this, a loss function can be empirically derived

$$f(\mathbf{x}, \mathbf{r}_s) = -\mathbf{r}_s^T \mathbf{x} \tag{22}$$

By taking $\gamma = \text{VaR}_\beta$, we can now approximate our CVaR_β with the following

$$\text{CVaR}_\beta \approx \gamma + \frac{1}{(1-\beta)S} \sum_{s=1}^{S} \max(-\mathbf{r}_s^T \mathbf{x} - \gamma, 0) \tag{23}$$

Now the term in the summation is not a linear function, but a simple manipulation can lead to the following optimization model for which the objective function converges to CVaR.

$$\begin{aligned} \min_{\mathbf{x},\mathbf{z},\gamma} \quad & \gamma + \frac{1}{0.05S} \sum_{s=1}^{S} z_s \\ st \quad & z_s \geq -\mathbf{r}_s^T \mathbf{x} - \gamma \quad \forall s = 1 \ldots S \\ & z_s \geq 0 \quad \forall s = 1 \ldots S \\ & \mathbf{1}^T \mathbf{x} = 1 \\ & x_i \geq 0 \quad \forall i = 1 \ldots n \end{aligned}$$

Auxiliary Variable and Constraints: The decision-variable $\mathbf{z} \in \mathbb{R}^{S \times 1}$ was introduced to accommodate the maximum function in equation 23 into the linear program. By doing this, we introduce two sets of additional constraints to ensure that the $\mathbf{z}$ is equivalent to the maximum function.

Short-Selling Prohibited: In this to ensure stability in our solution, we do not permit short-selling.

Budget Constraint: Always present.

$\boldsymbol{\beta}$, Time-interval and VaR: For our CVaR optimization, we take a $\beta = 95\%$ confidence level and we consider our simulations over a six month period. We are therefore optimizing a six month $\text{CVaR}_{95\%}$. Also, since we are looking to minimize γ in our objective function, we are elegantly able to extract what our six-month $\text{VaR}_{95\%}$ is as well.

2.6 Transaction Costs

Noticeably absent in all of our optimization models is the cost to transact. The transactional costs varies depending on the investor – large institutional firms pay a reduced price per transaction when compared to a lay investor because institutions have larger trading volumes.

Depending on the motive, transaction costs may be included in the budget allocated for investing. This penalizes high frequency re-balancing of the portfolio as this reduces the potential gain from the portfolio investment strategy. Additional develops to the optimization framework could include constraints limiting the fees incurred from transaction costs

For this project, we shall consider having access to an express account without limit. We will model our transactional cost based on 0.5% of the trading volume at the stocks current price p_i. This is calculated

$$c_i = 0.005 p_i \Delta_i \tag{24}$$

Where $\Delta_i \in \mathbb{R}$ is size of the transaction. Our comparison of the three optimization models will include monitoring the incurred transactional costs on a per-period basis.

3 Results

Table 1 has the list of stocks tracked under each portfolio optimization method. Weekly adjusting closing prices were used from the start of 2012 to the end of 2015. The simulation began in 2013 and ran over the two years, with the portfolios being rebalanced every six months. The calibration of the factor model was done during the start of each investment period and used a years worth of prior weekly adjusted closing prices.

Table 1. 20 Stocks from the S&P500 tracked

S&P500 Sector	**Company Tickers**		
Consumer Discretionary	F	MCD	DIS
Consumer Staples	KO	PEP	WMT
Financials	C	JPM	WFC
Healthcare	PFE	JNJ	
Industrials	CAT		
Energy	MRO	XOM	
Information Technology	AAPL	IBM	
Materials	NEM		
Utilities	ED		
Telecommunications	T	VZ	

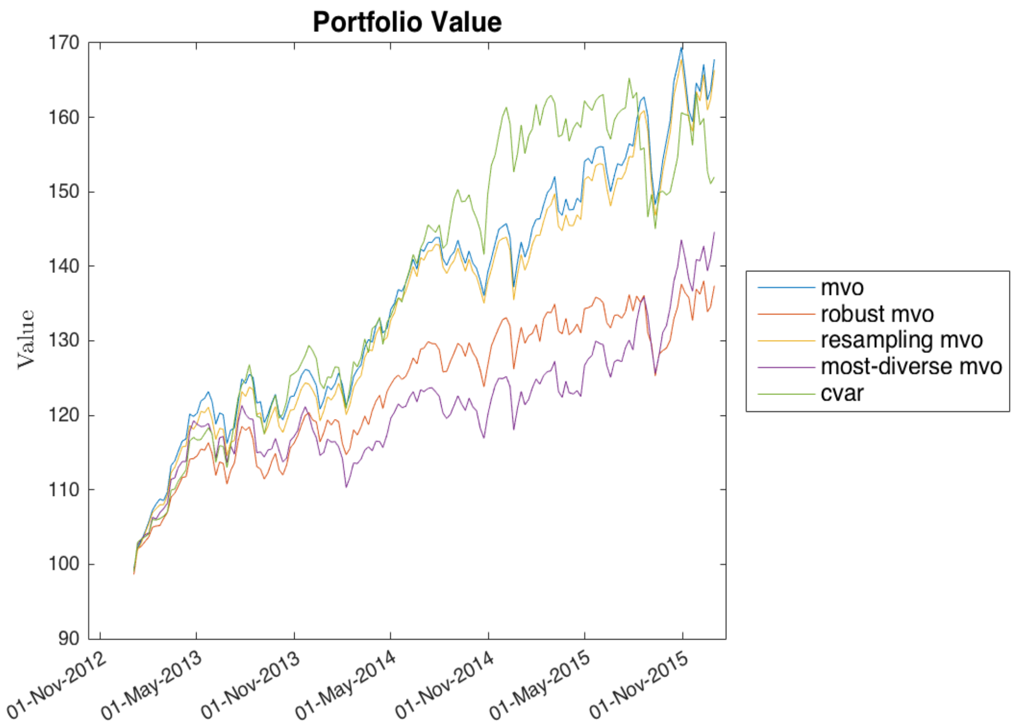

Fig. 2. Portfolio values for each optimization model over the course of the tracking period.

3.1 Portfolio Statistics

Table 2. Weekly Portfolio Value Metrics for each optimization model

	mvo	**robust**	**re-sample**	**diverse**	**cvar**
Maximum Value ($)	169	138	168	145	165
Week # for Maximum Value	148	154	148	157	134
Minimum Value ($)	99.4	98.6	99.2	99.2	99.0
Week # for Minimum Value	1	1	1	1	1
Largest Weekly Gain ($)	5.48	4.22	5.39	4.31	8.05
Week # for Largest Weekly Gain	121	151	121	28	95
Largest Weekly Loss ($)	7.71	5.77	7.79	5.85	9.31
Week # for Largest Weekly Loss	139	102	139	102	138
Longest Winning Streak (lws)	8	18	15	10	10
Week # When lws Ended	114	19	16	17	17
Longest Losing Streak (lls)	4	4	4	4	4
Week # When lls Ended	51	36	51	34	36

A summary of the average realized return of each portfolios at each time time is presented below in Table 3.

Table 3. Average Returns of each portfolios per period

	nominal	**robust**	**re-sample**	**diverse**	**cvar**
Investment Period 1	0.135	0.093	0.126	0.115	0.109
Investment Period 2	0.053	0.045	0.052	0.023	0.089
Investment Period 3	0.071	0.032	0.074	0.019	0.062
Investment Period 4	-0.004	-0.001	-0.008	-0.014	0.052
Investment Period 5	0.062	0.026	0.061	0.035	0.023
Investment Period 6	0.050	0.001	0.051	0.061	-0.022

3.2 Portfolio Weights

The following are the area plots for each of the portfolios, which show the percentage of the portfolio's total wealth allocated to each of the 20 market assets at each of the re-investment periods as the distance between the two lines that contain the region whose colour corresponds to each asset. Where applicable, comments will be made on note-worthy trends in the graph.

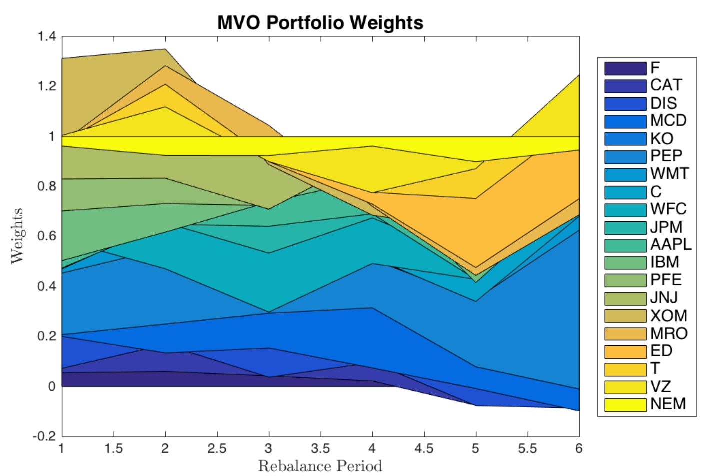

Fig. 3. Area plot for the nominal MVO problem.

The first plot is for the nominal MVO problem, and the first thing to note is the allowance of short-selling in this problem. This is shown by the fact that the base and peak values, which should be 0 and 1 respectively for a portfolio containing only long asset positions, are instead seen to be outside that range. This is a trend that should be common to the area plots of all problems that allow and are optimized by short-selling.

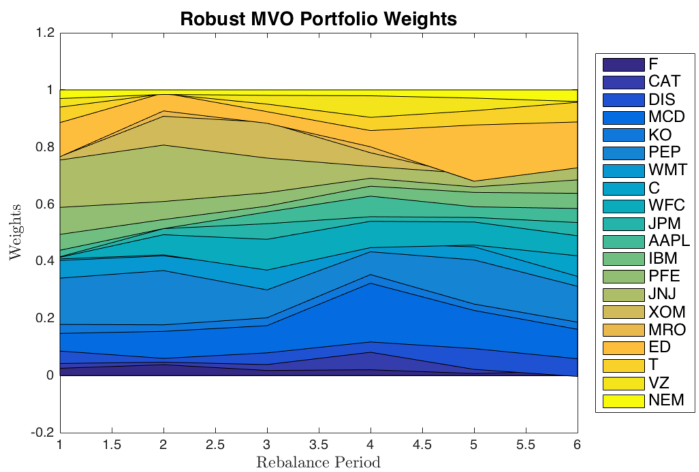

Fig. 4. Area plot for the robust MVO problem.

This next plot is for the robust MVO problem, and the first thing we see that even though the robust MVO problem should allow for short-selling, the optimal portfolio doesn't short-sell any assets for the duration of the portfolio's lifespan. This could only be a property of the model that was used, which would imply that imposing an uncertainty set on the returns of the market assets will favor taking a long position in all assets.

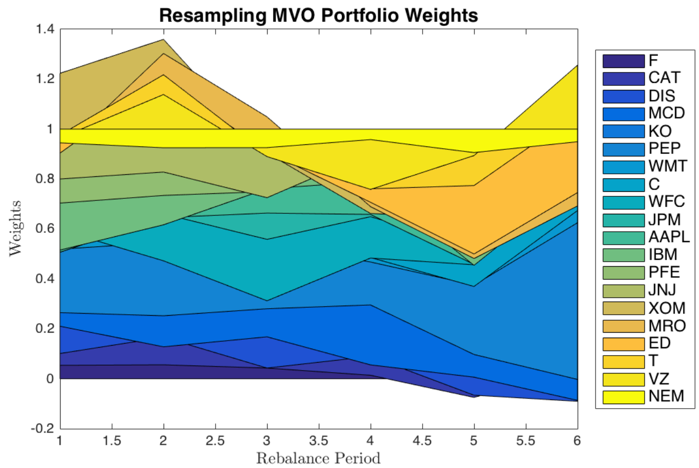

Fig. 5. Area plot for the resampling MVO problem.

In the graph of the portfolio values over the total testing period, we saw that the resampled MVO portfolio very closely followed the value of the nominal MVO portfolio. We see from the area plot that this was, as suspected, because their weights were very similar for every investment period. This shows that the resampling addition to the nominal MVO problem didn't affect the results very much.

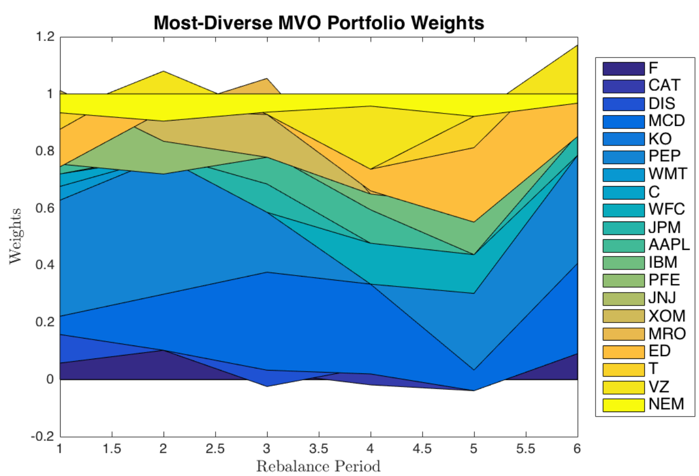

Fig. 6. Area plot for the diverse MVO problem.

The diverse portfolio was created so that it only held the 12 assets that would best track the market for any given investment. This results in the possibility of much larger perecentages of the total wealth being held in a single asset. Since these assets are chosen again after every investment, an asset that held a large portion of the portfolio wealth for one investment period can be liquidated for the next investment period and vice versa.

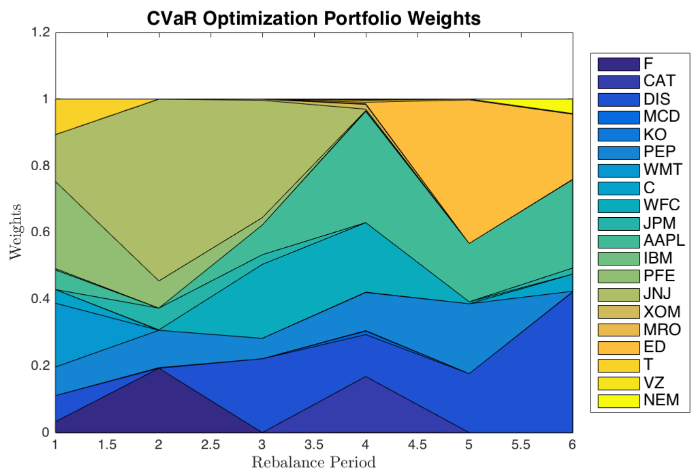

Fig. 7. Area plot for the CVaR Optimization problem.

The most note-worthy fact about the CVaR portfolio's plot of weights is that the portfolio never holds more than half the assets in the market at a time, which was often done in MVO to diversify away the idiosyncratic risk of the portfolio. Instead, the portfolio optimization, which focuses on minimizing the mean shortfall, will produce a portfolio with fewer assets, presumably all of which have lower values of mean shortfall for that investment period.

3.3 Re-sampling Behavioral Analysis

In order to select appropriate parameters for the re-sampling MVO model, the behavioral of the portfolio value based on variables 'T' and 'NoEpisodes' are analyzed by fixing one variable and changing the other. Specifically the variables are defined as such:

- $'T'$ - Number of samples generated and averaged to obtain the aforementioned geometric mean, of which is the input to the nominal MVO optimizer.

– $'NoEpisodes'$ - Number of cycles the sampling process is performed.

To compare the performance of these portfolios with different parameters, a baseline portfolio is derived and the differences in portfolio values are noted. The baseline portfolio is constructed by averaging the portfolio values of those with the largest varying parameter. For example, if $'T'$ is the varying parameter, then the portfolio values with the largest $'T'$'s are arithmetically averaged, and this average is set as the baseline performance for all other portfolios. This would indeed provide an acceptable baseline because the randomness inherent in the procedure will eventually regress to its mean by the law of large numbers.

The following results were obtained fixing one parameter while varying the others based on the set described in the legend.

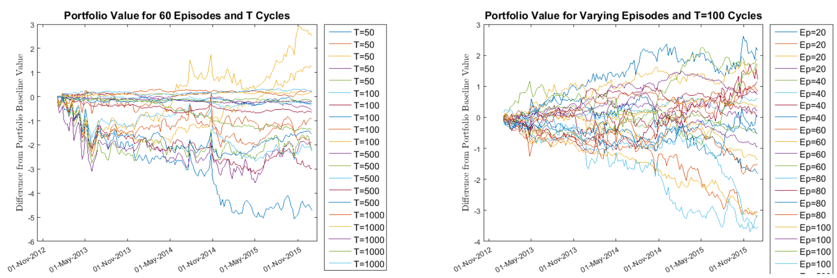

Fig. 8. (left) Resampling with varying T and 60 episodes. Optimized with MVO. (right) Resampling with varying episodes and T = 100. Optimized with MVO.

The left figure shown above depicts the aggregate behavior of the portfolio value with a fixed 60 iterative cycles and a varying number of samples (T samples). By inspection, those with smaller T show a randomized path with high volatility, while those with larger T are much closer to 0 with lower deviation. This is indeed expected as with a smaller T, the geometric mean and covariance of the generated returns are much more likely to inherent a large bias from the true sample mean and covariances. On the other hand, the geometric mean and covariance for larger T are much more likely to inherent a smaller bias from the true sample mean and covariances. To obtaining a more robust portfolio through resampling, the parameter T should neither be too small, as the mean is more likely to deviate from the true mean by too large of an amount, or too large, as the mean is more likely to affix on the true sample mean and defeat the intention of robustness. Hence, a compromising $T = 100$ is selected for the reminder of the experiment.

The right figure shown above depicts the aggregate behavior of the portfolio value with a fixed $T = 100$ and a varying number of iterative episodes

($NoEpisodes$ samples). By inspection, those with smaller $NoEpisodes$ had a more randomized path while larger ($NoEpisodes$ had a more linear path. While this behavior is similar to the above variant, it is worth noting that while in the previous analysis, portfolio values with larger T generally outperformed those with smaller T, portfolio values of larger $NoEpisodes$ does not necessarily imply the same behavior. Approximately the same number of values outperformed the baseline as the number of portfolio values that underperformed. This suggests that the behavior of the portfolio appear much more sensitive to the $NoEpisodes$ than to T.

3.4 Transaction Costs

Shown below are metrics for the transaction costs for each of the 5 portfolios. They give the most important statistics concerning transaction costs incurred for each investment strategy.

Table 4. Transaction Cost Metrics for each optimization model

	mvo	**robust**	**re-sample**	**diverse**	**cvar**
Maximum Incurred Cost ($)	1.70	0.40	1.59	1.19	0.78
Rebalancing Period for Max Incurred Cost	5	3	5	1	2
Minimum Incurred Cost ($)	1.03	0.26	0.90	0.94	0.56
Rebalancing Period for Min Incurred Cost	1	5	1	4	5
Average Transaction Cost ($)	1.37	0.33	1.35	1.05	0.66
Standard Deviation in Cost ($)	0.25	0.06	0.27	0.10	0.09
Total Transaction Costs ($)	6.87	1.67	6.74	5.24	3.32
Total Costs as % of Final Value	4.1	1.2	4.1	3.5	2.2

3.5 Sharpe Ratios

The two graphs below compare the Sharpe ratios of the different portfolios for each investment period.

For the Ex-Ante plot, the investment period refers to the period for which the Sharpe ratio is being predicted. So, the first period refers to the initial prediction made at time t = 0 prior to the first period, and the last period refers to the final prediction made just after the final re-balancing of the portfolios.

For the Ex-Post plot, the investment period refers to the same period, but since the actual portfolio return and variance are being used to calculate the Sharpe ratio, the value is calculated at the end of its corresponding. So, the first value is calculated at the same time as the first portfolio re-balancing, and the last value is calculated after the final liquidation of the portfolios.

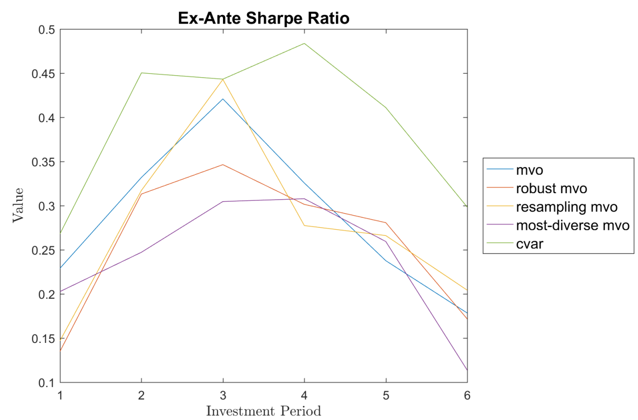

Fig. 9. Ex-Ante Sharpe Ratios for each portfolio before each investment period.

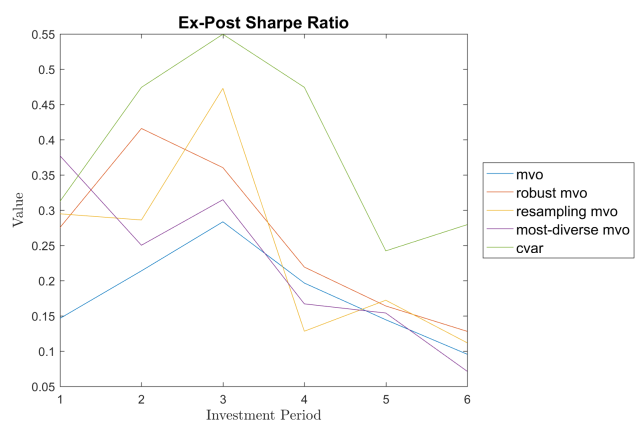

Fig. 10. Ex-Post Sharpe Ratios for each portfolio after each investment period.

We see immediately that the Sharpe ratios of the CVaR-optimized portfolio exceed those of all other portfolios for almost every investment period. As a portfolio that is seeking to minimize the expected shortfall, there's no direct objective of the portfolio, so this result is not immediately obvious. The most reasonable explanation for this would be that the other portfolios, all being created with a variation of the MVO problem, are averse to higher levels of risk that usually come from higher levels of return, so as a result, their Sharpe ratios all stay relatively low.

In addition, the ratios of the different portfolios tend to move together. This sort of behavior is expected, since the performance of the market affects each portfolio's ability to earn excess returns per unit of risk. Therefore we can say that the market, as modeled by our 20 representative assets, showed its strongest growth and stability during the third investment period, when most of the Sharpe ratios (both Ex-Ante and Ex-Post), achieved their maximum values.

Lastly, when comparing the two different Sharpe ratio values for each portfolio individually, there's no conclusive relationship that could suggest a way of using one of the two Sharpe ratios to predict the other. The fact that one is a predictive value while the other is a performance metric means there is no guarantee that one will be indicative of the other. For example, the nominal MVO portfolio generally produced Ex-Post values that were lower than their respective Ex-Ante values, but in general this is not true. The most that can be said is that like the values for different portfolios, the two different Sharpe ratios for a portfolio will generally move together, showing that the Ex-Ante Sharpe ratio for our choice of period length is able to predict movements in the ratio of returns to risk.

3.6 Portfolio $\mathbf{VaR_{95\%}}$ and $\mathbf{CVaR_{95\%}}$

As outlined in section 1 we have an analytical framework to calculate $\mathrm{VaR}_{95\%}$ and $\mathrm{CVaR}_{95\%}$. However as we saw in the CVaR optimization model development, using the analytical framework is cumbersome. Rather, we can approximate $\mathrm{VaR}_{95\%}$ and $\mathrm{CVaR}_{95\%}$ for each portfolio using the results from out Monte Carlo simulation (we do this with the exception of the CVaR optimized portfolio, which is optimized to be exact).

Given each investment period, 2000 scenarios were generated using a Monte Carlo simulation and under each scenario, the expected returns were calculated. With a loss function taken to be $f(\mathbf{x}, \mathbf{r}_s) = -\mathbf{r}_s^T\mathbf{x}$, the VaR and CVaR for each portfolio is approximated based on the following:

$$\mathrm{VaR}_{95\%} \approx P_{95\%}(\{-\mathbf{r}_s^T\mathbf{x} : s = 1 \dots 2000\}) \tag{25}$$

$$\mathrm{CVaR}_{95\%} \approx VaR_{95\%} + \frac{1}{2000(1-0.95)} \sum_{s=1}^{2000} \max(-\mathbf{r}_s^T\mathbf{x} - \mathrm{CVaR}_{95\%}, 0) \tag{26}$$

Where $P_{95\%}(A)$ is the 95^{th} percentile of set A. Note that shown above are approximations, and yet we find that the results from the above equations yield VaR and CVaR values which reasonably accurate. We first compare the approximated VaR and CVaR values and then we present the approximated VaR and CVaR values for each portfolio model over each investment period.

Table 5. Comparison of the approximated $\text{VaR}^a_{95\%}$ and $\text{CVaR}^a_{95\%}$ against the corresponding optimized $\text{VaR}_{95\%}$ and $\text{CVaR}_{95\%}$ for the CVaR optimized portfolio

	$\text{VaR}^a_{95\%}$	$\text{VaR}_{95\%}$	$\text{CVaR}^a_{95\%}$	$\text{CVaR}_{95\%}$
Investment Period 1	0.026	0.033	0.054	0.058
Investment Period 2	-0.048	-0.048	-0.023	-0.023
Investment Period 3	-0.044	-0.036	-0.019	-0.016
Investment Period 4	-0.056	-0.050	-0.026	-0.024
Investment Period 5	-0.021	-0.012	0.005	0.019
Investment Period 6	0.004	0.035	0.040	0.059

Table 6. Predicted portfolio $\text{VaR}_{95\%}$ values for each time period from the Monte Carlo simulated returns

	nominal	**robust**	**re-sample**	**diverse**	**cvar**
Investment Period 1	0.0257	0.0528	0.0424	0.0363	0.0252
Investment Period 2	-0.0030	0.0018	0.0034	0.0092	-0.0461
Investment Period 3	-0.0275	-0.0039	-0.0299	0.0030	-0.0439
Investment Period 4	0.0017	0.0047	0.0101	0.0044	-0.0540
Investment Period 5	0.0202	0.0120	0.0116	0.0197	-0.0197
Investment Period 6	0.0368	0.0485	0.0295	0.0570	0.0214

Table 7. Predicted portfolio $CVaR_{95\%}$ values for each time period from the Monte Carlo simulated returns

	nominal	**robust**	**re-sample**	**diverse**	**cvar**
Investment Period 1	0.0468	0.0743	0.0615	0.0558	0.0489
Investment Period 2	0.0149	0.0284	0.0211	0.0313	-0.0163
Investment Period 3	-0.0061	0.0226	-0.0087	0.0233	0.0238
Investment Period 4	0.0158	0.0251	0.0238	0.0200	-0.0254
Investment Period 5	0.0384	0.0403	0.0309	0.0389	0.0050
Investment Period 6	0.0536	0.0695	0.0472	0.0782	0.0446

4 Next Steps and Improvements

This section will briefly highlight some areas which can be the subject for further study to improve on or expand the given models.

More Pricing Data: This project relied on weekly adjusted closing prices over the span of three years. It was shown that this provided a good measure of the the portfolio's performance. Since several parameters were based on these historical prices, it would be ideal to collect as many sample points as possible in order to try and best capture any patterns, trends and variations. This clearly comes at the trade-off of computational time, but over short time periods such as three years this should be easily managed.

One may also try to increase the available pricing data by increasing the calibration window. Recall that in this project, factor models were calibrated using the data from the previous year. This will achieve the task of increasing sample data, but this may not be appropriate. Sampling over a long period of time without justification introduces the possibility of working in "old" data that may not be relevant to the current market.

Expanding the Factor Model: The factor model used in this report was the Fama-French Three Factor model. This model offers an improved outlook on say the CAPM model since it includes more factors. If one can identify additional factors that are relevant, their inclusion into a regression model can improve the parameter estimates used in MVO formulations.

Reducing Transaction Costs: This was not of a primary concern for this project but an investor is interested in paying as little as possible in transaction fees. This would directly impact how frequently and to what extent the investor re-balance his/her portfolio. If that investor is contend that it is preferable to reduce the transaction costs (perhaps these costs come from your portfolio budget) then it is possible to include a transaction constraint in the MVO to ensure that the total transaction costs per period are capped at some upper threshold.

Resampling Parameters: For the purpose of comparison, the values of T and number of episodes ran to optimize MVO were fixed at 100 and 60 respectively. However, while these figures used commonly in industry, it does not imply that they are optimal parameters. In fact, as shown in Figure 5, the weights generated from resampling MVO appeared very similar to to those from the nominal MVO. This similarity in weights could become problematic, if the resampling MVO was used in the same context as the Robust MVO. Hence, more noise in the weights could be introduced to the model in future to examine the robustness of portfolio.

Most-Diverse MVO Parameters: Similar to the resampling MVO parameters discussion aforementioned, the parameters k = 12 assets was selected arbitrarily. While the selection yielded fruitful results, it remains to be seen how these portfolios would behave under different parameters. While it can be easily proven that the optimization with fewer clusters would not outperform optimization with more clusters, the transaction cost saved through this process could be motivating.

Shorter Investment Periods: Shorter investment periods correspond to increased opportunities to re-balance your portfolios. Clearly this comes at a cost from transaction fees, but if you are confident in your model's ability to adapt to the market, it may be preferable to have the freedom to quickly adjust the allocations of your assets. Pushed to the extreme, this motivates algorithmic trading based trading portfolios which are high-frequency in their re-balancing.

Modeling Changes in Outstanding Shares In this report, we assumed that the shares outstanding for each company was constant. This was a reasonable simplification to our simulation since the number of shares outstanding does not frequently change for a small group of stocks. The only significant actions that would alter the number of shares outstanding would be for instance, a stock split. To better accommodate this freedom, we could have incorporated the historical number of shares outstanding per company, as we did with the pricing data. This would have a provided a truer sense of the market performance of the optimization methods. Yet for comparisons, we expect the results to be similar.