Uncertainty Radius Selection in Distributionally Robust Portfolio Optimization

Alice Liu
yl7797@stern.nyu.edu

Sharon Wang
jw8643@stern.nyu.edu

1 Introduction

Portfolio optimization involves the systematic allocation of assets to maximize expected portfolio returns for a given level of risk or, alternatively, to minimize risk for a given level of expected return. The classical formulation of portfolio optimization is the Markowitz model, which utilizes a mean-variance optimization framework.

The classical Markowitz model relies on precise estimates of expected returns, variances, and covariances of asset returns. However, these parameters are often estimated from historical data and are subject to estimation error, which can lead to optimized portfolios that are optimal in-sample but suboptimal out-of-sample. Robust portfolio optimization seeks to mitigate the risk that an asset's true characteristics deviate from their estimates by incorporating uncertainties in asset returns, covariances, and other parameters. It considers the worst-case scenario within a specified uncertainty set of all possibilities a market can give. Rather than optimizing for a point estimate of expected returns and covariances, robust optimization takes into account the entire range of possible values that these parameters can take.

Several published work, such as Blanchet et al. and Olivares-Nadal and DeMiguel, have introduced distributional robustness into the classical Markowitz model. The robustification of Markowitz model intends to address the discrepancies between the empirical distribution calculated based on historical data and the real distribution. We transform the robust optimization problem to non-robust minimization problem in terms of the empirical probability distribution adding regularization term.

Blanchet et al. measures the discrepancy between the empirical and true distribution by incorporating the size of ambiguity set δ into the objective function and leveraging the minimal target return threshold α to define the feasible set that includes the optimal solution.

Our results for Blanchet et al. approach on daily stock data is as strong as the authors originally achieved with monthly stock data. Blanchet et al. outperforms classical Markowitz model. It tends to have much higher return and risk than equally weighted portfolio, but similar risk adjusted return.

2 Problem Definition and Algorithm

2.1 Task

Our goal is to find an optimal portfolio weighting vector within a feasible set determined by δ and α to minimize the risk of the worst possible environment within an uncertainty set of empirical distribution. To achieve robust portfolio optimization, we compute δ, the size of the ambiguity set measuring the discrepancy between the empirical distribution and true distribution, and α, the minimal target return threshold. We first solve the traditional Markowitz mean-variance problem based on empirical data for initial weighting vector ϕ Then, we utilize simulated return matrix R generated by bootstrapping techniques and the initial weighting vector ϕ to calculate δ and α. Finally, we solve the robust optimization problem for an optimal portfolio weighting vector.

2.2 Algorithm

1. Choose target return rate $\rho = 0.07$ for daily horizon, $\rho = 0.0005$ for hourly horizon, $\rho = 0.00001$ for minute horizon
2. Collect return data $\{R_i\}_{i=1}^{n}$ for n periods of time and each $R_i \in R^d$
3. Calculate sample mean $\mu_n = E_{P_n}(R)$ and the second moment matrix $\sum_n = E_{P_n}(RR^T)$ to approximate μ^* and $\sum^*$. $\sum_n$ must be positive semidefinite.
4. Solve traditional Markowitz portfolio optimization model for ϕ_n^*

$$\min_{1^T\phi=1} \left(\phi^T E_{P_n}(RR^T)\phi\right)$$
$$\text{subject to } \phi^T E_{P_n}(R) = \rho$$

5. Compute δ
    - $g(x) = x + 2(xx^T\phi_n^* - (\phi_n^*)^T xx^T\phi_n^*)/\lambda_1$
      where $\lambda_1$is the Lagrange multiplier of the constraint $\phi^T E_{P_n}(R) = \rho$
    - $\Upsilon_g \sim Var_{P^*}(n^{-1/2}\sum_{k=1}^{n} g(R_k))$
      where R_k are bootstrap samples of the return matrix, n is sample size, P^* is empirical distribution
    - $L_0 = \frac{\|Z\|_2^2}{4(1-\mu_n^T\sum_n{}^{-1}\mu_n)}$ where $Z \sim N(0, \Upsilon_g)$
    - $\delta_\text{n} = \frac{(1-\delta_0) \text{ quantile of } L_0}{n}$
      where $(1 - \delta_0)$ is a predefined confidence interval, typically 95%

6. Compute α
    - $\Upsilon_{\phi_n^*} \sim Var_{P^*}(n^{-\frac{1}{2}}\sum_{k=1}^{n}(\phi_n^*)^T R_k / \|\phi_n^*\|_2)$
      where R_k are bootstrap samples of the return matrix, n is sample size
    - $\left\{\frac{(\phi_n^*)^T[\mathbb{E}_{\mathbb{P}_n}(R) - \mathbb{E}_{\mathbb{P}^*}(R)]}{\|\phi_n^*\|_2}\right\} \Rightarrow n^{-\frac{1}{2}}N(0, \Upsilon_{\phi^*})$
      where $\Upsilon_{\phi_n^*} \to \Upsilon_{\phi^*}$ when $n \to \infty$
    - Find v_0 such that the following inequality holds with confidence level $(1 - \delta_0)$
      $\left\{\frac{(\phi_n^*)^T[\mathbb{E}_{\mathbb{P}_n}(R) - \mathbb{E}_{\mathbb{P}^*}(R)]}{\|\phi_n^*\|_2}\right\} \geq \sqrt{\delta}(1 - v_0')$
    - Compute v_0'' such that $(\phi_n^*)^T\mathbb{E}_{\mathbb{P}_n}(R) - \sqrt{\delta}\|\phi_n^*\|_2 = \rho - \sqrt{\delta}\|\phi_n^*\|_2 v_0''$
    - $v_0 = \max(v_0', v_0'')$

- $\alpha = \rho - \sqrt{\delta}\|\phi_n^*\|_2 v_0$

7. Robust optimization

$$\min_{\phi \in R^d} \phi^T Var_{P_n}(R)\phi + \sqrt{\delta}\|\phi\|_2$$

$$\text{subject to } \phi^T 1 = 1, \phi \geq 0,\ \mathbb{E}_{\mathbb{P}_n}(\phi^T R) \geq \ \alpha + \sqrt{\delta}\|\phi\|_2$$

3 Experimental Evaluation

3.1 Data

Our experiment takes place in three different time frame settings: one is on a daily basis and the other is on an hourly basis. The dataset utilized for daily horizon comprises the daily adjusted close price of 100 largest stocks in S&P 500 covering the period from Jan 1st, 2021 to December 31st, 2023 and was sourced from Yahoo Finance. We calculated daily returns based on adjusted close price. We set the training window size to one year with the initial training set extending from Jan 1st, 2021 to December 31st, 2022. The test data is the subsequent trading day. We then incrementally shift the training data by one day at a time until the test data reaches the final day in the dataset.

The dataset utilized for hourly horizon is QuantQuote's minute market data. It contains high-frequency minute-level data on NASDAQ, NYSE, and AMEX stocks over a three-month period, specifically spanning December 2020, January 2021, and February 2021. We have price and volume data for each of the 627 stocks included in the dataset. For all trading days, data from 9:30am to 4pm is available, and before and after hour quotes are also provided. We limit our data sample to the hours that the market is open. The dataset includes Date, Time, Open, High, Low, Close, Volume, Split Factor, Earnings, and Dividends, although only Date, Time, and Close are utilized in the experiment. We sampled only 100 stocks with the highest average return in the first 50 data samples for our experiment, similar to the portfolio size experimented in Blanchet et al. This selection makes it more likely to achieve a positive semidefinite covariance matrix of our portfolio returns. We replace NaN values in the stock price with the last non-null value that appeared before it and assume that missing values can be reasonably estimated by their immediate predecessor. We also dropped the stocks that have equal to or more than 6 concessive NaN values of the half-hour return data to filter out the stocks that are no longer publicly traded in the stock markets.

The dataset used for minute horizon is QuantQuote's minute market data of December 2020. We limit our data sample to the open market, which is 9:30am to 4pm. We selected the 100 stocks with the highest average return on the first trading day of the dataset. We used 20% of the dataset as training data and rolled the training data to test the weight vector on the subsequent return data point.

3.2 Methodology

Our experiment aims to evaluate how robust portfolio optimization model, such as Blanchet et al., performs compared with other non-robust models. We hypothesize that the distributionally robust mean variance (DRMV) portfolio optimization model outperforms other non-robust models because it handles the optimization over a set of possibilities the market may give rather than just handle the optimization under the most possible market condition. Additionally, when we use Wasserstein Distance to model the uncertainty in the distribution of stock returns, we naturally generate a regularization term, which effectively addresses the issue of overfitting.

The detailed algorithm is specified in Section 2.2. We utilized CVXPY package to solve the robust optimization problem with size of the ambiguity set δ and minimal target return threshold α. Then, we calculated the portfolio return by weighting the individual stock return with the optimal weighting vector. The distributionally robust mean variance (DRMV) model is implemented in a rolling time horizon to account for market dynamics. In the first iteration, we train δ and α and compute the robust optimal weight with the data in the first training set. Then we apply the optimal weighting vector to individual stock's return on the one data point after the training set, getting the portfolio return. In the second iteration, the training set moves one-step ahead to the future and the size of the training set does not change. We train δ and α based on this new training set and got an optimal weighting vector. We then apply this weighting vector to the next data point subsequent to this new training set. We repeated this process until reaching the last data point in the testing set.

For evaluation purposes, we also trained classical non-robust Markowitz model and equally weighted model on the same dataset with rolling horizon.

3.3 Comparison Model

3.3.1 Classical Markowitz model

$$\begin{aligned}&\text{Minimize } \phi^T \Sigma \phi - \mu_n^T \phi \\ &\text{subject to } 1^T \phi = 1,\ \phi > 0\end{aligned}$$

3.3.2 Olivares-Nadal–DeMiguel Model

1. Use 10-fold cross validation to divide the empirical returns into 10 intervals. For each of the ten intervals, we remove an interval and use the remaining nine intervals to estimate parameters and obtain the corresponding portfolio. Then, we evaluate the portfolio on the one interval that is not used for training.
2. Compute ϕ_M: Solve portfolio optimization model for ϕ_M where τ is the trading volume restriction.

$$\begin{aligned}&\text{Minimize } \phi^T \Sigma \phi \\ &\text{subject to } 1^T \phi = 1,\ \|\phi - \phi_0\| \le \tau \\ &\tau = \{0\%, 0.5\%, 1\%, 2.5\%, 5\%, 10\%\}\end{aligned}$$

3. Compute τ: Apply ϕ_M to out-of-sample returns and compute the variance of the out-of-sample returns for different τ from the set $\{0\%, 0.5\%, 1\%, 2.5\%, 5\%, 10\%\}$. Then choose the τ that corresponds to the portfolio with the smallest variance.
4. Compute κ: Solve for κ by plugging in τ and ϕ_M where κ is the transaction cost parameter.

$$\kappa = \left(\frac{\| \phi_M - \phi_0\|_1}{\tau}\right) - 1$$

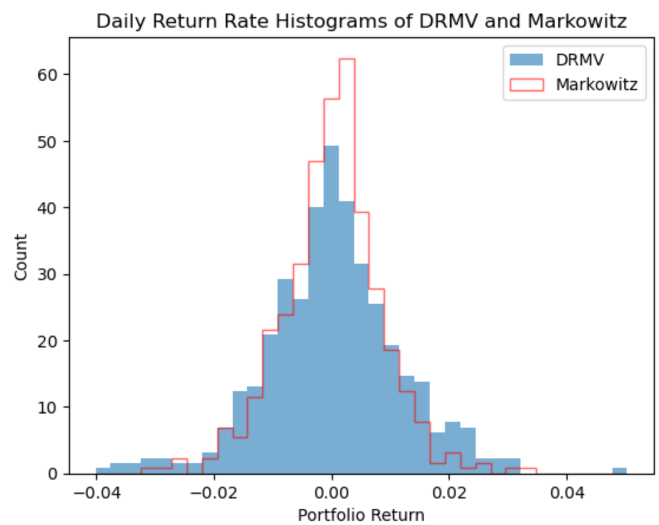
5. Compute ϕ_0: Since we do not have initial optimal portfolio in the first iteration, we use classical Markowitz method to solve for initial optimal portfolio in the first iteration to serve as a starting point.
6. Compute ϕ^*: Solve for optimal portfolio ϕ^* where ϕ_0 is the initial optimal portfolio.

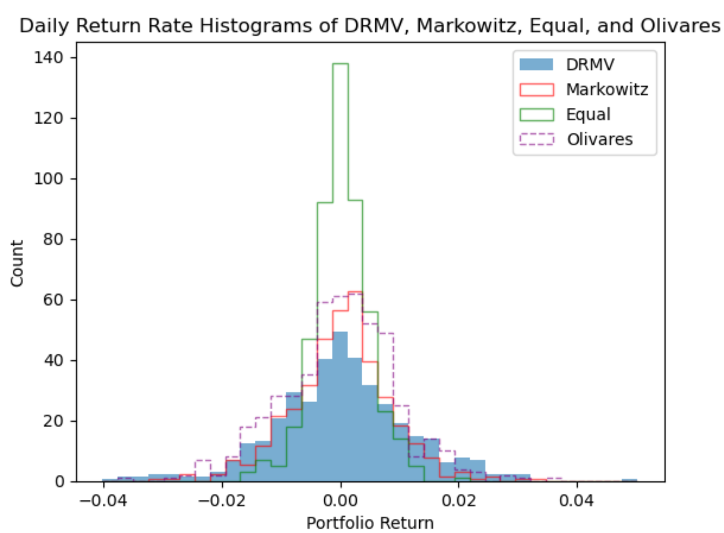
$$\phi^* = \frac{1}{1+k}\phi_M + \frac{k}{1+k}\phi_0$$

4 Result

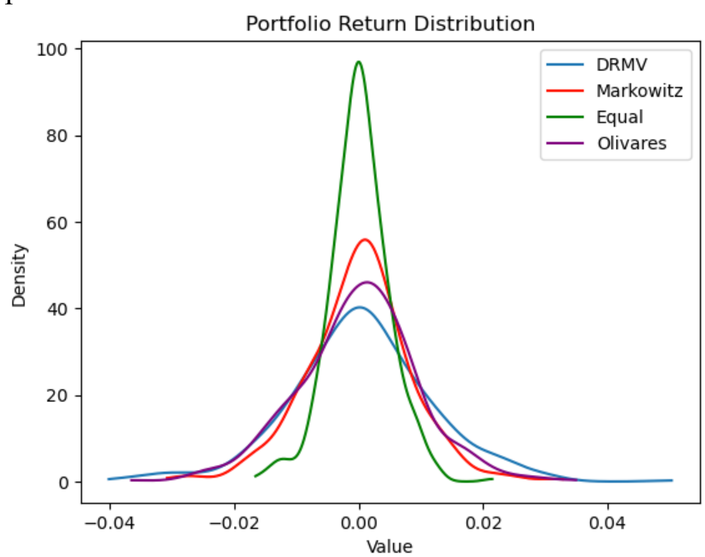
4.1 Daily Result

Comparing with Markowitz, Equal-weighting, and Olivares-Nadal-DeMiguel portfolio, the daily portfolio return of DRMV portfolio tends to have the highest annualized return, highest variance, and second highest Sharpe ratio. Below are histograms of averaged daily return rate of 20 different experiments on the four types of portfolios.





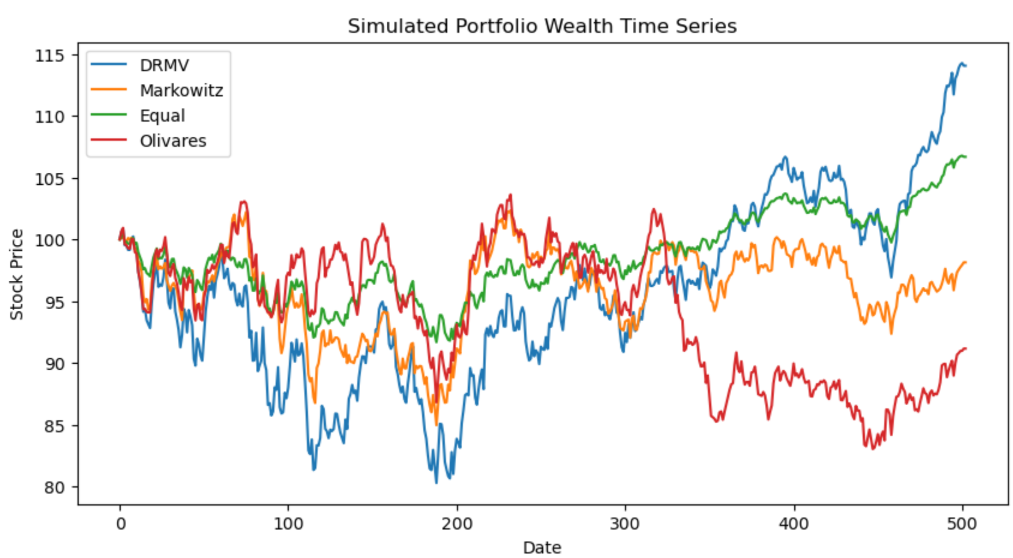
Below is a distribution plot of averaged daily return rate of 20 different experiments on the four types of portfolios.



Stock	Weight
VZ	0.155860
LMT	0.091980
PFE	0.084246
AMZN	0.080052
MCD	0.074951
PG	0.072023
JNJ	0.071296
ABT	0.054779
MRK	0.049051
WMT	0.037594

Top 10 holdings at the end of the holding period are shown above, corresponding to Verizon, Lockheed Martin Corp, Pfizer, Amazon, McDonald's Corp, Procter & Gamble Co, Johnson & Johnson, Abbott Laboratories, Merck & Co Inc, Walmart Inc.

We generate the wealth process for the period January 1st, 2022 to December 31st, 2023 under each of the four models- DMRV, Markowitz, Olivares, and equal-weighting model. Then we repeated the experiments 20 such sets of stocks and obtain the averaged realized wealth process for each model. We initialize the initial portfolio wealth to 100. The portfolio wealth evolves based on the portfolio returns gained from each model. We can see that DMRV generates the lowest portfolio wealth in the first half of the time series and then gradually surpasses the other two models, achieving the highest portfolio wealth in the end. It is obvious from the graph that DRMV generates a portfolio with the highest volatility.



Below is a comparison of annualized return rate and annualized Sharpe ratio for the four portfolios based on the averaged results of 20 experiments. Annualized return rate is presented below as a decimal rather than percentage.

Name	Annual Return	Annual Sharpe Raio
DRMV	0.064995	0.355548
Markowitz	-0.010022	-0.074498
Equal	0.031650	0.418256
Olivares	-0.048487	-0.317536

	DRMV annualized return	DRMV annualized sharpe	Markowitz annualized return	Markowitz annualized sharpe	Equal annualized return	Equal annualized sharpe	Olivares annualized return	Olivares annualize sharpe
0	0.065220	0.356764	-0.008467	-0.063062	0.03165	0.418256	-0.048129	-0.315226
1	0.064994	0.355554	-0.009801	-0.072937	0.03165	0.418256	-0.048206	-0.315819
2	0.065225	0.356774	-0.011274	-0.083767	0.03165	0.418256	-0.048946	-0.320350
3	0.064897	0.355011	-0.007457	-0.055610	0.03165	0.418256	-0.048302	-0.316355
4	0.064844	0.354718	-0.010455	-0.077700	0.03165	0.418256	-0.048612	-0.318294
5	0.064776	0.354356	-0.010543	-0.078350	0.03165	0.418256	-0.047847	-0.313364
6	0.064954	0.355308	-0.010520	-0.078167	0.03165	0.418256	-0.048203	-0.315744
7	0.064774	0.354351	-0.010713	-0.079602	0.03165	0.418256	-0.048892	-0.320113
8	0.065244	0.356891	-0.010594	-0.078731	0.03165	0.418256	-0.049238	-0.322322
9	0.065265	0.357002	-0.010743	-0.079821	0.03165	0.418256	-0.048186	-0.315639
10	0.064599	0.353413	-0.010420	-0.077385	0.03165	0.418256	-0.048929	-0.320320
11	0.064669	0.353808	-0.009716	-0.072160	0.03165	0.418256	-0.048321	-0.316477
12	0.064907	0.355089	-0.009857	-0.073251	0.03165	0.418256	-0.048943	-0.320462
13	0.065026	0.355749	-0.010478	-0.077870	0.03165	0.418256	-0.048220	-0.315765
14	0.065558	0.358612	-0.010311	-0.076628	0.03165	0.418256	-0.048443	-0.317221
15	0.064949	0.355246	-0.010129	-0.075290	0.03165	0.418256	-0.048961	-0.320722
16	0.065001	0.355600	-0.010479	-0.077878	0.03165	0.418256	-0.048381	-0.316812
17	0.065083	0.356044	-0.008312	-0.061763	0.03165	0.418256	-0.047945	-0.314191
18	0.065017	0.355716	-0.009603	-0.071438	0.03165	0.418256	-0.048945	-0.320511
19	0.064889	0.354944	-0.010569	-0.078542	0.03165	0.418256	-0.048085	-0.315018

The annualized return is calculated by $\left(\frac{\text{ending wealth}}{\text{beginning wealth}}\right)^{\frac{trading\ periods\ in\ a\ year}{trading\ periods\ in\ testing\ set}} - 1$. We first calculated the return rate over the entire testing period and then compounded it based on how many such testing period we have in a year. The annualized standard deviation was calculated by the daily standard deviation times $\sqrt{252}$, where 252 represents the number of trading days in a year.

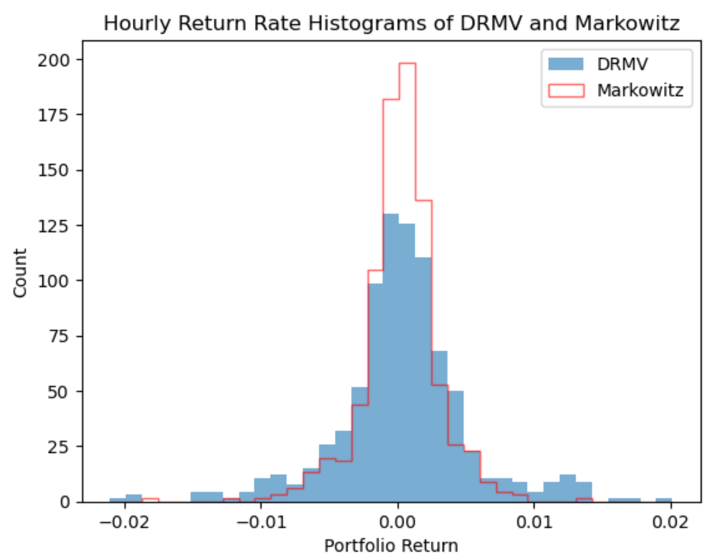
DRMV model is designed to perform well across a range of possible future scenarios, not just the most likely one. They incorporate uncertainties in asset returns, covariances, and other parameters.

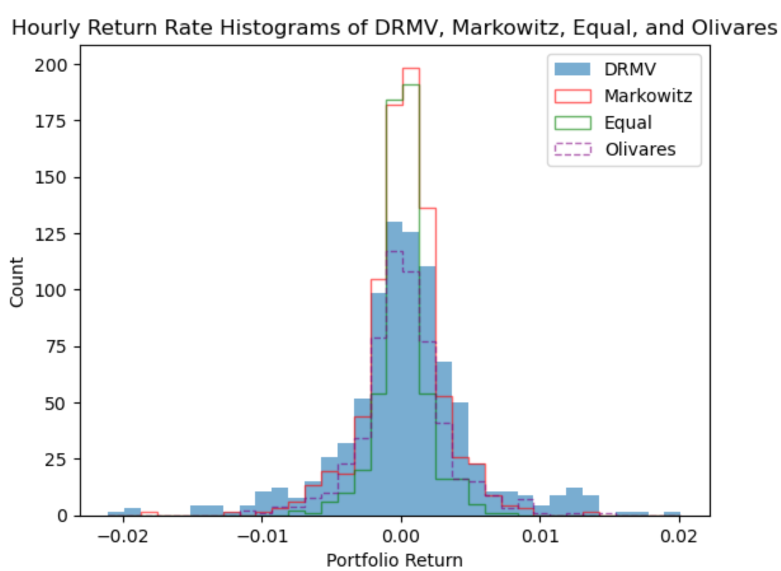
This leads to a more conservative approach that can account for worst-case scenarios, which results in higher variance in the portfolio to hedge against extreme outcomes. Therefore, although it has the highest annualized return in the case of daily horizon, the return is partially offset by the highest variance, resulting in the second highest Sharpe ratio with equal-weighting portfolio coming with the highest Sharpe ratio.

4.2 Hourly Result

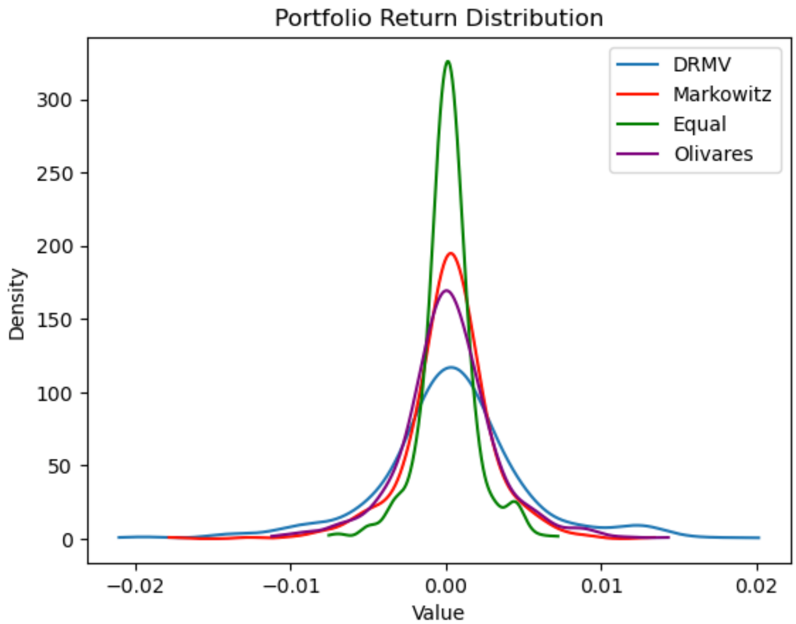
4.2.1 Hourly Horizon Base Case

The hourly horizon yields different results from daily horizon. In hourly horizon, DRMV portfolio has clear advantage over Markowitz, Equal-weighting, and Olivares-Nadal-DeMiguel portfolios. The portfolio return of DRMV model tends to have the highest average return, highest variance, and highest Sharpe ratio compared with the other three models. Below are histograms of averaged hourly return rate of 20 different experiments on the four types of portfolios.

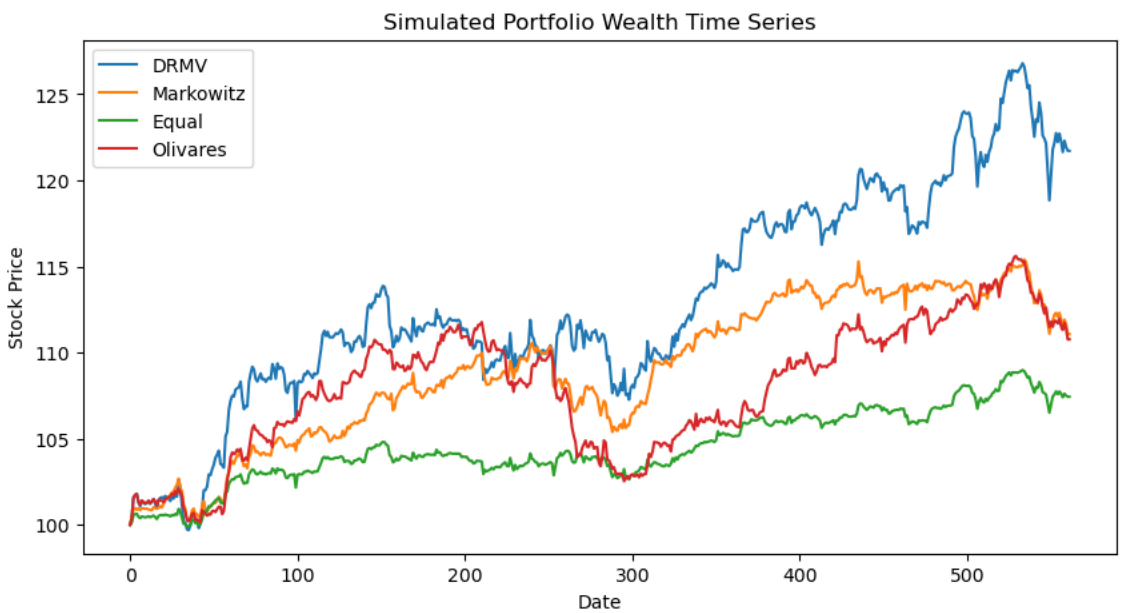




Below is a distribution plot of averaged hourly return rate of 20 different experiments on the four types of portfolios.



Below is a time series plot of averaged simulated wealth process of 20 different experiments on the four types of portfolios.



Below is a comparison of annualized return rate and annualized Sharpe ratio for the four portfolios based on the averaged results of 20 experiments. Annualized return rate is presented below as a decimal rather than percentage.

	Name	Annual Return	Annual Std	Annual Sharpe Raio
0	DRMV	2.364865	3.456871	8.175032
1	Markowitz	0.876389	5.980400	5.241157
2	Equal	0.557718	9.660743	5.387973
3	Olivares	0.860688	5.460592	4.699863

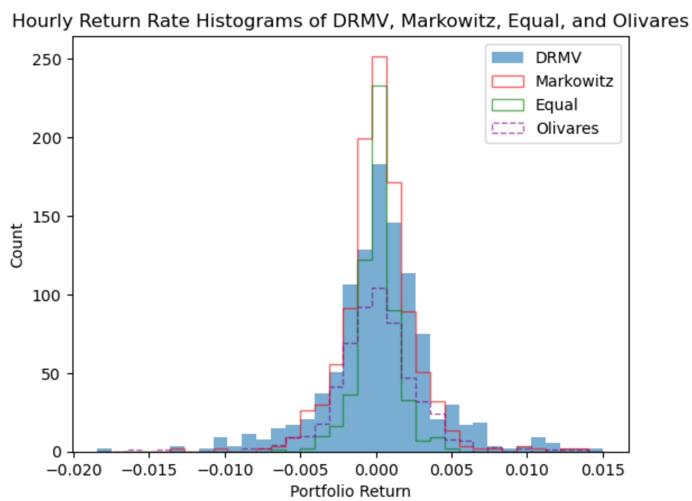
	DRMV annualized return	DRMV annualized sharpe	Markowitz annualized return	Markowitz annualized sharpe	Equal annualized return	Equal annualized sharpe	Olivares annualized return	Olivares annualize sharpe
0	2.366942	8.181514	0.940660	5.641300	0.557718	5.387973	0.764479	4.232349
1	2.363656	8.171194	0.886285	5.278592	0.557718	5.387973	0.791777	4.403072
2	2.366356	8.180006	0.823608	4.906474	0.557718	5.387973	0.827492	4.587927
3	2.366371	8.180383	0.881162	5.272583	0.557718	5.387973	0.918162	5.002922
4	2.363780	8.170922	0.878101	5.253559	0.557718	5.387973	0.837033	4.579827
5	2.364832	8.174980	0.818688	4.892023	0.557718	5.387973	0.796731	4.375397
6	2.364618	8.174245	0.840707	5.014129	0.557718	5.387973	0.823888	4.514037
7	2.365472	8.176926	0.898525	5.377908	0.557718	5.387973	0.855302	4.662068
8	2.362612	8.167545	0.874444	5.232529	0.557718	5.387973	0.843109	4.616580
9	2.364505	8.174387	0.878101	5.253559	0.557718	5.387973	0.884005	4.831806
10	2.365385	8.176491	0.866465	5.192664	0.557718	5.387973	0.840712	4.650782
11	2.364824	8.174053	0.855742	5.119561	0.557718	5.387973	0.893806	4.840342
12	2.362794	8.167754	0.951164	5.749669	0.557718	5.387973	0.852709	4.631585
13	2.364883	8.174612	0.877105	5.247717	0.557718	5.387973	0.904170	4.891264
14	2.364884	8.175426	0.926316	5.501203	0.557718	5.387973	0.853947	4.665184
15	2.367259	8.184025	0.831985	4.951410	0.557718	5.387973	0.910210	4.947384
16	2.364481	8.174389	0.858068	5.138117	0.557718	5.387973	0.856253	4.662992
17	2.364303	8.172847	0.882624	5.275861	0.557718	5.387973	0.885082	4.787482
18	2.362694	8.167235	0.852161	5.096975	0.557718	5.387973	0.974474	5.195961
19	2.366651	8.181703	0.905871	5.427301	0.557718	5.387973	0.900410	4.918310

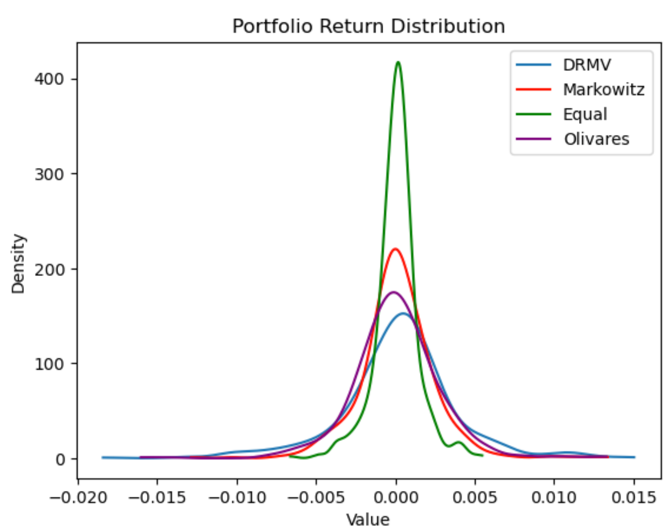
The annualized return was calculated by $\left(\frac{\text{ending wealth}}{\text{beginning wealth}}\right)^{\frac{trading\ periods\ in\ a\ year}{trading\ periods\ in\ testing\ set}} - 1$. The annualized standard deviation was calculated by the hourly standard deviation times $\sqrt{252 * 14}$, where 252 represents trading days and 14 represents trading periods in a day. Annualized Sharpe ratio was calculated by annualized return divided by annualized standard deviation. We can see that DRMV outperforms Markowitz, Equal-weighting, and Olivares-Nadal-DeMiguel portfolio in terms of annualized Sharpe ratio.

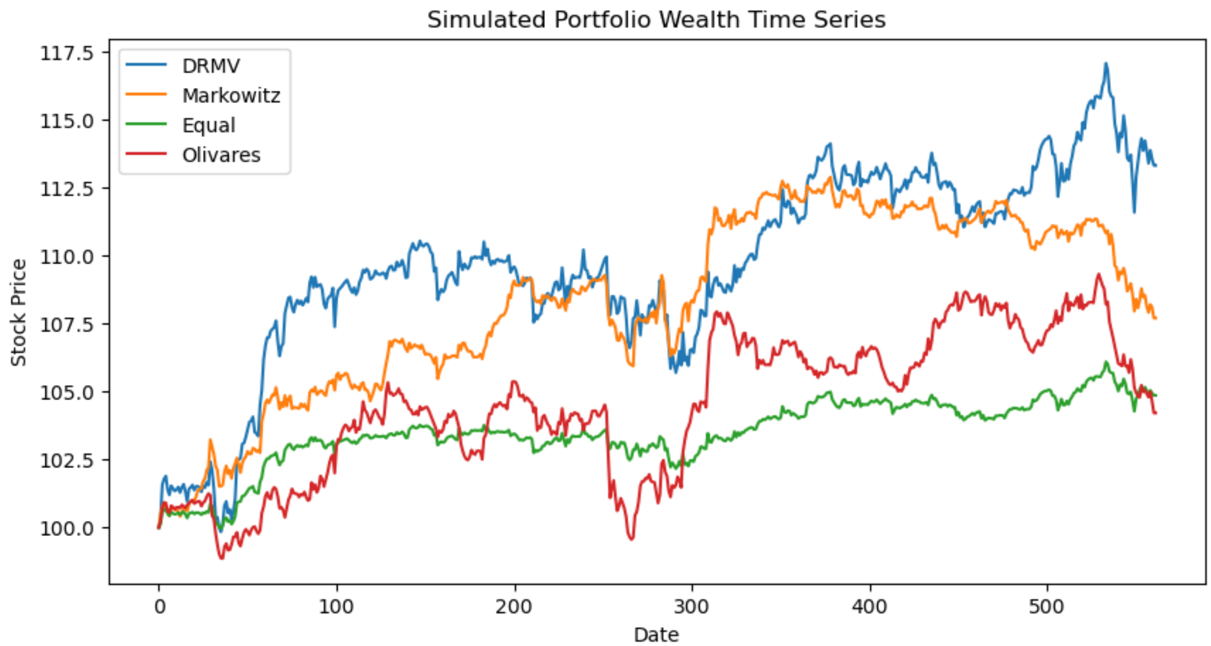
The annualized return is quite high for all four models, much higher than the annualized return Blanchet et al. got in their research paper. This is mainly because we use different trading frequencies on different time interval. Our high return rate is reasonable for the following two reasons. First, since we handle hourly data roughly than monthly data in Blanchet et al.'s research, we are able adjust portfolio allocation much more frequently and react to the market fluctuation immediately, resulting in more optimal portfolio at a given time and higher annualized return. Second, the equal-weighting portfolio return over the three-month period spanning from December 1st, 2020 to February 26th, 2021 is 7.3%, which leading to an annualized return rate of 55.77%. This indicates that the time period we select inherently comes with high return rate. To serve as benchmark, S&P price increased from 3662.45 to 3811.15 over this time period, resulting in period return rate of 4.06% and annualized return rate of 32.27%. Since we selected the 100 stocks with the highest average return of the first 50 trading periods, roughly 4 days, from NASDAQ, NYSE, and AMEX, it is reasonable that our equal weight portfolio can outperform S&P 500. Therefore, from the time series plot, we can see that DRMV portfolio increases by 21.28% over the three-month period, leading to an annualized return of 236.49%.

4.2.1 Hourly Horizon with Stock Reselection

Instead of choosing the 100 stocks with the highest average of the first 50 data points and sticking with these 100 stocks, we select a new set of 100 stocks from a pool of 626 stocks in each iteration.







	Name	Annual Return	Annual Std	Annual Sharpe Ratio
0	DRMV	1.145465	4.562792	5.226518
1	Markowitz	0.545491	6.906746	3.767565
2	Equal	0.336903	12.477646	4.203761
3	Olivares	0.276999	5.939180	1.645148

	DRMV annualized return	DRMV annualized sharpe	Markowitz annualized return	Markowitz annualized sharpe	Equal annualized return	Equal annualized sharpe	Olivares annualized return	Olivares annualize sharpe
0	1.143910	5.219206	0.495613	3.420898	0.336903	4.203761	0.272614	1.616046
1	1.145559	5.227313	0.542506	3.744368	0.336903	4.203761	0.294704	1.744167
2	1.145858	5.228167	0.557191	3.847411	0.336903	4.203761	0.222700	1.339758
3	1.146672	5.231635	0.571707	3.955922	0.336903	4.203761	0.292355	1.735439
4	1.145326	5.226270	0.560437	3.869228	0.336903	4.203761	0.302624	1.790329

All four models have lower annualized Sharpe ratio than the base case, but the ranking of the models does not change. DRMV portfolio still has the highest Sharpe ratio, while equal weighting portfolio is the second highest. The lower Sharpe ratio is due to both the lower annualized return and higher annualized standard deviation. One possible explanation for the lower annualized return is that frequent changes in the portfolio composition might lead to missing out on recovery periods or holding stocks during their underperformance phases. Another possible reason for lower annualized return is that changing stocks frequently can reduce the effectiveness of compounding returns over time. On the other hand, constantly changing stocks introduces volatility, increasing the annualized standard deviation. Therefore, reselecting stocks every time before the trades may not be a good choice comparing with sticking with the initial choice of stock based on our experiments.

Interestingly, in the distribution plot of portfolio return, DRMV is more shifted to the right than all the other three models, indicating the former model outperforms the other three.

5 Discussion

Our hypothesis that DRMV model outperforms non-robust models on out-of-sample data is supported to some extent. In all the scenarios and time horizons, DRMV approach results in higher volatility than non-robust Markowitz, equal weighting, and Olivares portfolio. In hourly horizon, DRMV portfolio has the highest Sharpe ratio, while in daily horizon, DRMV portfolio has the second highest Sharpe ratio, but still outperforming classical non-robust Markowitz portfolio. The one possible reason for the different outcomes is that high-frequency trading has more random fluctuations and noise. DRMV portfolio does not commit to a single probability distribution and are inherently better at handling the uncertainty noise. Furthermore, in high-frequency trading, the accuracy of a model's assumptions regarding distribution parameters are important due to the short timeframe for decision-making. DRMV portfolio does not depend on the exact specification of those parameters can effectively mitigate the risk of misspecification.

Our daily time horizon has overlapping results with Blanchet et al.'s results to some extent. We both find that in terms of the final realized wealth, DRMV and equal-weighting portfolio lead by a substantial margin. However, different from the finding of Blanchet et al., we find the average performances of DRMV and equal weighting are not close. DRMV is much more volatile and generates much higher final realized wealth. One possible explanation for the different results is that equal-weighting portfolio can be seen as an extreme case of the DRMV model when the uncertainty size $\delta = \infty$. In the monthly horizon tested by Blanchet et al., the uncertainty size might be close to infinity because of the long forecasting period. However, in our case, no matter it is daily horizon or hourly horizon, the forecasting period is much shorter than one month, so the uncertainty size would be smaller and would no longer be infinity. Therefore, in monthly horizon, the averaged performance of DRMV and equal weighting are close, but not close in daily and hourly horizon.

6 Conclusions

In conclusion, the goal of this replication study is to apply the distributionally robust mean variance portfolio optimization model to a high-frequency dataset. We find DRMV portfolio has higher volatility than non-robust Markowitz, equal weighting, and Olivares portfolio. DRMV portfolio outperforms the other three models in hourly horizon but is outperformed by equal weighting portfolio in daily horizon.

One area for future work could be comparing this method to additional models, such as Fama-French modeling and Goldfarb–Iyengar Robust Model. It can also be interesting to explore multi-period applications. Another area could be adjusting the constraints of the model to explore how the model adapt to investment with leveraging and how transaction fees and trading costs impact the optimal portfolio allocation.

7 Acknowledgements

We would like to thank Professor Zhengyuan Zhou for his mentorship and guidance on this project.

8 References

[1] Jose Blanchet, Lin Chen, and Xun Yu Zhou. Distributionally robust mean-variance portfolio selection with wasserstein distances. Management Science, 68(9):6382–6410, 2022.

[2] Alba V. Olivares-Nadal and Victor DeMiguel. Technical note—a robust perspective on transaction costs in portfolio optimization. Operations Research, 66(3):733–739, 2018.

[3] Harry Markowitz. Portfolio selection. Journal of Finance, 7(1):77–91, 1952.