Automated Investment Strategy Competition

Pulkit Gaur
Hongyi (Henry) Wu
Jiaqi (Key) Feng

Date: 19 May 2020

Contents

1 Introduction

This report discusses the process to design an automated asset management system using quantitative optimisation techniques. We use adjusted closing price data of 20 stocks and factor returns from Jan–2002 to Dec–2016 to test and validate multiple models and then choose one model which we deem to believe will perform the best under all market conditions.

We look to maximise the Sharpe ratio of a portfolio while minimising the turnover. So, we test factor models and simple empirical historical estimates with various techniques like **mean variance optimisation(robust and vanilla), risk parity optimisation, cVaR optimisation, Maximum Sharpe Ratio optimisation and sub-optmisation techniques**

2 Factor Models

Asset returns are usually strongly related to the market financial factors like capital size and profitability. To define, **factor models** are financial models that incorporate factors (macroeconomic, fundamental and statistical) to determine the market equilibrium and calculate the required rate of return. Such models associate the return of a security to single or multiple risk factors in a linear model.

In this analysis, the returns of eight financial factors are utilized to estimate the expected returns and covariance matrix of twenty assets. This method can capture those assets' systematic return and risk while reducing the noises. The estimated returns and covariance will be used as input parameters for further portfolio optimization.

2.1 OLS

OLS is short for **Ordinary Least Square**. In statistics, it is used to estimate the unknown parameters or coefficients of a linear regression model by minimizing the sum of the squares of the differences between the observed values and those predicted by the linear function.

In our problem, the dependent variables are assets returns and the independent variables are returns of the eight factors. This can be expressed as below :

$$r_{it} = \alpha_i + \sum_{k=1}^{8} \beta_{ik} f_{kt} + \varepsilon_{it}$$

where $\boldsymbol{r}_{it}$ stands for the return for asset $\boldsymbol{i}$ at time $\boldsymbol{t}$, $\boldsymbol{f}_{kt}$ stands for the return for factor $\boldsymbol{k}(k = 1, 2, ..., 8)$ at time $\boldsymbol{t}$. $\boldsymbol{\alpha}_i$ and $\boldsymbol{\varepsilon}_{it}$ are intercept and residuals.

Meanings of factors $\boldsymbol{f}_k$ are displayed below:

Table 1: Factor list of OLS model

$\boldsymbol{f}_1$	$\boldsymbol{f}_2$	$\boldsymbol{f}_3$	$\boldsymbol{f}_4$	$\boldsymbol{f}_5$	$\boldsymbol{f}_6$	$\boldsymbol{f}_7$	$\boldsymbol{f}_8$
Mkt RF	SMB	HML	ST Rev	RMW	CMA	Mom	LT Rev

In vector notation, let $\boldsymbol{r}_i$ be the return vector for asset $\boldsymbol{i}$, $\boldsymbol{X}$ be $[1, f_1, f_2, ..., f_8]$ and $\boldsymbol{B}_i$ be $[\alpha_i, \beta_{i1}, \beta_{i2}, ..., \beta_{i8}]$. For OLS method, we have the following unconstrained minimization problem,

$$\min_{\boldsymbol{B_i}} \quad \left\| \boldsymbol{r}_i - \boldsymbol{X}\boldsymbol{B_i} \right\|_2^2$$

Solving for B_i, we get :

$$\boldsymbol{B}_i^* = (\boldsymbol{X^T}\boldsymbol{X})^{-1}\boldsymbol{X_i^T}\boldsymbol{r_i}$$

2.2 Fama French 3

FF3 model is quite similar to OLS model. The only difference is that only three factors, which are Market, Size, and Value factors, are selected in the regression as shown below :

Table 2: Factor list of FF3 model

$\boldsymbol{f}_1$	$\boldsymbol{f}_2$	$\boldsymbol{f}_3$
Mkt RF	SMB	HML

According to the Fama-French three-factor model, over the long-term, small companies outperform large companies, and value companies beat growth companies. In this model, $\boldsymbol{X}$ will be $[1, f_1, f_2, f_3]$ and $\boldsymbol{B}_i$ be will $[\alpha_i, \beta_{i1}, \beta_{i2}, \beta_{i3}]$. For FF method, we have the same unconstrained minimization problem as OLS,

$$\min_{\boldsymbol{B_i}} \quad \left\| \boldsymbol{r}_i - \boldsymbol{X}\boldsymbol{B_i} \right\|_2^2$$

The process afterwards would be the same as with the OLS model.

2.3 Fama French 5

Fama and French added profitability (stocks with a high operating profitability perform better) and an investment factor (stocks of companies with the high total asset growth have below average returns) to the FF3 model. Both new factors are concrete examples of what are popularly known as **quality factors**. The process to calculate the coefficents is same as OLS and FF3 model.

2.4 LASSO

LASSO stands for **"Least Absolute Shrinkage and Selection Operator"**. The term is often used synonymous with 1-norm regularization. In its most basic form it refers to 1-norm regularized linear regression. The decisive property of LASSO regression is that the one-norm term enforces sparseness of the solution. We used the following formulation of LASSO to get our results.

$$\min_{\boldsymbol{B_i},\boldsymbol{y_i}} \quad \left\|\boldsymbol{r}_i - \boldsymbol{X}\boldsymbol{B_i}\right\|_2^2 + \lambda\left\|\boldsymbol{B_i}\right\|_1$$

As the ℓ_1 norm is not smooth everywhere, we use the penalised form of LASSO and introduce the auxiliary variables $\boldsymbol{y_i}$ to move the ℓ_1 norm from the objective to the constraints.

$$\begin{aligned} \min_{\boldsymbol{B_i},\boldsymbol{y_i}} \quad & \left\|\boldsymbol{r}_i - \boldsymbol{X}\boldsymbol{B_i}\right\|_2^2 + \lambda\boldsymbol{1^T y} \\ \text{s.t.} \quad & \boldsymbol{y_i} \geq \boldsymbol{B_i} \\ & \boldsymbol{y_i} \geq -\boldsymbol{B_i} \end{aligned}$$

The structure of the problem ensures that, at optimality, the auxiliary variable $\boldsymbol{y_i}$ will be equal to the absolute value of the coefficients, i.e., $\boldsymbol{y_i} = |\boldsymbol{B_i}|$. We choose λ by the same way as Project 1 by dynamically searching for the best lambda for each stock.

2.5 BSS

Best Subset Selection is the technique of choosing **K** variables out of available independent variables to minimize the RSS in a linear regression. Let r_i denote i^{th} stock's return, F denote factor's return and $B_i = (b_i^j)_{9\times 1}$ to be the coefficient of the regression against ith stock. Similar to the **OLS**, our objective is to minimize the following:

$$\|\boldsymbol{r_i} - \boldsymbol{F}\boldsymbol{B_i}\|_2^2$$

which is equivalent to minimizing

$$-2\boldsymbol{r_i^T}\boldsymbol{F}\boldsymbol{B_i} + \boldsymbol{B_i^T}\boldsymbol{F^T}\boldsymbol{F}\boldsymbol{B_i}$$

under the constraint of

$$\|\boldsymbol{B_i}\|_0 = \mathbf{K}$$

To write this problem into the standard form, we introduce binary variables $\boldsymbol{Y_i} = (y_i^j)_{9\times 1}$, then we have **Mixed-Integer Quadratic Programming** problem :

$$\min_{\boldsymbol{B_i}} \quad -2\boldsymbol{r_i^T F B_i} + \boldsymbol{B_i^T F^T F B_i}$$
$$\text{s.t.} \quad -1000\boldsymbol{Y_i} \leq \boldsymbol{B_i} \leq 1000\boldsymbol{Y_i}$$
$$-\boldsymbol{1^T Y_i = K}$$
$$\boldsymbol{Y_i} \in \{0,1\}^9$$

2.6 Equal Weighted and Historical Estimates

Though these are not factor based models, we also explore naive techniques like equal weight portfolio and using the sample mean & covariance (without incorporating the factor data). The advantage of equal weighted models is that there is no turnover and we are diversified completely in terms of dollar value. Whereas, while using the sample mean and covariance, we reduce the complexity of the models as there is no training of data required.

3 Portfolio Optimization Techniques

In this section, we discuss the different optimisation techniques we use with different objectives and constraints. The techniques offer a different blend of estimation risk and assumption risks. So, an investor can choose any kind of portfolio as per her active views on return and correlations.

3.1 Mean Variance Optimisation(MVO)

The **Mean Variance Optimisation** problem, which takes μ and Q parameters as inputs, is shown below

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{x}^T \boldsymbol{Q}\, \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{\mu}^T \boldsymbol{x} \geq R \\ & \boldsymbol{1}^T \boldsymbol{x} = 1 \\ & (x_i \geq 0, \quad i = 1,\ ...,\ n) \end{aligned}$$

Our objective is to minimise the portfolio variance while achieving a target return R. The above formulation is for a long only portfolio. If we change the last constraint ,($x_i \geq 0, \quad i = 1,\ ...,\ n$), and allow x_i to be unrestricted we allow short selling.

3.2 Robust MVO

Under the vanilla MVO assumptions, we make our investment decisions based solely on the estimated parameters, ignoring estimation errors and their impact during optimization. To improve stability of our estimates, we create an uncertainty set around the estimated returns and optimize around it.

We can use an ellipsoid uncertainty set,

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

ε_2 is the radius that bounds the standardized distance between μ and μ^{true} , $\Theta \in \mathbb{R}^{n \times n}$ is the measure of uncertainty that serves to standardize our estimated exp. return. The parameter ε_2 is a measure of distance, and it is related to the confidence level around our estimate. If we wish to have a confidence level α, then we can define ε_2 as

$$\varepsilon_2^2 = \chi_n^2(\alpha)$$

where χ_n^2 is the inverse cumulative distribution function of the chi-squared distribution with n degrees of freedom. So, the robust MVO problem after penalising the return constraint is :

$$\begin{array}{ll} \min_{\boldsymbol{x}} & \boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} \\ \text{s.t.} & \boldsymbol{\mu}^T \boldsymbol{x} - \varepsilon_2 \left\| \boldsymbol{\Theta}^{1/2} \boldsymbol{x} \right\|_2 \geq R \\ & \mathbf{1}^T \boldsymbol{x} = 1 \\ & (\boldsymbol{x} \geq 0) \end{array}$$

3.3 Risk Parity

Risk parity is about balancing a portfolio's risk exposures to attain a greater chance of investment success than what is offered by traditional, equity-centric approaches to asset allocation. The main advantages of it are that the portfolios are fully diversified from a risk perspective and does not require noisy estimated expected returns. We use the following formulation for the optimisation :

Consider the following function

$$f(y) = \frac{1}{2} \boldsymbol{y}^T \boldsymbol{Q} \boldsymbol{y} - \kappa \sum_{i=1}^{n} \ln y_i$$

where $\boldsymbol{y} \in \mathbb{R}^n$ is our decision variable and $\kappa > 0$ is a constant. We find the minimum y^* by setting the gradient to zero

$$\nabla f(\boldsymbol{y}) = \boldsymbol{Q} \boldsymbol{y} - \kappa \boldsymbol{y}^{-1} = 0$$

where $\boldsymbol{y}^{-1} = [1/y_1, 1/y_2, \ldots, 1/y_n]^T$ Thus, we have that

$$(\boldsymbol{Q}\boldsymbol{y})_i = \frac{\kappa}{y_i} \Rightarrow y_i (\boldsymbol{Q}\boldsymbol{y})_i = y_j (\boldsymbol{Q}\boldsymbol{y})_j \forall i, j$$

3.4 CVaR Optimization

CVaR,or Expected Shortfall, measures the **expected** loss we are likely to suffer with probability 1-α. We used the following convex reformulation of the cVaR:

$$F_\alpha(\boldsymbol{x},\gamma) = \gamma + \frac{1}{1-\alpha}\int (f(\boldsymbol{x},\boldsymbol{r})-\gamma)^+ p(\boldsymbol{r})d\boldsymbol{r}$$

where, $f(\boldsymbol{x},\boldsymbol{r})$ is the loss function, $p(r)$ is the density and γ is an auxiliary variable.

We use a scenario-based representation, $\hat{r}_s$ for s=1, ...,S, where $\hat{r}_s$ is the realization of scenario s. We can approximate $F_\alpha(\boldsymbol{x},\gamma)$ as

$$\tilde{F}_\alpha(\boldsymbol{x},\gamma) = \gamma + \frac{1}{(1-\alpha)S}\sum_{s=1}^{S}\left(f\left(\boldsymbol{x},\hat{r}_s\right)-\gamma\right)^+$$

And finally, our cVaR optimisation looks like:

$$\begin{array}{lll}
\min_{\boldsymbol{x},\boldsymbol{z},\gamma} & \gamma + \frac{1}{(1-\alpha)S}\sum_{s=1}^{S} z_s & \\
\text{s.t.} & z_s \geq 0, & s=1,\ldots,\mathcal{S} \\
& z_s \geq f\left(\boldsymbol{x},\hat{\boldsymbol{r}}_s\right)-\gamma, & s=1,\ldots,S \\
& \boldsymbol{x}\in\mathcal{X} &
\end{array}$$

where z_s are auxillary variables and $x \in X$ includes our budget, target return, and any other constraints pertaining to our portfolio weights x.

3.5 Maximum Sharpe Ratio

We optimize Sharpe Ratio with an n-year back testing window. We use the historical data to estimate the future return and volatility, and we do not allow short-selling.

Define $\boldsymbol{x}$ as the investment weight, $\boldsymbol{r}$ as stocks' estimated return and $\boldsymbol{Q}$ as covariance. Originally, the problem is,

$$\begin{array}{ll}
\max\limits_{\boldsymbol{x}} & \dfrac{\boldsymbol{r^T x}}{\sqrt{\boldsymbol{x^T Q x}}} \\
\text{s.t.} & \mathbf{1}^T\boldsymbol{x} = 1 \\
& \boldsymbol{x} \geq \mathbf{0}
\end{array}$$

We define $\hat{\boldsymbol{r}}$ as stock's expected return minus the risk-free rate. If there exists $\boldsymbol{x}$ s.t.

$$\boldsymbol{r}^T\boldsymbol{x} - r_f \quad \geq 0$$

We can solve the problem by solving

$$\begin{aligned} \min_{\boldsymbol{y}} \quad & \boldsymbol{y^T Q y} \\ \text{s.t.} \quad & \hat{\boldsymbol{r}}^T \boldsymbol{y} = 1 \\ & \boldsymbol{y} \geq \boldsymbol{0} \end{aligned}$$

and the optimal $\boldsymbol{x}$ for the original problem $\boldsymbol{x^*}$ satisfies

$$\boldsymbol{x^*} = \boldsymbol{y^*}/(\boldsymbol{1^T y^*})$$

where $\boldsymbol{y^*}$ is the solution to the transformed problem.[1]

3.6 Kurtosis and Skewness Optimization

With this technique, we optimize Sharpe ratio with a linear penalty on skewness and kurtosis. The intuition it that lower skewness and kurtosis will improve the portfolio's maximum drawdown. We use n-year backtesting window and use the historical data to estimate the future return and volatility, and we do not allow short-selling. We applied the naive global minimization since there is no proper convex interpretation of the 3^{rd} or 4^{th} order moment.

We define $\boldsymbol{x}$ as the investment weight, $\boldsymbol{r}$ and $\boldsymbol{Q}$ are stocks' estimated return and covariance, $M_3(\boldsymbol{x})$ and $M_4(\boldsymbol{x})$ are the estimated 3^{rd} or 4^{th} central moments of the portfolio built with the weight $\mathbf{x}$, and positive α_3 and α_4 are corresponding penalty coefficients. We want higher Sharpe ratio, lower skewness, and lower kurtosis, and thus, the optimization is,

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & -\frac{\boldsymbol{r^T x}}{\sqrt{\boldsymbol{x^T Q x}}} - \alpha_3 M_3(\boldsymbol{x}) + \alpha_4 M_4(\boldsymbol{x}) \\ \text{s.t.} \quad & \mathbf{1}^T \boldsymbol{x} = 1 \\ & \boldsymbol{x_i} \geq 0 \end{aligned}$$

3.7 Sub Portfolio Optimisation

In sub portfolio optimisation, we try to create a decision based ensemble model. We subset the pool of n stocks into k and $n-k$ baskets based on a certain decision parameter like momentum, value, volatility etc. The stocks within the two subsets have similar characteristics like stocks with high momentum vs stocks with comparatively low momentum. As the properties of the two buckets of stocks are different, we then use a separate portfolio optimisation technique for each of the bucket. At last, we combine both the sub-portfolios by adjusting the weights in the right proportion so that

[1]Proof: http://people.stat.sc.edu/sshen/events/backtesting/reference/maximizing%20the%20sharpe%20ratio.pdf

the sum of the weights equals to 1.

This technique being more dynamic requires fine tuning of various parameters like, classification algorithm for splitting the n stocks, splitting ratio/deciding k(can be naive or SVM), optimisation techniques for each basket, proportion of weight/investment to be given to each basket.

3.7.1 Momentum Based

The momentum factor refers to the tendency of winning stocks to continue performing well in the near term. Momentum is categorized as a "persistence" factor i.e., it tends to benefit from continued trends in markets.

The academic research response is to focus on so-called, "12-2 momentum," which measures the total return to a stock over the past twelve months, but ignores the previous month. (e.g., Ken French data). The reason why we skip the most recent month relates to the short-term reversal effect associated with momentum. Using this definition, we split our **20 stocks into two equal sized baskets of 10 stocks each**. Then, we apply all the available optimisation techniques and then reallocate weights based on different proportions.

3.7.2 Realised Volatility Based

We also split the stocks into highly risky and low risk stocks. So, that we can use the optimal technique for each of the stock bucket. We use 24 month window to calculate the realised volatility/standard deviation for each stock and then split the stocks into two equal size buckets.

4 Methodology

We combined all the factor models & historical estimates with the relevant portfolio optimization techniques and we seek to find the best strategy by maximising the Sharpe ratio while simultaneously trying to achieve a reasonable turnover level.

For each of the investment period, our first step is to estimate the mean returns $\boldsymbol{\mu}$ / $\boldsymbol{r}$ and covariance of the stocks' returns $\boldsymbol{Q}$ with 6 models: **OLS, FF3, FF5, LASSO, BSS,** and **historical estimates**. After estimation, we optimize the portfolio with 7 optimisation techniques described in the previous section namely : MVO, Robust MVO (both MVOs' returns are targeted to the mean of factors' returns), Risk Parity, Maximum Sharpe, Maximum Sharpe with penalty on skewness and kurtosis, cVaR, risk parity and sub-portfolio optimization.

With the different combinations of estimaton and optimisation techniques(if we take rebalancing frequency to be 6 months), we do the backtesting and get relevant metrics including return, volatility, CAGR, Sharpe, skewness, maximum draw down, and average turnover. We focus on Sharpe ratio, maximum drawdown and average turnover while choosing among the portfolios, and keep the 3 best strategies according to our metrics.

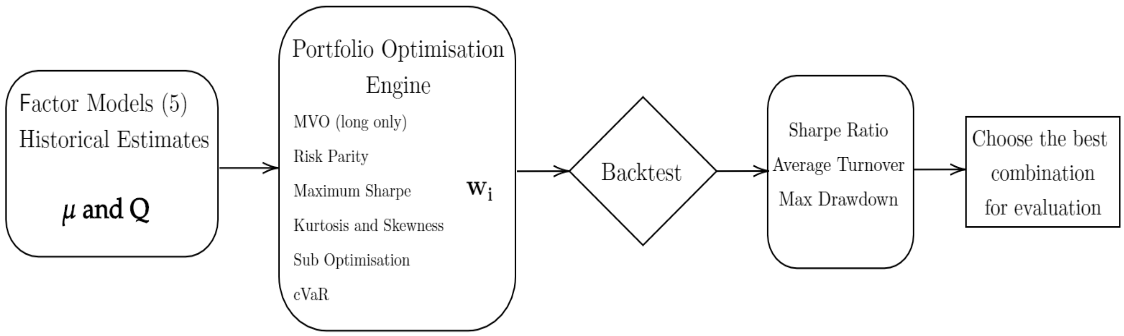


Figure 1: Process of Strategy Selection

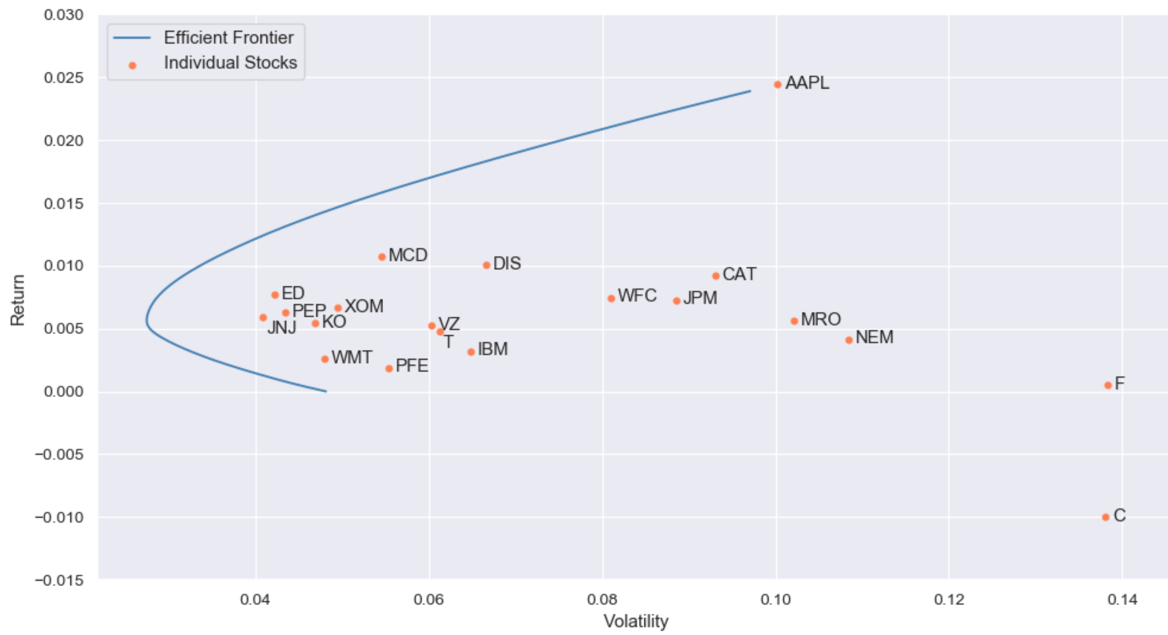
5 Analysis: Training, Validation and Testing

5.1 A glimpse at the data

In our project, we used the monthly price data of 20 stocks from Dec 31, 2001 till Dec 31, 2016, and we have monthly return data for 8 factors as well as the risk free rate from Jan 31, 2002 to Dec 31, 2016.

In figure 2, we have plotted the return-risk profile of individual stocks through the entire period. **APPL** and **C** are obvious outliers, and the distances between stocks and the efficient frontier are quite different, which means stocks have significantly different risk efficiency. Thus, we divided stocks into two subsets in some of our optimizations.

Figure 2: Efficient Frontier



5.2 Observations and Results

We use 5 years of data to callibrate our models at every rebalancing period and then validate the results on the subsequent available years.

a) **Factor Models, Equal Weighted and Historical Estimates**

MVO Long only:

For the factor models, we utilized MVO, Risk Parity, CVaR and Robust MVO for portfolio weights optimization. Rebalancing frequency was tuned to get better outcome. Four statistics were applied to measure the model performance, which were **Sharpe Ratio, Skewness, Maximum Drawdown, Average Turnover**. One thing worth mentioning is that for MVO, target return was also tuned, but in order to keep the conciseness, only two of them will be displayed here.

Here, we just used the geometric mean of the market excess return as the target return (same as project 1).

Table 3: MVO metrics with 6 months rebalancing frequency (Long only)

	OLS	FF	LASSO	BSS	FF5	Historical	EW
Sharpe Ratio:	0.1847	0.1785	0.1747	0.1848	0.1772	0.1583	0.1747
Skewness:	-0.7224	-0.7823	-0.6903	-0.5522	-0.7129	-0.6764	-0.4919
Max Drawdown:	0.2743	0.2869	0.2987	0.2617	0.283	0.2606	0.4189
Mean Turnover:	0.3127	0.293	0.256	0.3145	0.3159	0.3485	0

Table 4: MVO metrics with 12 months rebalancing frequency (Long only)

	OLS	FF	LASSO	BSS	FF5	Historical	EW
Sharpe Ratio:	0.1943	0.1903	0.1736	0.18	0.1896	0.1616	0.1748
Skewness:	-0.5944	-0.6539	-0.6681	-0.5601	-0.6157	-0.4768	-0.4176
Max Drawdown:	0.2607	0.2673	0.2869	0.2616	0.2685	0.2584	0.4187
Mean Turnover:	0.4188	0.3755	0.3372	0.4237	0.3966	0.4554	0

MVO allowed Short

Table 5: MVO metrics with 6 months rebalancing frequency (Short allowed)

	OLS	FF	LASSO	BSS	FF5	Historical	EW
Sharpe Ratio:	0.2042	0.2081	0.1767	0.1767	0.2032	0.1143	0.1747
Skewness:	-0.5316	-0.4774	-0.5477	-0.4228	-0.4327	-0.5801	-0.4919
Max Drawdown:	0.2086	0.2042	0.2584	0.2099	0.2075	0.3152	0.4189
Mean Turnover:	0.4258	0.3448	0.2992	0.3983	0.3949	0.8864	0

Table 6: MVO metrics with 12 months rebalancing frequency (Short allowed)

	OLS	FF	LASSO	BSS	FF5	Historical	EW
Sharpe Ratio:	0.1974	0.2079	0.1754	0.1697	0.2007	0.0796	0.1748
Skewness:	-0.5557	-0.5046	-0.5375	-0.4856	-0.4903	-0.6161	-0.4176
Max Drawdown:	0.2172	0.2112	0.2467	0.2171	0.2197	0.3491	0.4187
Mean Turnover:	0.5771	0.4682	0.4147	0.5764	0.5062	1.2081	0

According to the results above, it is clear that for factor models, MVO allowing short would have better performance. And for short-allowed method, 6-month rebalancing frequency would outperform. Thus, for the sake of simplicity of the result display below, we will only show the metrics table for short-allowed method with 6-month rebalancing.

Next, we kept short permission and the 6-month rebalancing, then tuned the target return. In the MVO below, instead of using the market excess return, we take the average return of all the interested assets returns, and then get the same geometric mean. Some constants like 0.01 or 0.005 was added to that number. For this particular example, we chose 0.01.

Table 7: MVO metrics with 6 months rebalancing frequency (Short allowed and new target return)

	OLS	FF	LASSO	BSS	FF5	Historical	EW
sharpRatio:	0.2333	0.2418	0.2071	0.1638	0.2334	0.1178	0.1747
skewness:	-0.4936	-0.473	-0.5235	-0.7035	-0.4584	-0.3722	-0.4919
maxDrawdown:	0.1741	0.1588	0.2536	0.3056	0.174	0.3352	0.4189
meanTurnover:	0.8565	0.7558	0.7502	1.0827	0.819	1.4276	0

Risk Parity:

Table 8: RP metrics with 6 months rebalancing frequency (Short allowed)

	OLS	FF	LASSO	BSS	FF5	Historical	EW
Sharpe Ratio:	0.1857	0.1861	0.1841	0.1848	0.1853	0.1872	0.1747
Skewness:	-0.7739	-0.7902	-0.7731	-0.7609	-0.7756	-0.7542	-0.4919
Max Drawdown:	0.3516	0.3503	0.356	0.3477	0.3512	0.339	0.4189
Mean Turnover:	0.1182	0.1168	0.1173	0.1188	0.1185	0.1233	0

CVaR:

Table 9: CVaR metrics with 6 months rebalancing frequency (Short allowed)

	OLS	FF	LASSO	BSS	FF5	Historical	EW
Sharpe Ratio:	0.1318	0.1318	0.1318	0.1318	0.1318	0.1318	0.1747
Skewness:	-0.0015	-0.0015	-0.0015	-0.0015	-0.0015	-0.0015	-0.4919
Max Drawdown:	0.4275	0.4275	0.4275	0.4275	0.4275	0.4275	0.4189
Mean Turnover:	2.9898	2.9898	2.9898	2.9898	2.9898	2.9898	0

Robust MVO

Table 10: Robust MVO metrics with 6 months rebalancing frequency (Short allowed)

	OLS	FF	LASSO	BSS	FF5	Historical	EW
Sharpe Ratio:	0.1911	0.1909	0.1917	0.1789	0.1909	0.1906	0.1747
Skewness:	-0.835	-0.8407	-0.8227	-0.8417	-0.8384	-0.845	-0.4919
Max Drawdown:	0.3434	0.3437	0.3437	0.3631	0.3436	0.3444	0.4189
Mean Turnover:	0.1338	0.1367	0.1285	0.1523	0.1354	0.139	0

Maximum Sharpe Optimisation:

To keep things concise, we will show the maximum sharpe ratio techniques' results in the suboptimisation section as the individual results are very similar to cVaR.

b) **Sub Optimisation**

In this part, all the stocks would be divided into two sub groups according to their momentum or volatility. Each group will have 10 stocks, and will be optimized using different optimization algorithms. For example, the high-momentum group can be optimized by Risk Parity model while the low-momentum group might be optimized by Conditional VaR model.

In addition, parameters like proportion would be tuned in order to get better results. Plus, for kurtosis-skewness optimizing method, $\boldsymbol{\alpha_3}$ and $\boldsymbol{\alpha_4}$ should be adjusted as well. Since we can plug 4 models (Risk Parity, CVaR, maximum sharpe and kurtosis-skewness optimizing) into each group, there can be 16 combinations in total. In order to save space and keep the paragraph concise, we only displayed some of them with better performance.

Momentum

Table 11: Risk Parity for high momentum and CVaR for low momentum

proportion	0.4	0.45	0.5	0.55	0.6	0.65	0.7
Sharpe Ratio:	0.2939	0.2961	0.2963	0.2944	0.29	0.2832	0.274
skewness:	-0.5403	-0.5807	-0.6102	-0.63	-0.6432	-0.6542	-0.668
Max Drawdown:	0.136	0.1268	0.1331	0.1393	0.1456	0.1519	0.1593
Mean Turnover:	1.5772	1.4404	1.3054	1.1733	1.0512	0.9417	0.8471

Table 12: Skewness-Kurtosis optimization for high momentum and CVaR for low momentum ($\boldsymbol{\alpha_3}, \boldsymbol{\alpha_4} = 1, 1$)

proportion	0.4	0.45	0.5	0.55	0.6	0.65	0.7
Sharpe Ratio:	0.3136	0.313	0.3093	0.3028	0.2937	0.2828	0.2706
skewness:	-0.5811	-0.557	-0.5057	-0.4361	-0.3593	-0.286	-0.2239
Max Drawdown:	0.1271	0.1341	0.1411	0.1481	0.1556	0.1705	0.2023
Mean Turnover:	1.8851	1.7989	1.7186	1.6474	1.5843	1.5246	1.4774

Table 13: Maximum Sharpe for high momentum and CVaR for low momentum

proportion	0.4	0.45	0.5	0.55	0.6	0.65	0.7
Sharpe Ratio:	0.2705	0.2666	0.2605	0.2523	0.2421	0.2305	0.2176
skewness:	-0.6302	-0.678	-0.7121	-0.7332	-0.7441	-0.7482	-0.7493
Max Drawdown:	0.1841	0.1977	0.2112	0.2246	0.2379	0.2511	0.2652
Mean Turnover:	1.7852	1.6895	1.5986	1.5111	1.4278	1.3511	1.2851

Volatility

Table 14: Risk Parity for high vol. and CVaR for low vol.

proportion	0.4	0.45	0.5	0.55	0.6	0.65	0.7
Sharpe Ratio:	0.1937	0.1928	0.1911	0.1887	0.1858	0.1823	0.1785
skewness:	-0.8449	-0.8445	-0.8279	-0.7958	-0.75	-0.6927	-0.6265
Max Drawdown:	0.3674	0.3764	0.3853	0.3941	0.4028	0.4115	0.42
Mean Turnover:	1.0069	0.9424	0.88	0.8178	0.7575	0.7009	0.6532

Table 15: Skewness-Kurtosis optimization for high vol. and CVaR for low vol. ($\boldsymbol{\alpha_3}, \boldsymbol{\alpha_4} = 1, 1$)

proportion	0.4	0.45	0.5	0.55	0.6	0.65	0.7
Sharpe Ratio:	0.2169	0.2177	0.2175	0.2166	0.2151	0.213	0.2106
skewness:	-0.6347	-0.6008	-0.558	-0.5087	-0.4554	-0.4002	-0.3448
Max Drawdown:	0.3576	0.3653	0.3729	0.3805	0.3881	0.3957	0.4032
Mean Turnover:	1.1921	1.1513	1.1117	1.0767	1.042	1.0072	0.9726

Table 16: Maximum Sharpe for high vol. and CVaR for low vol.

proportion	0.4	0.45	0.5	0.55	0.6	0.65	0.7
Sharpe Ratio:	0.2383	0.2406	0.2417	0.2419	0.2413	0.2401	0.2383
skewness:	-0.3165	-0.2963	-0.2655	-0.2266	-0.1824	-0.1351	-0.0866
Max Drawdown:	0.3751	0.3839	0.3927	0.4015	0.4101	0.4186	0.4271
Mean Turnover:	0.8625	0.8243	0.7861	0.7496	0.7146	0.6795	0.6444

Evaluation of 5 Champion Candidates

In this section, we would pick one champion candidate from each of the models mentioned above (MVO, RP, Robust MVO, Sub Optimization Momentum and Sub Optimization Volatility), the list is as below:

Table 17: List of the best choice within each models

MVO	FF with 6 months rebalancing frequency (Short allowed and new target return)
RP	FF with 6 months rebalancing frequency (Short allowed)
Robust MVO	LASSO with 6 months rebalancing frequency (Short allowed)
Momentum	Risk Parity for high momentum and CVaR for low momentum (proportion=0.6)
Volatility	Maximum Sharpe for high vol. and CVaR for low vol. (proportion=0.65)

Moreover, we displayed additional metrics for these selected ones. Also, a trend plot of their portfolio values is provided for comparison.

Table 18: Metrics of the 5 candidates (CAGR is yearly data)

	MVO	RP	Robust MVO	Momentum	Volatility
Average Return:	0.0097	0.0076	0.0075	0.0102	0.0128
Return Standard Deviation	0.0358	0.0353	0.0336	0.0316	0.0485
CAGR:	0.1129	0.0858	0.085	0.1213	0.1473
Sharpe Ratio:	0.2418	0.1861	0.1917	0.29	0.2428
skewness:	-0.473	-0.7902	-0.8227	-0.6432	-0.1981
Max Drawdown:	0.1588	0.3503	0.3437	0.1456	0.4175
Cumulative Return	1.9408	1.2927	1.277	2.1723	2.9983
Mean Turnover	0.7558	0.1168	0.1285	1.0512	0.7714

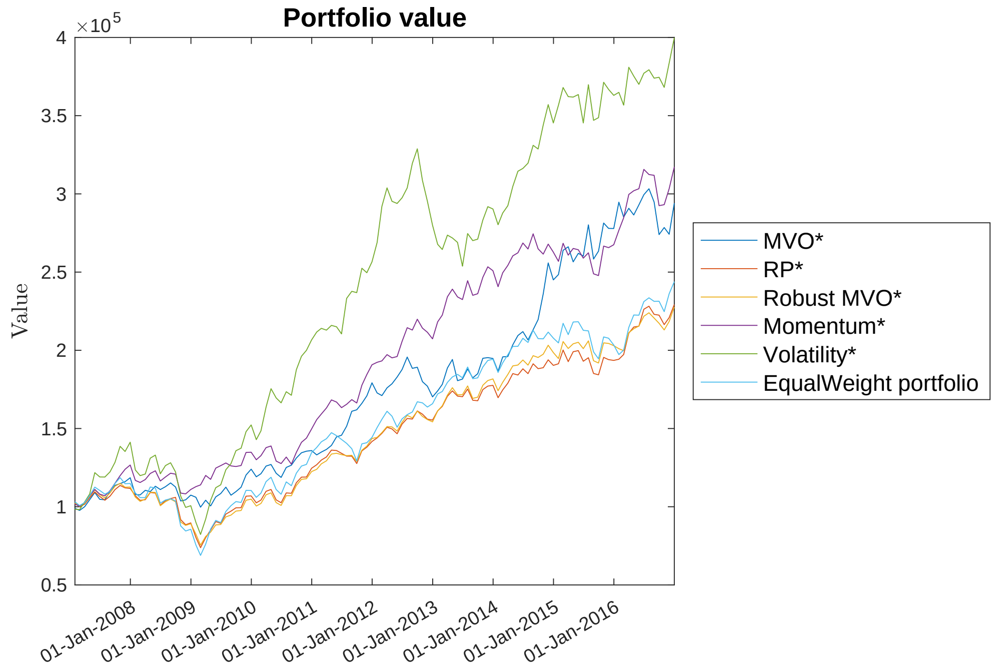


Figure 3: Cumulative Returns for All Candidates

From the diagram above, it is obvious that Momentum and Volatility rank the top 2. To shed more light on the two best methods, weights trend graphs are shown in the following:

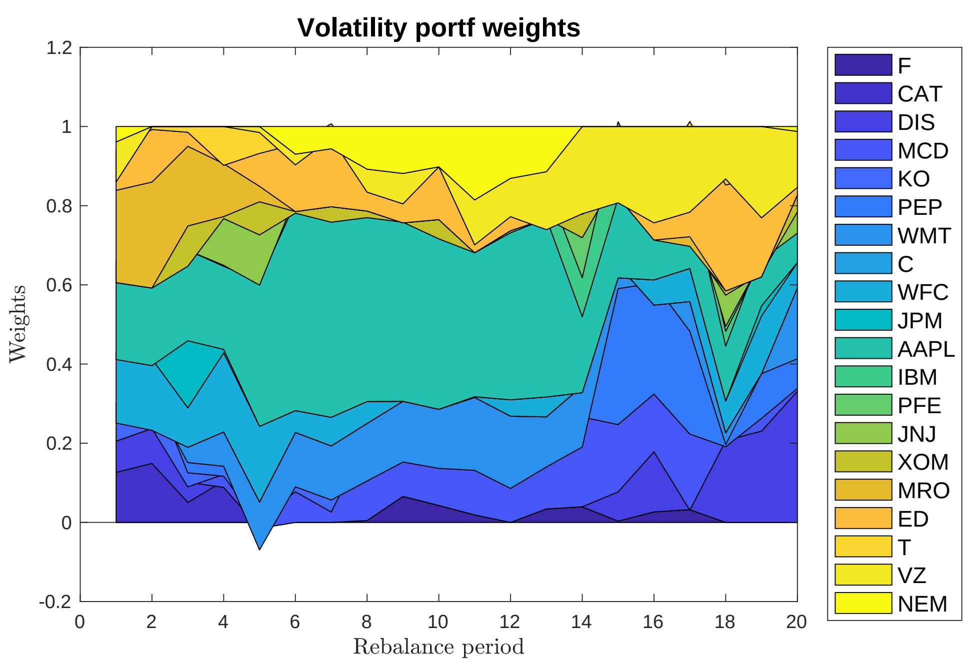


Figure 4: Weights Trend of Volatility Sub Optimizing Method

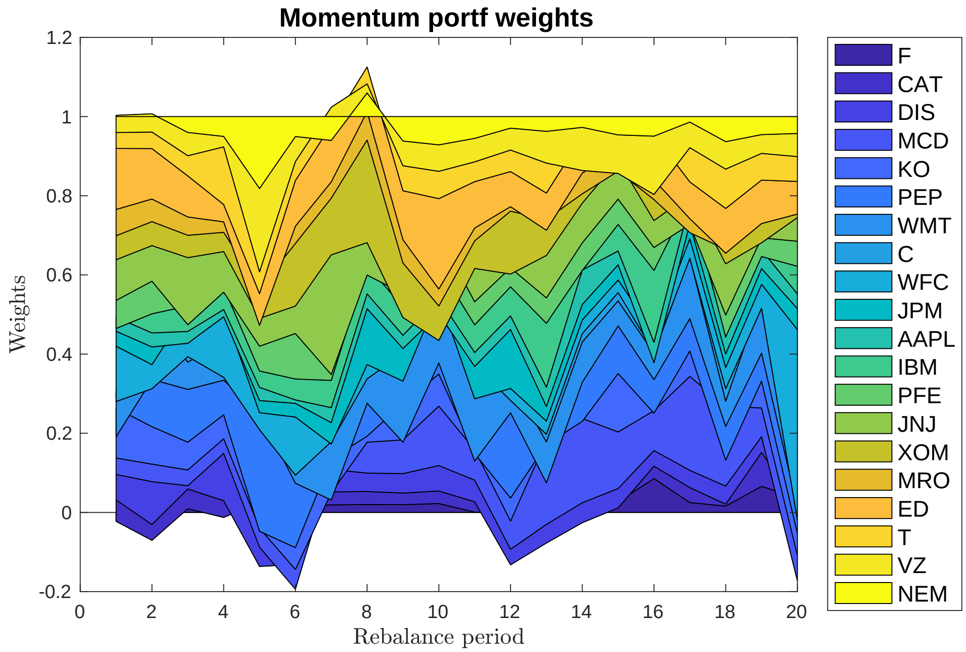


Figure 5: Weights Trend of Momentum Sub Optimizing Method

6 Conclusion and Discussion

After analysing our results, we chose **the momentum portfolio** with the sub optimisation technique, where we use **risk parity for high momentum stocks and cVaR for low momentum stocks** to be our model to be evaluated for the competition.

The **suboptimisation** technique works well as we can actively choose among the two baskets of stocks and then use the appropriate technique of optimisation. It showed us the best results because of the gap between winners and losers. On a hindsight perspective, the losers are far from the efficient frontier and we are supposed to optimize them in a way that the weights are allocated appropriately to the winners. The proportion parameter also helps with this, as we are adding more money to our high momentum stocks.

In a sufficiently diversified portfolio, most of the strategies can achieve similar returns, and the objective of **maximizing Sharpe ratio will regress to minimizing the volatility**. We thus chose MVO, Robust MVO, and RP. For two types of MVOs, the objective itself is to minimize the volatility , and risk parity has the established reputation of controlling risk. The significantly different return-risk profiles of individual stocks, as well as the financial intuition, are our drivers for the sub-portfolio optimization. Generally, the winner stocks will consistently outperform, while losers always lose; and those companies with highly volatile market prices has more uncertain outlook than others, and we might need to treat them differently.

In our best-Sharpe strategies, we have **high average($\geq$1.05) turnover rate**, which means we will on average sell about 53% of our portfolio and buy another batch of stocks when we rebalance. We consider the number acceptable because :

a) In our strategy, stocks are constantly migrating from one subset to another, and we apply different optimization methods in two subsets where individual stocks can be treated very differently. Thus, our turnover tend to be much higher than other optimization method;

b) Comparing to the strategy return, annually 210% of turnover will not cost us much in the actual investment. Most of the brokers charge around 5 bps for equity trading, and it is reasonable to assume the execution cost is also around 5 bps since all of the stocks in the portfolio are very liquid. Besides, the 10 bps penalty on turnover is also an industry standard of backtesting in Canadian buy side companies. The total transaction cost will be 0.21% per year, and it will not have a substantial impact on the return-risk profile of our portfolio.

Eugene Fama referred to momentum as the biggest challenge to his theory of financial market efficiency which implies that past prices of securities are irrelevant in the determination of future prices. Also, momentum does behave like a risk factor. This means that in some time periods, following a momentum-driven strategy will produce a negative return relative to the market. Of course, what happened in the past is no guarantee such tendencies will persist into the future. Hence, value and momentum are frequently combined, backed by reasonable arguments. Buying cheap stocks that are performing strongly might indicate corporate turnarounds and helps to avoid value traps. Inversely, buying stocks that have performed well, but are priced reasonably, aids to bypass highly speculative companies. Also, it makes sense to use momentum for this project, as our investment universe will be stocks or ETF's only. In future research, we should also explore a more important factor in momentum investing — rebalance frequency, and we can try to optimise this for better average turnover.