MMF1921: Operations Research

Project 1 (Summer 2023)

Group Members: Joshua Kim, Ian Lee, Liam Wu

Abstract: The rise of quantitative investment movement took off when Nobel Prize-winning economist Harry Markowitz published "Portfolio Selection" in March 1952. Since then the quantitative investment analysis has widely studied a portfolio theory, and many financial investors have focused in the portfolio return, risk and optimal methods of combining risky and risk-free assets in a particular portfolio. In this project, we will study and compare four different factor models such as Ordinary Least Squares model, Fama-French three-factor model, Least Absolute Shrinkage and Selection Operator model, and Bast Subset Selection model, and then we will utilize these models to estimate the parameters required for portfolio optimization. Lastly, we will build portfolios using mean–variance optimization (MVO) as our investment strategy.

Contents

1 Introduction

The goal of this project is to examine the effects of different factor models on the optimal choice for constructing portfolios using the Mean Variance Optimization investment strategy. The time frame investigated in this report is December 31, 2005, to December 31, 2016, with data concerning monthly-adjusted closing prices. Note, the adjustment refers to taking into account dividends, stock splits, risk-free rate, etc. The 20 assets used throughout this report are summarized below in *Table 1*.

Table 1: List of 20 assets by ticker

F	CAT	DIS	MCD	KO	PEP	WMT	C	WFC	JPM
T	IBM	PFE	JNJ	XOM	MRO	ED	AAPL	VZ	NEM

The historical monthly-adjusted closing prices for the 20 given assets are used to determine observed monthly asset returns. These, along with given factor returns for January 31, 2006, to December 31^{st}, 2016, will be used in the following factor models:

- Ordinary least squares regression on all eight factors (OLS model)
- Fama-French three factor model (FF model)
- Least absolute shrinkage and selection operator model (LASSO model)
- Best subset selection model (BSS model)

The given factors are summarize below in *Table 2*. Using the factor models and monthly-adjusted returns, covariance between the assets and covariance between the factors are determined, which combined are used as input for our portfolio optimization.

Table 2: List of eight factors

Market	Size	Value	Short-term Reversal
Profitability	Investment	Momentum	Long-term Reversal

2 Factor Models

All of the following factor models are based on the following assumptions of an ideal multi-factor environment:

- $\text{cov}(f_i, \epsilon_j) = 0$ for all i, j
- $\text{cov}(\epsilon_i, \epsilon_j) = 0$ for all $i \neq j$
- $\text{cov}(f_i, f_j) = 0$, for all $i \neq j$

where f_i is a factor and ϵ_i is the corresponding stochastic error term of f_i.

Note that the factor models used in this report do not respect the ideal environment, as a number of factors used are in fact correlated. For example, Size and Value, both being subsets of the market, are correlated.

2.1 OLS Model

Ordinary least squares (OLS) method is a common statistical technique for estimating the relationship between one or more independent quantitative variables and a dependent variable by the principle of least squares, i.e. minimizing the sum of the squares of the differences between the observed dependent variable and the linear output from the independent variables.

This analysis begins by examining a multi-factor OLS model, using all eight factors provided in *Table 2*. For the case of the OLS model, the return of asset i can be described by the following relationship:

$$r_i - r_f = \alpha_i + \textstyle\sum_{k=1}^{8} \beta_{ik} f_k + \epsilon_i$$

for $i \in \{1, \ldots, 20\}$ where r_i is the return of asset i, r_f is the risk-free rate, f_k is the the return of factor k, α_i and β_{ik} are the intercept from regression and corresponding factor loading, respectively, and ϵ_i is the stochastic error term of the asset. We can then calibrate the OLS model by choosing a set of optimal regression coefficients that satisfy the minimization problem seen below.

$$\begin{aligned}
\min_{\mathbf{B}_i} \quad \|r_i - XB_i\|^2 &= \min_{\mathbf{B}_i} \quad (r_i - XB_i)^\top (r_i - XB_i) \\
&= \min_{\mathbf{B}_i} \quad r_i^\top ri - 2r_i^\top XB_i + B_i^t opXB_i + B_i^\top X^\top XB_i \\
&= \min_{\mathbf{B}_i} \quad f(B_i)
\end{aligned}$$

where $\mathbf{B}_i = [\alpha_i \quad \mathbf{V}_i^\top]^\top$ is the vector of regression coefficients, $\mathbf{V}_i = [\beta_{i1} \quad \beta_{i2} \quad \cdots \quad \beta_{i8}]^\top$ is the factor loadings, and $\mathbf{X} = [\mathbf{1} \quad \mathbf{f}_1 \quad \mathbf{f}_2 \quad \cdots \quad \mathbf{f}_8]$.

This is a convex quadratic unconstrained minimization problem, thus only the first order necessary conditions must be satisfied to ensure a global minimum. Therefore, the optimal regression coefficients will satisfy the following system of equations.

$$\begin{aligned}
\bigtriangledown f(B_i) = -2X^\top r_i + 2X^\top XB_i = 0 \\
2X^\top XB_i = 2X^\top r_i \\
B_i = (X^\top X)^{-1} X^\top r_i
\end{aligned}$$

This implies a closed form solution to the OLS is possible with optimal regression coefficients B^* given by:

$$\therefore B^* = (X^\top X)^{-1} X^\top r$$

Finally, the residuals resulting from the optimal regression coefficients are:

$$\epsilon_i = \mathbf{r}_i - \mathbf{X}\mathbf{B}_i^*$$

which can be used to calculate the unbiased estimate of the residual variance:

$$\sigma_{\epsilon_i}^2 = \frac{1}{T - p - 1} \|\epsilon_i\|_2^2$$

where we divide by the appropriate number of degrees of freedom $T - p - 1$. This is a result from having T observations but calculating $p + 1$ coefficients (one α and p number of β).

2.2 FF Model

Next, we implemented a three-factor Fama-French model to be a subset of the already used OLS model. The Fama-Frech model considers only the Market, Size and Value factors, and the expected return for asset i is shown below:

$$r_i - r_f = \alpha_i + \beta_{im}(f_m - r_f) + \beta_{is} f_s + \beta_{iv} f_v + \epsilon_i$$

for $i \in \{1, \ldots, 20\}$ where r_i is the return of asset i, r_f is the risk-free rate, f_m, f_s and f_v are the Market, Size and Volume factors, respectively, α_i and β_{ik} are intercept from regression and corresponding factor loading respectively, and ϵ_i is the stochastic error term of the asset.

The optimal regression coefficients for the Fama-French model are derived using a similar approach that is described above in the OLS subsection. In other words, finding the optimal Fama-French coefficients becomes an identical optimization problem as the OLS model with only three factors.

2.3 LASSO Model

One problem with including a large pool of factors is that we may run into risk of overfitting, thus making the models very sensitive to noise generated from estimating the parameters. Additionally, it reduces the interpretability of the model as accurate risk attribution is difficult when we have a very large number of factors in the model. One approach to avoid this is by introducing regularization to promote sparsity and reduce statistical overfitting.

Here, we implemented a penalized form of the LASSO (least absolute shrinkage and selection operator) model, one of the two sparse models in this report seen below:

$$\min_{\mathbf{B}_i} \quad \|\mathbf{r}_i - \mathbf{X}\mathbf{B}_i\|_2^2 + \lambda\|\mathbf{B}_i\|_1$$

where $\mathbf{r}_i$ is the return of asset i, λ is the penalty parameter of our choice, and $\mathbf{B}_i$ is the vector of regression coefficients. Note that λ is selected by hyperparameter tuning and the optimal value of λ is determined to be 0.029, which is discussed in further detail in *Section 4*.

The ℓ_1 norm used by the LASSO model is continuous and convex but is not smooth everywhere. In order to address this issue that arises from the absolute value of $\mathbf{B}_i$, we make the following substitutions:

$$\begin{aligned} \mathbf{B}_i &= \mathbf{B}_i^+ - \mathbf{B}_i^- \\ \|\mathbf{B}_i\|_1 &= \mathbf{B}_i^+ + \mathbf{B}_i^- \end{aligned}$$

Using the substitutions, the penalized objective function of the LASSO model becomes solvable:

$$\begin{aligned} & \|\mathbf{r}_i - \mathbf{X}\mathbf{B}_i\|_2^2 + \lambda\|\mathbf{B}_i\|_1 \\ = \quad & (r_i - \mathbf{X}\mathbf{B}_i)^\top (r_i - \mathbf{X}\mathbf{B}_i) + \lambda\|\mathbf{B}_i\|_1 \\ = \quad & r_i^\top r_i - 2r_i^\top \mathbf{X}\mathbf{B}_i + \mathbf{B}_i^\top \mathbf{X}^\top \mathbf{X}\mathbf{B}_i + \lambda\|\mathbf{B}_i\| \\ = \quad & r_i^\top r_i + \left[\begin{pmatrix} -2r_i^\top \mathbf{X} \\ 2r_i^\top \mathbf{X} \end{pmatrix} + \begin{pmatrix} \lambda \\ \vdots \\ \lambda \end{pmatrix}\right]^\top \begin{bmatrix} \mathbf{B}_i^+ \\ \mathbf{B}_i^- \end{bmatrix} + \begin{bmatrix} \mathbf{B}_i^+ \\ \mathbf{B}_i^- \end{bmatrix}^\top \begin{bmatrix} \mathbf{X}^\top\mathbf{X} & -\mathbf{X}^\top\mathbf{X} \\ -\mathbf{X}^\top\mathbf{X} & \mathbf{X}^\top\mathbf{X} \end{bmatrix} \begin{bmatrix} \mathbf{B}_i^+ \\ \mathbf{B}_i^- \end{bmatrix} \end{aligned}$$

Let $\mathbf{H} = \mathbf{2}\begin{bmatrix} \mathbf{X}^\top\mathbf{X} & -\mathbf{X}^\top\mathbf{X} \\ -\mathbf{X}^\top\mathbf{X} & \mathbf{X}^\top\mathbf{X} \end{bmatrix}$, $\quad \mathbf{c}^\top = \left[\begin{pmatrix} -2r_i^\top\mathbf{X} \\ 2r_i^\top\mathbf{X} \end{pmatrix} + \begin{pmatrix} \lambda \\ \vdots \\ \lambda \end{pmatrix} \right]^\top$.

The original problem is now in the standard form of MATLAB *QuadProg*:

$$\begin{aligned} \min_{\mathbf{B}_i^+, \mathbf{B}_i^-} \quad & \frac{1}{2}\begin{bmatrix} \mathbf{B}_i^+ \\ \mathbf{B}_i^- \end{bmatrix}^\top \mathbf{H} \begin{bmatrix} \mathbf{B}_i^+ \\ \mathbf{B}_i^- \end{bmatrix} + \mathbf{c}^\top \begin{bmatrix} \mathbf{B}_i^+ \\ \mathbf{B}_i^- \end{bmatrix} \\ \text{s.t.} \quad & \mathbf{B}_i^+, \mathbf{B}_i^- \geq \mathbf{0}. \end{aligned}$$

Note that $r_i^\top r_i$ is a constant term so it can be omitted in the objective function of the minimization problem. This standard form of *QuadProg* gives an equivalent solution to the original quadratic programming problem with linear constraints containing absolute values [1].

2.4 BSS Model

The best subset selection (BSS) model serves a similar purpose as the LASSO method where both models promote true sparsity. However, the BSS model introduces a cardinality constraint instead of penalizing $\mathbf{B}_i$ in ℓ_1 norm. In this project, we consider the constrained form of the BSS Model using all eight factors as inputs. The model is formulated as follows:

$$\begin{aligned} \min_{\mathbf{B}_i} \quad & \|\mathbf{r}_i - \mathbf{X}\mathbf{B}_i\|_2^2 \\ \text{s.t.} \quad & \|\mathbf{B}_i\|_0 \leq K \end{aligned}$$

where $K = 4$ is the cardinality constraint parameter of our choice.

Given the discontinuity and non-convexity of the ℓ_0 norm, solving the model is computationally difficult. However, we can introduce auxiliary binary variables to address the ℓ_0 norm constraint, converting the problem into a mixed-integer quadratic program (MIQP). Employing the work of Bertsimas [2], the following is the derivation:

$$\begin{aligned} & \|\mathbf{r}_i - \mathbf{X}\mathbf{B}_i\|_2^2 \\ = \quad & (r_i - \mathbf{X}\mathbf{B}_i)^\top (r_i - \mathbf{X}\mathbf{B}_i) \\ = \quad & r_i^\top r_i - 2r_i^\top \mathbf{X}\mathbf{B}_i + \mathbf{B}_i^\top \mathbf{X}^\top \mathbf{X}\mathbf{B}_i \end{aligned}$$

The original problem is now in the standard form of MATLAB *QuadProg*:

$$
\begin{aligned}
\min_{\mathbf{B}_i} \quad & \|\mathbf{r}_i - \mathbf{X}\mathbf{B}_i\|_2^2 \\
s.t. \quad & \|\mathbf{B}_i\|_0 \le K \\
\implies \min_{\mathbf{B}_i, y_i} \quad & \frac{1}{2}\|r_i - \mathbf{X}\mathbf{B}_i\|_2^2 \\
s.t. \quad & lb \cdot y_i \le B_i \le ub \cdot y_i \\
& y_i \in \{0,1\}, \\
& \sum y_i \le K. \\
\implies \min_{\mathbf{B}_i, y_i} \quad & \frac{1}{2}\begin{bmatrix}\mathbf{B}_i \\ y_i\end{bmatrix}^\top \mathbf{H}\begin{bmatrix}\mathbf{B}_i \\ y_i\end{bmatrix} + \mathbf{c}^\top \begin{bmatrix}\mathbf{B}_i \\ y_i\end{bmatrix} \\
\text{s.t.} \quad & \mathbf{A}\begin{bmatrix}\mathbf{B}_i \\ y_i\end{bmatrix} \le \mathbf{b} \\
& y_i \in \{0,1\}
\end{aligned}
$$

Where $\mathbf{H} = 2\begin{bmatrix}\mathbf{X}^\top\mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{bmatrix}$, $c = \begin{bmatrix}-2r_i^\top\mathbf{X} \\ \mathbf{0}\end{bmatrix}$, $\mathbf{A} = \begin{bmatrix}-\mathbf{I}_d & lb \cdot \mathbf{I}_d \\ \mathbf{I}_d & -ub \cdot \mathbf{I}_d \\ \mathbf{0}_d & \mathbf{1}_d\end{bmatrix}$, $\mathbf{b} = \begin{bmatrix}0 \\ \vdots \\ 0 \\ K\end{bmatrix}$, $\mathbf{0}_d = \begin{bmatrix}0 \\ \vdots \\ 0\end{bmatrix}^\top$, $\mathbf{1}_d = \begin{bmatrix}1 \\ \vdots \\ 1\end{bmatrix}^\top$, $lb = -30$, $ub = 30$, and d is one greater than the number of parameters.

An upper bound and lower bound, ub and lb respectively, are selected to be sufficiently large but do not encapsulate the noise from integer programming in MATLAB. See [2] for further details.

3 Portfolio Optimization

As stated above, mean-variance optimization is used to calculate the optimal portfolio by minimizing variance subject to our target expected return. Mean-variance optimization is defined as below:

$$
\begin{aligned}
\min_{x} \quad & x^\top \mathbf{Q} x \\
\text{s.t.} \quad & \mu^\top x \ge R \\
& \mathbf{1}^\top x = 1 \\
& x \ge \mathbf{0}
\end{aligned}
$$

where x is our portfolio weights, $\mathbf{Q}$ is our covariance matrix, μ is the expected returns of the assets, R is our target return, and $\mathbf{1}$ and $\mathbf{0}$ are the ones vector and zeroes vector respectively. Here, the target return is calculated as the geometric mean of the market factor for the calibration period. The investment horizon simulated is a five year period, from January 2012 to December 2016. Our portfolio is rebalanced at the start of every year, with the previous four years used as the calibration period.

Transaction cost has not been considered during optimization; however this does not resemble today's financial markets where transaction costs are associated with every trading activity. To examine how each portfolio would perform while incurring trading fees, we create a scenario where transaction cost is determined by 0.5% the notional value of stock being bought or sold, and it is calculated every time the portfolio is rebalanced. Note that we do not include transaction cost associated with the initial construction of our portfolio; transaction cost is considered every other rebalancing period.

This quadratic program was solved using MATLAB *QuadProg*. μ and $\mathbf{Q}$ are:

$$\mu = \alpha + V^\top \overline{f}$$
$$Q = V^\top F V + D$$

where

$$V_i = \begin{bmatrix} \beta_{i1} \\ \beta_{i2} \\ \\ \beta_{pi} \end{bmatrix} \in \mathbb{R}^p, \quad D = \begin{bmatrix} \sigma^2_{\epsilon_1} & 0 & \cdots & \cdots & 0 \\ 0 & \sigma^2_{\epsilon_2} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & 0 & \sigma^2_{\epsilon_{N-1}} & 0 \\ 0 & \cdots & \cdots & 0 & \sigma^2_{\epsilon_N} \end{bmatrix}$$

and $\sigma^2_{\epsilon_i}$ is the residual variance of asset i, $\sigma^2_{\epsilon_i} = \frac{1}{T-P-1}\|\epsilon_i\|_2^2$, $\epsilon_i = r_i - \mathbf{X}\mathbf{B}_i$.

4 In-Sample Analysis

4.1 Factor Model Measure of Fit

Beginning with the in-sample section of the analyses, the adjusted R^2 was calculated to compare the different factor models, and the equation is given below.

$$R^2_{adj} = 1 - (1 - R^2)\frac{n-1}{n-p-1}$$

$$\mathrm{R}^2 = 1 - \frac{SS_{res}}{SS_{tot}}, \quad SS_{tot} = \sum(y_i - \overline{y})^2, \quad SS_{res} = \sum(y_i - f_i)^2 = \sum \epsilon_i^2$$

where f_i is the predicted return of the asset at time i, $\overline{y}$ is the mean return for the asset over the investment period, and y_i is the actual return for the asset at time i. In general, R^2 either increases or remains the same when new predictors are introduced to the model, but high R^2 can indicate a problem in the model. This can cause misinterpretation of the model that is explained by variables. On the other hand, adjusted R^2 increases only when independent variable is significant and affects dependent variable. It penalizes for introducing extra predictor variables that don't improve the existing model.

The R^2_{adj} is calculated for every asset, over every period and for all factor models. The data is shown below, in *Table 3*. Note that for the remainder of this section, R^2_{adj} is taken to be the asset-weighted average over every period for each factor model: $\overline{R}^2_{adj} = \mathbf{x}^\top R^2_{adj}$, where $\mathbf{x}$ is the portfolio weight vector.

Table 3: $\overline{R}^2_{adj}$, averaged over 20 assets, for all factor models and periods

Factor Model	Period 1	Period 2	Period 3	Period 4	Period 5
OLS Model	0.3023	0.2584	0.2851	0.3351	0.4354
FF Model	0.2691	0.1879	0.1841	0.1504	0.3371
LASSO Model	0.2335	0.1561	0.1293	0.0966	0.2410
BSS Model	0.1947	0.1201	0.1263	0.0987	0.3202

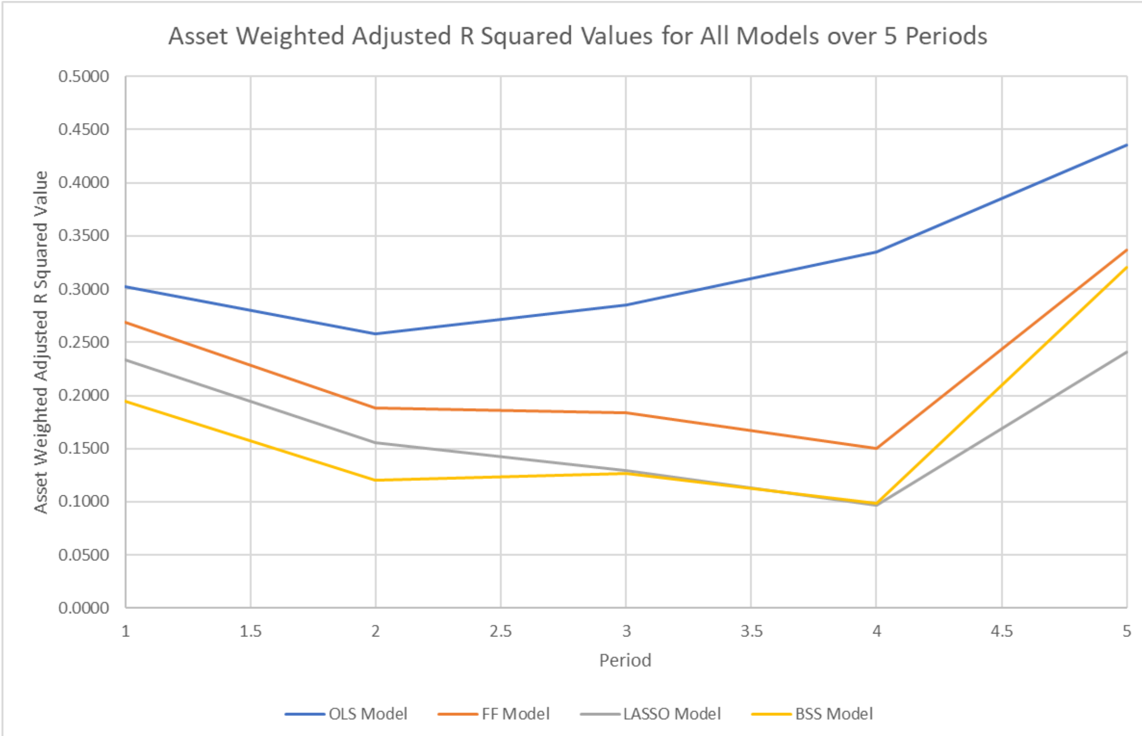


Figure 1: $\overline{R}^2_{adj}$ for all models, over 5 periods

The OLS model achieved the highest average $\overline{R}^2_{adj}$, averaged over all 5 periods, with a value of 0.3233. Next followed the FF Model with an value of 0.2257, then the BSS Model with a value of 0.1720 and finally the LASSO model with a value of 0.1720. The fact that the OLS model has the highest in-sample correlation is expected since it has the most number of features, i.e. factors, to estimate the asset returns. This could however also indicate that it may have overfitted to its calibration data and an out-of-sample analysis is required to verify this. In fact, we often run into the risk of OLS model overfitting when $p >> T$, where p is the number of factors and T is the number of observations. Similarly, it is natural for the Fama-French model and the BSS model to have lower in-sample average $\overline{R}^2_{adj}$ since they are constrained to select fewer factors. The LASSO model also has lower correlation values as its objective function is simply a penalized version of the OLS method, resulting in a worse in-sample fit.

$\overline{R}^2_{adj}$ is plotted for all models, over all periods, and can be seen below in *Figure 1*.

4.2 LASSO Model Hyperparameter Tuning

In *Section 2.3*, we introduce the ℓ_1 norm penalty parameter λ and leave it undetermined. In this section, we address this by tuning the hyperparameter based on in-sample $\overline{R}^2_{adj}$ values. The effect of varying λ and its impact on $\overline{R}^2_{adj}$ is summarized in *Table 4*. The general trend is that $\overline{R}^2_{adj}$ decreases and converges to 0 as λ increases as high λ encourages coefficients to shrink, making the model less sensitive to noise.

The λ value chosen for our LASSO model was 0.029, for a number of reasons. Primarily, we sought to have between 2 and 5 factors selected by the LASSO model across all assets, which is achieved by λ between 0.029 and 0.031. Empirically speaking, having more than 5 factors has risk of capturing spurious patterns which do not generalize on out-of-sample data; having less than 2 factors simplifies the model too much, leading to poor performance. Consequently, we examined the $\overline{R}^2_{adj}$ values of the LASSO models within this range of λ values, and $\lambda = 0.029$ yielded the highest average $\overline{R}^2_{adj}$ of 0.1713.

	Period 1	Period 2	Period 3	Period 4	Period 5
$\lambda = 0.1$	0.1394	0.0487	0.0690	0.0181	0.0094
$\lambda = 0.2$	0.0291	0.0179	0.0058	-0.0239	-0.0244
$\lambda = 0.3$	0.0170	-0.0030	-0.0258	-0.0273	-0.0277
$\lambda = 0.4$	-0.0047	-0.0152	-0.0280	-0.0312	-0.0322
$\lambda = 0.5$	-0.0124	-0.0251	-0.0302	-0.0357	-0.0377
$\lambda = 0.6$	-0.0178	-0.0310	-0.0321	-0.0405	-0.0434
$\lambda = 0.7$	-0.0207	-0.0372	-0.0352	-0.0453	-0.0489
$\lambda = 0.8$	-0.0211	-0.0442	-0.0380	-0.0485	-0.0548
$\lambda = 0.9$	-0.0238	-0.0528	-0.0411	-0.0507	-0.0604
$\lambda = 1.0$	-0.0267	-0.0576	-0.0437	-0.0427	-0.0657
$\lambda = 1.1$	-0.0265	-0.0619	-0.0453	-0.0472	-0.0694
$\lambda = 1.2$	-0.0282	-0.0654	-0.0498	-0.0320	-0.0806
$\lambda = 1.3$	-0.0300	-0.0659	-0.0547	-0.0573	-0.0909
$\lambda = 1.4$	-0.0277	-0.0702	-0.0600	-0.0674	-0.1055
$\lambda = 1.5$	-0.0278	-0.0741	-0.0657	-0.0690	-0.1138

Table 4: $\overline{R}^2_{adj}$ for LASSO models for varying λ values

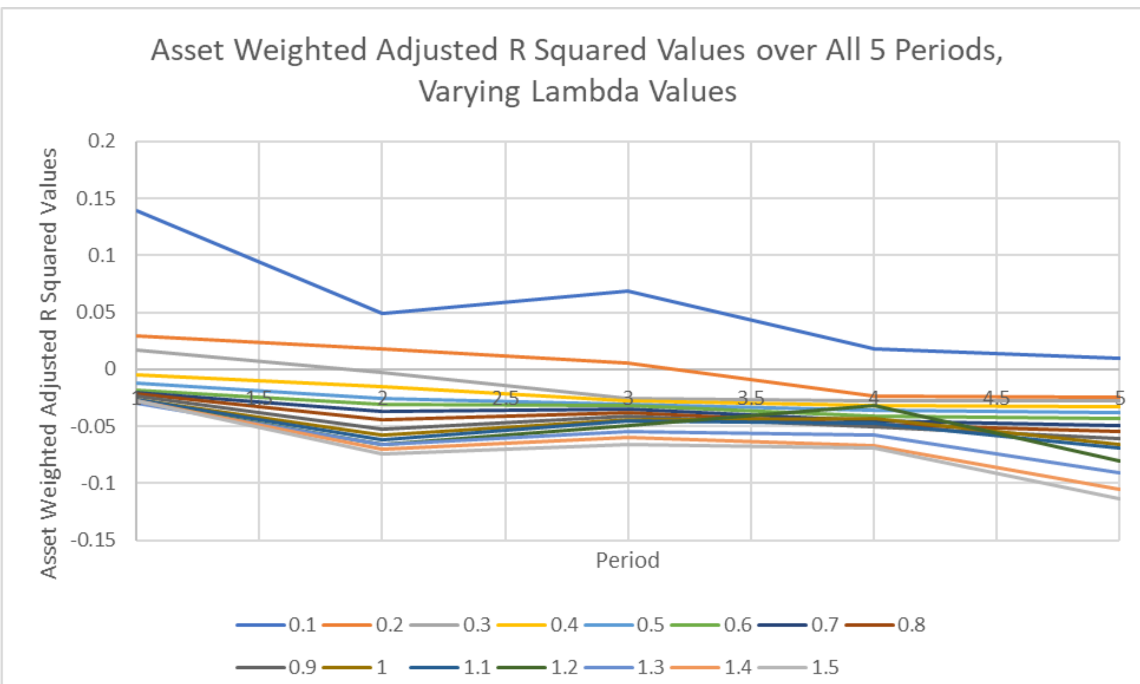


Figure 2: $\overline{R}^2_{adj}$ for LASSO models for varying λ values

4.3 BSS Model Hyperparameter Tuning

Although we specify the K value of the BSS model as 4, we took a step further and examined the effect of varying K values on the asset-weighted average $\overline{R}^2_{adj}$; the results are summarized in *Table 5* and *Figure 3*, shown below.

Table 5: $\overline{R}^2_{adj}$ for BSS models with varying K values

	Period 1	Period 2	Period 3	Period 4	Period 5
$K = 1$	0.1770	0.4344	0.4385	0.3911	0.3287
$K = 2$	0.2003	0.1361	0.4469	0.4141	0.4100
$K = 3$	0.2032	0.1275	0.5059	0.1190	0.3277
$K = 4$	0.1947	0.1201	0.1263	0.0986	0.3201

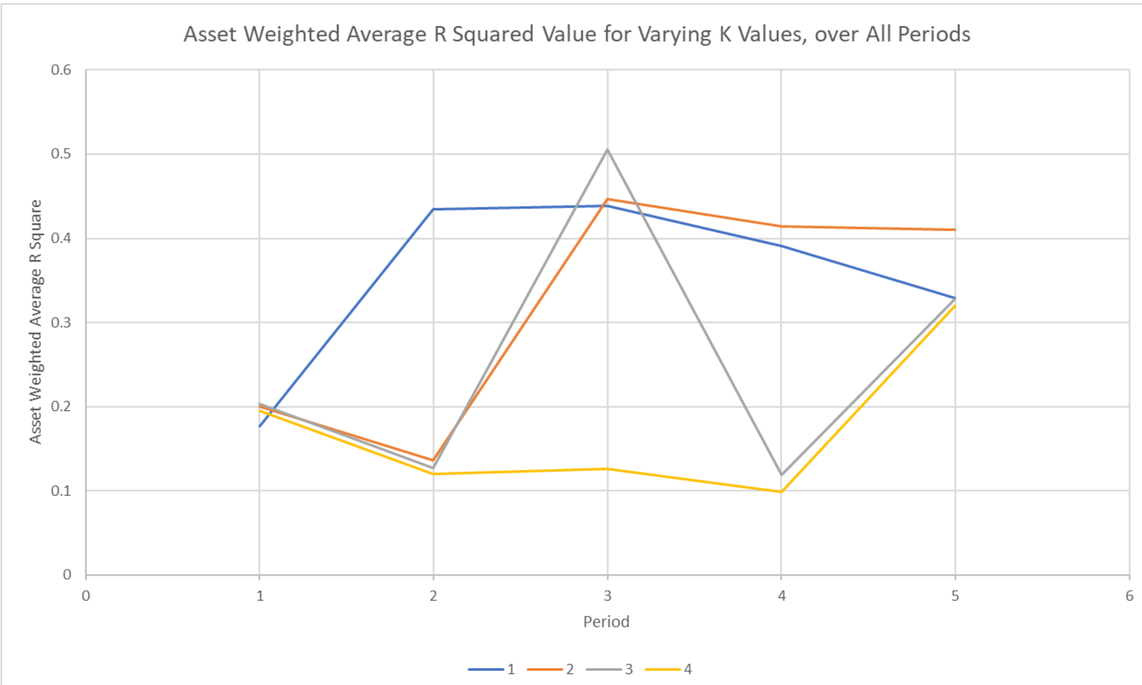


Figure 3: $\overline{R}^2_{adj}$ for BSS models for varying K values

By observing *Figure 3*, we can see that $K = 1$ plot provides a general upper bound for the rest of the asset-weighted average $\overline{R}^2_{adj}$ plots with larger K values. Similarly, $K = 4$ plot provides a general lower bound for the rest of the asset-weighted average $\overline{R}^2_{adj}$ plots with smaller K values. That $K = 4$ plot does not necessarily dominate the others can be attributed to the notion that we take the number of features, i.e. the degrees of freedom, into account when computing $\overline{R}^2_{adj}$.

5 Out-of-Sample Analysis

5.1 Portfolio Performance

First, we begin by calculating the portfolio average monthly returns and volatilities sustained over the entire period. The results can be seen summarized in *Table 6, 7* seen below.

Table 6: Average monthly realized portfolio returns

Factor Model	Average Monthly Return
OLS Model	1.00696559426219
FF Model	1.00687232002119
LASSO Model	1.00707858175983
BSS Model	1.00814027626078

Table 7: Monthly realized portfolio standard deviations over 5 periods

Factor Model	Standard Deviation over 5 Periods
OLS Model	0.0261136724681486
FF Model	0.0262353074567235
LASSO Model	0.0265747411988063
BSS Model	0.0485108999269709

Evidently, the BSS model achieved the highest average monthly portfolio return over the period from 01-Jan-2012 to 31-Dec-2016. The BSS model was then followed narrowly by the LASSO, OLS and FF models. However, the BSS model also maintained the highest standard deviation over five periods. The BSS model had a calculated deviation nearly twice as large as the other three models.

Next, the Sharpe ratio (risk-adjusted return), of the four portfolios was determined. The calculated Sharpe ratios can be seen in the table below.

Table 8: Portfolio Sharpe ratios

Factor Model	Sharpe Ratio
OLS Model	2.55405930846781
FF Model	2.50085335623305
LASSO Model	2.55951333610139
BSS Model	1.66730926358282

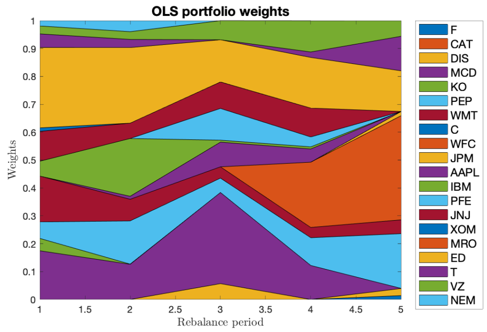
The calculated Sharpe ratios are consistent with our expectations. The similar values for the OLS, FF and LASSO models can be explained by the nearly identical returns and volatilies across these three models. On the other hand, we expect the BSS model to have a lower Sharpe ratio because of its relatively higher standard deviation over the five periods.

Additionally, the values of the four portfolios were plotted against time to help visualize the development of our portfolios over the given investment period. The plot can be seen below in *Figure 5*. The OLS, Fama-French, and LASSO models seem to follow similar trends as they progressed through time. However, the BSS model evidently did not follow the same trend. Furthermore, in the BSS model there is a large peak present in the first half of 2015.

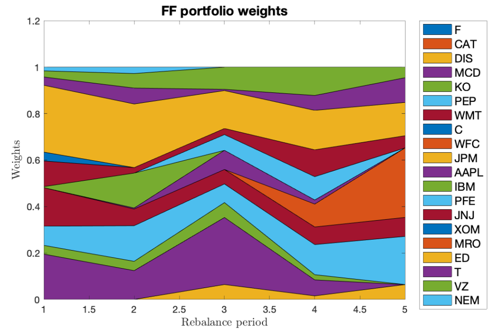
More insight into the performance of the portfolios can be gained by examining how their composition changes over time. The area plots below depict the per-period changes in composition corresponding to the four portfolio models. Over the given investment horizon, the OLS, Fama-French and LASSO models all maintained diverse holdings throughout the entire period. More specifically, these three portfolios maintained a minimum of a third of the provided stocks in every period.

The BSS model, however, did not maintain diversity over the investment period. Although the BSS model maintained some level of diversity up until the end of year two, the subsequent rebalance resulted in a 100% holding of AAPL in year 3. Again, the BSS model's holdings were not diverse in the final year of the investment period where the model sold its entire portfolio of AAPL to purchase an equivalent holding of DIS.

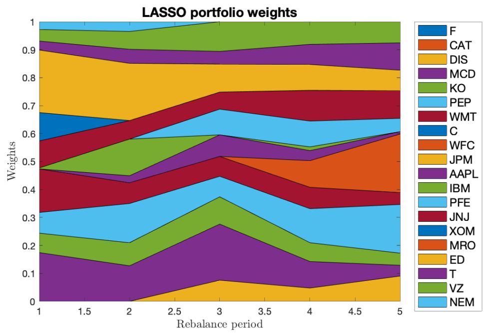
Evidently, there is evidence of instability within the BSS model. Some of this instability can be explained by the existence of the linear 'no short selling' constraint [3]. This constraint appears to dominate the BSS model in particular. To verify the effect of short-selling we reran the model by allowing for short sales in *Section 5.2*.



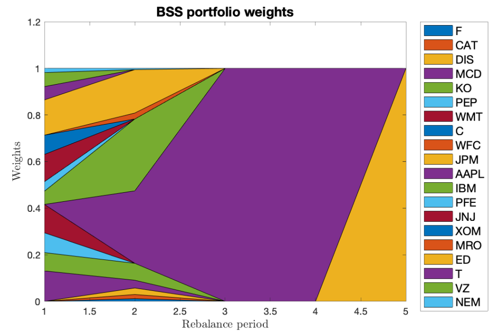
(a) OLS



(b) Fama-French Model



(c) LASSO Model



(d) BSS Model

Figure 4: Portfolio composition, all models, short selling disallowed

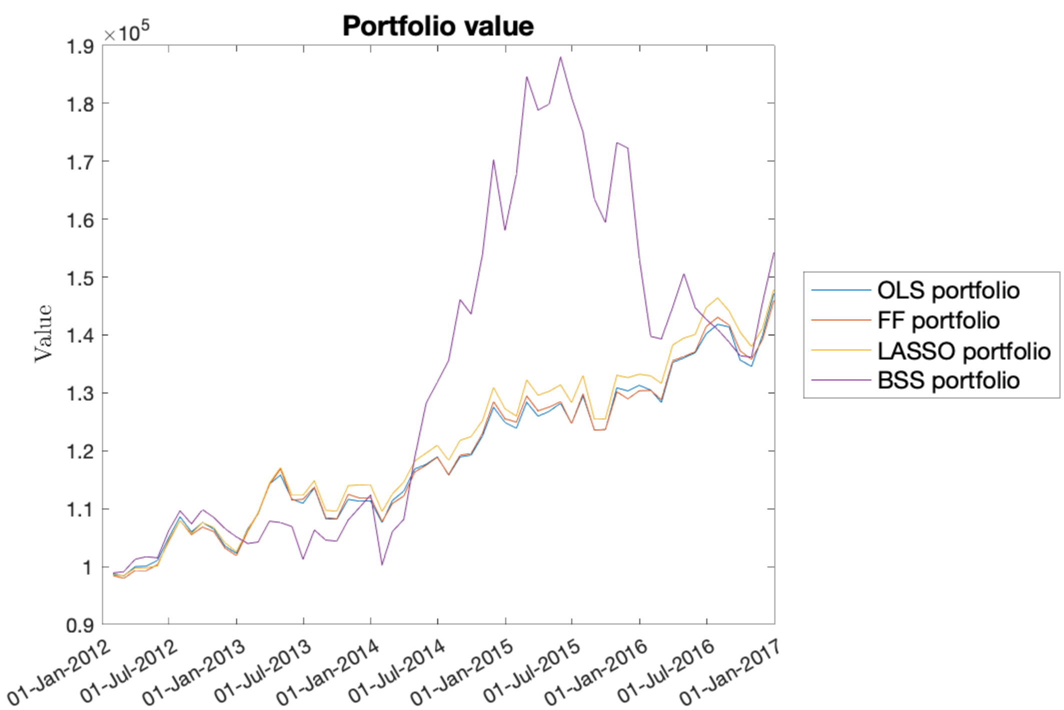
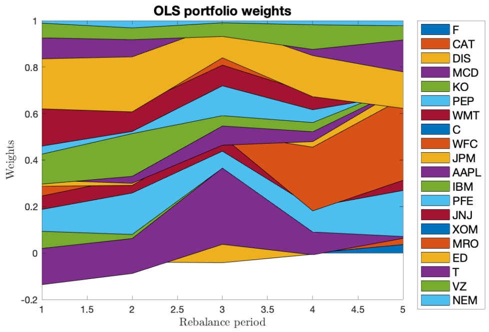
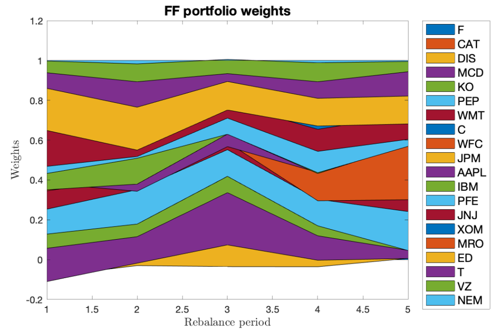


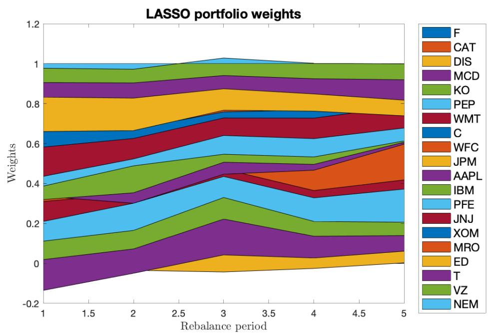
Figure 5: Value of portfolio over all models and periods with short selling disallowed



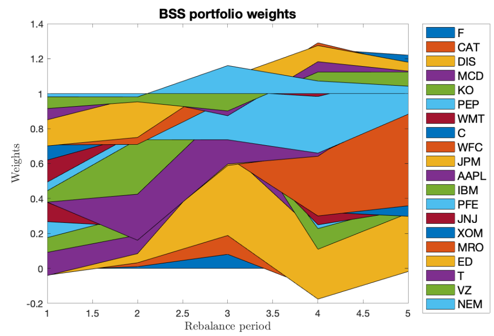
(a) OLS



(b) Fama-French Model



(c) LASSO Model



(d) BSS Model

Figure 6: Portfolio composition of all models, short selling allowed

5.2 Short-Selling

The area charts for the four portfolios with short-selling allowed have been included in *Figure 6*. When short sales are allowed, the OLS, Fama-French and LASSO models continue to demonstrate a strong degree of diversity. However, unlike when short sales were disallowed the BSS model is able to maintain a much more diverse portfolio over the entire investment horizon.

Additionally, we computed the Sharpe ratios across the four portfolios with short-selling allowed. Allowing for short sales improves the ratio for the OLS, Fama-French and LASSO models and decreases the ratio for the BSS model. The summarized Sharpe ratios and the time-values of the short-sales portfolios have been provided in *Table 9* and *Figure 7*.

Table 9: Portfolio Sharpe ratios with short selling allowed

Factor Model	Sharpe Ratio
OLS Model	3.05333632291607
FF Model	2.84522533640213
LASSO Model	2.61463291957453
BSS Model	1.20918274691080

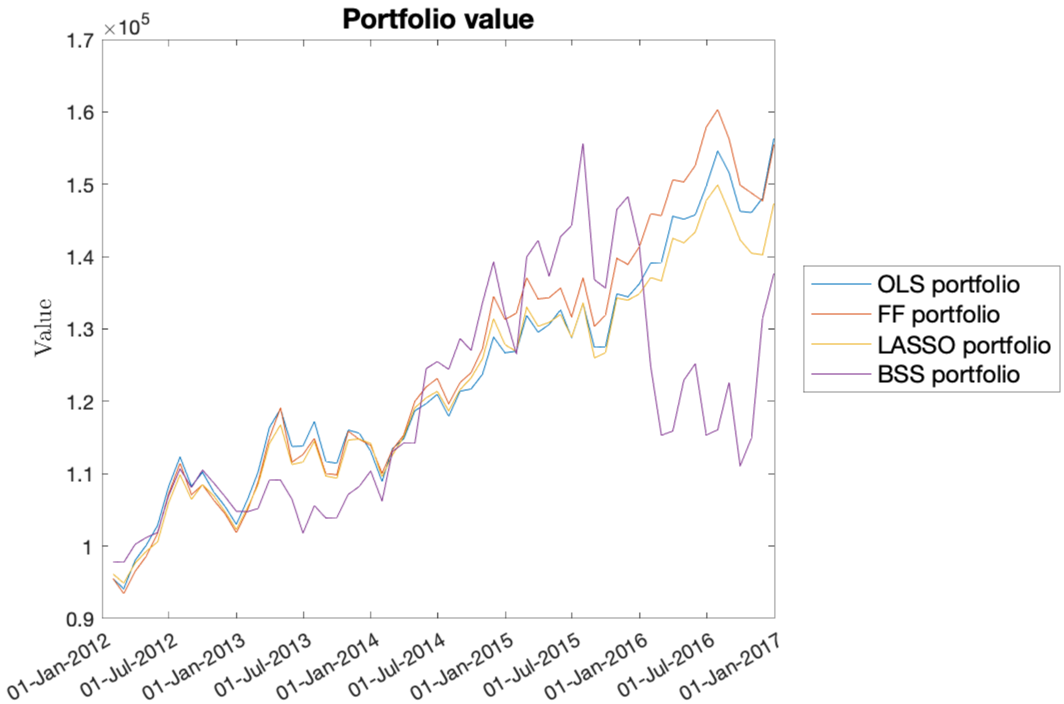


Figure 7: Value of portfolio over all models and periods with short selling allowed

5.3 Transaction Cost Analysis

In this section, we examine a scenario where there exist transaction costs associated with the rebalancing of the four portfolios in each period with short-selling disallowed. The calculated costs have been provided in the table below.

Table 10: Yearly transaction costs – note that the transaction costs of the first period are set to be 0

Factor Model	Period 2	Period 3	Period 4	Period 5
OLS Model	2957.590412975	5664.526475571	4631.892555779	5538.984919355
FF Model	3092.850675816	4695.478713997	4762.389623185	5067.936656269
LASSO Model	2787.259519422	3872.975510871	3062.429644305	3023.257627187
BSS Model	7326.643170966	7707.123812668	0.1083078460085	15326.31280947

The results above suggest that three of the four models all share very similar transaction cost. Specifically, the OLS, Fama-French and LASSO models whom all share transaction costs of the same order magnitude of 10^3. Again, the BSS model differs significantly with much higher transaction costs in years one, two and four and nearly negligible costs in year three. This is expected since we observe large portfolio turnovers, especially in periods 3 and 5.

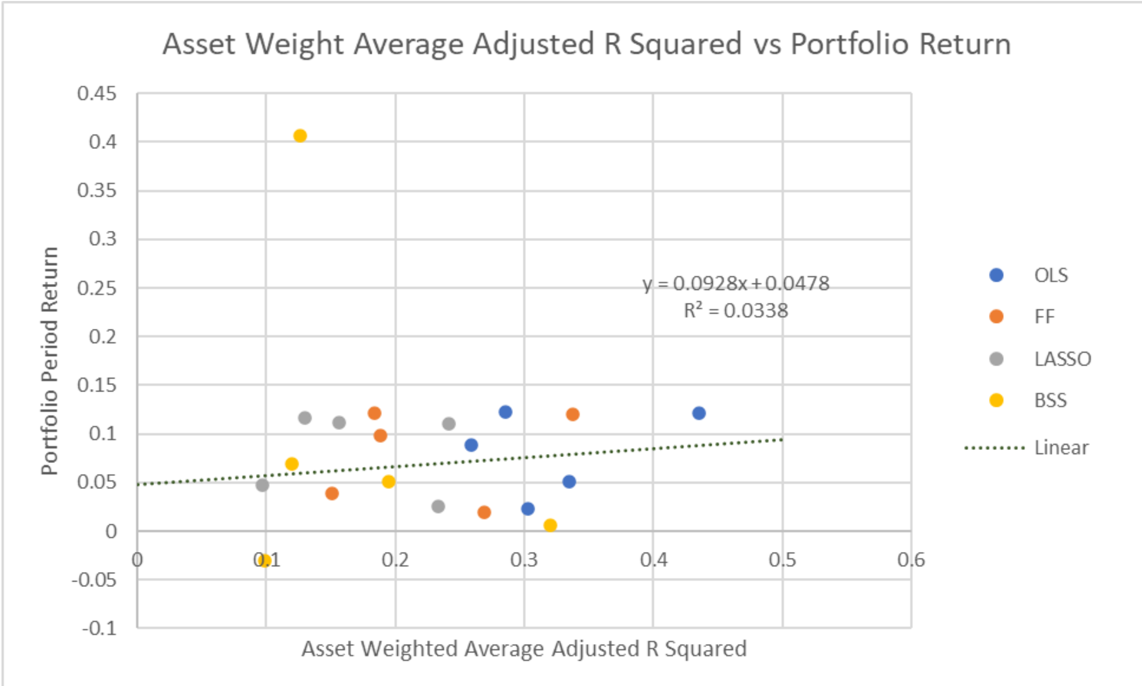


Figure 8: $\overline{R}^2_{adj}$ vs portfolio return

5.4 Portfolio Measure of Fit

Finally, to determine if the positive linear relation between $\overline{R}^2_{adj}$ and portfolio excess return is significant, null-hypothesis testing should be done over the slope of the line of best fit, with data summarized above in *Figure 8*. Using the student-t distribution, the p-value of the slope is calculated to be 0.4515. With the significance benchmark as 0.05, the calculated p-value is greater than the benchmark, thus the null hypothesis of the slope being 0 cannot be rejected. This implies that the linear relation is not significant and we cannot conclude that there exists a positive correlation between $\overline{R}^2_{adj}$ and portfolio excess return. A possible next step is to collect more data points and rerun the hypothesis testing.

6 Conclusion

The goal of this analysis was to study and compare the OLS, Fama-French, LASSO and BSS models. We then used the four models to estimate the inputs required for portfolio optimization. Finally, we built four portfolios for each of the factor models using MVO as our investment strategy which minimizes volatility with a targeted return. The aim of this investigation was achieved; four portfolios were constructed corresponding to the four factor models. The BSS model achieved the highest average monthly return however, when inspecting Sharpe ratios, the OLS, FF and LASSO models appear to perform stronger because of their relatively lower volatilities. The results achieved within this analysis have led to certain areas we would like to examine further. First, it would be insightful to investigate the effects of increasing the frequency at which we rebalance the portfolios. More specifically, we would examine how increasing rebalance-frequency impacts the stability of the BSS model. Furthermore, it would be interesting to incorporate transaction costs within the model and to estimate this cost more precisely. In practice we cannot assume the transaction costs to be a flat percentage and must rather account for factors such as the bid/ask spread, buying in round lots and the inability to buy fractional shares.

References

[1] J. Busa. “Solving Quadratic Programming Problem with Linear Constraints Containing Absolute Values”. In: *Acta Electrotechnica et Informatica* 12.3 (2012), pp. 11–18. DOI: 10.2478/v10198-012-0024-4.

[2] D. Bertsimas, A. King, and R. Mazumder. “Best subset selection via a modern optimization lens”. In: *The annals of statistics* (2016), pp. 813–852.

[3] N. Gulyás and I. Kondor. “Portfolio instability and linear constraints”. In: (Jan. 2007).