Mean-Variance Portfolio Optimization

Minjun Shuai (1003782042), Yiwei Zeng (1003612496), Betty Qin (1009336489)

May 2023

Abstract

The objective of our paper is to explore and compare four different factor models and use them to derive the necessary parameters for optimizing our portfolio. Our initial investment budget is $\$100,000$ and our investment universe comprises 20 stocks (n = 20), as listed in Figure 1. We have access to monthly adjusted closing prices spanning from 31–Dec–2005 to 31–Dec–2016, which we use to calculate the monthly returns of our assets. In addition, we have monthly factor returns for eight different factors, including the monthly risk-free rate, as listed in Figure 2. These factors cover the period from 31–Jan–2006 to 31–Dec–2016.

Our methodology involves using the factor models to estimate the expected returns and covariance matrix of the assets, which we will then use as inputs for portfolio optimization. For this project, we have chosen the rigorous mean-variance optimization (MVO) approach as our investment strategy. The goal of this approach is to evaluate and compare the performance of the portfolios generated using the different factor models in out-of-sample analysis.

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

Figure 1: List of Assets

Market ('Mkt_RF')	Size ('SMB')	Value ('HML')	Short-term reversal ('ST_Rev')
Profitability ('RMW')	Investment ('CMA')	Momentum ('Mom')	Long-term reversal ('LT_Rev')

Figure 2: List of Factors

Contents

1 Introduction

We will explore four different factor models: (a) Ordinary Least Square regression on all eight factors (OLS model), (b) Fama–French three-factor model (FF model), (c) Least Absolute Shrinkage and Selection Operator model (LASSO model), and (d) Best Subset Selection model (BSS model).

Our approach involves several steps. First, we will derive the mean and variance-covariance estimates for each of the 20 stocks in our investment universe. Next, we will apply the mean-variance optimization technique to each set of estimates derived from the four different factor models. To demonstrate the accuracy of our models, we will conduct an in-sample analysis. Finally, we will assess the performance of our portfolio over time by conducting an out-of-sample analysis, which will include an evaluation of risk-to-reward ratios and a presentation of the portfolio's value evolution over time.

The report will be structured as follows: In Section 2, we will explain the methodologies used for both the factor models and the portfolio optimization model. The section will be divided into five subsections, with each subsection dedicated to one model. We will provide a detailed description of each model and its implementation.

Subsequently, we will present the outcomes of the four factor models and the portfolio optimization model. In Section 3, we will outline the performance metrics of the n-sample analysis and our out-of-sample analysis. In Section 4, we will present the results of our in-sample analysis and our out-of-sample analysis. Section 5 is the discussion and conclusion.

2 Methodology

Our project involves modeling the monthly excess return of 20 assets with respect to the risk-free rate and computing the covariance between the assets based on this excess return.

To accomplish this, we are provided with 8 different monthly factor returns spanning the period from 31-Dec-2005 to 31-Dec-2016, covering a total of 132 months. We are also provided with the historical data of monthly adjusted closing prices of 20 stocks in that period. The adjusted closing price is preferred over the regular closing price because it provides a more precise indication of a stock's value and performance over time. As a result, investors and analysts often rely on this metric when evaluating a stock's historical performance and projecting future trends. As the correlations among these factors are an important factor in the modeling process, we will also include the factor covariance matrix in our calculations of asset covariance.

In this study, the investment horizon covers a period of one year, spanning from 2012 to 2016. To train the model, we utilized four years of monthly return data preceding the investment period, optimizing the weights based on this historical data. Subsequently, we applied these optimized weights to the investment period to conduct an out-of-sample analysis, with the goal of assessing the performance of the portfolio.

Now, let's delve into the methodologies of the factor models and the portfolio optimization model, providing a detailed description and implementation for each.

2.1 Ordinary Least Square Regression Model

The Ordinary Least Square (OLS) model is expressed in the following form and incorporates the following variables:

$$r_{it} = \alpha_i + \sum_{k=1}^{p} \beta_{ik} f_{kt} + \epsilon_{it} \tag{1}$$

where:

- r_{it} represents the rate of return of i^{th} asset at time t. For $i = 1, \ldots, I$ and $t = 1, \ldots, T$.
- α_i represents the intercept of regression for i^{th} asset.
- β_{ik} represents the k^{th} factor loading/coefficient on i^{th} asset. For $k = 1, \ldots, p$.

- f_{kt} represents the k^{th} factor monthly return at time t.
- T represents the total number of time steps.
- I represents the total number of asset types.
- p represents the total number of factor types.

The objective of the OLS model is to find the estimate of all α_i and β_{ik}. To achieve this, we need to define a series of matrices beforehand. First, we construct a matrix named A as follows:

$$A = [1, f_1, f_2, f_3, \ldots, f_p] = [1, f]_{T\times(p+1)} \tag{2}$$

where $[f_i]$ represents a $T\times 1$ column that contains the i^{th} factor return over T time points. 1 is a $(T\times 1)$ matrix that consists of 1 in every element, and it corresponds to each regression intercept.

Furthermore, we need to define a vector of regression coefficients named B in the following form:

$$B = \begin{bmatrix} \alpha_i \\ V_i \end{bmatrix} = \begin{bmatrix} \alpha_i \\ \beta_{i1} \\ \vdots \\ \beta_{ip} \end{bmatrix} \tag{3}$$

Moreover, we also need to obtain the expected factor return on monthly basis by using geometric mean return:

$$\bar{f} = \left(\prod_{t=1}^{T} 1 + f_{kt} \right)^{\frac{1}{T}} - 1 \tag{4}$$

Our objective is to minimize the sum of $L-2$ norms of the residuals of regression, which can be transformed into the following format:

$$\min_{B_i} ||r_i - XB_i||_2^2 \tag{5}$$

where $||\cdot||$ represents the $L-2$ norm operator.

By basic calculus, the closed form of solution (B_i^*) to the OLS model is:

$$B_i^* = (X^T X)^{-1} X^T r_i \tag{6}$$

Thus, the regression residuals (ϵ_i^*) will be the following form:

$$\epsilon_i^* = r_i - X B_i^* \tag{7}$$

The expected asset return (μ) and asset covariance matrix (Q) take the following form:

$$\mu = \alpha + V^T \bar{f} \tag{8}$$

where $\bar{f}$ represents the vector of excepted monthly factor returns. and

$$Q = V^T F V + D \tag{9}$$

where F represents the factor covariance matrix and D represents the diagonal matrix of residual variance of $\sigma_{\epsilon_i}^2$

2.2 Fama–French Three-Factor Model

The Fama–French three-Factor (FF) model is essentially an OLS regression model that consumes 3 factors to explain the asset returns. It takes the following form:

$$r_{it} = \alpha_i + \beta_{im} f_M + \beta_{is} f_{SMB} + \beta_{iv} f_{HML} + \epsilon_{it} \tag{10}$$

where:

- r_{it} represents rate of return of i^{th} asset at time t.

- α_i represents intercept of regression for i^{th} asset.
- β_{im} represents market factor loading on i^{th} asset.
- f_M represents monthly market factor return premium (risky - risk-free) at time t.
- β_{is} represents size factor loading on i^{th} asset.
- f_{SMB} represents monthly size factor return premium (small Cap-Ex - large Cap-Ex) at time t.
- β_{iv} represents value factor loading on i^{th} asset.
- f_{HML} represents monthly value factor return premium (high value - low value) at time t.
- ϵ_{it} represents residual return of i^{th} asset at time t.

Since the FD factor model and OLS model share some similarities as linear regression models, their implementation and methodologies are close to identical.

2.3 Least Absolute Shrinkage and Selection Operator Model

The Least Absolute Shrinkage and Selection Operator (LASSO) is a regularization technique that is commonly used in statistical modeling. This method involves adding a penalty term to the objective function, which helps to regularize the model's parameters. The penalty term penalizes the addition of non-performing and underperforming variables into the OLS model. Compared to other models, LASSO provides models with high prediction accuracy by reducing variance and minimizing bias. The strength of the regularization is determined by the value of the penalty term, denoted as λ. The formulation of the LASSO model is expressed below:

$$
\begin{aligned}
\text{minimize}_{B_i,\lambda} \quad & ||r_i - XB_i||_2^2 + \lambda ||B_i||_1 \\
\text{subject to:} \quad & y_i \geq B_i \\
& y_i \geq -B_i \\
\text{where:} \quad & y_i = |B_i| \\
& B_i = \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix}
\end{aligned}
\tag{11}
$$

In order to incorporate this equation into the quadprog function, we transformed the minimizing problem into the following form:

$$
\begin{aligned}
\text{minimize} \quad & X^T Q X + c^T X \\
\text{subject to:} \quad & AX \leq b \\
& A_{eq} X = b_{eq} \\
& lb \leq X \leq ub
\end{aligned}
\tag{12}
$$

We started with transforming the objective function into matrix form:

$$
\begin{aligned}
||r_i - XB_i||_2^2 + \lambda ||B_i||_1 &= r_i^T r_i - 2r_i^T X B_i + B_i^T X^T X B_i + \lambda 1^T y_i \\
\text{where:} \quad X &= [1 \quad f]
\end{aligned}
\tag{13}
$$

In order to match the format of the objective function with the formulation derived above, we set the following parameters:

$$Q = X^T X \tag{14}$$

$$C^T = -2r_i^T X + \lambda 1^T \tag{15}$$

Since we need to incorporate the constraint $y_i = |B_i|$ into the minimizing problem, hence we will expand our matrices as follows:

$$B_i^{new} = \begin{bmatrix} B_i \\ y_i \end{bmatrix} \tag{16}$$

$$Q^{new} = \begin{bmatrix} X^T X & 0_{(p+1)\times(p+1)} \\ 0_{(p+1)\times(p+1)} & 0_{(p+1)\times(p+1)} \end{bmatrix} \tag{17}$$

$$X^{new} = \begin{bmatrix} 1_{T\times 1} & f_{T\times p} & 0_{T\times(p+1)} \end{bmatrix} \tag{18}$$

$$1_{new}^T := Z^{new} = \begin{bmatrix} 0_{(p+1)\times 1} \\ 1_{(p+1)\times 1} \end{bmatrix} \tag{19}$$

Then, the objective function can be revised into the following form (since $r_i^T r_i$ is constant, we can remove it from the objective function):

$$\begin{aligned} \text{minimize} \quad & B_i^{new^T} Q^{new} B_i^{new} + C^T B_i^{new} \\ \text{where:} \quad & C^T = -2r_i^T X^{new} + \lambda Z^{new} \end{aligned} \tag{20}$$

For the constraint, we need also to define:

$$\begin{aligned} \text{such that} \quad & B_i - y_i \leq 0 \\ & -B_i - y_i \leq 0 \\ \text{which we can transfer into:} \quad & A_{eq} B_i^{new} = b_{eq} \\ \text{where:} \quad & A_{eq} = \begin{bmatrix} I_{(p+1)\times 1} & -I_{(p+1)\times 1} \\ -I_{(p+1)\times 1} & -I_{(p+1)\times 1} \end{bmatrix} \\ & b_{eq} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}_{2(p+1)\times 1} \end{aligned} \tag{21}$$

To sum up, the finalized version of the minimizing problem is:

$$\begin{aligned} \text{minimize} \quad & B_i^{new^T} Q^{new} B_i^{new} + C^T B_i^{new} \\ \text{where:} \quad & A_{eq} B_i^{new} = b_{eq} \end{aligned} \tag{22}$$

In the later part, we will analyze different λ applied into the equation and how coefficient numbers and values will change.

2.4 Best Subset Selection Model

Best Subset Selection (BSS) is a method that involves considering all possible combinations of independent variables to determine the subset that provides the best prediction of the outcome variable. In this project, we will be using the constrained form of the BSS model, which includes all eight factors as inputs. By utilizing this approach, we aim to identify the most effective subset of variables that can accurately predict the outcome variable.

Consider the constrained form of the BBS model:

$$\begin{aligned} \text{minimize} \quad & ||r_i - XB_i||_2^2 \\ \text{where:} \quad & ||B_t||_0 \leq K \end{aligned} \tag{23}$$

We can transform the minimizing problem above into a mixed-integer quadratic program (MIQP) by incorporating an auxiliary binary variable y:

$$\begin{aligned} \text{minimize} \quad & ||r_i - XB_i||_2^2 \\ \text{where:} \quad & Ly_i \leq B_i \leq Uy_i \\ & 1^T y_i = K \\ & X = [1 \quad f] \end{aligned} \tag{24}$$

The auxiliary binary y is a $(p+1) \times 1$ vector with each element indicating whether its corresponding factor is included in the model or not.

L is the lower bound of coefficient and U is the upper bound. We set bounds for our β. We need to ensure that the magnitudes of both L and U are large enough so that the performing explanatory will not be cut off. Therefore, we set $U = 100$ and $L = -100$.

In order to incorporate this equation into the Gurobi function, we transformed the minimizing problem into the following form:

$$\text{minimize} \quad X^T Q X + C^T X \tag{25}$$

We started with transforming the objective function into matrix form:

$$\begin{aligned} \text{minimize} \quad & ||r_i - XB_i||_2^2 \\ \equiv\text{minimize} \quad & r_i^T r_i - 2r_i^T xB_i + B_i^T X^T XB_i \\ \equiv\text{minimize} \quad & -2r_i^T XB_i + B_i^T X^T XB_i \end{aligned} \tag{26}$$

Thus, we can set:

$$Q = X^T X \tag{27}$$

$$C^T = -2r_i^T X \tag{28}$$

Furthermore, we also need to incorporate the auxiliary binary variable into our final minimizing problem so we will expand our matrix as follows:

$$B_i^{new} = \begin{bmatrix} B_i \\ y_i \end{bmatrix} \tag{29}$$

$$Q^{new} = \begin{bmatrix} X^T X & 0_{(p+1)\times(p+1)} \\ 0_{(p+1)\times(p+1)} & 0_{(p+1)\times(p+1)} \end{bmatrix} \tag{30}$$

$$X^{new} = \begin{bmatrix} 1_{T\times 1} & f_{T\times p} & 0_{T\times(p+1)} \end{bmatrix} \tag{31}$$

Then, the objective function can be revised into the following form (since $r_i^T r_i$ is constant, we can remove it from the objective function):

$$\begin{aligned} \text{minimize} \quad & B_i^{new^T} Q^{new} B_i^{new} + C^T B_i^{new} \\ \text{where:} \quad & C^T = -2r_i^T X^{new} \end{aligned} \tag{32}$$

Consider the first constraint in the original minimizing problem:

$$\begin{aligned} & Ly_i \leq B_i \leq Uy_i \\ \text{which is equivalent to:} & \begin{cases} B_i - Uy_i \leq 0 \\ -B_i - Ly_i \leq 0 \end{cases} \end{aligned} \tag{33}$$

Transform it into matrix form:

$$\begin{aligned} AB_i^{new} &= b \\ \text{where:} \quad A &= \begin{bmatrix} I_{(p+1)\times 1} & -UI_{(p+1)\times 1} \\ -I_{(p+1)\times 1} & LI_{(p+1)\times 1} \end{bmatrix} \\ b &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned} \tag{34}$$

Consider the second constraint in the original minimizing problem:

$$1^T y_i = K \tag{35}$$

Transform it into matrix form:

$$\begin{aligned} A_{eq}B_i^{new} &= b_{eq} \\ \text{where:} \quad A_{eq} &= \begin{bmatrix} 0_{(p+1)\times 1} \\ 1_{(p+1)\times 1} \end{bmatrix} \\ b_{eq} &= K \end{aligned} \tag{36}$$

To sum up, the finalized version of the minimizing problem is:

$$\begin{aligned} \text{minimize} \quad & B_i^{new^T} Q^{new} B_i^{new} + C^T B_i^{new} \\ \text{where:} \quad & A_{eq}B_i^{new} = b_{eq} \end{aligned} \tag{37}$$

The initial model will begin with K=4, and the BSS model will be utilized to determine whether to include the intercept in the model. Various K values will be examined during the latter part.

2.5 Portfolio Optimization Model

Mean-variance portfolio optimization is a widely used methodology in finance for constructing optimal portfolios. It is based on the principle of balancing expected returns with portfolio risk. The process involves estimating the expected returns, variances, and covariances of each asset in the portfolio, and then using these estimates to calculate the optimal portfolio weights. The optimal portfolio is the one that maximizes expected returns while minimizing portfolio risk, subject to constraints on the weights of each asset.

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & x^T \Sigma x \\ \text{subject to} \quad & x^T \mu \geq r \\ & \sum_{i=1}^{n} x_i = 1 \\ & x_i \geq 0, \forall i = 1, 2, \ldots, n \end{aligned}$$

In our project, we use the derived mean and variance matrix from the models above in order to arrive at the optimal weights.With the Q and µ from each factor model as inputs, we use the gurobi in MATLAB to solve the problem.

3 Performance Metrics

3.1 In Sample analysis

R squared:

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} \tag{38}$$

where SS_{res} is the sum of squares of residuals (the difference between the actual and predicted values), and SS_{tot} is the total sum of squares (the difference between the actual values and the mean value).

Adjusted R squared:

$$\text{Adjusted } R^2 = 1 - \frac{(n-1)}{(n-p-1)} \frac{SS_{\text{res}}}{SS_{\text{tot}}} \tag{39}$$

where n is the number of observations, and p is the number of independent variables. The adjusted R squared penalizes the inclusion of irrelevant or redundant variables by adjusting the R squared value based on the number of variables used in the model. SS_{res}:

$$SS_{\text{res}} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{40}$$

where y_i is the actual value of the dependent variable for observation i, and $\hat{y}_i$ is the predicted value of the dependent variable for observation i.

SS_{tot}:

$$SS_{\text{tot}} = \sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{41}$$

where y_i is the actual value of the dependent variable for observation i, and $\bar{y}$ is the mean value of the dependent variable.

3.2 Out of Sample analysis

Average return:

$$\bar{R} = \frac{1}{n}\sum_{i=1}^{n} R_i \tag{42}$$

where $\bar{R}$ is the average return, n is the number of periods, and R_i is the return in period i.

Volatility:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(R_i - \bar{R})^2}{n-1}} \tag{43}$$

where σ is the volatility, $\bar{R}$ is the average return, n is the number of periods, and R_i is the return in period i.

Sharpe ratio:

$$\text{Sharpe ratio} = \frac{R_p - R_f}{\sigma_p} \tag{44}$$

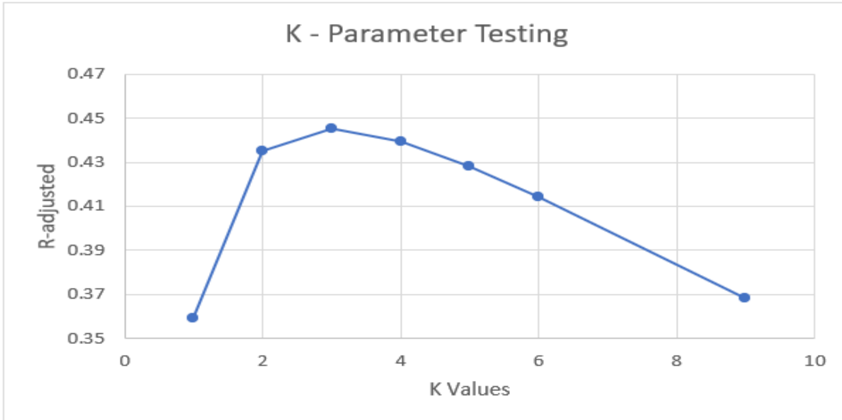
where R_p is the average return of the portfolio, R_f is the risk-free rate, and σ_p is the volatility of the portfolio. The Sharpe ratio measures the excess return earned by a portfolio per unit of volatility, after adjusting for the risk-free rate.

4 Results

4.1 In Sample Analysis

4.1.1 K selection

As shown in Figure 3, the results of this study suggest that when selecting parameters for a Best Subset Selection (BSS) model, an optimal number of independent variables to include is between 1 and 3. Specifically, our analysis shows that the adjusted R^2 increases as the



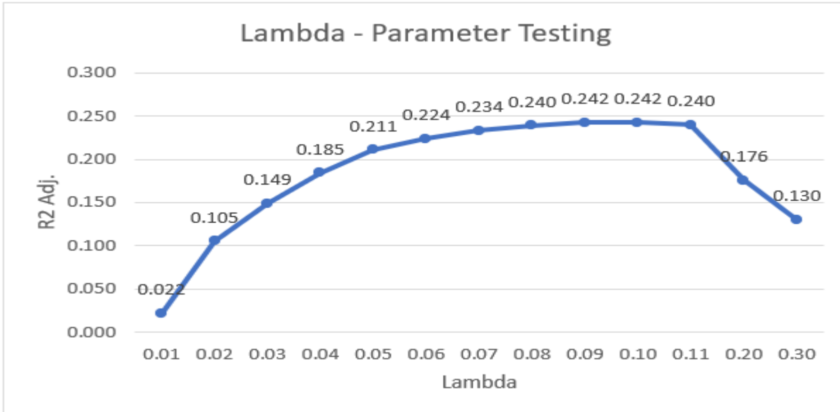
K	R2	R2_adj	Ret	Vol	Sharp
1	0.3725	0.3589	0.0078	0.0272	0.2871
2	0.4593	0.4352	0.0086	0.0263	0.3261
3	0.4808	0.4454	0.0087	0.0265	0.3271
4	0.4872	0.4395	0.0086	0.0265	0.3247
5	0.4888	0.428	0.0087	0.0263	0.3313
6	0.489	0.4143	0.0086	0.0262	0.3294
9	0.489	0.368	0.0087	0.0262	0.3301

Figure 3: K testing with adjusted R^2

number of variables increases from 1 to 3, indicating that the model's ability to explain the variance in the training data improves with the inclusion of additional variables. However, beyond three variables, the adjusted R^2 decreases, suggesting that including more variables does not significantly improve the model's ability to explain the variance in the data. Furthermore, the values for average return and volatility remain consistent, indicating that the inclusion of more variables beyond three does not result in any significant material changes. Therefore, we recommend selecting 3 independent variables when determining the optimal parameters for a BSS model.

4.1.2 λ selection

When selecting the λ parameter for a LASSO factor linear program, it is crucial to strike a balance between model complexity and performance. Setting λ too high can result in an overly simplified model that fails to capture important relationships in the data while setting λ too low can result in overfitting the model to the training data.



lambda	Avg. Num of Variables	R2	R2 Adj.	Avg. Ret	Vol	Sharpe ratio
0.01	5.2	0.1364	0.022	0.0088	0.0345	0.2537
0.02	3.5	0.1804	0.105	0.0089	0.0346	0.2563
0.03	2.8	0.2078	0.149	0.0088	0.0345	0.2547
0.04	2.3	0.2317	0.185	0.0088	0.0346	0.2559
0.05	2.1	0.2515	0.211	0.0087	0.0345	0.2527
0.06	2.0	0.2614	0.224	0.0085	0.0344	0.2466
0.07	1.8	0.2668	0.234	0.0082	0.0343	0.2405
0.08	1.6	0.2697	0.240	0.0081	0.0342	0.2383
0.09	1.5	0.2711	0.242	0.0081	0.0341	0.2366
0.1	1.4	0.2698	0.242	0.008	0.034	0.2354
0.11	1.3	0.2661	0.240	0.008	0.0339	0.2349
0.2	0.9	0.1971	0.176	0.008	0.0328	0.2439
0.3	0.9	0.1505	0.130	0.0083	0.032	0.2589

Figure 4: λ testing with adjusted R^2

As shown in Figure 3, to avoid overfitting, we tested various values of λ, ranging from 0.01 to 0.3, and determined that the optimal value was 0.05. By increasing the value of λ, we increased the penalty for having more non-zero coefficients in the model. This helped to avoid overfitting the model to the training data and improved its generalization performance on new data. At the same time, we needed to be careful not to set λ too high, as this would have resulted in an overly simplified model that fails to capture important relationships in the data.

4.1.3 *R^2 Adjusted*

Years	OLS	FF	LASSO	BSS
2012	57.4%	48.7%	17.7%	49.0%
2013	63.4%	49.9%	16.9%	49.8%
2014	47.6%	41.1%	23.4%	46.2%
2015	42.4%	33.7%	23.6%	41.7%
2016	39.2%	28.7%	23.8%	35.9%
R2 Adj. Avg:	50.0%	40.4%	21.1%	44.5%

Figure 5: Adjusted R Square through 2012 to 2016

OLS (Ordinary Least Squares) Model:

Strengths: The OLS model has the highest adjusted R-squared values, suggesting that it may be better at explaining the variance in asset returns compared to the other models. OLS is a simple and widely used method that is easy to implement and interpret.

Weaknesses: OLS assumes that the relationship between the factors and asset returns is linear and that there is no multicollinearity between the factors. OLS may not perform well if these assumptions are violated.

FF (Fama-French) Model:

Strengths: The FF model has the second-highest adjusted R-squared values and is a widely used method in finance. The FF model accounts for both the market risk and size risk factors, which makes economic sense and can improve its explanatory power.

Weaknesses: The FF model assumes that the relationship between the factors and asset returns is linear and that there is no multicollinearity between the factors. The FF model does not account for all of the possible risk factors that may affect asset returns.

LASSO (Least Absolute Shrinkage and Selection Operator) Model:

Strengths: The LASSO model applies a regularization penalty to the regression coefficients, which can result in a simpler model that performs better on out-of-sample data. LASSO can handle a large number of predictors and is effective at identifying a subset of important predictors.

Weaknesses: The LASSO model has the lowest adjusted R-squared values, suggesting that it may not be as effective at explaining the variance in asset returns compared to the other models. LASSO may exclude some important factors that are captured by the other models, resulting in a simpler but less accurate model.

BSS (Best Subset Selection) Model:

Strengths: The BSS model allows for the selection of the best subset of factors that can explain the variance in asset returns. BSS can handle a large number of predictors and can improve the accuracy of the model by selecting only the most important predictors.

Weaknesses: BSS can be computationally intensive and may not be practical for large datasets. BSS does not account for the interactions between the selected factors, which can limit its explanatory power. BSS may also suffer from overfitting if the number of predictors is too large relative to the sample size, as discussed in the previous K parameter testing.

4.2 Out of Sample Analysis

This section will examine the financial performance of four portfolios, each of which is supported by a factor model discussed earlier. To begin the analysis, we calculated the average monthly return, volatility, and Sharpe Ratio, as depicted in Figure 5.

	OLS	FF	LASSO	BSS
Avg. Return	0.861%	0.877%	0.879%	0.867%
Volatility	0.02648	0.02684	0.02875	0.02652
Sharpe Ratio	0.325	0.327	0.306	0.327

Figure 6: Out of Sample Key Statistics through 2012 to 2016

As per the table presented above, the average monthly returns for all models are centering around 0.87%. The LASSO model has the highest volatility (0.02875), which is significantly higher than the other three models. The OLS and BSS models have the lowest volatility (0.02648 and 0.02652, respectively), while the FF model has slightly higher volatility (0.02684). The difference in average return between the four models is relatively small, with the OLS model having the lowest return (0.861) and the LASSO and FF models having the highest return (0.879). However, the difference in returns may still be meaningful from a portfolio management perspective, especially when compounded over time. The LASSO model's higher volatility results in the lowest Sharpe Ratio among all the models. The OLS, FF, and BSS models all have similar Sharpe ratios (0.325, 0.327, and 0.327, respectively), while the LASSO model has a slightly lower Sharpe ratio (0.306). This suggests that the LASSO model may be taking on more risk relative to its return compared to the other models.

To further analyze the long-term performance of the portfolios, we have plotted the portfolio values for 5-year periods. The result is presented in Figure 6.

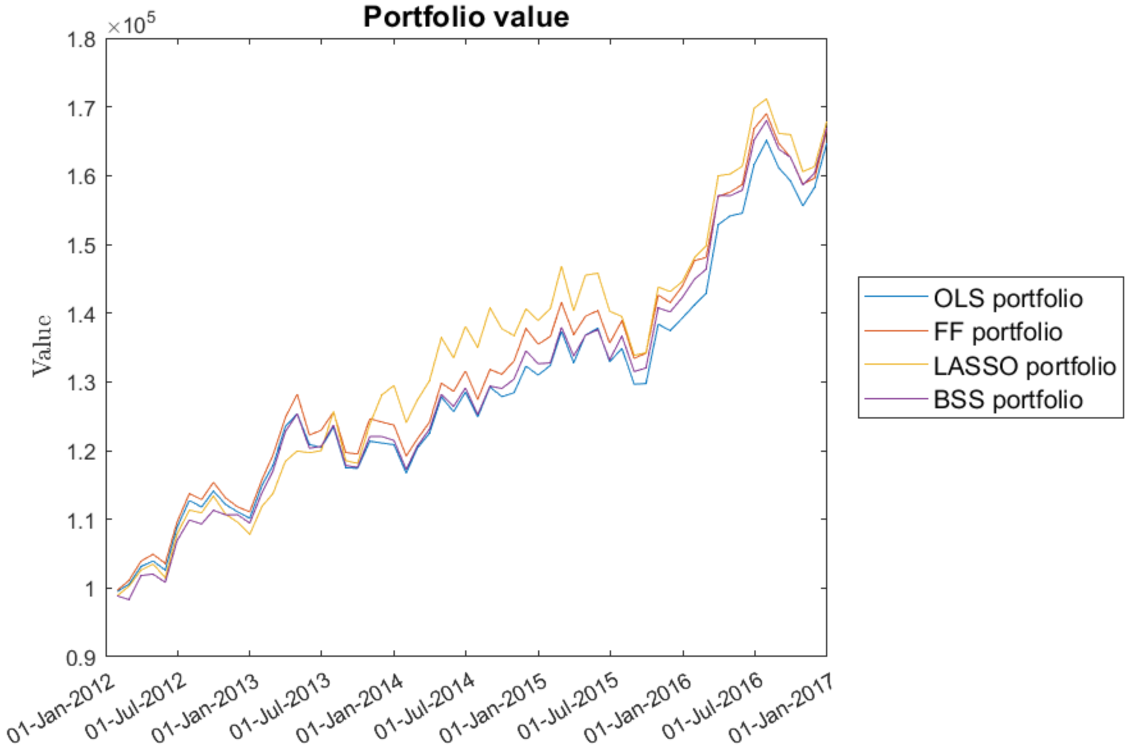


Figure 7: Portfolio Value through 2012 to 2016

Based on the plot presented above, the OLS, FF, and BSS portfolios exhibit comparable trends in the evolution of their portfolio values. However, the LASSO model initially underperforms from the start of 2012 to the start of 2014. Nevertheless, it begins to outperform the other models after that, exhibiting a noticeable increase in portfolio value and maintaining a leading position until the start of 2017. It is important to note that during the period when all four portfolios experience a loss such as the first half of 2013 and the whole year of 2015, the LASSO portfolio experiences a higher percentage of loss relative to the others. Both of the previously discussed findings align with the conclusion drawn from Figure 6, which indicates that the LASSO model has a higher average volatility than the other models. However, it is true that the LASSO model outperforms the other models in terms of wealth evolution, as evidenced by its higher portfolio value over time. This outperformance is likely due to the higher volatility associated with the LASSO model, which creates greater growth opportunities.

In addition to analyzing the portfolio value trend, we also conducted an assessment of the asset weight distribution for each portfolio, which is presented in Figures 8 to 11. This analysis provides further insight into how each model constructs the optimal portfolio.

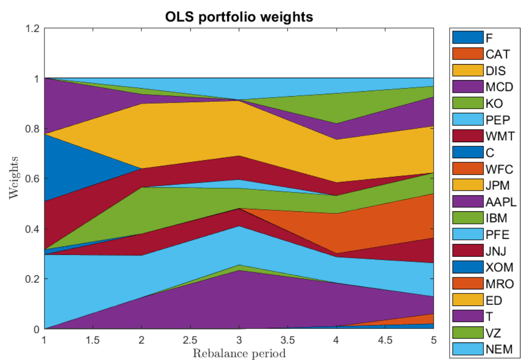


Figure 8: OLS

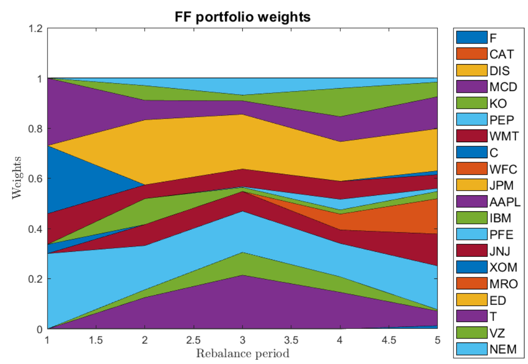


Figure 9: FF

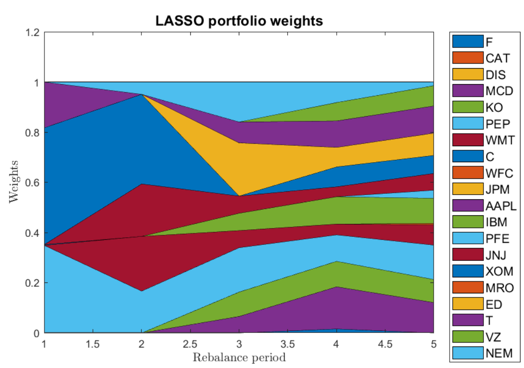


Figure 10: LASSO: $\lambda = 0.05$

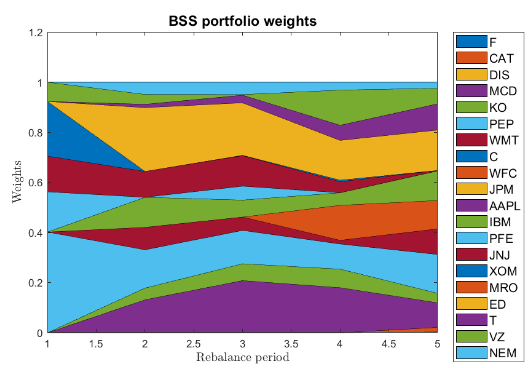


Figure 11: BSS: K = 3

Below are some highlights on stock selection from each model:

1. Stock MCD, PEP, WMT, JNJ, ED, T, NEM and IBM have a consistently high weight in all models and time periods, suggesting they are important stocks in the portfolio with strong performance.

2. Stock F, CAT and WFC are not significantly favored by any of the models until later periods (t=4, t=5), where it sees a slight increase in weight in the OLS and FF model. This may suggest a performance change from those two stocks in the later time period.

3. Stock OIS, JPM, AAPL, MRO and C are not generally not given high weight by any of the models across all time. This may suggest the stocks underperform comparing to other stocks.
4. Stock KO sees a significant weight in the BSS model and some weight in FF and LASSO models, which decrease over time. The OLS model, however, does not give it much weight. There may be an unfavorable correlation of this stock in relation to the rest of the position in the portfolio.
5. Stock XOM has a high weight in the LASSO and BSS models this may be because the correlation with the portfolio is strong.
6. Stock VZ has a variable weight across the models and over time. The weight in the OLS model stays consistent, suggesting it has a stable relationship with the other stocks in the portfolio from this model's perspective. In contrast, the LASSO and BSS models allocate a higher weight in the first time period, which decreases over time, indicating these models may have identified some initially positive feature or performance of this stock that lessened in later periods.

Based on the plots presented above, it appears that the weight distributions for the OLS, FF, and BSS models are relatively comparable across all rebalancing periods. The models may assign higher weights to stocks with higher expected returns. If a stock has shown strong performance during the time periods, a model might allocate a larger weight to it. If two stocks are highly positively correlated, they will tend to move in the same direction. To diversify the risk, the OLS and FF model might not allocate high weights to both. This finding suggests that these models tend to prioritize similar assets and exhibit similar levels of diversification in the portfolio. However, the LASSO model showcases a more rapid and aggressive change in shifting asset weighting. It appears that the LASSO model provides less diversification in the portfolio up until the fourth rebalance time point compared to the other models. This finding is likely due to the LASSO model's tendency to prioritize a smaller subset of assets with higher weightings, leading to a more concentrated portfolio.

5 Discussion and Conclusion

In conclusion, our research has been centered around the evaluation of four distinctive factor models: the OLS, FF, LASSO, and BBS models. The aim was to construct an investment portfolio for each and assess their performance, using both in-sample and out-of-sample performance analyses. Our objective was to discern the relative strengths and weaknesses of each model.

Our findings indicate that the OLS model is the top performer among the four models, demonstrating superior performance in both in-sample and out-of-sample analyses. The OLS model appears to lean towards a balanced portfolio, a conclusion corroborated by several metrics including the highest adjusted R^2, a strong Sharpe Ratio, and consistent asset diversification over time.

The FF model, though a subset of the OLS model with only three explanatory variables, displays commendable performance. Despite a lower adjusted R^2 in the in-sample data compared to the OLS model, the FF model showcases impressive out-of-sample returns and a solid Sharpe Ratio. It also maintains asset diversification over time, as demonstrated in Figure 9. Hence, due to its reduced complexity and fewer factors, the FF model can be considered a viable alternative to the OLS model.

The BBS model, with an adjusted R^2 significantly lower than that of the OLS model, yields results comparable to the FF model when evaluated using in-sample data. Despite this, the BBS model maintains similar levels of risk-adjusted excess returns and asset diversification as the OLS and FF models in out-of-sample analysis.

Contrastingly, the LASSO model underperforms, with the lowest in-sample adjusted R^2, a disappointing out-of-sample Sharpe Ratio, and poor asset diversification. We hypothesize that the BBS and LASSO models' lackluster performance may be due to under-fitting, possibly caused by a lack of sufficient training data.

In summary, while each model has its merits and drawbacks, our analysis reveals the OLS model as the most reliable, with the FF model presenting itself as a feasible alternative.