Factor Models and Portfolio Optimization

MMF1921 – Project 1 Report

Matthew Mamelak, Nicholas Carbonara, Josh Maier, Aubrey Tafadzwa

May 2025

Contents

1 Introduction

The purpose of this project is to investigate and compare four different factor models — the Ordinary Least Squares (OLS) multi-factor model, the Fama–French three-factor model, the Least Absolute Shrinkage and Selection Operator (LASSO) model, and the Best Subset Selection (BSS) model, to estimate expected returns and covariances for a portfolio of 20 U.S. stocks. Using these estimated parameters, we implement a Mean–Variance Optimization (MVO) strategy to construct optimal portfolios and evaluate their performance over a five-year out-of-sample testing period (2012–2016). This report details the methodology, presents computational results, and analyzes both in-sample and out-of-sample performance, ultimately providing insights into the strengths and weaknesses of the various modeling approaches.

2 Methodology

2.1 Factor Models

2.1.1 OLS Model

The Ordinary Least Squares (OLS) model estimates the relationship between asset excess returns and the eight provided factors by fitting a multi-factor linear regression. The regression equation for each asset i is:

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

where:

- r_i is the return of asset i,
- r_f is the risk-free rate,
- f_k is the return of factor k,
- α_i is the intercept,
- β_{ik} is the factor loading of factor k for asset i,
- ε_i is the residual (unexplained part of the return).

In the code, we first set up the data:

- T is the number of months; n is the number of assets.
- p is the number of factors (eight in this case).
- We construct the matrix X, which has an intercept column (ones) plus the factor returns, making it $T \times (p + 1)$.

For each asset:

- We solve for the OLS coefficients:

$$\hat{\beta} = (X^\top X)^{-1} X^\top y,$$

  where y is the vector of excess returns for the asset i.

- We calculate the predicted returns $\hat{y} = X\hat{\beta}$.

- We compute residuals $\varepsilon = y - \hat{y}$.

- The expected return μ_i is set as the mean of $\hat{y}$.

Finally, the residuals across all assets are used to estimate the asset covariance matrix:

$$Q = \text{cov}(\varepsilon),$$

where Q is the $n \times n$ covariance matrix of asset excess returns.

In summary, the OLS model provides:

- μ: the $n \times 1$ vector of expected excess returns.

- Q: the $n \times n$ covariance matrix, used for portfolio optimization.

2.1.2 Fama–French Three-Factor Model

The Fama–French (FF) three-factor model estimates asset excess returns using a reduced set of factors: the market excess return (Mkt–RF), the size factor (SMB), and the value factor (HML). The regression equation for each asset i is:

$$r_i - r_f = \alpha_i + \beta_{i,\text{mkt}}(f_{\text{mkt}} - r_f) + \beta_{i,\text{smb}} f_{\text{smb}} + \beta_{i,\text{hml}} f_{\text{hml}} + \varepsilon_i,$$

where:

- r_i is the return of asset i,

- r_f is the risk-free rate,

- f_{mkt}, f_{smb}, f_{hml} are the three Fama–French factors,

- α_i is the intercept,

- $\beta_{i,*}$ are the factor loadings,

- ε_i is the residual term.

In the code, we proceed as follows:

- We retain only the first three columns of the factor data, corresponding to the FF factors.

- We construct the matrix X by adding an intercept column to the factor matrix, making X of dimension $T \times (k+1)$, where $k = 3$.
- We solve for the OLS coefficients across all assets at once using:

$$\text{coef} = X \backslash R,$$

  where R is the $T \times n$ matrix of asset excess returns.
- We extract:
  - α (intercepts) from the first row,
  - B (factor loadings) from the remaining rows.

Next, we calculate:

- The average factor returns:

$$\bar{f} = \frac{1}{T} \sum_{t=1}^{T} f_t,$$

- The factor covariance matrix Σ_f using maximum likelihood estimation,
- The idiosyncratic variances σ^2 from the residuals.

Finally, the model outputs:

$$\begin{aligned} \mu &= \alpha + B\bar{f}, \\ Q &= B\Sigma_f B^\top + \text{diag}(\sigma^2), \end{aligned}$$

where μ is the $n \times 1$ vector of expected excess returns and Q is the $n \times n$ covariance matrix of asset returns.

This setup allows us to estimate expected returns and covariances under the Fama–French framework, focusing on a simpler and more interpretable factor model compared to the entire eight-factor OLS model.

2.1.3 LASSO Model

The Least Absolute Shrinkage and Selection Operator (LASSO) model extends the multi-factor regression by applying an ℓ_1-norm penalty to encourage sparsity in the factor loadings. For each asset i, we solve the following optimization problem:

$$\min_{\beta_i} \; \|y_i - X\beta_i\|_2^2 + \lambda \|\beta_i\|_1,$$

where:

- y_i is the vector of excess returns for asset i,
- X is the $m \times (p+1)$ design matrix with an intercept and eight factors,

- β_i contains the intercept and factor loadings,
- λ is the regularization parameter controlling the strength of the ℓ_1 penalty.

The ℓ_1-norm penalty encourages some coefficients to shrink to zero, effectively performing factor selection.

In the code:

- For each asset, we reformulate the problem using a positive–negative split:

$$\beta = \beta^+ - \beta^- \quad \text{with} \quad \beta^+, \beta^- \geq 0,$$

  and solve the equivalent quadratic program:

$$\min_{z \geq 0} \frac{1}{2} z^\top H z + f^\top z,$$

  where $z = [\beta^+; \beta^-]$.

- We use MATLAB's `quadprog` to efficiently solve this optimization for each asset.
- After recovering the intercept (α) and factor loadings (B) from the solution, we compute the residual variance σ^2.

Once all assets are processed, we compute:

$$\begin{aligned} \mu &= \alpha + B\,\mu_f, \\ Q &= B\,\Sigma_f\,B^\top + \text{diag}(\sigma^2), \end{aligned}$$

where:

- μ_f is the mean factor return vector,
- Σ_f is the factor return covariance matrix,
- σ^2 is the vector of residual variances.

This setup allows us to generate expected returns (μ) and covariances (Q) for portfolio optimization, with the benefit of automatic factor selection provided by the LASSO penalty.

2.1.4 Best Subset Selection (BSS) Model

The Best Subset Selection (BSS) model aims to find the optimal subset of factors that best explains each asset's excess returns. For each asset i, the BSS optimization problem is formulated as:

$$\min_{\beta_i} \|y_i - X\beta_i\|_2^2 \quad \text{subject to} \quad \|\beta_i\|_0 \leq K,$$

where:

- y_i is the $T \times 1$ vector of excess returns for asset i,

- X is the $T \times (p+1)$ design matrix (intercept + p factors),
- β_i is the coefficient vector (including intercept),
- K is the maximum number of nonzero factor loadings allowed (excluding intercept).

In practice, the code does the following:

- It generates all $\binom{p}{K}$ combinations of K-factor subsets.
- For each subset S, it performs an OLS regression:

  $$\hat{\beta}_S = (X_S^\top X_S)^{-1} X_S^\top y_i,$$

  where X_S includes only the intercept and the selected K factors.
- It computes the residual sum of squares (RSS) for each subset and keeps the subset with the minimum RSS.
- The best subset's coefficients are stored in:
  - α_i — the intercept,
  - B_i — the $1 \times p$ factor loadings, where only the selected factors have nonzero coefficients.

After looping over all n assets, the model computes:

$$\mu = \alpha + B\,\mu_f,$$
$$Q = B\,\Sigma_f\,B^\top + \mathrm{diag}(\sigma^2),$$

where:

- μ_f is the mean factor return vector,
- Σ_f is the factor return covariance matrix,
- σ^2 is the vector of idiosyncratic variances.

This approach allows the BSS model to balance explanatory power with model simplicity by explicitly controlling the number of included factors.

2.2 Implementation Setup

The main program begins by importing weekly adjusted stock price data and factor return data from the provided CSV files. The weekly excess returns for each stock are calculated by taking percentage price differences and subtracting the corresponding weekly risk-free rate. The data is aligned on a common date axis to ensure all asset and factor series are synchronized, dropping any initial rows where necessary to match dimensions.

We set an initial portfolio value of $100,000 and define a rolling window structure:

- **In-sample calibration window**: a four-year window used to calibrate the factor models (e.g., 2008–2011 for the first run).
- **Out-of-sample test window**: the following one year, used to test and track out-of-sample performance (e.g., 2012 for the first run).
- This process repeats five times, rolling forward one year each time, to cover the full out-of-sample period from 2012–2016.

2.2.1 Parameter Tuning

For the LASSO and Best Subset Selection (BSS) models, hyperparameters must be carefully selected:

- **LASSO (λ)**: The program constructs a grid of 40 logarithmically spaced λ values from 10^{-4} to 10^{1}. For each calibration period, it applies four-fold cross-validation: the calibration dataset is split into four equal blocks; in each fold, one block is used as the validation set while the other three blocks are used for training. For each λ candidate, the program fits the LASSO model, computes the average R^2 across all assets on the validation folds, and selects the λ with the highest cross-validated average R^2.
- **BSS (K)**: For each calibration period, the program tests different values of K (the number of nonzero factor coefficients), limited to values 1 through 3. For each K, it runs the BSS model and calculates the mean adjusted R^2 across all assets. The value of K that yields the highest average adjusted R^2 is selected for that period.

The program records the selected optimal λ and K values for each period and stores them for reporting.

2.3 Portfolio Optimization

Once the factor models are calibrated and we have estimates for the expected excess return vector μ and covariance matrix Q, we use a classical Mean–Variance Optimization (MVO) framework to construct the optimal portfolio. The goal is to find the set of portfolio weights x that minimize portfolio variance while achieving at least a specified minimum target excess return.

Formally, the optimization problem is defined as:

$$\begin{aligned} \min_{x} \quad & x^\top Q x \\ \text{subject to} \quad & \mu^\top x \geq \text{targetRet}, \\ & \mathbf{1}^\top x = 1, \\ & x \geq 0, \end{aligned}$$

where:

- x is the $n \times 1$ vector of asset weights,

- Q is the $n \times n$ covariance matrix of asset excess returns,
- μ is the $n \times 1$ vector of expected excess returns,
- targetRet is the target portfolio excess return,
- $\mathbf{1}^\top x = 1$ enforces full investment,
- $x \geq 0$ enforces long-only (no short-selling) positions.

This optimization is a convex quadratic programming problem and is solved in MATLAB using the `quadprog` solver, part of the Optimization Toolbox. Specifically:

- The objective function $x^\top Qx$ is reformulated in the quadratic program's standard form:
$$\frac{1}{2}x^\top Hx + f^\top x,$$
where we set $H = Q + Q^\top$ (to ensure symmetry) and $f = 0$.
- The expected return constraint $\mu^\top x \geq \text{targetRet}$ is rewritten as $-\mu^\top x \leq -\text{targetRet}$.
- The equality constraint $\sum x_i = 1$ ensures the portfolio is fully invested.
- The non-negativity bounds $x \geq 0$ enforce long-only weights.

Importantly, if the optimizer detects that the problem is infeasible (for example, if no portfolio exists that can achieve the specified target return under the given constraints), the program automatically drops the target return constraint and re-solves the problem as a pure minimum-variance optimization:

$$\min_x \quad x^\top Qx \quad \text{subject to} \quad \mathbf{1}^\top x = 1,\ x \geq 0.$$

This fallback ensures the system always produces a feasible, investable portfolio even under tight or stressed market conditions.

2.3.1 Implementation Details

In the MATLAB implementation:

- The function `MVO(mu, Q, targetRet)` takes as inputs the expected return vector, covariance matrix, and target return.
- It constructs the problem matrices H and f and sets up the linear and equality constraints.
- It solves the problem using `quadprog` with the `interior-point-convex` algorithm.
- If the optimizer returns an infeasibility flag, the function issues a warning and re-optimizes without the target return constraint, focusing solely on minimizing portfolio variance while respecting the long-only and fully invested constraints.

This robust optimization setup ensures that each model produces a well-defined and economically interpretable portfolio at every rebalancing point, allowing us to consistently evaluate out-of-sample performance across the five-year test window.

2.4 Rebalancing and Performance Tracking

At the start of each test year:

- The model recalibrates μ and Q using the latest four-year window.
- Optimal portfolio weights x are computed via MVO.
- The weights are converted into share quantities using current asset prices.
- The portfolio is held fixed (buy-and-hold) over the one-year out-of-sample test window.

At the end of each year:

- The program records the evolving portfolio value based on realized out-of-sample returns.
- It stores period-specific performance metrics, including adjusted R^2, optimal λ, optimal K, portfolio weights, and final wealth.

2.5 Evaluation Metrics

For each portfolio, the program computes:

- **In-sample adjusted** R^2: the average across all assets, reflecting the explanatory power of the factor model during calibration.
- **Out-of-sample annualized metrics**: including average return, standard deviation, and Sharpe ratio, derived from the multi-year out-of-sample performance.
- **Visualizations**: including cumulative portfolio value plots over time and area plots showing the evolution of asset weight compositions across rebalance periods.

These metrics and visual outputs enable a thorough comparison of both the statistical fit and economic performance of the different factor models under realistic, rolling rebalancing conditions.

3 Results

3.1 In-Sample Analysis

To assess the explanatory power of the factor models, we evaluated their in-sample fit using the adjusted R^2 metric over five calibration periods (2008–2011, 2009–2012, 2010–2013, 2011–2014, 2012–2015). Adjusted R^2 corrects for model complexity by penalizing the addition of explanatory variables, making it particularly suitable for comparing models with differing numbers of predictors.

Period	**Optimal λ (LASSO)**	**Optimal K (BSS)**
1	0.00106	3
2	0.00079	3
3	0.00018	3
4	0.00059	3
5	0.00033	3

Table 1: Optimal hyperparameters for LASSO and BSS models

3.1.1 Hyperparameter Selection

For the LASSO and Best Subset Selection (BSS) models, we tuned the regularization parameter (λ) and the number of selected factors (K) respectively, by choosing the configuration that maximized the average adjusted R^2 across assets. The optimal values per period were:

These results suggest that across all periods, a sparse specification using three factors (BSS) provided the best balance of fit and simplicity, while the LASSO model consistently favored very small regularization penalties, effectively retaining most factors.

3.1.2 Adjusted R^2 Results

The table below summarizes the average adjusted R^2 across all assets for each model and period:

Period	**OLS**	**Fama–French**	**LASSO**	**BSS**
1	0.4797	0.4358	0.4797	0.4358
2	0.4767	0.3984	0.4767	0.3984
3	0.4364	0.3473	0.4364	0.3473
4	0.3974	0.2800	0.3974	0.2800
5	0.4393	0.3411	0.4393	0.3411

Table 2: Average adjusted R^2 per model and period

3.1.3 Interpretation

Across all periods, the multi-factor OLS and LASSO models achieved the highest adjusted R^2, reflecting their ability to capture a greater share of asset return variation when using the full set of eight factors. The regularization of the LASSO model had a minimal effect due to the low optimal λ values, essentially replicating the OLS performance.

The Fama–French three-factor model and the BSS model (with $K = 3$) consistently produced lower adjusted R^2 values, reflecting their more constrained specifications. Although these models sacrifice explanatory power, they offer interoperability and simplicity, which can be advantageous in some settings.

In particular, adjusted R^2 values decreased over successive periods, likely reflecting increasing market complexity and structural changes over time. This highlights the importance of regularly recalibrating models and not assuming static relationships.

3.2 Out-of-Sample Analysis

We evaluate the out-of-sample financial performance of the four portfolios over the five-year test horizon (2012–2016). Performance is assessed using three key annualized metrics:

- **Annualized Return (AnnReturn)**: the geometric average annualized portfolio return,
- **Annualized Standard Deviation (AnnStdDev)**: the annualized portfolio volatility,
- **Annualized Sharpe Ratio (AnnSharpeRatio)**: the risk-adjusted excess return, computed as:

$$\text{Sharpe Ratio} = \frac{\text{AnnReturn}}{\text{AnnStdDev}}.$$

The table below summarizes the annualized performance metrics:

Portfolio	**AnnReturn**	**AnnStdDev**	**AnnSharpeRatio**
OLS portfolio	12.58%	10.42%	1.14
FF portfolio	9.09%	9.95%	0.88
LASSO portfolio	9.60%	9.86%	0.93
BSS portfolio	9.30%	9.83%	0.91

Table 3: Annualized out-of-sample performance metrics (2012–2016)

3.2.1 Interpretation

Among the models, the OLS portfolio achieves the highest annualized return (12.58%) and the highest Sharpe ratio (1.14), indicating the strongest overall performance in risk-adjusted terms. The LASSO portfolio also performs well, with a Sharpe ratio of 0.93, showing that its sparsity constraints do not meaningfully degrade performance relative to full OLS. In contrast, the Fama–French three-factor (FF) and Best Subset Selection (BSS) portfolios deliver slightly lower annualized returns (around 9.1–9.6%) and lower Sharpe ratios (below 1), reflecting their more constrained factor structures.

The results suggest that models utilizing the full factor set (OLS, LASSO) may better capture exploitable signals over this period, although simpler models like FF and BSS offer easier interpretability and potentially lower estimation risk.

3.2.2 Visualizations

To further explore portfolio dynamics, we provide:

- **Figure 1**: A portfolio value evolution plot showing the cumulative growth of wealth over time for each model.
- **Figures 2–5**: Area plots for the weight compositions of each portfolio (OLS, FF, LASSO, BSS) across the five rebalancing periods.

These visualizations help illustrate both the stability and turnover of the portfolios and reveal how different factor models drive changes in allocations over time.

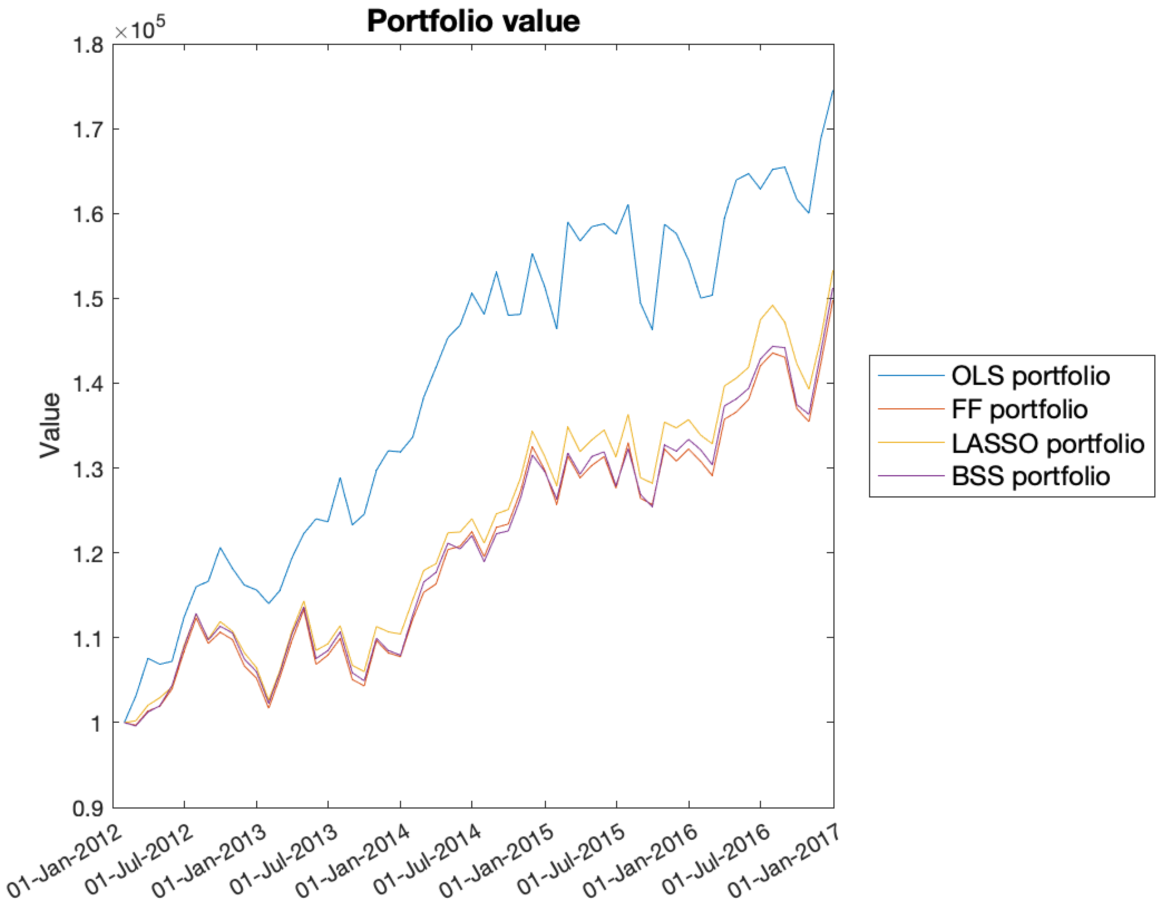

Figure 1: Cumulative portfolio value evolution over 2012–2016

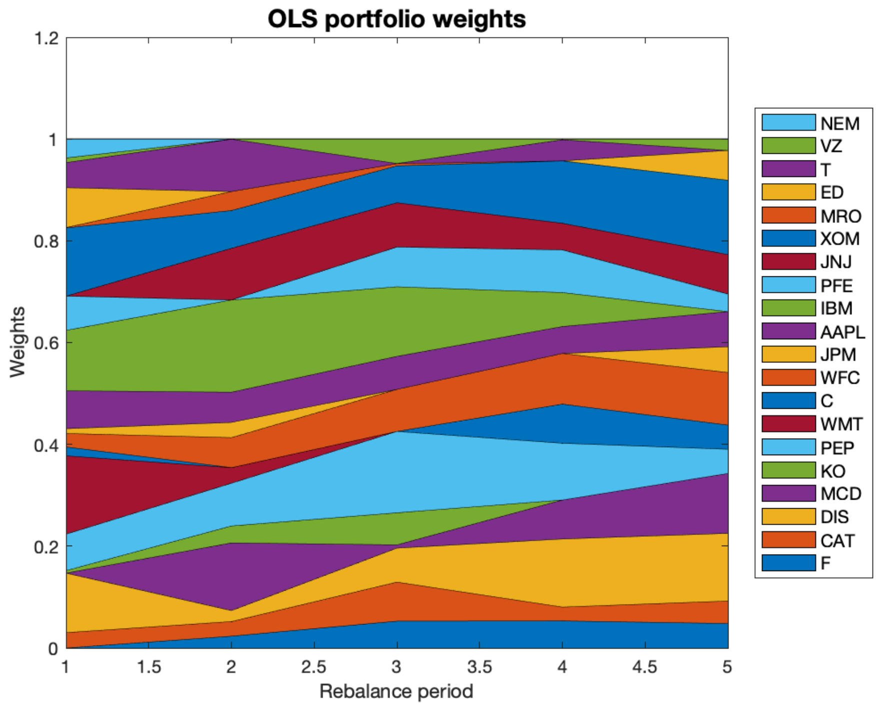

Figure 2: OLS portfolio weight composition across re-balancing periods

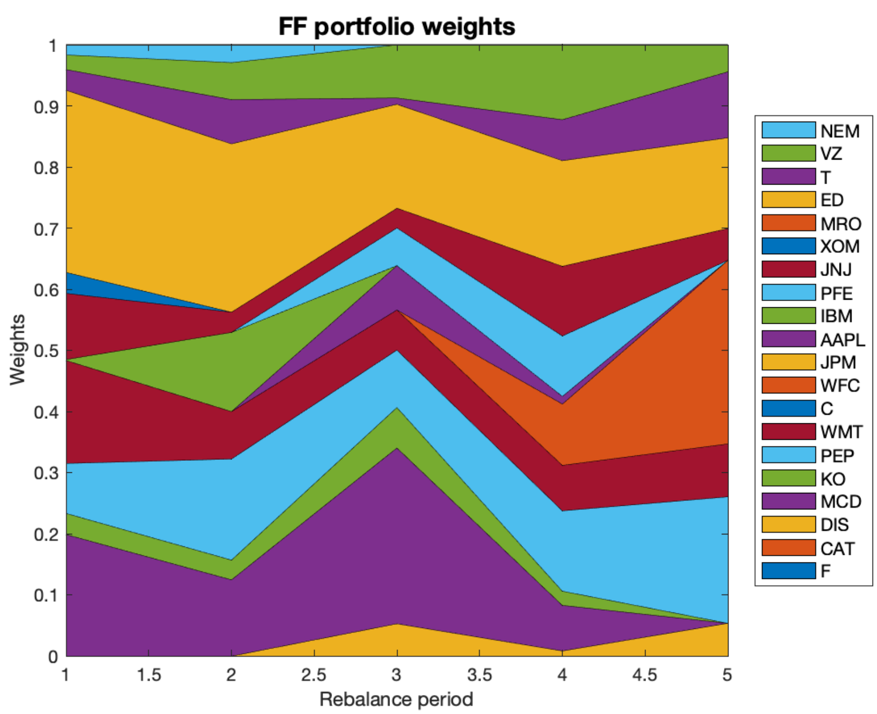

Figure 3: Fama–French (FF) portfolio weight composition across re-balancing periods

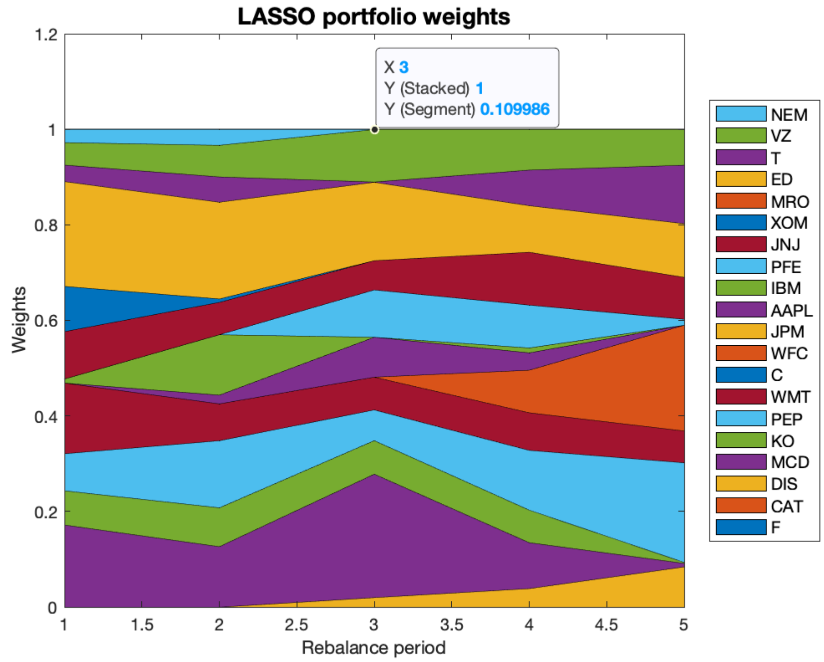

Figure 4: LASSO portfolio weight composition across re-balancing periods

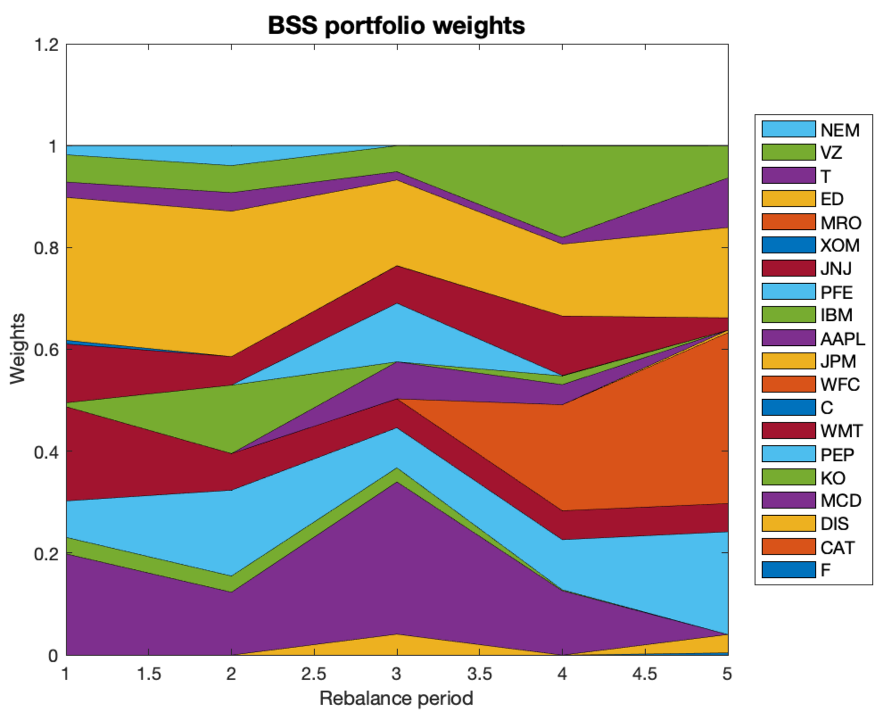

Figure 5: Best Subset Selection (BSS) portfolio weight composition across re-balancing periods

These plots provide insight into which stocks drive performance in each strategy, how diversified or concentrated the portfolios become over time, and how the models respond to changing market conditions across the five-year out-of-sample window.

4 Discussion and Conclusion

This project set out to implement and evaluate four factor modeling approaches—OLS, Fama–French, LASSO, and Best Subset Selection (BSS)—and to apply these models within a dynamic Mean–Variance Optimization (MVO) portfolio framework. By systematically comparing their in-sample fit and out-of-sample economic performance, we aimed to assess both their statistical robustness and practical investment value.

Key Findings

In-Sample Analysis The in-sample adjusted R^2 results revealed that the OLS and LASSO models consistently provided the highest explanatory power across all five calibration periods, reflecting their ability to fully utilize the available factor set. Interestingly, the LASSO's regularization penalty had little material effect, as the optimal λ selected via cross-validation was consistently very small, leading LASSO to effectively mirror OLS performance.

In contrast, the simpler Fama–French three-factor model and the specified BSS model (with $K = 3$) exhibited lower adjusted R^2 values. While these models sacrificed some explanatory power, they offered interpretability advantages: Fama–French relies on well-established factors, and BSS explicitly limits factor inclusion to control overfitting.

Overall, the in-sample analysis confirmed that more complex models fit the data better but raised the classic trade-off question of whether such complexity translates to better out-of-sample performance.

Out-of-Sample Performance The five-year out-of-sample performance analysis (2012–2016) provided striking results. The OLS portfolio achieved the highest annualized return (12.58%) and the best risk-adjusted performance with a Sharpe ratio of 1.14. The LASSO portfolio, despite its penalization structure, delivered similar performance (9.60% return, 0.93 Sharpe), slightly outperforming the simpler FF and BSS portfolios, both of which clustered around 9.1–9.3% returns and Sharpe ratios under 1.

The cumulative portfolio value plots visually reinforced these rankings, with the OLS portfolio pulling ahead over time, especially in later years, while the FF, LASSO, and BSS portfolios tracked each other more closely. The weight composition plots further revealed the differences in model behavior: OLS and LASSO exhibited more dispersed allocations across assets, while FF and BSS portfolios tended to concentrate weights more narrowly due to their reduced factor structures.

Strengths and Weaknesses of the Models

- **OLS**: The strongest performer in terms of both fit and economic outcome, but at the cost of overfitting risk, as it leverages all eight factors without penalization or selection.

- **Fama–French**: Provides a simple, interpretable framework, grounded in economic theory, but underperforms in both fit and returns relative to more flexible models.
- **LASSO**: Offers a theoretically attractive balance between fit and sparsity, but in this case, the optimal λ was so small that it essentially reproduced OLS behavior. This suggests that the data may not have required heavy regularization, or that stronger penalization could have been explored.
- **BSS**: Provides an explicit control over factor inclusion and reduces dimensionality, but its aggressive restriction to three factors likely limited its ability to capture the full breadth of return drivers, resulting in lower performance.

Insights and Lessons Learned

This project demonstrates that statistical fit does not always guarantee superior economic performance, but in this dataset, the better-fitting models (OLS, LASSO) indeed translated into better out-of-sample Sharpe ratios and cumulative returns. The results highlight several key lessons:

- Regularization and factor selection are not universally beneficial; their effectiveness depends on the underlying data-generating process and the signal-to-noise ratio.
- Simpler models, while easier to interpret, may miss important predictive signals, as shown by the underperformance of FF and BSS.
- Portfolio optimization frameworks like MVO are highly sensitive to input estimates, underscoring the importance of careful calibration and robust estimation of μ and Q.

Conclusion

Overall, the combination of flexible factor models and disciplined portfolio optimization can yield strong out-of-sample results, as demonstrated by the OLS and LASSO approaches in this study. Future extensions could explore alternative regularization methods, dynamic factor models, or more advanced optimization approaches (e.g., robust optimization or Bayesian methods) to further enhance portfolio performance.

This project has deepened our understanding of both factor modeling and portfolio construction, providing valuable insights into the practical trade-offs involved in building quantitative investment strategies.