FINANCIAL ANALYTICS

Final Assignment

Professor: I. Vrontos

26/06/2024
MSc in Statistics - AUEB
Grammenos Konstantinos
AM : f3612302
Final Project
Title: Analysis of the US fund data

Contents

List of Figures

List of Tables

Introduction

The objective of this report is to conduct a comprehensive analysis of USA mutual funds' performance and construct optimal portfolios based on various financial models. The analysis is divided into two main parts: performance evaluation and portfolio construction. For the analysis, we used an excel file with US fund data. This file includes data which are monthly returns of USA mutual funds (dependent variables) in the one sheet and several explanatory factors such as S&P500 excess returns, SMB, HML, RMW, CMA, MOM, BAB, and CAR in the other excel sheet. All the data refer to the period from July 1963 to July 2019.

In Part A, the performance evaluation is conducted using data from July 1963 to July 2015 as the initial estimation period, and from August 2015 to July 2019 as the out-of-sample period. Various performance metrics, including the Sharpe ratio, Treynor ratio, Jensen's alpha, and Sortino ratio, are used to assess the performance of equally weighted portfolios formed from the top 20% of selected funds. Jensen's alpha is estimated using several models: the single factor model, multiple regression models, multiple regression with GARCH-type models and more specifically GARCH (1,1) model, and finally by fitting multiple regression with GARCH (2,2) model and also a EGARCH (1,1) model.

Part B focuses on the construction of optimal minimum variance portfolios variance and optimal mean-variance portfolios. Using the same data periods as in Part A, the mean vector and covariance matrix of returns are estimated using: the sample estimate of mean vector and covariance matrix sample estimates, the Single Index Model, multivariate multiple regression models, and using the Constant Conditional Correlation for the variance-covariance matrix of the form. Additionally, we estimated the mean vector and the covariance matrix by fitting a Famma French 3 Factor model and a Famma French 4 Factor model. In the 3 factors models we use the factors "Mkt-RF", "SMB", "HML", while in the 4 factors model we used the same factors but we added also the momentum factor "MOM". We fitted those models in order to explore and may improve the accuracy of the expected returns and the covariance structure. All those constructed portfolios are evaluated based on their realized returns, cumulative returns, and Conditional Sharpe Ratio for the out-of-sample period.
Finally, the report provides a detailed methodology, results, and interpretations of the performance evaluation and portfolio construction, offering insights into the effectiveness of various financial models in analyzing and optimizing mutual fund portfolios.

Data

The analysis begins with loading the dataset US_FUND_DATA.xlsx, which contains monthly returns of USA mutual funds (dependent variables) and several explanatory factors (independent variables) from July 1963 to July 2019. The data is read from two sheets: US_Funds for mutual fund returns and Factors for the explanatory variables. Initially, we read the data from the US_Funds sheet. This dataset contained 679 rows and 251 columns. The first row refers to the number of each fund (1…251) while the first column represents the dates (from 7/31/1963 to 7/31/2019). For our analysis, we kept only the first 90 funds (1…90). Next, we extracted the dates from the first column into a variable called dates and we removed the column. We also observed that our last 6 rows were useless for us with no data so we removed them in order to end up with a dataset called Data, with 673 rows (the months) and 90 columns (the funds), where in each column there is the return of this fund at this specific month. Of course, as you may understand our dataset contains a significant number of missing values (NAs) because different funds start and end at various time periods. After doing all these steps and ensuring that the data is

clean and in the correct format, we convert the returns data into a time series object using the dates variable that we created before.
As regards now the data from the second sheet, named Factors, we loaded the data which was of dimension 673 rows and 9 columns. Also here the first column represents the dates so we also removed the first column and we ended up with a dataset named as factors, consisting of 673 rows (months) and 8 columns (factors), containing monthly returns of each factor. The names of the factors in our dataset are those: “Mkt-RF“, “SMB”, “HML”, “RMW” , “CMA”, “MOM”, “BAB”, “CAR”. More specifically:

1. The first factor represents the excess return of the market over the risk-free rate and it refers as the market factor,

2. the SMB (Small Minus Big) factor measures the performance difference between small-cap and large-cap stocks,

3. HML (High Minus Low) factor represents the difference in returns between high book-to-market (value) stocks and low book-to-market (growth) stocks, the

4. RMW (Robust Minus Weak): this factor measures the difference in returns between companies with robust (high) and weak (low) profitability

5. CMA (Conservative Minus Aggressive): this factor reflects the difference in returns between companies that follow a conservative investment policy and those that follow an aggressive investment policy.

6. MOM (Momentum): this factor measures the tendency of stocks that have performed well in the past to continue performing well in the near future, and vice versa for poorly performing stocks. This factor shows a strategy.

7. BAB (Betting Against Beta): This factor captures the return difference between portfolios with low-beta stocks and high-beta stocks and finally

8. The CAR factor represents the return associated with the carry trade strategy, which involves borrowing in a currency with a low interest rate and investing in a currency with a high interest rate.

After ensuring that also the factors data is clean and in the correct format, we convert the returns them into a time series object. In the factors data there were no Nas.

Part A : Performance Evaluation

Single Factor Model

As we said above, we want to analyze the mutual fund returns using the data of period 7/1963 – 7/2015 as an initial estimation period (In Sample) and evaluate the performance of the USA mutual funds for the Out Of Sample period 8/2015-7/2019 and construct equally weighted portfolios for the out-of-sample period based on the top 20% selected funds. To evaluate the performance of these funds we analyzed their returns using various performance evaluation measures based at first, on the single factor model where the S&P500 return index was used as the market factor. The market factor in our factors dataset is denoted as “Mkt-RF”, so we took a subset of the factors data extracting only the Mkt-RF column from in order to represent the market excess return. Next, we converted the extracted data (the subset) into a time series object aligned with a Time vector representing monthly dates from July 1963 to July 2019. We also divided the returns (Data) and factors (Factors) into in-sample and out-of-sample datasets and we calculated the number of top performing funds, that is the top 20%, so because in our Data there are 90 funds, the top ones are 18. So, first of all, to evaluate them we used the Sharpe measure, which is calculated using this type: $\frac{E(R_i)-r_f}{\sigma_i}$ and is a measure that does not take into account the characteristics of the data and considers that σ_i is constant over time. The top performing 18 funds using the Sharpe's measure are the funds: 39, 42, 35, 87, 48, 49, 14, 22, 64, 50, 72, 59, 71, 85, 33, 70, 9, 47. Next, we evaluated the funds using the Traynor's measure which is calculated as : $\frac{E(R_i)-r_f}{\beta_i}$, where βi is an exposure risk of the fund manager to the market. The top 18 funds derived from this measure are those : 43, 48, 49, 39, 22, 1, 64, 87, 72, 11, 71, 85, 57, 33, 59, 70, 9, 47. Furthermore, we evaluated the US mutual funds based on Jensen's Alpha, which measures the average return of a portfolio or investment above or below that predicted by the Capital Asset Pricing Model (CAPM), given the portfolio's beta and the average market return and it is calculated from this type: $R_{i,t}$-r_f = a_i + b_i* $(R_{m,t} - r_f) + \varepsilon_{i,t}$, where α is a constant that shows the return of fund i that does not depend on the market. It also indicates whether the portfolio manager is superior or inferior in market timing or stock selection, that is if the manager is skilful or not. The top 18 performing funds based on Jensen's Alpha and the Single Index Model are those: 35, 71, 36, 73, 48, 56, 87, 72, 38, 67, 8, 85, 32, 59, 33, 64, 11, 57. Finally, we evaluated our funds using the Sortino measure which is calculated as $\frac{E(R_i)-R_{MAR}}{\delta_i}$, where δ_i is a downside risk and R_{MAR} is a reference value called minimal acceptable return. For each performance measure (Sharpe Ratio, Treynor Ratio, Jensen's Alpha, and Sortino Ratio), we constructed equally weighted portfolios using the top performing funds identified by each measure. Larger values of these measures indicate a better portfolio. In the plot presented below, we display the cumulative returns of the portfolios constructed based on each performance measure. The plot provides a visual comparison of the growth of these portfolios over the out-of-sample period from August 2015 to July 2019. The x-axis represents time, while the y-axis represents the cumulative return. The different colored lines correspond to the cumulative returns of the portfolios constructed using the top funds based on Sharpe Ratio (blue), Treynor Ratio (yellow), Jensen's Alpha (green), and Sortino Ratio (red).

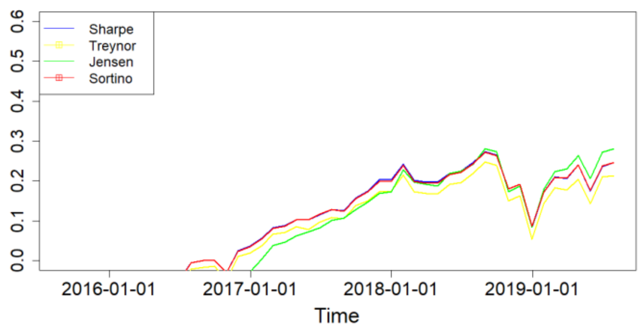

Figure 1: Cumulative Returns of Top Performing Funds

Multiple Regression Model

Beyond the Single Index Model we estimated the Jensen's alpha, based on a multiple regression model analysis, where we considered all the factors in the factors dataset, to evaluate the performance of mutual funds. We also used stepwise regression to select the most significant explanatory variables for each fund. With this analysis, we identified that the top performing funds are those: 52, 32, 14, 33, 89, 47, 38, 26, 42, 58, 85, 72, 63, 64, 48, 56, 11, and 57. Also here, we constructed equally weighted portfolios using these top funds and we plotted the cumulative returns of the this portfolio. This plot visually represents the growth of the portfolio over the out-of-sample period, demonstrating the effectiveness of estimating the Jensen's Alpha using multiple regression models of the form: $Y_t = a + \beta_1 X_{1,t} + \ldots + \beta_k X_{k,t} + \varepsilon_t$ where $\varepsilon_t \sim N(0, \sigma^2)$.

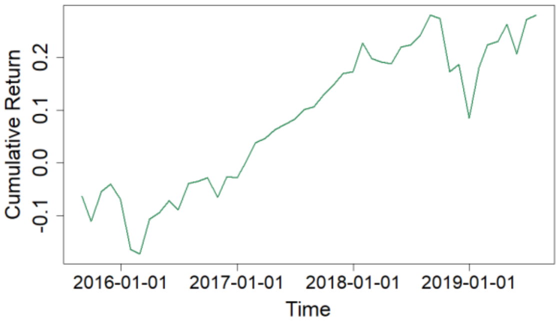

Figure 2: Cumulative Return of Portfolio based on Jensen's Alpha (Multiple Regression Model)

Multiple Regression – GARCH type models

In this section, we estimated the Jensen’s alpha, using multiple regression GARCH type models and more specifically the GARCH (1,1) model of the form:

$$Y_t = a + \beta_1 X_{1,t} + \ldots + \beta_k X_{k,t} + \varepsilon_t$$

where $\varepsilon_t \sim N(0, \sigma_t^2)$ and

$$\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + a_2 \sigma_{t-1}^2,$$

in order to evaluate the performance of mutual funds by incorporating the volatility clustering observed in financial time series. However, before fitting a model like this, we converted the returns data and factors data into matrices and more significantly we performed initial autocorrelation tests, ACF and PACF plots of residuals and of the squared residuals, histogramm and normality tests of the returns. We did this analysis for each of the 90 funds. For example, we can see below the plots of this analysis for the fund 1. From the Box-Ljung autocorrelation test of the residuals, we derived a p-value greater than 0.05, indicating that we do not reject the null hypothesis that the returns are uncorrelated so there is no significant autocorrelation in the returns, but from the results of the same test for the squared residuals we saw that there is significant autocorrelation in the volatility of the returns, indicating a volatility clustering phenomenon as p value was less than 0.05. We can confirm those results by checking the ACF and the PACF plots below. In the ACF and the PACF plot of the residuals all the lines are inside the limits while in the ACF and the PACF plots of the squared residuals we see that in many lags and especially in the first lags, the lines are outside of the limits which means that we reject Ho and that the term is statistically significant. Also both from Shapiro and Jarque Bera test we had the same result of non-normality in the returns. In the QQ-norm plot we see that the plot has fat tails.

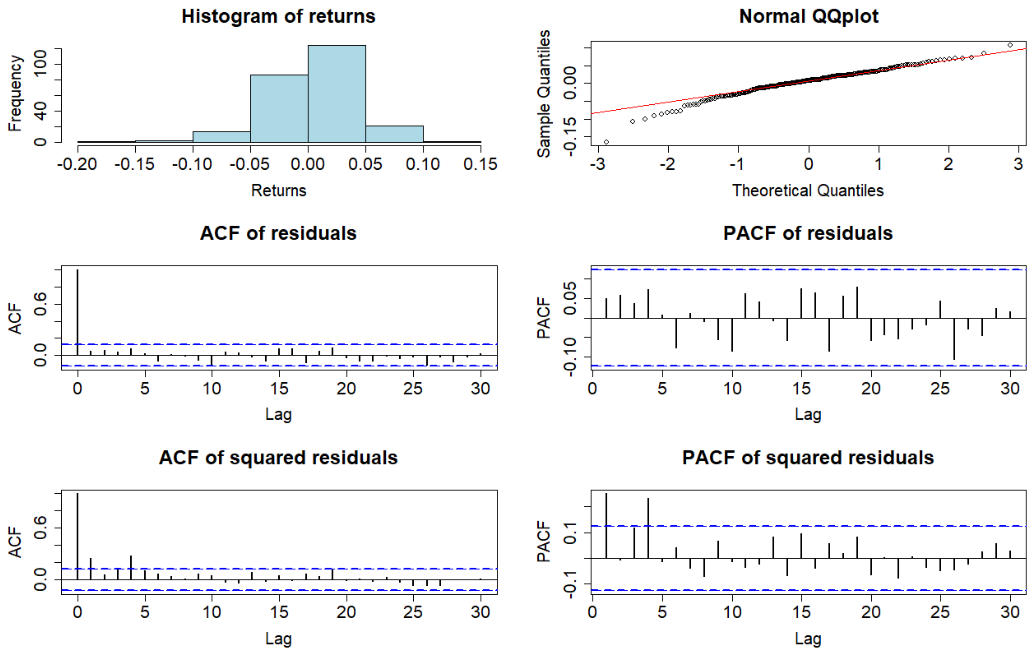

Figure 3: Analysis for Fund 1

Next of this analysis, we estimated the Jensen's Alpha by fitting a GARCH (1,1) model to evaluate the performance of mutual funds and we extracted the alpha (intercept) from each fitted model. Using this model, we identified that the top 18 performing funds are the 71, 42, 51, 12, 89, 46, 85, 11, 29, 26, 63, 32, 57, 31, 70, 50, 67 and the fund 47 and we constructed an equally weighted portfolio using them. We also calculated the mean return and cumulative return of this portfolio for the out-of-sample period. In the plot below we see the Cumulative Returns based on Jensen's Alpha (GARCH (1,1)).

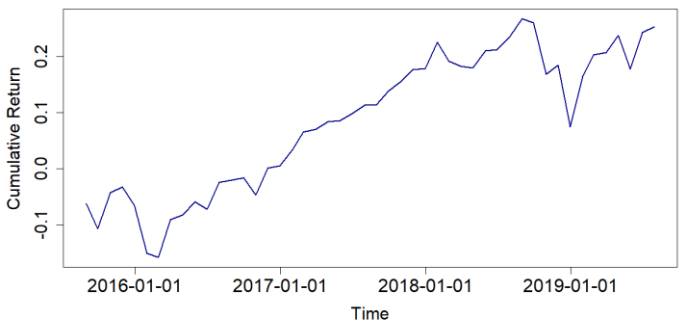

Figure 4: Cumulative Return of Portfolio based on Jensen's Alpha (GARCH (1,1)).

Alternative modelling approach

After testing all these models, we would like to explore two alternative modeling approaches for analyzing mutual fund returns: the GARCH(2,2) model and the EGARCH(1,1) model. These approaches are used to assess the performance of mutual funds based on Jensen's Alpha. The GARCH models capture the volatility clustering but do not capture the leverage effect while the EGARCH models capture both of them.
First, we estimated the Jensen's Alpha by fitting the GARCH(2,2) model which is an extension of the basic GARCH(1,1) model, allowing for more flexibility in capturing volatility dynamics. The model includes both past returns and past volatilities, with two lag terms for each. We specified a GARCH(2,2) model for each fund's returns, incorporating multiple factors as external regressors in the mean equation and we extracted the alpha coefficients. Next, we ranked the funds based on the extracted alphas, and the top 20% of

funds were selected. As a result we derived that the top 18 funds are the funds: 42, 29, 77, 71, 51, 31, 46, 85, 26, 32, 11, 57, 63, 67, 70, 47, 50 and the fund 89. Also here we constructed an equally weighted portfolio with these top performing funds. Finally, we generated a cumulative return plot to visualize the performance of the portfolio based on Jensen's Alpha using the GARCH(2,2) model.

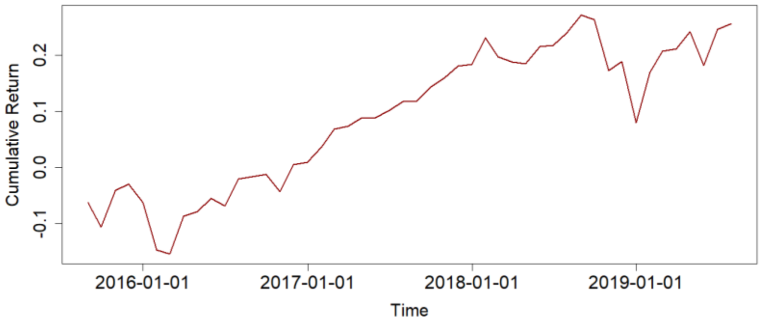

Figure 5:Cumulative Return of Portfolio based on Jensen's Alpha (GARCH(2,2))

The second approach was the EGARCH(1,1) model, or Exponential GARCH model, which allows for asymmetry in the volatility process. As we said, it captures the leverage effect, where negative returns increase future volatility more than positive returns of the same magnitude. Again we followed the same process and based on this model we found out that the top 18 are the funds 26, 46, 16, 42, 63, 86, 85, 32, 12, 27, 67, 31, 11, 70, 57, 89, 47 and the fund 9 and we construct an equally weighted portfolio where its cumulative return we can see in the plot below.

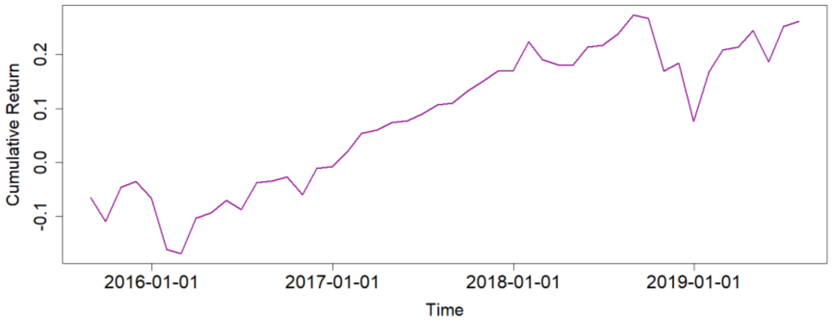

Figure 6:Cumulative Return of Portfolio based on Jensen's Alpha (EGARCH(1,1))

This comprehensive analysis helped us in understanding the dynamics of mutual fund returns and provides insights into better portfolio management strategies.

Comparison between different models

To compare the performance of the above fitted models used to construct and evaluate the returns of mutual fund portfolios we did 2 insightful plots. The models included in this comparison are the Single Factor model, Multiple Regression model, GARCH(1,1), GARCH(2,2), and EGARCH(1,1). The comparison focuses on two key metrics: cumulative returns (1^{st} plot) and Jensen's Alpha measure (2^{nd} plot).

The first plot displays the cumulative returns of portfolios constructed using various financial models over the out-of-sample period from August 2015 to July 2019. The models compared include GARCH(1,1), GARCH(2,2), EGARCH(1,1), Multiple Regression, and the Single Index Model. The x-axis represents the time period, while the y-axis shows the cumulative return of each portfolio. As we see all models show an upward trend in cumulative returns, indicating positive portfolio performance over the period with the GARCH(1,1) and Multiple Regression models to exhibit the highest cumulative returns towards the end of the period. Finally, the Single Index Model, represented by the yellow line, generally shows slightly lower performance compared to the other models but follows a similar trend.

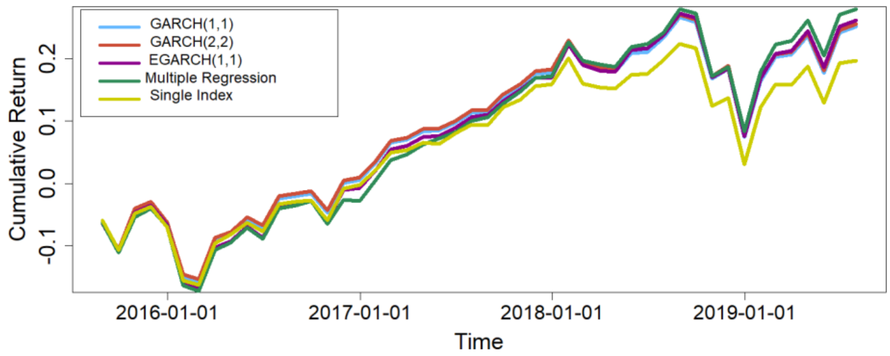

Figure 7: Cumulative Returns of the different models

The second plot is a bar plot comparing Jensen's Alpha for the top 18 funds across the five different models. Each bar represents a fund, and the height of the bar corresponds to the Jensen's Alpha value. We see that the EGARCH, GARCH(1,1) and GARCH(2,2) models consistently produces high Jensen's Alpha values for all the funds, indicating superior risk-adjusted performance. The Single Index Model and the Multiple Regression Model often results in lower or even negative Jensen's Alpha values, highlighting their less robust performance compared to the other models. Finally, fund 8 stands out with a notably high Jensen's Alpha under the EGARCH model, suggesting exceptional performance for this particular fund.

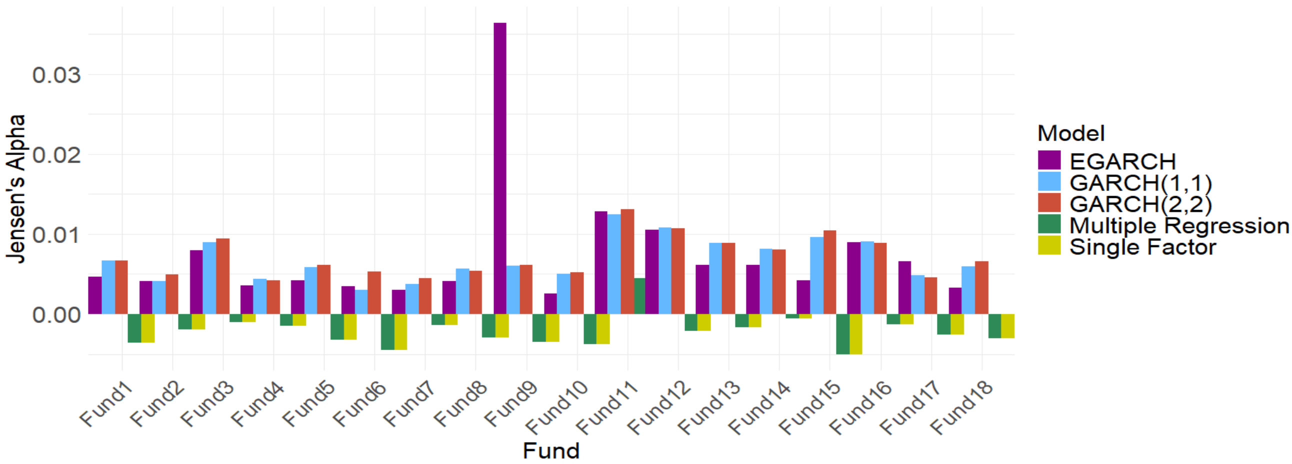

Figure 8: Comparison of Jensen's Alpha Across Models (Top 18 Funds)

Part B : Portfolio Construction

In Part B of this assignment, the primary goal was to construct optimal minimum variance portfolios of the form :

$$min_w \frac{1}{2} \mathrm{V}R_{p,t} = \frac{1}{2} w' \Sigma_t w, \text{ where } w_i \geq 0 \text{ and } \sum_{i=1}^{n} w_i = 1$$

And mean variance portfolios of the form:

$$min_w \frac{1}{2} \mathrm{V}R_{p,t} = \frac{1}{2} w' \Sigma_t w, \text{ with } w_i \geq 0,\ \sum_{i=1}^{n} w_i = 1$$

and $\mathrm{E}(R_{p,t}) \geq r_{target}$, where r_{target} is a target return on a monthly basis.

The analysis was conducted using the same initial estimation period as in the part A and that is from July of 1963 up to July 2015, and the portfolios were constructed for the out-of-sample period from August 2015 to July 2019. The mean vector and the covariance matrix of returns were estimated using a sample estimate of mean vector and covariance matrix, based on the single index mode, based on multiple regression models based on models with Constant Conditional Correlation for the variance-covariance matrix and finally based on an alternative modelling approach which is the Famma French Models.The performance of the constructed portfolios was evaluated based on realized returns, cumulative returns, and the Conditional Sharpe Ratio.

Sample Estimate of Mean Vector and Covariance Matrix

First of all, we would like to construct optimal portfolios based on the sample estimates of the mean vector and the covariance matrix of returns. The very first problem that we faced in this section, was the correct calculation of the covariance matrix of the returns. As we have said also above our dataset contained plenty of NA values so using the default function in R in order to compute the covariance matrix did not work. So, to address the issue of missing values in the data, a custom covariance function, was implemented. This function is designed to compute the covariance matrix while appropriately handling the NA values. By handling NA values effectively, we ensure that the covariance matrix is accurately estimated, which is critical in order to proceed in the construction of the optimal portfolios. After setting up this function, we extracted the In Sample data of returns from the overall dataset (Data) to calculate the sample means and covariances. We calculated the mean vector (m_vec) by taking the average of each fund's returns over the in-sample period and by ignoring NA values. Next, we defined a target return of 0.01 based on the fact that the maximum return in the mean vector was 0.0127. Then, we started our process to construct optimal mean variance portfolios, so we set up the matrices required for the quadratic programming (QP) problem, that is the D.mat, the d.vec vector, A.mat and the b.vec vector and we used the solve.QP function to solve this problem and finding the optimal weights. After finding those, we estimated the expected portfolio return which was 0.0099 almost the same as our target return, and we also calculated the expected risk of the optimal mean variance portfolio and that was about 0.34. Finally, because we want to evaluate the performance of the constructed portfolio in the out-of-sample period, we first computed the Out Of Sample covariance matrix and then we calculated the Out-of-Sample return and risk. Of course, we computed also the realized returns, the cumulative returns, and the Conditional Sharpe Ratio over the out-of-sample period as those will be our evaluation and performance metrics.

We followed the same process to create a minimum variance optimal portfolio. Now, we aimed to find a portfolio that minimizes risk (standard deviation) without explicitly targeting a specific return. This approach contrasts with the mean-variance portfolio, where we aimed to balance both return and risk. Again, we solved the quadratic programming problem with the only difference being the A.mat matrix which includes only the constraints of the sum of the portfolio weights should be equal to 1 and the Non-Negativity Constraint, that each weight should be non-negative and not the constraint of the target return. Here, the In Sample Portfolio return was much lower than in the mean var optimal portfolio as it is only 0.00486, but also the Portfolio risk is much lower and it is about 0.109. We derived the same key out-of-sample metrics to consider and evaluate both mean variance and min variance optimal portfolios. In the table below we can see all those measures for both portfolio construction mechanisms.

Metric	Mean Variance Portfolios	Min Variance Portfolios
Mean Realized Return	0.0042	0.0054
Mean Portfolio risk	0.344	0.112
Conditional Sharpe Ratio	0.0122	0.048

Table 1: Out Of Sample Evaluation Metrics of Mean-Var and Min-Var Optimal Portfolios

To evaluate the optimal portfolios generated from each method we plotted the cumulative Portfolio returns over the Out Of Sample period from 8/2015-7/2019. In the plot below, we observe that both portfolios demonstrate positive cumulative returns over the period but the Minimum Variance Portfolio (red line) generally achieves higher cumulative returns compared to the Mean-Variance Portfolio (blue line). This suggests that the Minimum Variance Portfolio construction method provided better performance during the observed period and that's why it seems to be the better option between the two.

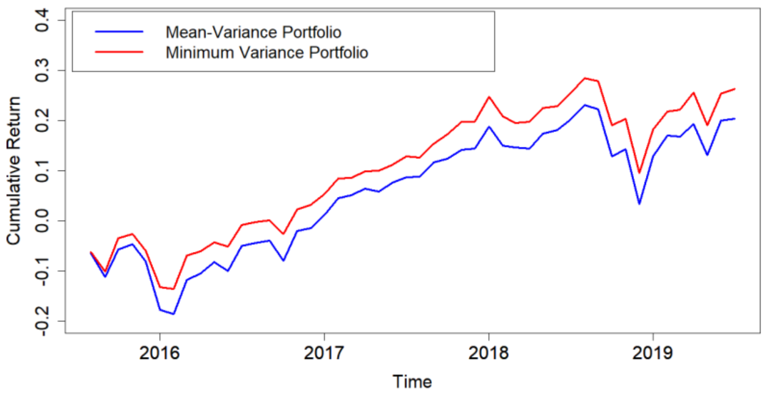

Figure 9: Cumulative portfolio returns over the Out Of Sample period

Single Index Model

The next step in our analysis involves the application of the Single Index Model of the form:

$$R_{i,t} = a_i + b_i R_{m,t} + \varepsilon_{i,t} \,,\ \varepsilon_{i,t} \sim N(0, \sigma_{i,\varepsilon}^2) \text{ όπου i=1,…,n and t=1,…,T}$$

for the estimation of the mean vector and the covariance matrix of returns. This model considers the relationship between each mutual fund's return and a common market factor. First we construct Mean-Variance Portfolios using this model. The process is described below.

For each fund, we fitted the Single Index Model with the market factor and we estimated the parameters of the model, that is the alpha, beta, and the residual variance. These parameters allow us to compute the mean return and covariance matrix of the funds using the Single Index Model approach. Specifically, the mean return of each fund is calculated

by adjusting its alpha with the product of its beta and the mean market return, while the covariance matrix is constructed considering both the market factor's variance and the residual variances. Once the mean vector and covariance matrix are estimated, we proceed to construct the mean-variance portfolio. This involves solving again a quadratic programming problem to determine the optimal portfolio weights that maximize the return for a given level of risk. We also set the target return at 0.009, because now the maximum value of the mean vector is 0.0102. Then, we calculated the in-sample performance metrics, including the expected portfolio return and risk. The In Sample Portfolio return of the Mean Variance optimal portfolios is about 0.0090, while the In Sample Portfolio risk is about 0.0537. As regards the Out Of Sample performance of this portfolio, first we estimated the out-of-sample covariance matrix using the Single Index Model and then we computed the realized return, the Out Of Sample Portfolio risk and the Condiotal Sharpe Ratio.
We also constructed Minimum Variance Portfolios using the Single Index Model. We solved a similar quadratic programming problem as we did for the mean-variance portfolio, but with different constraints. Again as we have already said, the key difference here is that the target return constraint is removed. So, to achieve this, we set up the optimization problem with the same covariance matrix estimated from the Single Index Model, so we do not need to calculate it again and solved for the portfolio weights that minimize the overall portfolio variance. When the optimal weights were determined, we calculated the In Sample Portfolio return and Risk which were 0.0024 and 0.0313 respectively. For the out-of-sample evaluation, we calculated the same performance metrics as with the mean-variance portfolio. We can see all these metrics in the table below.

Metric	Mean Variance Portfolios	Min Variance Portfolios
Mean Realized Return	0.00254	0.00502
Mean Portfolio risk	0.0473	0.02734
Conditional Sharpe Ratio	0.053	0.183

Table 2: Out Of Sample Evaluation Metrics of Mean-Var and Min-Var Optimal Portfolios based on Single Index Model

We also made the same plot of the cumulative Portfolio returns over the Out Of Sample period from 8/2015-7/2019, but based on the Single Index Model, this time. Notably, the Minimum Variance Portfolio consistently outperforms the Mean-Variance Portfolio, demonstrating higher cumulative returns throughout the entire period. This indicates the effectiveness of the Minimum Variance Portfolio strategy in capturing returns while maintaining lower risk in comparison to the Mean-Variance Portfolio, under the Single Index Model framework.

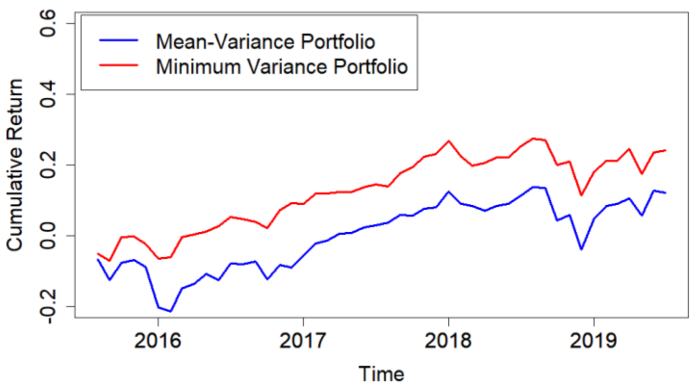

Figure 10: Cumulative portfolio returns over the Out Of Sample period based on the Single Index Model

Multivariate Multiple Regression Models

In this section, the mean vector and the covariance matrix of returns are estimated based on multivariate multiple regression models of the form:

$$R_t = X_t \Gamma' + E_t \,, \qquad E_t \sim N(0, \Sigma) \,, \qquad t = 1, \dots, T$$

For each fund, we employed a multivariate multiple regression model to capture the relationship between fund returns and all the factors in the factors dataset. During this process we firstly removed the NA values of this fund and then we fitted the multivariate multiple regression model. Next, we estimated the alphas and betas and the residual variances. These coefficients and variances are then used to compute the mean return vector and the covariance matrix of the fund. After obtaining the mean vector and covariance matrix, we proceed to construct the Mean-Variance Portfolio and we also set up a target return of 0.01 and thus we found out an In Sample portfolio return of 0.01 and In Sample portfolio risk about 0.056.
Next step is to construct Minimum Variance Portfolios using the multivariate multiple regression model. We solve a quadratic programming problem similar to the mean-variance portfolio but without the target return constraint. This results in the optimal portfolio weights that minimize the overall portfolio variance. The in-sample performance metrics are calculated, with the In-Sample portfolio return and risk being 0.002 and 0.0336, respectively. For the out-of-sample evaluation for both construction methods, we first estimate the out-of-sample covariance matrix using the multivariate regression model. Finally, we compute the realized return, the Out-of-Sample Portfolio Risk, and the Conditional Sharpe Ratio. The results are shown in the table below.

Metric	Mean Variance Portfolios	Min Variance Portfolios
Mean Realized Return	0.000599	0.005210
Mean Portfolio risk	0.045	0.029
Conditional Sharpe Ratio	0.013	0.178

Table 3:Out Of Sample Evaluation Metrics of Mean-Var and Min-Var Optimal Portfolios based on Multivariate Multiple Regression

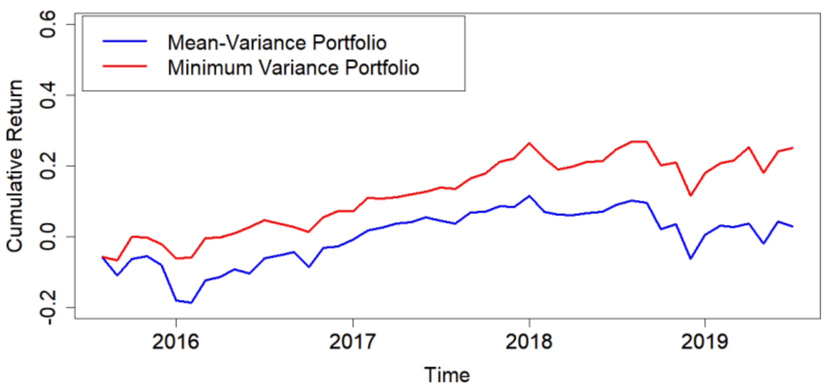

Figure 11: Cumulative portfolio returns over the Out Of Sample period based on Multivariate Multiple Regression

In the table 3, we observe that the Minimum Variance Portfolio achieves a higher return compared to the Mean-Variance Portfolio and also a lower risk. Minimum Variance

Portfolio shows a significantly better risk-adjusted performance.
As regards the figure 11, the Cumulative Portfolio returns over the Out Of Sample period based on Multivariate Multiple Regression, the Minimum Variance Portfolio consistently outperforms the Mean-Variance Portfolio, demonstrating superior cumulative returns throughout the Out Of Sample period 8/2015-7/2019.

Constant Conditional Correlation (CCC) model for the variance-covariance matrix

The next step in our analysis applies the Constant Conditional Correlation (CCC) model to estimate the mean vector and the covariance matrix of returns. This approach modifies the covariance matrix to incorporate a constant correlation structure between the funds. The CCC model assumes that while the variances of the individual funds may change over time, the correlation between them remains constant. This is expressed mathematically as:

$$\Sigma_t = D_t \, \mathrm{R} \, D_t$$

where D_t is a diagonal matrix of the time-varying standard deviations of the funds, and R is the constant correlation matrix.
To construct the Mean-Variance Portfolios, we first calculate the diagonal matrix D_t using the square root of the diagonal elements of the covariance matrix estimated from the multivariate multiple regression model. Then, we estimate the constant correlation matrix R using the residuals from the regression models and with these components, we construct the new covariance matrix of the above form.
We then solve the quadratic programming problem to determine the optimal portfolio weights that maximize return for a given level of risk. The in-sample performance metrics, including the expected portfolio return and risk, are calculated based on the CCC model and are 0.01 and 0.038 respectively. The target return is set at 0.01. For the out-of-sample evaluation, we again the same metrics as in every other model.
Similarly, for the Minimum Variance Portfolios, we use the same covariance matrix that we estimated from the CCC model. The optimization problem is solved to find the portfolio weights that minimize the overall portfolio variance and we compute once more the in-sample and out-of-sample performance metrics for these portfolios as well. The In Sample return is 0.0037 while the In Sample risk is 0.005. Below we see the table of the Out Of Sample evaluation metrics and the plot of the Cumulative portfolio returns.

Metric	Mean Variance Portfolios	Min Variance Portfolios
Mean Realized Return	0.00035	0.00527
Mean Portfolio risk	0.0387	0.00502
Conditional Sharpe Ratio	0.0090	1.049

Table 4:Out Of Sample Evaluation Metrics of Mean-Var and Min-Var Optimal Portfolios based on CCC

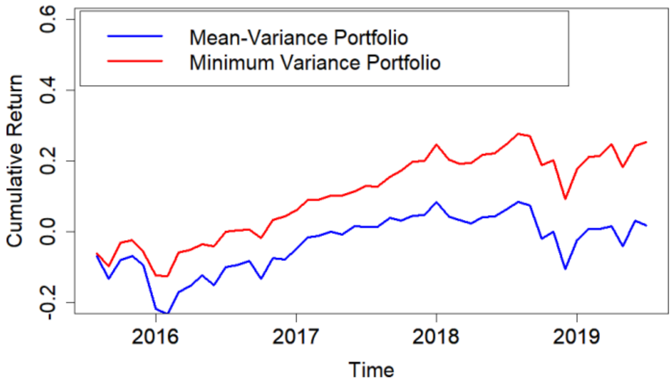

Figure 12:Cumulative portfolio returns over the Out Of Sample period based on CCC

Alternative modeling approach- Fama-French models

Finally we propose an alternative modeling approach for estimating the expected returns and the covariance structure of the analyzed series. Our proposal is the Fama-French model and in order to check its efficiency we did not one but 2 Fama-French models with different numbers of factors each time.

Initially we fitted the Fama-French three-factor model which extends the single index model by incorporating two additional factors: "SMB" and "HML". This model captures more dimensions of risk and return, providing a more comprehensive framework for portfolio construction. To construct the Mean-Variance Portfolios, we fit the Fama-French three-factor model for each mutual fund. This involves regressing each fund's returns on the market factor, the size factor (SMB), and the value factor (HML). The estimated alphas and betas and the residual variances are used to compute the mean return and covariance matrix of the funds. Furthermore, the covariance matrix is constructed considering both the factor covariances and the residual variances. Finally the target return was set up as 0.007, so the In Sample return was 0.007 and the risk 0.0393.

Similarly, for the Minimum Variance Portfolios, we use the same covariance matrix estimated from the Fama-French three-factor model. The optimization problem is solved to find the portfolio weights that minimize the overall portfolio variance. We calculate the In-Sample and Out-Of-Sample performance metrics for these portfolios as well. The In Sample return here is 0.0027 and the risk is 0.0312.

Metric	Mean Variance Portfolios	Min Variance Portfolios
Mean Realized Return	0.00278	0.00502
Mean Portfolio risk	0.0372	0.0286
Conditional Sharpe Ratio	0.0747	0.1756

Table 5:Out Of Sample Evaluation Metrics of Mean-Var and Min-Var Optimal Portfolios based on Fama-French 3 Factor

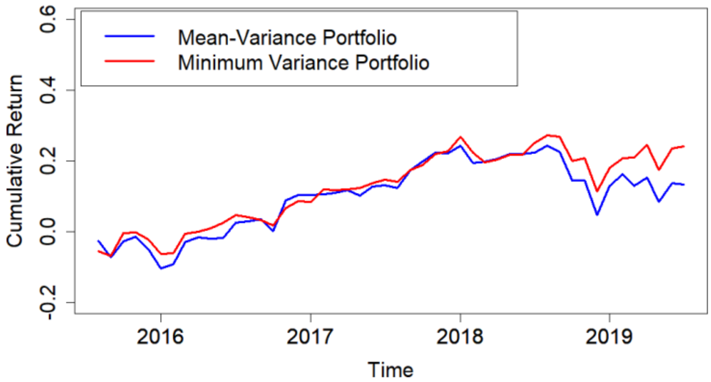

Figure 13:Cumulative portfolio returns over the Out Of Sample period based on Fama-French 3 Factor

At figure 13 we observe that the cumulative returns from the two portfolios are pretty close together and only at the end of the Out Of Sample period we see that the Min Variance Portfolio has a bit higher return.

Additionally, we propose the Fama-French four-factor model as an extension to the three-factor model. This model incorporates an additional momentum factor (MOM), which is a momentum factor. We followed the same process as in the Fama-French 3 factor and we derived an In Sample portfolio return of 0.009 and a risk of 0.0635 at the Mean Variance method and 0.0029 and a risk of about 0.0314 at the Min Variance. Lastly, in the table and the figure below, we can see the Out Of Sample metrics and the cumulative returns of the portfolios.

Metric	Mean Variance Portfolios	Min Variance Portfolios
Mean Realized Return	0.0015	0.00499
Mean Portfolio risk	0.053	0.0286
Conditional Sharpe Ratio	0.0297	0.1744

Table 6:Out Of Sample Evaluation Metrics of Mean-Var and Min-Var Optimal Portfolios based on Fama-French 4 Factor

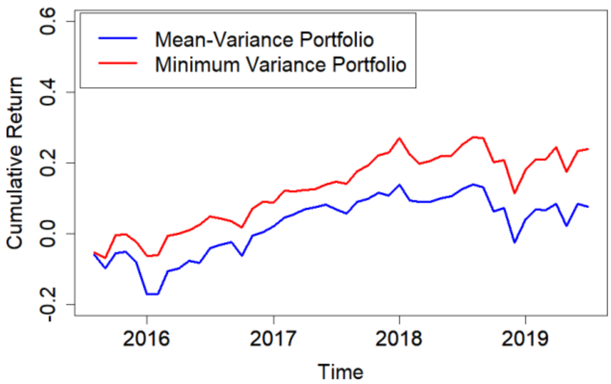

Figure 14:Cumulative portfolio returns over the Out Of Sample period based on Fama-French 4 Factor

Conclusions and further research

This report aimed to evaluate the performance of various financial models in constructing optimal portfolios using US mutual fund data. The analysis utilized a range of models, including the Single Index Model, Multivariate Multiple Regression Models, GARCH models, Constant Conditional Correlation (CCC) model, and Fama-French Three-Factor and Four-Factor Models.

In Part A, the sample estimate method provided a baseline for comparison against more sophisticated models. The Single Index Model, Multiple Regression Model, and GARCH-type models demonstrated improved performance over the sample estimates, highlighting the benefits of accounting for additional factors and volatility dynamics. Among these, the GARCH(1,1) and Multiple Regression models exhibited the highest cumulative returns, while the Single Index Model generally showed slightly lower performance. Jensen's Alpha analysis across different models indicated that the EGARCH, GARCH(1,1) and GARCH(2,2) models consistently produced high Jensen's Alpha values, underscoring its superior risk-adjusted performance.

In Part B, the Sample Estimate of Mean Vector and Covariance Matrix approach highlighted the limitations of relying solely on historical data without incorporating more dynamic modelling techniques. The Minimum Variance Portfolio constructed using this method achieved higher cumulative returns compared to the Mean-Variance Portfolio, suggesting its effectiveness in minimizing risk. The Single Index Model demonstrated that the Minimum Variance Portfolio outperformed the Mean-Variance Portfolio in terms of cumulative returns and lower risk during the out-of-sample period, indicating the effectiveness of the Minimum Variance Portfolio strategy in capturing returns while maintaining lower risk. The Multivariate Multiple Regression Models approach provided a comprehensive understanding of the relationships between fund returns and multiple factors. The Minimum Variance Portfolios consistently showed lower risk and higher cumulative returns compared to the Mean-Variance Portfolios, demonstrating superior risk-adjusted performance. Incorporating a CCC structure for the variance-covariance matrix improved the estimation accuracy, and portfolios constructed using this model demonstrated better risk-adjusted performance, as indicated by higher Conditional Sharpe Ratios. The Fama-French Three-Factor Model provided a robust framework for capturing different dimensions of risk, with the Minimum Variance Portfolio exhibiting superior performance in terms of realized return and Conditional Sharpe Ratio. Adding the momentum factor (MOM) in the Fama-French Four-Factor Model further enhanced the model's capability to explain returns. Similar to the three-factor model, the Minimum Variance Portfolio demonstrated better performance.

Overall, the analysis showed that more sophisticated models that incorporate multiple factors, volatility dynamics, and constant conditional correlation structures provide better risk-adjusted returns compared to simpler models based on sample estimates.

Now as regards the future research someone can explore several avenues to build on the findings of this report especially in the part of the construction of Mean Variance and Min Variance portfolios. A choice can be to investigate Fama-French models with other or more factors like BAB or RMW and CAR factors, in order to provide more insights and improve the accuracy of return and risk predictions. Also, someone could explore non-linear models, machine learning techniques, or hybrid models that combine traditional econometric approaches with advanced algorithms could enhance predictive performance or even extend the Out-Of-Sample evaluation period that may provide a more robust assessment of the models' predictive capabilities and their performance over different market cycles.

References

[1] Ioannis Vrontos, Financial Analytics Lecture Notes, Department of Statistics, Athens University of Economics and Business, Athens, Feb, 2024.

[2] Hamilton, James D. Time Series Analysis. Princeton, New Jersey: Princeton University Press,1994.

[3] Enders, Walter. Applied Econometric Time Series. New York: Wiley, 2010.

[4] Tsay, Ruey S. (2010). Analysis of Financial Time Series, New York: Wiley.

[5] Cowpertwait, Paul S.P., and Metcalfe V. Andrew. Introductory Time Series with R. New York:Springer Texts in Statistics, 2009.

[6] Cryer, Jonathan D., and Chan Kung-Sik. Time Series Analysis with Applications in R. Springer Texts in Statistics, 2010.

[7] Elton, E.J., Gruber, M.J., Brown, S.J., and Goetzmann W.N. (2014). Modern Portfolio Theory and Investment Analysis, 9th edition, Wiley.

[8] Sharpe, W.F., Alexander, G.J., and Bailey, J.V. (1999). Investments, 6th edition, Prentice-Hall.

[9] Bodie, Zvi, Alex Kane, and Alan J. Marcus. Essentials of Investments, 12th Edition. New York: McGraw Hill, 2022.