Financial Analytics

kostas botsi

June 2024

1 Introduction

The purpose of this analysis is to identify the top 20% of US mutual funds and construct equal-weighted portfolios from these funds. We will evaluate US fund returns over the period from July 1963 to July 2019, focusing on the monthly returns of these funds.

Our goal is to distinguish fund managers with strong performance skills from those with weaker skills based on their returns. To aid in identifying the top 20% funds, we will use the following independent variables:

- S&P 500 excess returns
- SMB (Small Minus Big)
- HML (High Minus Low)
- RMW (Robust Minus Weak)
- CMA (Conservative Minus Aggressive)
- MOM (Momentum)
- BAB (Betting Against Beta)
- CAR (Carry)

These factors will be used to explain the variations in the monthly returns of the US funds over the specified period.

1.1 Methodology

The methodology we will follow to analyze the returns of mutual funds is as follows:

1. **Initial Estimation Period**: We will analyze mutual fund returns using data from July 1963 to July 2015.

2. **Out-of-Sample Period**: We will evaluate the performance of US mutual funds for the period from August 2015 to July 2019, covering the last four years.

3. **Portfolio Construction**: For the out-of-sample period, we will construct equal-weighted portfolios based on the top 20% of the selected mutual funds.

This approach allows us to assess the predictive power of our models and the robustness of our findings over different time periods.

2 Performance evaluation measures

In this chapter, we will present performance measures that will help us build equal-weighted portfolios consisting of the best 20% of mutual funds. These performance measures will enable us to identify and select the top-performing funds based on their monthly returns over the evaluation period.

2.1 Sharpe ratio

The first performance measure is the Sharpe ratio,where larger Si values indicate a better portfolio

$$S_i = \frac{E(R_i) - r_f}{\sigma_i} = \frac{\bar{R}_i - r_f}{\sigma_i}$$

Here,

- S_i is the Sharpe ratio for asset i.
- $E(R_i)$ denotes the expected return of asset i.
- r_f is the risk-free rate.
- σ_i represents the standard deviation of asset i's returns.
- $\bar{R}_i$ denotes the average return of asset i.

Following are the top-performing funds based on the Sharpe Ratio where we invest in these funds during the next out-of-sample period. It's important to note that there will be no rearrangement in the portfolio. The selection of top funds may vary in each period based on their performance metrics.

16	33	23	32	5	40	1	72	85
65	22	2	71	24	30	42	8	15

Top 20% Best Performing Funds based on Sharpe Ratio

In the plot we present the equally weighted portfolio's base on the sharper ratio top 20% funds.

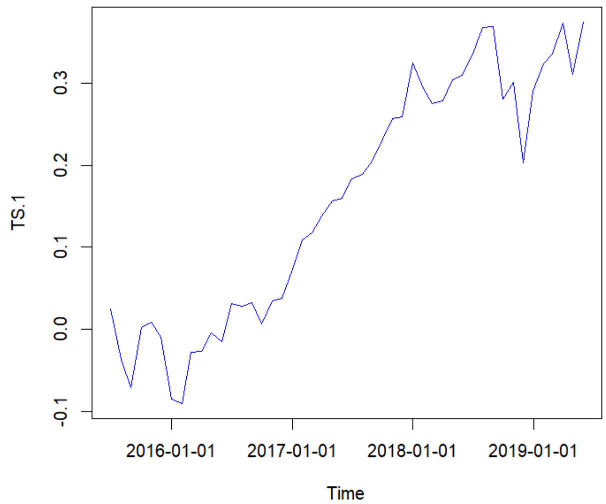

Figure 1: Equally weighted portfolios based on sharpe ratio

From the graph we observe that the portfolio exhibits an upward trend and demonstrates strong performance during the out-of-sample period.

2.2 TREYNOR RATIO

Next, we use the Treynor performance measure, which takes into account the expected return as well as the systematic risk β_i, to identify the top 20% of funds. Below, we present the Treynor ratio:

$$T_i = \frac{E(R_i) - R_f}{\beta_i}$$

Where:

- T_i is the Treynor ratio for fund i,
- $E(R_i)$ is the expected return of fund i,
- R_f is the risk-free rate,
- β_i is the systematic risk (beta) of fund i.

42	33	4	14	59	75	68	74	13
8	40	81	31	1	57	11	41	5

Top 20% Best Performing Funds based on treynor Ratio

This are the top 20% Funds base on treynor ratio that we will construct equally weighted portofolios for the out of sample period.The following graph shows this

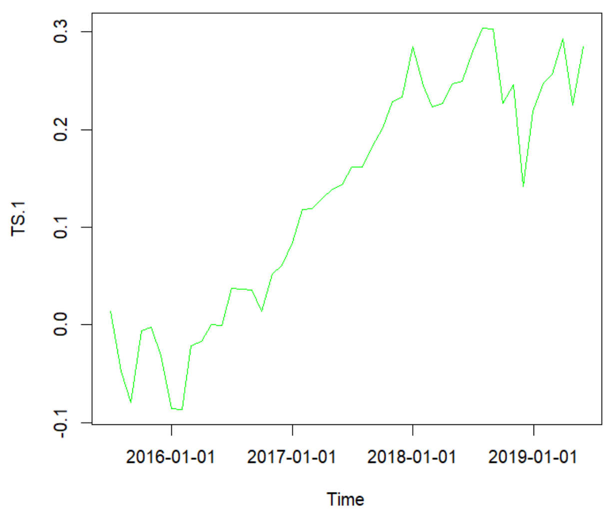

Figure 2: Equally weighted portfolios based on Treynor ratio

The plot is similar with the Sharpe ratio construction portofolio,we see also an upward trend, around 2018-31-01 and after we see a variability in the portofolio.

2.3 Sortino performance measure

The Sortino ratio (SoR) is computed as follows:

$$\text{SoR}_i = \frac{E(R_i) - \text{RMAR}}{\delta_i}$$

where:

- $E(R_i)$ is the expected return of fund i,
- RMAR is the reference value (e.g., mean return, risk-free rate, or zero),

- $\delta_i = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\min(0, R_{i,t} - \text{RMAR})^2}$ represents the downside risk measure.

A higher value of SoR_i indicates better portfolio performance.

33	22	57	32	24	30	1	72	8
16	64	23	71	5	40	42	15	85

Top 20% Best Performing Funds based on sortino Ratio

2.4 Jensen performance measure - Single factor model

Jensen (1968) performance measure is based on the Capital Asset Pricing Model (CAPM). This model assumes that a fund's investment behavior can be approximated by using a single market index:

$$R_{i,t} - r_f = \alpha_i + \beta_i(R_{M,t} - r_f) + \epsilon_{i,t}, \quad \epsilon_{i,t} \sim N(0, \sigma_i^2) \tag{1}$$

- $R_{i,t}$ is the return of investment i at time t
- $R_{M,t}$ is the market return at time t
- r_f is the risk-free rate
- α_i is the model constant; α_i values indicate whether the portfolio manager is superior or inferior in market timing and/or stock selection. Positive α_i indicates a superior or skilled manager, while negative α_i indicates an inferior fund manager.

69	29	72	24	50	40	86	7	5
42	8	77	1	30	85	15	2	79

Top 20% Best Performing Funds based on Jensen sigle model

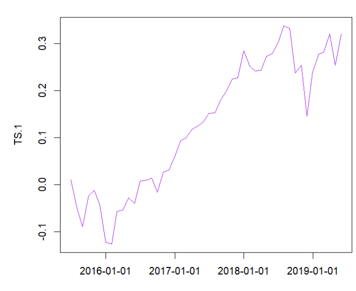

Figure 3: Equally weighted portfolios based on sigle index model

2.5 Jensen performance measure - Three factor model

The three-factor model (Fama and French, 1993) is given by the following equation:

$$R_{i,t} - r_f = \alpha_i + \beta_{i,1}(R_{M,t} - r_f) + \beta_{i,2}\mathrm{SMB}_t + \beta_{i,3}\mathrm{HML}_t + \epsilon_{i,t} \tag{2}$$

$R_{i,t}$ is the return of investment i at time t

$R_{M,t}$ is the market return at time t

SMB$_t$ is the difference between small stocks portfolio returns and large stock portfolio returns

HML$_t$ is the difference between stocks portfolio returns with high book-to-market ratios and low book-to-market ratios

Positive α_i indicates a superior or skilled manager, while negative α_i indicates an inferior fund manager.

$$\begin{bmatrix} 7 & 16 & 22 & 5 & 69 & 24 & 23 & 25 & 40 \\ 72 & 1 & 30 & 42 & 85 & 50 & 15 & 2 & 79 \end{bmatrix}$$

Top 20% Best Performing Funds based on Jensen three factor model

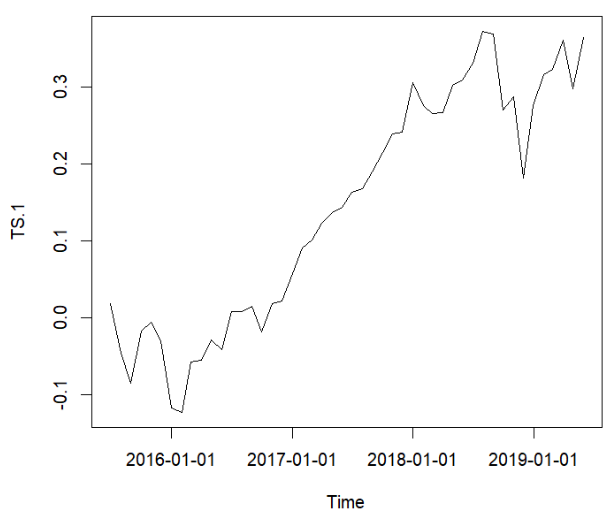

Figure 4: perfroming funds base on Jensen measure

In the plot we see the comulative portofolio returns base on top 20% funds,selected by each performing measure

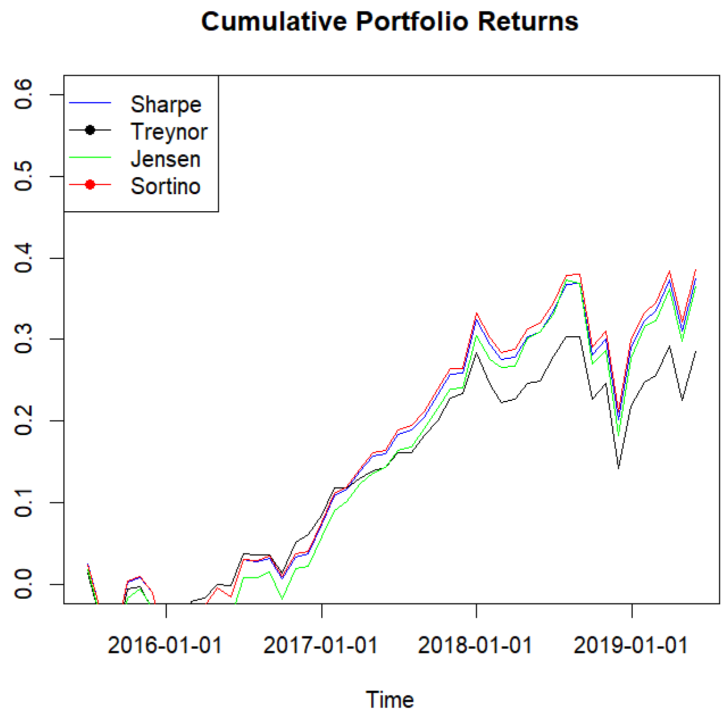

Figure 5: Top perfroming funds

2.6 Multiple regression models

In this section i present a multiple regrassion model for each fund.As it follows

$$\begin{aligned} R_{i,t} = & \alpha_i + \beta_{i,1}(SMB)_{i,t} + \beta_{i,2}(HML)_{i,t} + \beta_{i,3}\mathrm{HML}_t \\ & + \beta_{i,4}(RMW)_{i,t} + \beta_{i,5}(CMA)_{i,t} + \beta_{i,6}(MOM)_{i,t} + \beta_{i,7}(BAB)_{i,t} \\ & + \beta_{i,8}(CAR)_{i,t} + \epsilon_{i,t} \end{aligned} \tag{3}$$

$R_{i,t}$ is the expected return of investment fund i at time t

SMB$_t$ is the difference between small stocks portfolio returns and large stock portfolio returns

HML$_t$ is the difference between stocks portfolio returns with high book-to-market ratios and low book-to-market ratios

$\mathbf{RMW}_t$ is the Robust Minus Weak

$\mathbf{CMA}_t$ is the onservative Minus Aggressive

$\mathbf{MOM}_t$ is the Momentum

$\mathbf{BAB}_t$ is the Betting Against Beta

$\mathbf{CAR}_t$ is the Carry

Positive α_i indicates a superior or skilled manager, while negative α_i indicates an inferior fund manager.

This model can also writen in this way

$$R_{i,t} - rf = \alpha_i + \sum_{k=1}^{K} \beta_{i,k} f_{k,t} + \epsilon_{i,t} \tag{4}$$

$r_{i,t}$**:** Return of investment i at time t.

r_f**:** Risk-free rate.

α_i**:** Intercept term (alpha) specific to investment i.

$\beta_{i,k}$**:** Coefficients representing the sensitivity of investment i's return to the k-th factor $f_{k,t}$.

$f_{k,t}$**:** k-th factor at time t.

$\epsilon_{i,t}$**:** Error or residual term for investment i at time t.

To have a accurate prediction for the expected return for each fund we need to consider some assumption.

- Errors Normally Distributed
- Errors to be Homoskedastic
- Errors Don't Have Autocorrelation Problem

For example we see in the plot that the multiple regression model for the fund2 have autocorellation problem because it exceeds the decision boundary.

To fix the autocorrelation problem we need to put AR models or MA models,it depends if have autocorellation or partial autocorelation problem.Or if we see both problems then we combine this models for ARMA models.

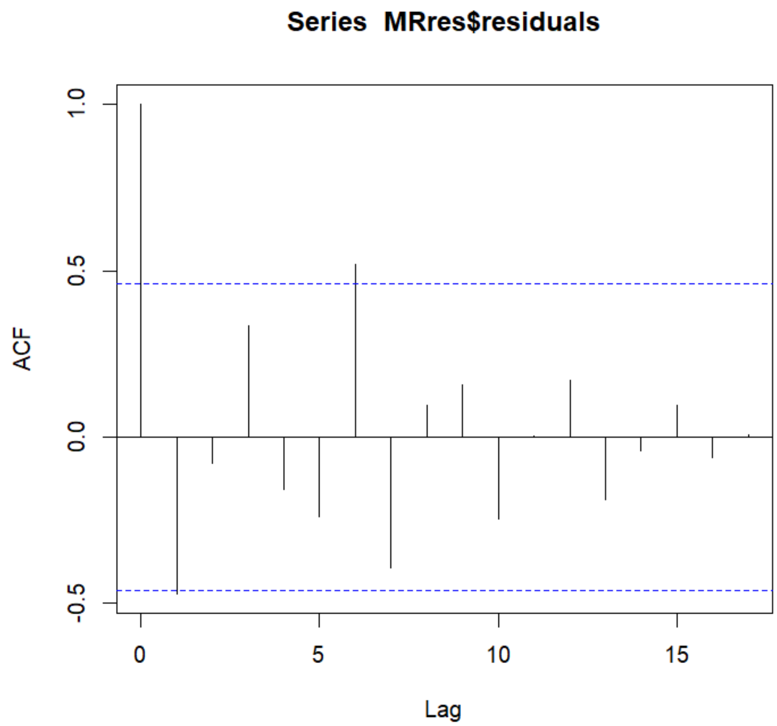

Figure 6: Autocorelation plot for the investent fund 2

The model that we use tha captures the autocorellation problem is the above

$$y_t = \alpha + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \ldots + \beta_k x_{k,t} + u_t \tag{5}$$

where:

- y_t: Dependent variable at time t
- $x_{1,t}, x_{2,t}, \ldots, x_{k,t}$: Independent variables at time t
- $\alpha, \beta_1, \beta_2, \ldots, \beta_k$: Parameters to be estimated
- u_t: Error term at time t

The error term u_t is specified as:

$$u_t = \delta + \phi_1 u_{t-1} + \epsilon_t \tag{6}$$

where:

- δ: Constant term in the error process

- ϕ_1: Autoregressive coefficient
- ϵ_t: White noise error term at time t, with $\epsilon_t \sim N(0, \sigma^2)$

The second assumption to face is the homoskedastisity problem,this means that the variace is not constant over time and need a model to capture this fainomeno.For this we introduce the garch models in following. The variance of the error term at time t is modeled as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \sigma_{t-1}^2 \tag{7}$$

where:

- σ_t^2: Variance of the error term at time t
- $\alpha_0, \alpha_1, \alpha_2$: Parameters to be estimated
- ϵ_{t-1}^2: Squared error term at time $t-1$
- σ_{t-1}^2: Variance of the error term at time $t-1$

For example, we observe in Hedge Fund 28 that there is a heteroskedasticity problem because the squared residuals exceed the decision boundary. Therefore, we need to model the variance using GARCH models.

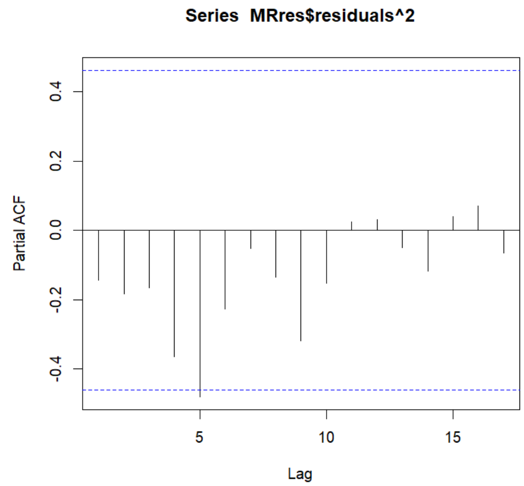

Figure 7: Squared residuals plot for Hedge Fund 28

We see from the plot that at lag 5 we have heteroskedastisity problem and we need to fix it with garch model.
To check the normality erros we need to test the residuals if it is normaly distributed,to test this we do a Jarque Bera Test for each fund and if the p-value is lower than 0.05 we reject the HO and the assumption it is not valid.Only for the hedge fund 49,43 it is not valid the normal.

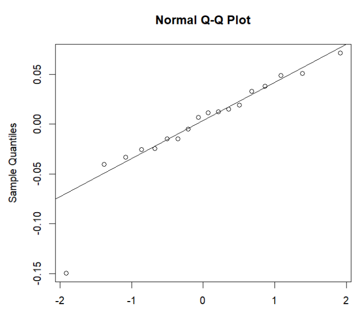

Figure 8: Residual errors model for fund 49

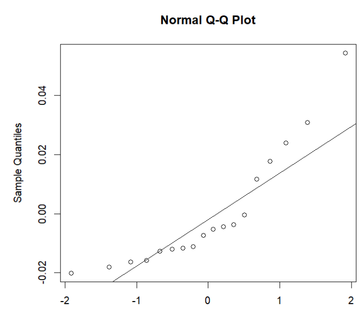

Figure 9: Residuals errors for fund 43

3 Portfolio Construction

In this section we Construct optimal portfolios based on Markowitz (1952) framework, known as Mean-Variance approach. Basic idea is to Maximize the expected portfolio return for a given level of risk or minimize portfolio risk (standard deviation) for a specified target return.

3.1 Minimum Variance Portfolios

The Minimum Variance Portfolios can be constructed based on the following optimization problem:

$$\min_{\mathbf{w}} \frac{1}{2}\text{Var}(R_{p,t}) = \frac{1}{2}\mathbf{w}'\Sigma_t\mathbf{w}$$

subject to

$$\sum_{i=1}^{n} w_i = 1$$

where:

- $\text{Var}(R_{p,t}) = \mathbf{w}'\Sigma_t\mathbf{w}$ is the variance of the portfolio
- $\mathbf{w} = (w_1, w_2, \ldots, w_n)'$ is the vector of portfolio weights
- Σ_t is the $n \times n$ covariance matrix of the returns
- Portfolio weights w_i can be either positive (long position) or negative (short position)

3.2 Mean-Variance Portfolio

The Mean-Variance Portfolios can be constructed based on the following optimization problem:

$$\begin{aligned}
\min \quad & \frac{1}{2}\text{Var}(R_{p,t}) = \frac{1}{2}\mathbf{w}'\Sigma_t\mathbf{w} && (8)\\
\text{s.t.} \quad & w_i \geq 0, \quad i = 1, 2, \ldots, n && (9)\\
& \sum_{i=1}^{n} w_i = 1 && (10)\\
& E(R_{p,t}) \geq r_{\text{Target}} && (11)
\end{aligned}$$

where:

- $\text{Var}(R_{p,t}) = \mathbf{w}'\Sigma_t\mathbf{w}$ is the variance of the portfolio.
- $E(R_{p,t})$ is the expected return of the portfolio.

- $\mathbf{w} = (w_1, w_2, \ldots, w_n)'$ is the vector of portfolio weights.
- Σ_t is the $n \times n$ covariance matrix of the returns.
- r_{Target} is the target portfolio return.
- Portfolio weights w_i are non-negative (long position).

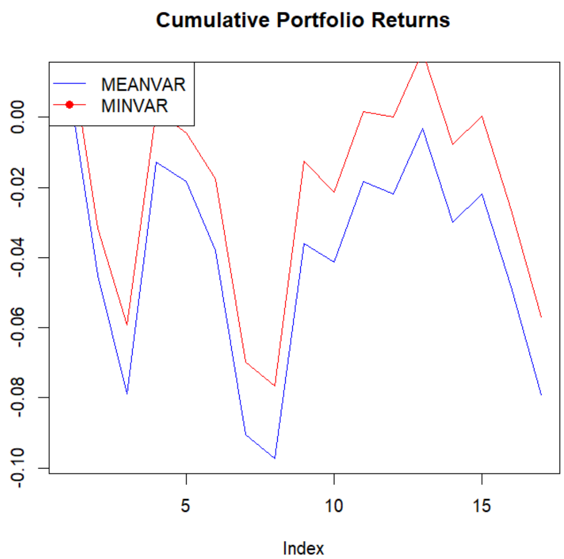

Figure 10: Cumulative return of Mean-variance and Minimum variance portofolio

From the plot we see that the portofolio base on MINVAR approach performing better that the MEANVAR.Because it has lower loses that MeanVAR approach and greater mean returns.