

# MATH5010 Final Project

*Jongyoon Sohn, Daniel Lee, Kitae Keum, Dongoh Shin*

*5/9/2019*

## 1. Introduction

History has shown that the stock market as well as economy moves repeatedly in cycles over and over. However, Investors do not need to radically change their long-term plans, rather implement simple tweaks to their portfolio to help cushion catastrophic losses or extra gains. Adjusting their portfolio allocations along with the market's flow like the bull or the bear, there are chances to improve outputs compared to well-known indices such as S&P 500. Surrounded by all the euphoria of a bull market, investors are inclined to feel confident and take more risks. Some industry sectors tend to particularly do well during bull markets and some others do not. Thus, it can be a good strategy to revise the industry weights of a portfolio.

We obtain historical returns of 11 sector ETFs and S&P 500, our benchmark, from 2007 to 2018 at Yahoo Finance. The sector ETFs, made by Vanguard, are all value-weighted to emphasize an industry allocation effect rather than a stock selection effect. In order to analyze the impact from market traits on optimal portfolio, we focus on certain periods as our in-sample periods: 2007, 2008-2009, 2016-2017, and 2017. Through this project, we see if sector allocation portfolio can outperform the market.

## 2. Main objectives and plan

The primary objectives are to derive optimal industry weights by applying the mean-variance optimization, to observe if the portfolios can perform well in out-of-sample periods, and to test their legitimacy by implementing a time series model. The performance of the weights derived under conditions of different periods can provide us with directions about how to construct a portfolio in the future.

Our study is mainly focused on finding a low-risk portfolio which can lead that the quadratic programming is applied to minimize their variance of the daily returns. We set the benchmark's average daily returns in the four periods as the constraints of the quadratic programming. After finding optimal portfolios, we apply them to the market in our test period of 2018. Their performance with respect to the benchmark in 2018 helps us determine the effects of our portfolio and it is also helpful to figure out the difference in Sharpe ratios between each portfolio and the benchmark. We assume that a naked short-selling is allowed and there is no trading cost accompanied by rebalancing portfolios to adjust weights to our optimal weights everyday.

### 3. Analysis description

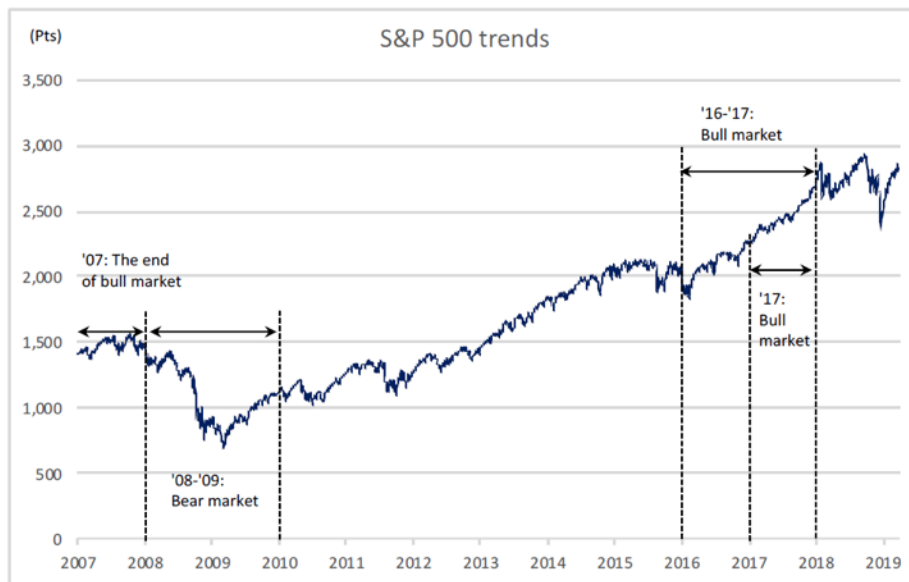
#### 3.1. Assumptions for market data selection

First of all, we divide the market into 11 different sectors based on WRDS (Wharton Research Data Services)'s GICS industry classification. These include: Communication, Consumer discretionary, Consumer staples, Energy, Financials, Healthcare, Industrials, Information technology, Materials, Real estate and Utilities.

After then, we decided to use an ETF to represent each sector. An Exchange-Traded Fund (ETF) is type of an investment fund. Just like stocks, it is traded on stock exchanges. There could be different types of assets under an ETF, such as stocks, commodities and bonds.

The reason why we decide to use ETFs rather than select individual stocks is the sector ETFs represent the market's characteristics better. Through our time period (2007-2018), the numbers/types of stocks constantly changes due to delisting or a IPO. Thus, if the top 5% stocks are selected manually from each sectors, some of the data will have missing values. However, ETFs automatically control the change of stocks in each sector and represent the overall trend throughout the time period. Moreover, ETFs keep the consistency so that the Vanguard ETFs are chosen for our analysis.

#### 3.2. Assumptions for market periods selection



We chose to use four different time periods for two reasons. First, we wanted to see the weight difference based on the market situation. As we can see from the S&P 500 plot above, bull market includes 2007, 2016-2017 and the bear market includes 2008-2009 and 2018. We assumed that the weight allocations of bear vs. bull market would be vastly different. Second, we also wanted to see the effect of length of time. Thus, we tried optimization in both 2 year periods and a 1 year period.

### 3.3. Weight allocation

Traditional weight allocations in the portfolio management is done by managing the weights between different types of assets such as stocks, fixed income and cash. For example, a typical portfolio includes stocks that have high beta and low beta in its combination. However, for the purpose of weight allocations between the sectors of the stock market, the quadratic programming is used to optimize the weight allocation by minimizing its risk. Through the quadratic programming, a portfolio with lower volatility can be achieved and it gains at least the same returns as the benchmark, S&P 500.

For the optimization process, we have set several assumptions below.

- i) Short-selling is allowed
- ii) Portfolio's return is same as the market return
- iii) Portfolio's volatility is lower than the market volatility
- iv) Sum of weights should be equal to 1
- v) Sum of weight multiplied by returns of each sector should be equal to total portfolio return

### 3.4. Time series analysis for returns

The time series forecasting model, ARIMA, is applied to predict future portfolio returns based on the past performances. A prediction using a non-seasonal ARIMA(p,d,q) model is the most appropriate since no seasonal trend has found. The result of Augmented Dickey-Fuller(ADF) test with a p-value 0.01 shows that it is a stationary time series process which allows us to let d be 0. In order to choose parameters, the number of autoregressive (p) and moving average (q), which remove any autocorrelation remainings in the residuals, the number of combinations of these parameters is iterated. Also, the Akaike information criterion (AIC) is used to evaluate the model.

To see if the residuals have no significant patterns and are normally distributed, it is proceeded in making one-step out-of-sample forecast by splitting the whole dataset into 100 samples for training and 1 sample for testing each time, and then calculating the prediction interval.

## 4. Results of statistical analysis and interpretation

### 4.1. Prediction of portfolio returns in 2018

Based on portfolio optimization process, the combinations of weights in four groups are plugged into the 2018 market. To compare the performance, we choose to see the sharpe ratio different between portfolios. Sharpe ratio is also calculated because it is a fair key metric

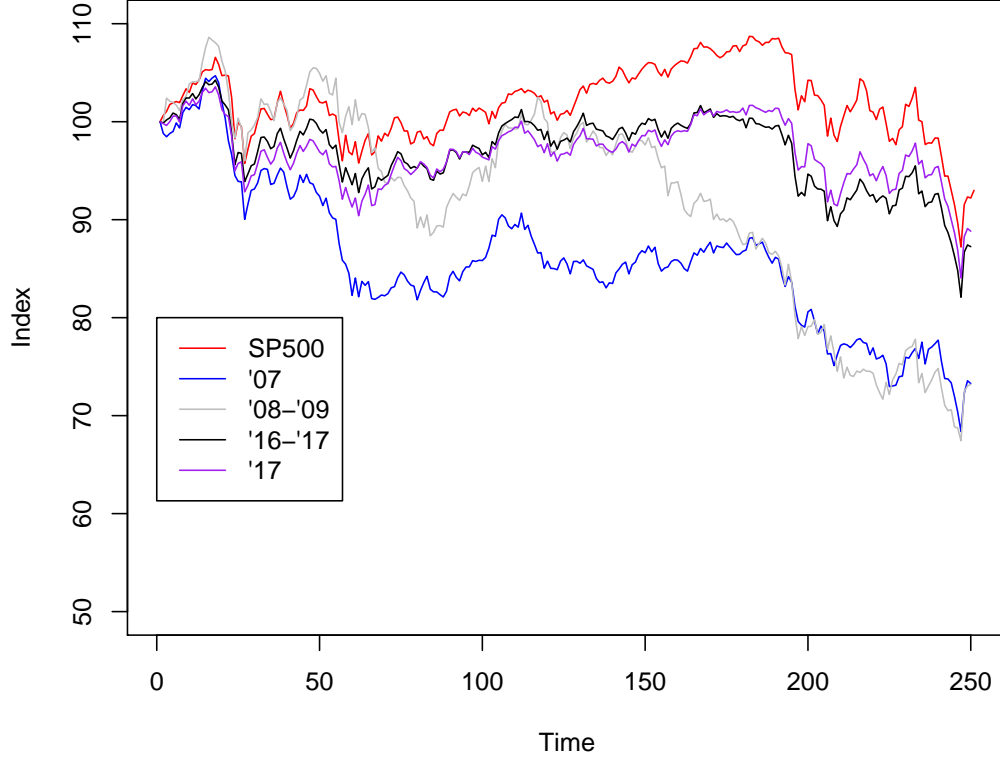


Figure 1: Index Movement

to show the return of an investment compared to its risk. The ratio represent the average return earned in excess of the risk-free

As we can see from the table, all Sharpe ratios are negative because the 2018 market was a bear market and the market performance was negative. However, we should not just focus on the magnitude of Sharpe ratio, but should focus on each portfolio's volatility and return. Although returns of 2017 and 2016-2017 portfolios were smaller than the S&P 500, the volatility was lower. We can also see that the performance of 2016 and 2017 weights are much better than that of 2007 and 2008-2009. Both return and volatility underperformed compared to the benchmark. This is because the market characteristics have drastically changed a lot from the past years. From this, we can see that more recent weights fits better for the future. Even though both 2008-2009 and 2018 markets were bull, the performance was not good. This is because the market driver was vastly different. In 2008-2009, every sector did not perform well because of the macro economical reason, but in 2018, the main reason was the fall of FAANG and IT stocks. The heatmap beneath illustrates clearly the change of weights during the chosen periods.

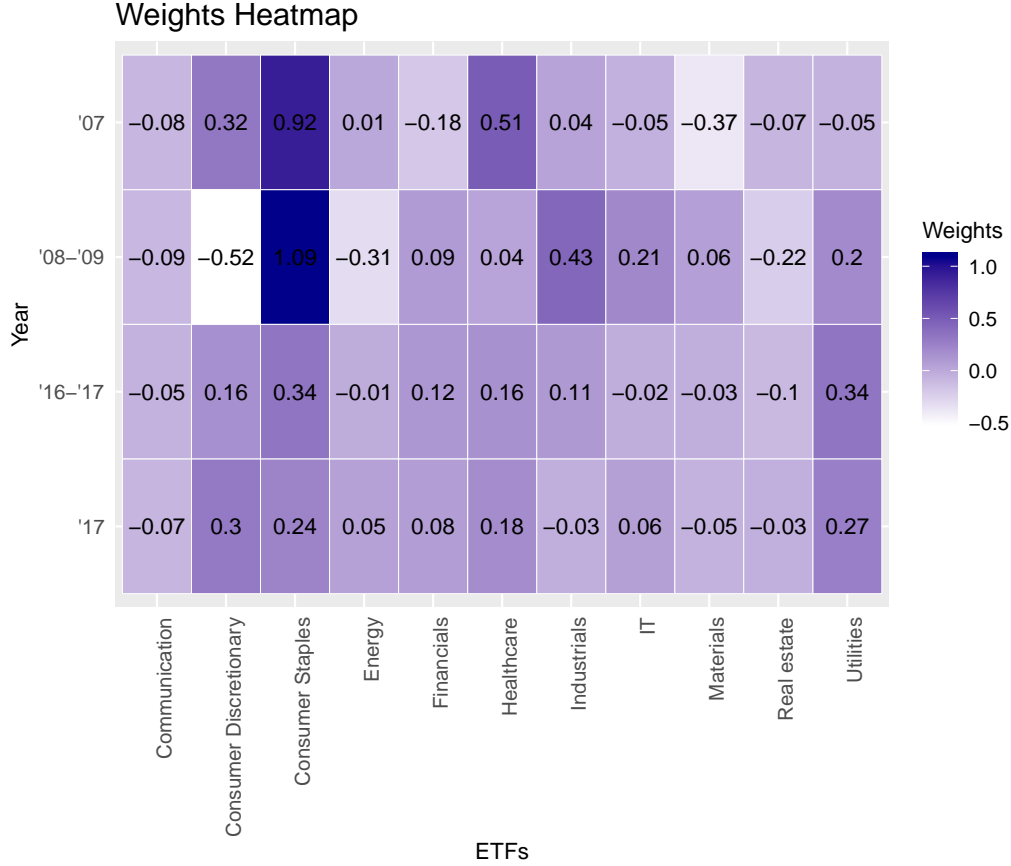


Figure 2: Weights Heatmap

'18 S&P500 :	$\frac{-0.0503 - \mu_f}{0.17}$	= -0.451
'17 Bull weights :	$\frac{-0.094 - \mu_f}{0.157}$	= -0.767
'16 -'17 Bull weights :	$\frac{-0.11 - \mu_f}{0.161}$	= -0.847
'07 Bull weights :	$\frac{-0.2909 - \mu_f}{0.1865}$	= -1.701
'08 -'09 Bear weights:	$\frac{-0.2842 - \mu_f}{0.229}$	= -1.356

where  $\mu_f = 0.0264$

## 4.2. Result of Time Series Analysis

The ARMA(1, 1) is used to predict the future portfolio based on the past performance by using the weights from 2017 data and to get the prediction for the return of 2018 market. Optimal weights from 2017 are used in the prediction because it is assumed that the weights in 2017, which is the most close period to the 2018 market for the prediction, perform better than other market weights. Hence, in the figure of ARIMA Model Prediction on 2018 above, its movement depicts the result of the prediction for 2018 market using 2017 market weights.

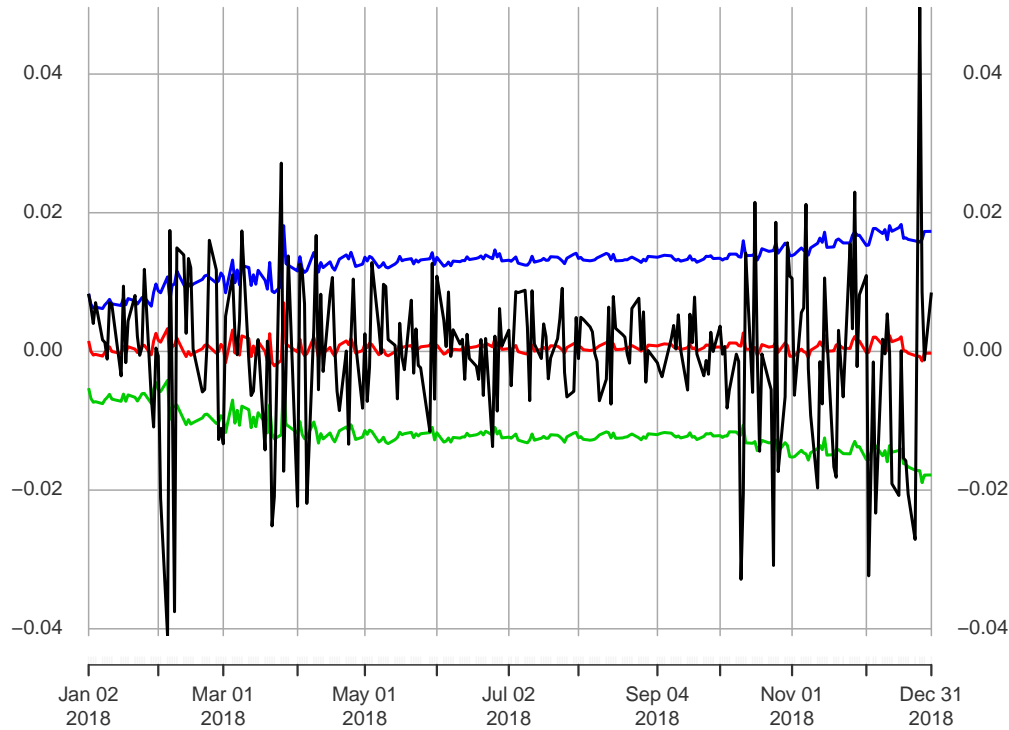


Figure 3: Prediction

In detail, the S&P 500 returns in 2018 in black is plotted to compare our prediction with the actual market trend. The red line of our prediction of returns with weights from 2017 is also plotted with its 95% confidence interval. The confidence interval covers around 85.4% of actual data.

```
##
## Augmented Dickey-Fuller Test
##
## data: data[, -1]
## Dickey-Fuller = -9.4487, Lag order = 7, p-value = 0.01
## alternative hypothesis: stationary
```

## 5. Conclusions

Overall, it has been revealed that the most recent optimal weight works relatively better than the weights from the weights of the further past. Interestingly, it is believed that there is a reason for this because the current market characteristics are similar to that of the previous year. In fact, the price trend of each industry does not change a lot within one year. On the contrary, the weights of the sample are not performed as good as we expected in the 2008-2009 market data. It may be a hasty generalization, but it seems like constructing a optimized portfolio by using past returns is a huge demanding work due to the change of

main driver of economic growth.

Moreover, portfolio performance relies more on market drivers than on market direction. What it means is that it is observed that the markets in both 2008-2009 and 2018 have a decreasing pattern but its reason wouldn't be identical. For example, some decreasing trend in 2008-2009 did not perform well in 2018, due to the different reasons for sluggish markets. This is because in 2008, almost all of the industries did not perform well under the financial crisis, whereas the market drops in 2018 are mainly followed by the plunge of IT sector which outperformed in 2017.

Rather than bringing weight allocations from the past, it seems to work better if we update the weights more frequently. Even the time periods of 1-year and 2-year demonstrate the difference of weight optimization so that we, for the future research, would like to compare the performance of portfolios by differentiating the period. In other words, if the optimization is studied in each 6 months or each quarters, the portfolio might perform better than longer update period ones.

Lastly, after a bunch of time series modeling, we found ARMA(1,1) performs the best for our prediction of return analysis by plotting in the 95% confidence interval with the weight from 2017. This result supports the accuracy and preciseness of our return forecasting as well.

## 6. Reference

Bielstein, P and M. Hanauer, (2018), "Mean-variance optimization using forward-looking return estimates"

Rodriguez-Lujan, I., R. Huerta, C. Elkan, and C. Santa Cruz, (2010), "Quadratic Programming Feature Selection"

Fama, E. and French, K. (1996). Multifactor explanations of asset pricing anomalies, *Journal of Finance* Best MJ, Grauer RR (1991) On the sensitivity of mean-variance efficient portfolios to changes in asset means: some analytical and computational results. *Rev Financ Stud* 4(2):315–342