1.0 Introduction

Portfolio selection and portfolio optimization is something that is a popular area of research in the financial industry. Both for fund managers and academic researchers (Huo and Fu, 2017). We want to select the best combination of assets to receive the highest possible return with the lowest risk. New research and methods of doing this could potentially yield a great amount of reward, as fund managers are constantly trying to maximize the return for their investors. Because of this complexity, investors normally adopt a mean-variance framework to reduce risk rather than merely maximizing expected returns.

In recent years researchers have been trying to use agent-based modeling (ABM) to discover new possibilities for managing a portfolio. ABM consists of agents who you can train to optimize a certain problem or situation in a specified environment. Previous economic models such as the DSGE assume that the people interacting in the environment are very sophisticated and are always able to solve difficult computational problems. However, we have learned that the environment that we are in is too complex for anyone to understand (Hamill & Gilbert, 2016). ABM's are very useful in this context because you are able to take into account heterogeneity, add dynamics, and model interactions between different agents.

In this assignment, we will try to optimize a set of equity portfolios and implement a crisis with the use of the multi-armed bandit problem. The MAB problem can be compared to facing multiple slot machines in a casino where each slot machine is configured with a certain possibility to yield a reward. The goal here is to maximize the possible cumulative payout by using certain slot machines (Zhu Et al. 2019).

There have been different attempts to turn the portfolio selection problem into a multi-armed bandit problem. Zhu Et.al, 2019, tried to do this. They used traditional portfolio techniques as strategic arms to create a dynamic portfolio strategy with various cycles that could adjust to different periods. They also implemented a reward function that was based on the user's preference to risk to select the optimal "arm" of each period. They managed to create an algorithm that could balance the benefits and risk in a good manner and achieve higher returns by controlling risk (Zhu Et al. 2019).

Many variations of the MAB in portfolio selection are only based on the best possible machine to play at each trail, or the best portfolio to select at each time step. To deal with this issue Xiaoguang Huo and Feng Fu, implemented a model which first selects a basked of assets to generate the first portfolio. Then, at each time step, they use the optimal multi-armed bandit algorithm to combine a single-asset portfolio that globally minimizes the risk and value at risk (VaR). By doing this they were able to achieve the balance by maximizing the reward and minimizing the risk. Because of the risk selection at each time-step they were also able to minimize the risk when the market was volatile (Huo and Fu, 2017).

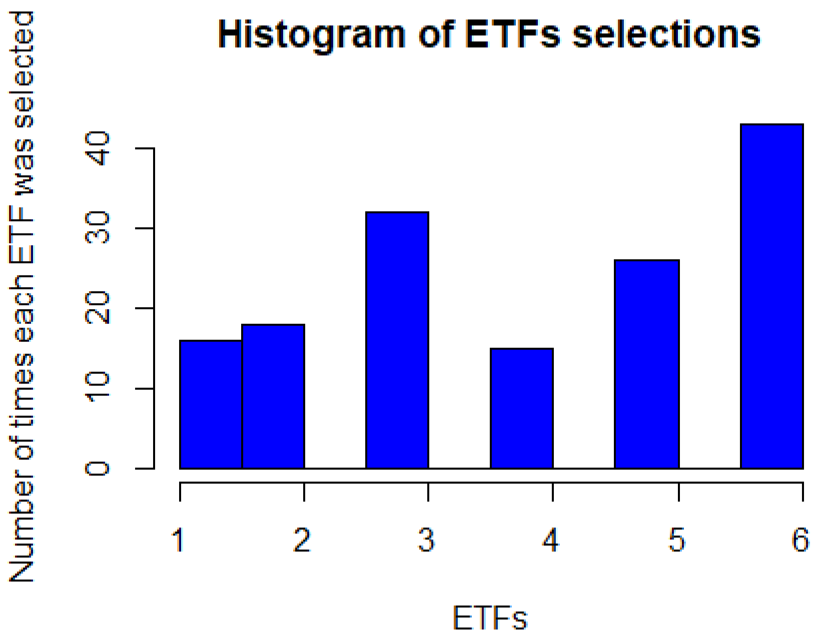
The first thing one should do when trying to select different assets to create the optimal portfolio is to explore each assets' potential yield and risk. Dependent on the particular investor we want the best possible trade-off between risk and return. Some investors are very risk-averse. Their portfolio would most likely consist of assets with very low risk. With low risk, one should expect a lower expected return as well, as when you invest in risky assets you demand a higher return (risk premium). With this in mind, one uses their individual selection criteria to end up with the portfolio that suits the investor's preferences the best. Finally, we use the MAB to evaluate what assets that can create this optimal portfolio.

In this particular problem, our dataset of assets contains ETF's, which stands for Exchange-traded funds. ETFs are investment companies that are traded on the public exchange. In opposition to mutual funds, ETFs can be traded continuously throughout the trading day on the public exchange where their major aim is to track a security index (Itzhak Ben-David et Al., 2017).

Optimization of equity portfolios using MAB

For the optimization of the equity portfolio we used the R code that was provided that modeled the problem as a multi armed bandit problem and used the upper confidence bounds approach. In this kind of approach at the beginning, a confidence limit for each arm will be created and an arm will be selected randomly, knowing that all arms have the same observed average value. If the arms return a mistake, the average value observed by the arm will decrease, otherwise the average will increase and the confidence limit will also increase. The goal is to select the stock with the maximum return for the highest number of observations.

In this case we used the same dataset that was used on the book chapter provided so in the end the result was that the sixth stock called GLD was the most selected.

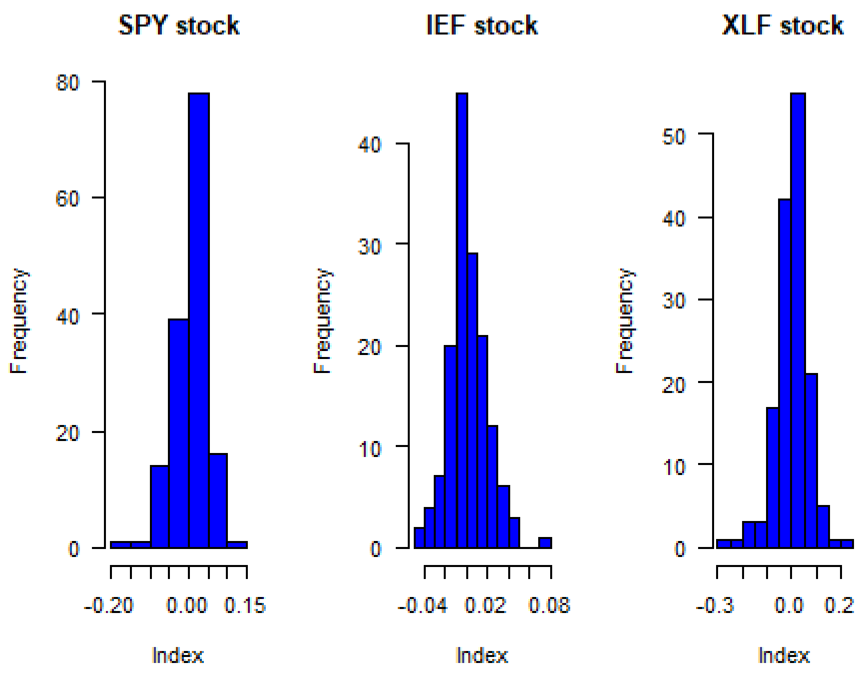


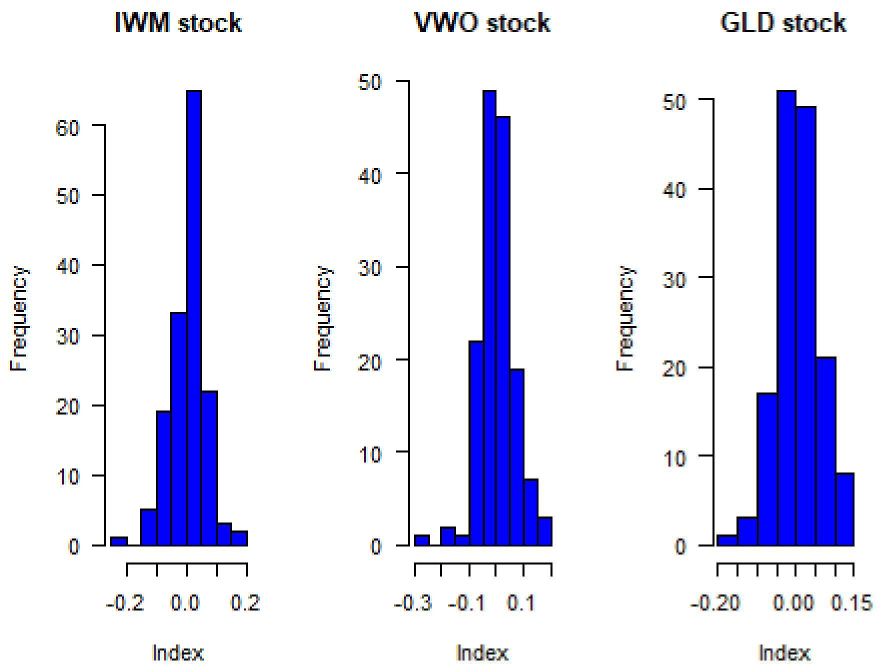
After this we introduced a shock/crisis on the system to see the effects on the pricing strategy and for that we created two scenarios, the first when only one stock suffers the shock to see the effect on the pricing strategy and the second scenario is when half the stocks are hit with the crisis to see the effect on the pricing strategy of all the portfolio.

Based on the work of an economist named Harry Markowitz that wrote his dissertation on "Portfolio Selection" that transformed the landscape of portfolio management that gave him the Nobel Prize in Economics and was the base for the Modern Portfolio Theory that offers a an approach which aims to maximize a portfolio's expected return for a given amount of risk by selecting the proportions of various stocks and it also tries to minimize the risk for a given level of expected return.

Having all investments two important aspects that are risk, where the investors look for the lowest possible risk, and return, where investors look for the highest possible return, knowing if the stocks follow a normal distribution is important because we can link the mean for the returns and the standard deviation for the risk.

So analyzing the histograms of all the stocks we can see that all of them follow this distribution. This will help us to know the returns and to expect that around 68% of the values will be within 1 standard deviation and 95% within 2 standard deviations.





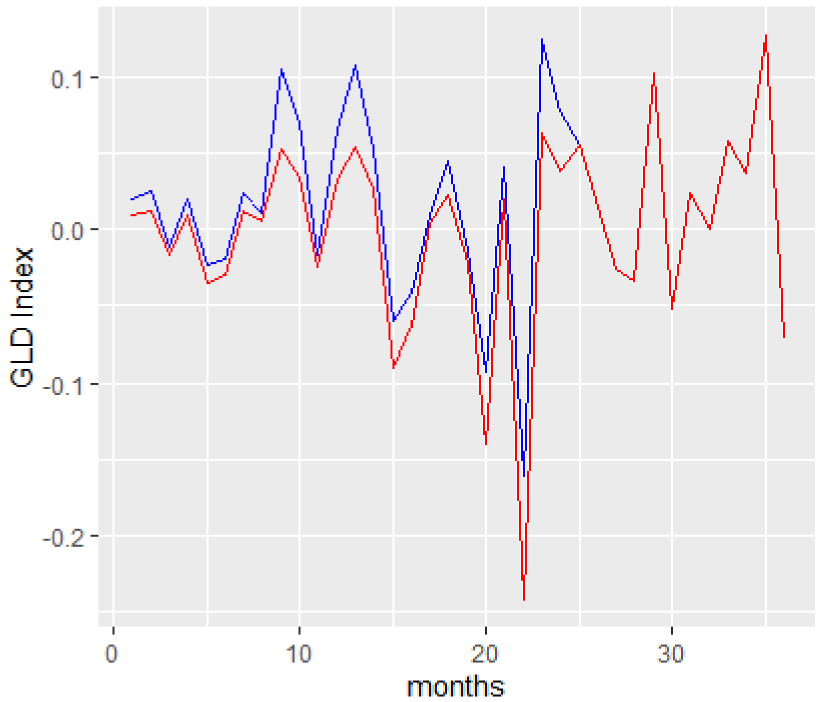
In the first scenario we only provoked the shock on one of the stocks and we chose the one that was chosen the most on the normal scenario that was stock GLD.

For this we created a new dataset of the EFTs with the modifications on the periods that we will apply the shocks, and knowing that a stock crash is normally for values over 10%, we chose to decrease the stock index on the first 24 months on 50% of their value.

For that we created the code below code that decreased the values on 50% for positive values and negative ones.

```
#shock for variable GLD
for(i in 1:24){
  if(dataset2$GLD[i]>0){
    dataset2$GLD[i] <- dataset2$GLD[i]*(1-0.5)
  }else{
    dataset2$GLD[i] <- dataset2$GLD[i]*(1+0.5)
  }
}
```

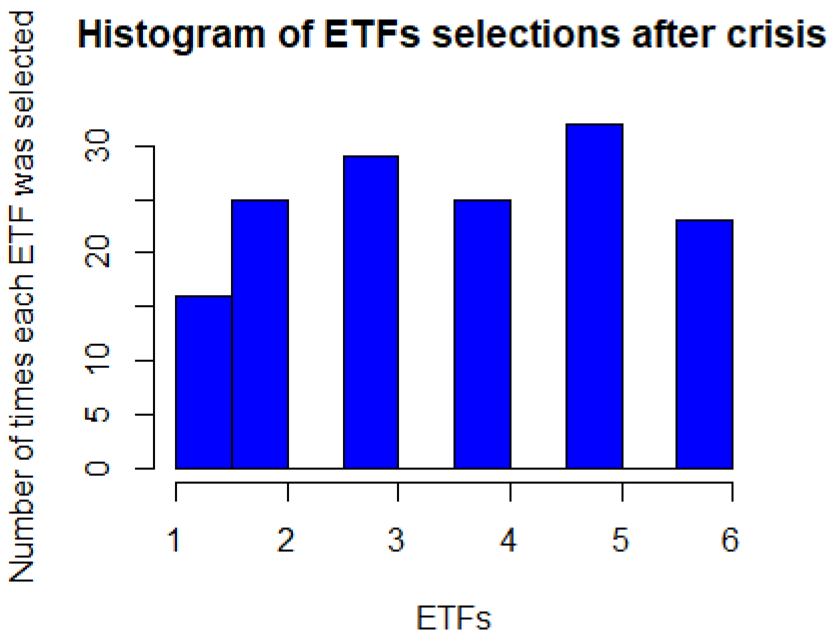
And we can clearly see the shock, represented by the red line, in the first 24 months and converging to the same line after that.



Knowing that the standard deviation is the best way to measure risk, below we can see that stock GLD got a little increase on the standard deviation value.

```
> sd(dataset$GLD)
[1] 0.05103855
> sd(dataset2$GLD)
[1] 0.05167097
>
```

Analyzing the results on the pricing strategy we could see a clear change on the stock that got chosen more times. Now is stock 5, that is VWO, that was the stock more times chosen.

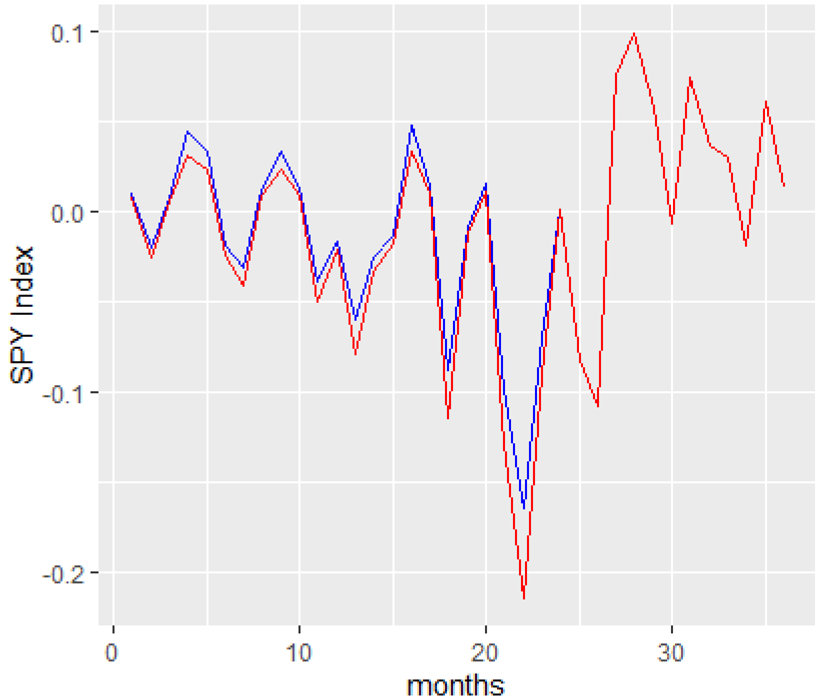


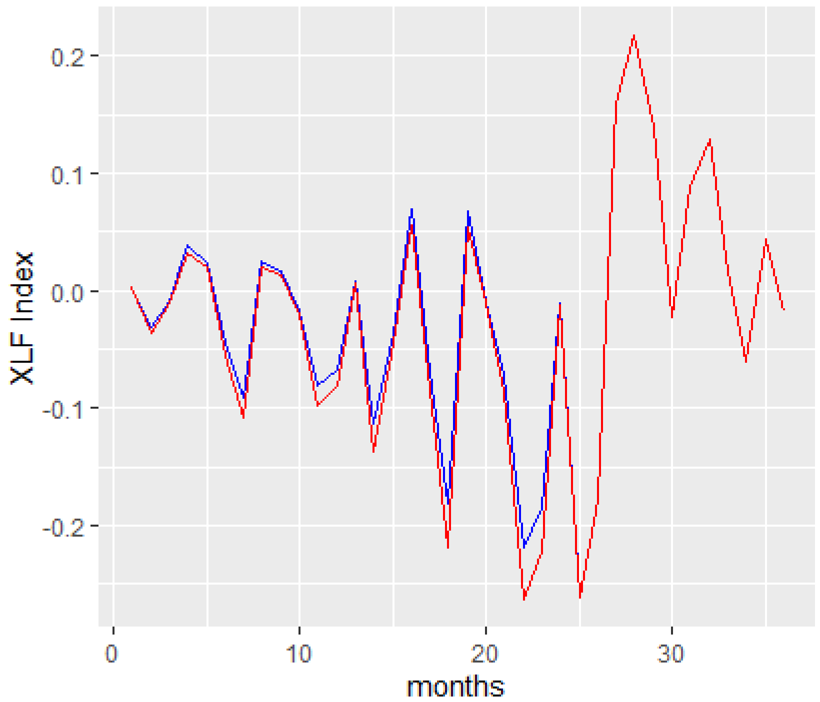
For the second scenario we chose to apply the shock to three stocks, that were SPY, XLF and GLD. With different decreasing percentages, SPY decreased 30%, XLF 20% and GLD was decreased for the same 50% as before.

Below we can see the plots that show the decreasing of the indexes and the increase on the standard deviations for SPY and XLF, since for GLD we already did it before.

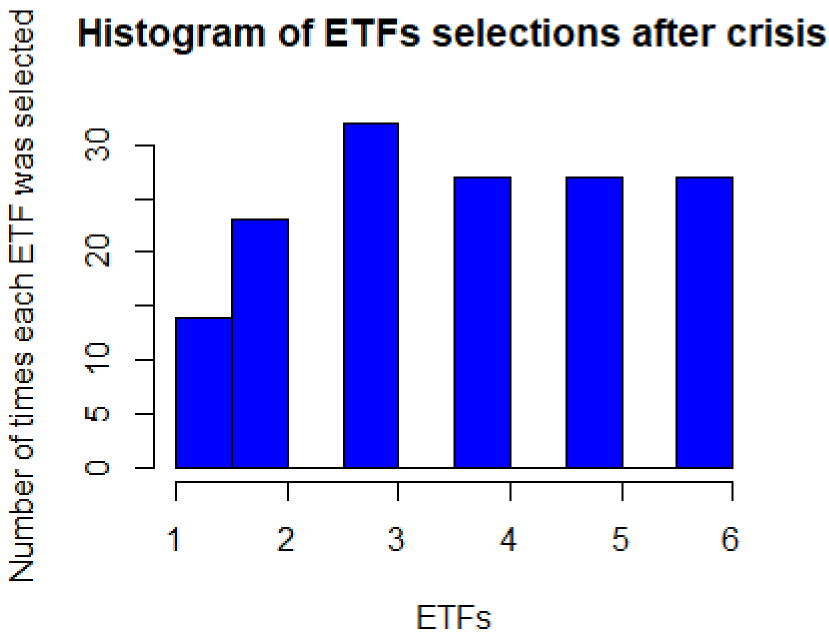
```
> sd(dataset$SPY)
[1] 0.04262596
> sd(dataset2$SPY)
[1] 0.04556684
>
```

```
> sd(dataset$XLF)
[1] 0.06756594
> sd(dataset2$XLF)
[1] 0.07083026
```





In this case we can see that was stock 3 or XLF that was chosen the most, but stocks 4, 5 and 6 had good values that were similar. Despite the shock, stock XLF was the most chosen that shows that its value in relation to the other was good even after its index got lowered in the first 24 months.



References:

Huo X, Fu F. (2017) Risk-aware multi-armed bandit problem with application to portfolio selection.*R. Soc. open sci. 4*: 171377. http://dx.doi.org/10.1098/rsos.171377

Mengying Zhu et al. (2019) Adaptive Portfolio by Solving Multi-armed Bandit via Thompson Sampling.

https://www.investopedia.com/articles/investing/100714/using-normal-distribution-formula-optimize-your-portfolio.asp

https://analystprep.com/cfa-level-1-exam/equity/price-return-total-return-index/

https://en.wikipedia.org/wiki/Stock_market_crash

https://www.investopedia.com/terms/s/stock-market-crash.asp

https://www.investopedia.com/ask/answers/021915/how-standard-deviation-used-determine-risk.asp

https://www.investopedia.com/articles/investing/100714/using-normal-distribution-formula-optimize-your-portfolio.asp

https://en.wikipedia.org/wiki/Modern_portfolio_theory

https://www.guidedchoice.com/video/dr-harry-markowitz-father-of-modern-portfolio-theory/