Imperial College London

Department of Electrical and Electronic Engineering

Final Year Project Report 2024

Project Title: **Robust Portfolio Optimization in Uncertain Markets**

Student: **Pranav Viswanathan**

CID: **01744658**

Course: **EIE3**

Project Supervisor: **Dr Imad Jaimoukha**

Second Marker: **Dr Adria Junyent-Ferre**

Final Report Plagiarism Statement

I affirm that I have submitted/will submit an electronic copy of my final year report to the provided EEE link.

I affirm that I have submitted/will submit an identical electronic copy of my final year report to the provided Blackboard module for plagiarism checking.

I affirm that I have provided explicit references for all the material in my Final Year Report that is not authored by me but represented as my own work.

I have not used linear learning models as an aid in the preparation of my report.

Acknowledgements

My journey through this final project, and by extension, my time at Imperial has been full of ups and downs, but extremely enlightening and rewarding. I'd like to thank my supervisor Dr. Imad Jaimoukha for not only his help ideating and navigating the direction of this project, but also teaching me several values for both learning and self-appreciation. I'd also like to thank my mother, father and little sister. Sindhu, Sajeev and Leya, this would not have been possible without you. Thank you for for being a constant in my life.

Pranav

Abstract

Throughout the history of investing, Investment Managers have either made or broke fortunes in anticipation of certain stock investing opportunities seeming ludicrous. Warren Buffet immediately comes to mind with his Berkshire Hathaway now amassing an astonishing 892.60 Billion USD Market Cap [1]. Multitudes claim to use advanced portfolio selection and optimization techniques to maximize returns yet still, Investing and portfolio management is a very experience heavy field where confidence is inspired more by the vintage of an Investment Manager rather than the capabilities of the strategy he employs. Nevertheless, there are many statistical engineering concepts that can be applied to traditional investing strategies that can theoretically mitigate risk and increase returns when choosing a portfolio. With that thought in mind, this paper evaluates the performance of certain data sampling techniques that were replicated in Matlab code to try and examine how much data analytics and manipulation can affect investor returns. The findings are promising yet intriguing, and there are many avenues in which this can be explored further.

Contents

1.1 Introduction:

Public Stock Markets are heavily volatile and tumultuous, driven by public demand which is influenced by a variety of aspects, be it technological, economic or creative breakthroughs, or even international crises such as wars and health pandemics. Investment Managers are tasked with constantly finding the most optimal portfolio of assets to either maximize returns or minimize risk to appease stakeholders in Indexes. It is a strenuous job and requires constant data upkeep and monitoring.

Considering this, Harry Markowitz, widely regarded as the pioneer of portfolio optimization, put forth his nobel prize winning paper 'Portfolio Selection' in 1952. In it, the mean-variance portfolio theory introduced by Markowitz remains influential in both research and practice. This theory provides a straightforward formula for the mean-variance efficient portfolio, which relies solely on two key population characteristics: the mean and the covariance matrix of asset returns. Ideally, when these parameters are known, investors can easily calculate the optimal portfolio weights according to their desired risk level or target return. However, in reality, the actual value of these parameters are unknown and subject to multiple conditions. Instead, investors use the sample means of assets and sample covariance matrices of the same as substitutes, creating what is known as the "plug-in" portfolio. This method is supported by classical statistics since the plug-in portfolio is a maximum likelihood estimate (MLE) of the optimal portfolio. Nevertheless, as highlighted by Michaud in 1989 and others, the plug-in portfolio performs poorly out-of-sample. The situation becomes even more problematic as the number of assets increases. This issue, referred to as the "Markowitz Optimization Enigma" by Michaud, has been further explored by researchers like Best and Grauer (1991), Green and Hollifield (1992), Chopra and Ziemba (1993), Britten-Jones (1999), Kan and Zhou (2007), and Basak et al. (2009), who document the challenges of constructing the mean-variance efficient portfolio using sample estimates.

The aim of this Final Year Project is to identify and implement various data manipulation techniques to achieve more representative estimators that will, to some extent, alleviate concerns of uncertainty and robustness. An additional aim of the experiment is to maximise returns and minimize risk on a created, controlled set of stock price data of companies from the Dow-Jones Index from the 1st of January 2010 to the 1st of January 2020.

This is a very well documented subject, with over nine hundred research papers being published on this topic and this project will evaluate how changing certain parameters in a robust portfolio optimisation algorithm affect returns and risk in a well-documented financial case study. This project is investigative and experimental and should yield a greater understanding of how

Investment Managers can utilize engineering concepts to improve estimators regarding risk and returns and maximise their profitability in uncertain markets considering the mean-variance framework.

In order to achieve this, three statistical approaches were considered. The Markowitz Mean-Variance Optimisation, Principal Component Analysis, Bayesian Mean-Variance Optimization and by extension, a combination of both PCA and Bayesian Updating. The Markowitz model would be considered the basis and baseline for the experiment, whilst Principal Component Analysis (PCA) and Bayesian Inference would be supplemental to data fed to the Mean Variance Optimization Model.

This report will provide an overview and analysis of the data set used, a mathematical explanation of concepts used, and how they might affect the Mean Variance Optimization model. Also presented is how these mathematical concepts were translated in code to suit the experiment design and relevance of the data set. This report will also present the experiment design and an evaluation on both the experiment's design and the models performance. Conclusions will be presented at the end and rationalized based on the experiment's results. Furthermore, avenues for further study will be identified and built upon.

All experiments were developed and run on MATLAB. Microsoft Excel was used for the preliminary and evaluation-based data analysis. Appendix A holds a link and QR code to the code used in the project.

1.2 Definitions:

Considering the scope of the project, it would be beneficial to properly define and elaborate on key words related to the project title:

Markets: [5]. These are forums that trade forex, stock (both publicly and privately traded) or bonds.

Portfolio: A (financial) portfolio refers to a collection of assets an individual/corporate entity owns for the purpose of wealth management and generation over a period and/or maintenance of a trading dossier. [6] In this project, Assets will refer to a publicly traded stock of the financial managers choosing.

Robust: Seeing as public markets are extremely volatile due to constantly being traded and influenced by foreign factors such as world economies and technological/political breakthroughs, a financial portfolio generation model will always have to be adaptable, that is, perform effectively under constant change of variables and assumptions.

Returns: Any net gain or loss on an investment in an asset over a period.

Risk: Chance that an outcome or investment's gain differs from an expected return. In the context of the project, and in line with Markowitz's theory, it will be considered the variance of an asset, or the covariance of a portfolio.

Uncertain: As indicated above, when an expected return is indicated on a specific asset, it is essentially an estimate based on historical trends which a financial manager can never really know whether it will hold true for periods beyond historical data as the market as a whole is affected by many qualitative and random factors. Thus, the historical data estimates must be manipulated to represent more recent trends in the historical data.

2.1 The Data Set

In order to evaluate the performance of the various Optimization algorithms and by extension, their estimators, the monthly closing price of ten stocks from the Dow-Jones Index were chosen. The closing prices were taken from 01/01/2010 to 01/12/2020.

This range was selected as it falls directly between two global crises: The 2008 US housing crisis and the 2020 Covid-19 Pandemic. In these periods, global economies widely fluctuated to due reduction in consumer buying power or consumer buying restrictions. Stock trends in these periods where heavily attuned by human response, inducing extremely volatile stock prices. These responses are inherently hard to quantify and unexpected, and modern portfolio theory cannot account for such economic meltdowns. The market had recovered from the US housing Crisis in 2008 by the start of 2010, and the Market began to fall by March 2020 due to Covid, right after the project data set bounds. Thus, the data set was not affected by these crises, and trendlines can be inferred via proposed estimators explored below.

The following stocks were selected:

Table 1: Companies Selected

Index Name	**Company Name**
LMT	Lockheed Martin Corporation
MSFT	Microsoft Inc.
APPL	Apple inc.
AMZN	Amazon inc.
V	Visa Inc.
MCD	McDonald’s Corporation
WFC	Wells Fargo and Co
CMCSA	Comcast Corp
MS	Morgan Stanley
C	Citigroup Inc

These ten stocks were chosen as they branch from infrastructure (such as Amazon and Visa) to Consumables (McDonalds’s) to Technology (such as Microsoft and Apple.) Having a wide range of type of company allows for the optimisation analysis to be more holistic and not wholly affected by global trends at the time.

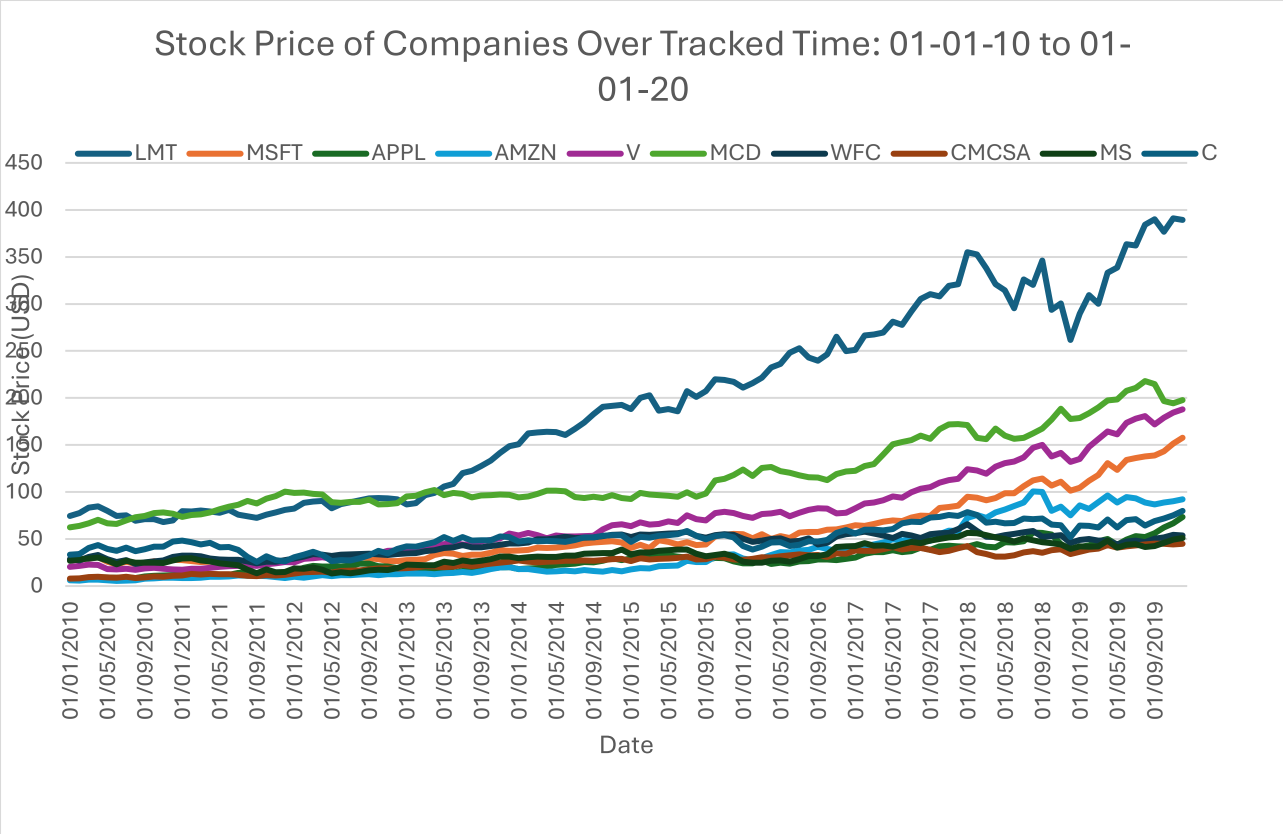


Figure 1: Stock price evolution of the 10 companies from 2010 to 2020

Above, figure 1 presents the evolution of stock prices from the 1st of January 2010 to the end of 2020. Largely, with the exception of LMT, the stock prices of assets progress together. It was debated whether LMT should be retained in the data set as it showed severe variation, with the stock price ballooning from 140 dollars in the end of 2013 to 357 in March 2018. It was ultimately retained in order to explore how the optimization approaches would handle such a case.

This data was taken from Yahoo Finance Historical Data webpages [4] and collated into excel. From there, the monthly returns were calculated and plotted:

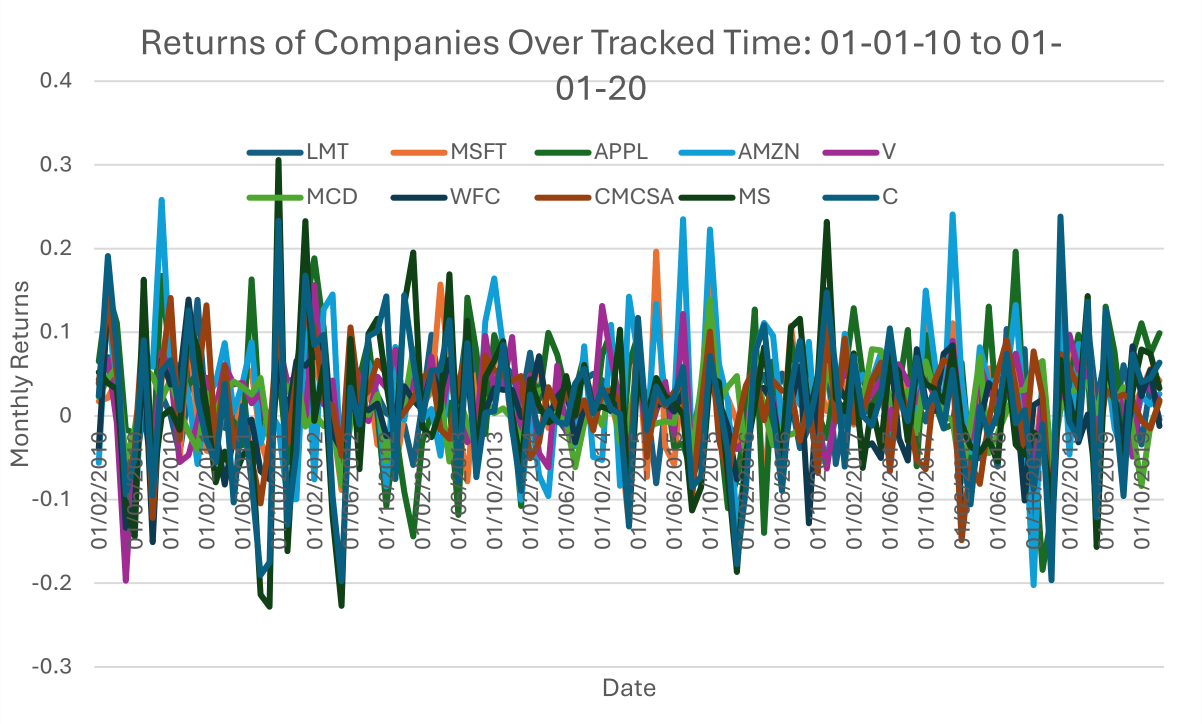


Figure 2: Monthly returns of Assets from 2010 to 2020

Ultimately, not much can be inferred visually from the frankly messy plot. However, via Matlab, the returns were handled efficiently and mean returns and variance over various periods were assessed through mean-variance optimization logic.

The next section details the mathematical background of all concepts used throughout the project.

3.1.1 Mathematical Background: Markowitz Mean Variance Optimization

Markowitz in his Nobel Prize winning paper 'Portfolio Selection' developed a mean variance framework that took a portfolio target return, portfolio target risk, asset expected returns (a sample estimator) and a sample covariance matrix (representing risk of the assets across sample data) both representing estimators grounded in historical data to achieve portfolio weights that would return the above targets. This section will go over how Markowitz's portfolio optimization works.

To begin with, his standard mean-variance optimisation framework should be considered as a base. For a portfolio consisting of n assets, a matrix x is defined which corresponds to the budget weightage split across the n assets (in the project's case, a distribution of an investment manager's budget across the n assets.)

$$x = \{x_{1,}x_{2,}x_{3,}x_{4,} \dots x_n\} \tag{1}$$

$$subject\ to: 0 \ \leq x_i \leq 1$$

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

This scenario assumes that the investor cannot borrow money, or short sell. Alongside the weightage matrix, a matrix μ is retrieved, representing the expected (or estimated) returns of all the n assets across a time period:

$$\mu \ = \{\mu_{1,}\mu_{2,}\ \mu_{3,}\mu_{4,} \dots \mu_n\} \tag{2}$$

From this, a basic return maximisation problem can be constructed:

$$maximise\ \mu^T x$$

$$subject\ to: 0 \ \leq x \leq e$$
$$e^T x = 1$$

Where $\mu^T x$ represents the expected return of the portfolio with asset weights x. e is a (nx1) matrix representing ones. This ensures the budget is maintained, and bars short selling by ensuring x_i remains between 0 and 1 inclusive. This maximisation problem however does not consider the associated volatility of the n portfolio assets. An investment manager would ideally aim to maximise the return whilst minimising the volatility of his portfolio. To incorporate this, Markowitz stated that a ϵ matrix representing the covariances of the n portfolio asset returns is retrieved. This takes the form of symmetric, positive, semi-definite covariance matrix:

$$\epsilon = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{1n} & \cdots & \sigma_{nn} \end{bmatrix} \tag{3}$$

And subsequently, the risk is defined for an individual portfolio with asset weights x as:

$$risk = x^T \epsilon x$$

From this, a minimization program can be created:

$$minimize\ x^T \epsilon x$$

$$subject\ to: 0 \leq x \leq e$$
$$e^T x = 1$$

Considering the maximisation and minimisation problem, the two can be combined to create a program that aims to attain a return μ_p whilst minimising the risk of the portfolio

$$minimize\ x^T \epsilon x$$

$$subject\ to: 0 \leq x \leq e$$
$$e^T x = 1$$

$$\mu^T x = \mu_p$$

And conversely, to obtain a risk ${\sigma^2}_p$ whilst maximizing the return:

$$maximise\ \mu^T x$$

$$subject\ to: 0 \leq x \leq e$$
$$e^T x = 1$$

$$x^T \epsilon x = {\sigma^2}_p$$

The two minimization and maximization problems could even be combined to create a minmax 'cost' function:

$$maximise\ \mu^T x - k({\sigma^2}_p), where\ {\sigma^2}_p = x^T \epsilon x \tag{4}$$

$$subject\ to: 0 \leq x \leq e$$

$$0 \leq k$$

$$e^T x = 1$$

$$u^T x = \mu_p$$

Where k (the risk aversion factor) represents a scaling factor for the importance of risk whilst attaining the maximum return for a given level of risk. Adjusting k gives the maximum return at a considered risk position (dependent on k.)

Going forward, these minmax optimization problems will be considered as the Mean-Variance Framework Problems:

Solving these problems can be done via various ways:

1. **Monte-Carlo Simulation:** the Monte-Carlo Model is very computationally heavy and rather than mathematically attempt to find the 'optimal' portfolio, the model runs through all possible solutions of the weightage matrix and outputs returns based on the historical returns of each of the assets stocks prices. The steps are as follows [7]:
   a. Identify the input variables. In this case, it is the weightage of the n assets in the portfolio. This is the variable that the Monte Carlo Model will vary throughout its repeated runs.
   b. Specify probability distributions of the independent variables. Use historical data to define a range of estimated returns and assign a probability distribution to the expected return for each asset.
   c. Using a random number generator (still subject to the weight and budget conditions,) generate random weights for each asset and then predict based on probability distribution of said asset the return obtained after a defined period. This is done multiple times to obtain a representative cumulative probability distribution for all assets, presenting multiple cases of portfolio weights and their respective global returns to the manager.

   The Monte-Carlo Simulation is extremely computationally intensive, and only becomes more representative of how the weights of each asset in a portfolio would affect the overall return and risk with more and more runs of the simulation. Furthermore, since the probability distribution of the expected return must be calculated for each asset based on historical data available, it is still very tedious for a financial manager or computationally expensive for a system.

2. **Analytical Approach**: with the help of Lagrange Multipliers, the minimization problem can be differentiated down to partial differentiations of the Lagrange constants in the

minmax problem. These are then set to 0 (equations now known as optimality conditions) and solved to find values for each Lagrange Multiplier. This is not in the scope of the paper, but still explored in Section B of the Appendix.

Returning to equations (1). The paper defines how the expected returns vector and the estimated covariance matrix are calculated. Markowitz [5] defined the return of an asset across unit time i to be:

$$R_i = \frac{r_{i+1} - r_i}{r_i}$$

Where r represents the stock price (close price in the project's case) of the asset at time i. This is done for all time periods for each of the n stock prices:

$$R_i = \{R_1, R_2, R_3, \dots R_{n-1}\}$$

Thus, the expected return estimator for each asset according to Markowitz:

$$\mu_i = E(R_i) = Mean(R_i)$$

This estimator is purely based on historical sample data of an asset, which is often not very representative of future trends on the return of the asset.

This is then collated in to a vector μ representing expected returns for all n assets:

$$\mu = \{\mu_1, \mu_2, \mu_3, \dots\ \mu_n\}$$

Returning to equation (3) The estimator for risk is considered to be the Covariance of the Returns Matrix:

$$Cov(R) = Cov\{R_1, R_2, R_3, \dots R_{n-1}\}$$

Where:

$$\sigma_{ij} = E\{[R_i - E(R_i)]\left[R_j - E\left(R_j\right)\right]\}$$

Resulting in:

$$\epsilon = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1i} \\ \vdots & \ddots & \vdots \\ \sigma_{j1} & \cdots & \sigma_{ij} \end{bmatrix}, \max(i) = \max(j) = N$$

Where N is the total number of assets. This is used as the risk estimator for subsequent Markowitz optimizations. Once again. The risk estimator is based solely on historical trends. There are many issues related to using the expected returns and covariance based on historical sample data, and this is explored below in section 3.1.2.

3.1.2 Mathematical Background: Limitations of Markowitz Mean Variance Optimization:

While Markowitz's portfolio optimization strategy formed the framework to modern portfolio strategies, There are several limitations associated with it:

Assumption of Normally Distributed Returns: MVO assumes that asset returns are normally distributed. In reality, asset returns often exhibit skewness and kurtosis, meaning they have fat tails and asymmetry that the normal distribution does not capture.

Sensitivity to Input Estimates: The optimal portfolio is highly sensitive to the estimates of expected returns, variances, and covariances. Small errors or changes in these inputs can lead to significantly different portfolios, making the optimization unstable and potentially leading to suboptimal investment decisions. The process of estimating the expected returns, variances, and covariances involves a degree of uncertainty and potential error (estimation risk). This can affect the reliability of the optimized portfolio. Furthermore the estimates for the expected returns and the risk are based purely on historical data which is only representative of historical market trends. Although it is nearly impossible to accurately predict future trends, Markowitz Mean Variance optimization essentially produces an averaged historical trend, which more often than not, is hardly indicative of where an asset will be at some point in the future.

Static Model: MVO is a single-period model and does not take into account the dynamic nature of markets and investments over time. It does not consider changing economic conditions, market trends, or the investor's changing circumstances and preferences.

Considering this, the project evaluates how PCA can make the system more robust, and how Bayesian inference alleviates some uncertainty and provide a more realistic estimate with more importance given to more recent historical trends.

3.1.2 Mathematical Background: Efficient Frontier

Returning to the minimum risk optimization problem mentioned in Section 3.1.1:

$$minimize\ x^T \epsilon x$$

$$subject\ to: 0 \ \leq x \leq e$$
$$e^T x = 1$$

$$\mu^T x = \mu_p$$

Markowitz explained that for every expected portfolio return μ_p, there was a minimum variance that could be achieved. Thus, by varying the desired expected return, the minimum variance achievable could be solved for, and logged and plotted in an efficient frontier plot:

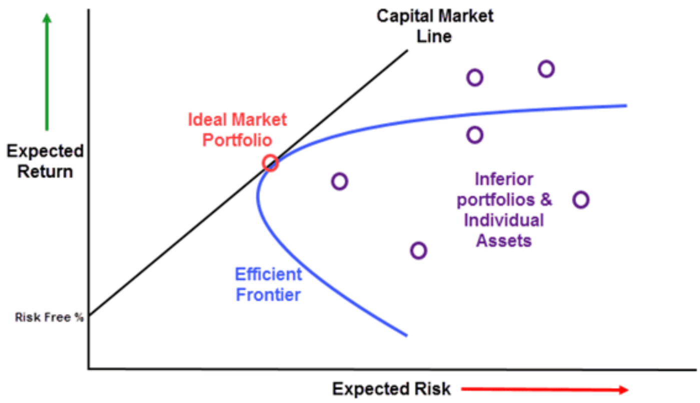


Figure 3: Efficient frontier example [5]

The Efficient frontier line (in blue) in figure 1 represents "efficient" portfolios that offer a minimum risk for an expected return. The ideal market portfolio, or the tangential portfolio, refers to the portfolio that offers the highest Sharpe Ratio. Also note that the Efficient Frontier is typically plotted with the expected return as the y axis and standard deviation representing risk as the x axis. The standard deviation for a particular portfolio is:

$$\sigma_p = \sqrt{{\sigma_p}^2} = \sqrt{x^t \epsilon x}$$

The Sharpe ratio is defined as below:

$$\frac{R_p - r_f}{\sigma_p}$$

Where R_p refers to the expected return of the portfolio and r_f refers to the risk-free rate of a risk free asset. A risk-free asset is defined as an investment with a guaranteed future return and negligible risk of loss. U.S. Department of the Treasury debt obligations, including bonds, notes, and particularly Treasury bills, are considered risk-free due to the backing of the U.S. government. Consequently, the yield on risk-free assets closely aligns with prevailing interest rates, reflecting their high safety and reliability. [6] However, since it is essentially a constant, for the purpose of this paper it is omitted:

$$\frac{R_p}{\sigma_p}$$

This updated sharpe ratio will represent the basis on which an investor will choose his portfolio. Typically investors are of risk averse or return oriented nature. In order to appeal to both profiles, only the portfolio with the highest sharpe ratio will be considered. This portfolio, known as the ideal market portfolio, offers the most return per unit risk to the investor and so should represent the needs of a risk-adjusted returns oriented investor. The next section evaluates how Bayesian Inference can be used to more accurately predict historical trends to alleviate uncertainty concerns in returns/risk.

3.4 Mathematical Background: Bayesian Mean Variance Optimization

Bayesian inference is a method of statistical inference that combines prior knowledge with observed data to update the probability estimates for a prior distrubution. This approach is grounded in Bayes' theorem, which relates current evidence to prior beliefs to form an updated belief.

This technique could be very beneficial to the mean-variance framework. Recall that return and risk estimators are only averaged returns and covariance matrices in Markowitz Mean Variance Optimization. As explained above, this often misrepresents historical trends as it may encompass events that are anomalous due to periods of economic misfortune not directly related to the assets stock price evolution. With Bayesian Inference, the project hopes that as new observed data becomes more available, the prior beliefs, which are the sample return and risk estimators used in Markowitz Mean Variance Optimization (recall that these are merely averaged sample historical data,) are updated or even skewed to more accurately incorporate trends in the new observed data. While this does not necessarily estimate future trends, it does skew the estimators in favor of more recent, and by extension, more relevant historical trends.

This section explores how the Covariance Matrix (risk estimator) and Expected Returns can be estimated based on the presence of new data to be more perceptive of how trends in both returns and risk have evolved given the new data. This should ideally provide less uncertainty in return/risk predictions.

3.4.1 Mathematical Background: Baye's Formula

To begin with, The Baye's Formula is considered:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

The formula states that the Probability of two events happening, $P(A \cap B)$, is the probability of A occurring multiplied by the probability of B given that A has occurred [7]:

$$P(A \cap B) = P(A)P(B|A)$$

However, it is also true for the probability of B multiplied by the probability of A given B occurs:

$$P(A \cap B) = P(B)P(A|B)$$

Therefore:

$$P(B)P(A|B) = P(A)P(B|A)$$

And so:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Or:

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

3.4.2 Mathematical Background: Bayesian Inference

Considering the scope of the paper, Baye's Theorem is rewritten as follows:

$$\mathrm{p}(\mu \mid \mathrm{X}) = \frac{p(\mathrm{X} \mid \mu) \cdot p(\mu)}{p(X)}$$

Where:

$p(\mu)$ represents the probability density function of the prior mean and variance (standard Markowitz Estimators.)

$p(X)$ represents the probability density function of the observed data mean and variance.

$\mathrm{p}(\mu \mid \mathrm{X})$ represents the probability density function of the posterior distribution estimate given the prior distribution and observed distribution that occurs.

$p(\mathrm{X} \mid \mu)$ represents a likelihood function of how probable it is to observe the data given a specific value. It quantifies the compatibility between the observed data and the parameter values.

Since $p(X)$ does not depend on $p(\mu)$ we can rewrite $\mathrm{p}(\mu \mid \mathrm{X}) = \frac{p(\mathrm{X}|\mu) \cdot p(\mu)}{p(X)}$ as:

$$p(\mu \mid \mathrm{X}) \propto p(\mathrm{X} \mid \mu) \cdot p(\mu)$$

Similarly, assuming $p(\mu)$ has data following a multivariate distribution with mean μ_0 and covariance: ϵ_0:

$$p(\mu) = \mathcal{N}(\mu_0, \epsilon_0)$$

Yielding:

$$p(\mu) = \frac{1}{(2\pi)^{d/2} \mid \epsilon_0 \mid^{1/2}} \exp\left(-\frac{1}{2}(\mu - \mu_0)^{\top} \epsilon_0^{-1} (\mu - \mu_0)\right)$$

Where d represents the dimension of $\boldsymbol{\mu}_0$.

As $\frac{1}{(2\pi)^{d/2}|\epsilon_0|^{1/2}}$ does not depend on $\boldsymbol{\mu}$, proportionality can be inferred and $\frac{1}{(2\pi)^{d/2}|\epsilon_0|^{1/2}}$ can be omitted:

$$p(\mu) \propto \exp\left(-\frac{1}{2}(\mu - \mu_0)^{\top} \epsilon_0^{-1} (\mu - \mu_0)\right)$$

Considering $p(\mathrm{X} \mid \mu)$, it is asserted to be the maximum likelihood that the observed data $\mathrm{X} = \{\mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_n\}$ follows an identical and independent distribution equivalent to $\mathcal{N}(\mathrm{x}_i \mid \mu, \epsilon_1)$:

$$p(\mathrm{X} \mid \mu) = \prod_{i=1}^{n} \mathcal{N}(\mathrm{x}_i \mid \mu, \epsilon_1)$$

Expanding using the probability density function for a multivariate normal distribution:

$$p(\mathrm{X} \mid \mu) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{d/2} \mid \epsilon_1 \mid^{1/2}} \exp\left(-\frac{1}{2}(\mathrm{x}_i - \mu)^{\top} \epsilon_1^{-1} (\mathrm{x}_i - \mu)\right)$$

Similar to the process used for $p(\mu)$, proportionality is introduced:

$$p(\mathrm{X} \mid \mu) \propto \prod_{i=1}^{n} \frac{1}{(2\pi)^{d/2} \mid \epsilon_1 \mid^{1/2}} \exp\left(-\frac{1}{2}(\mathrm{x}_i - \mu)^{\top} \epsilon_1^{-1} (\mathrm{x}_i - \mu)\right)$$

Removing the product operator:

$$p(\mathrm{X} \mid \mu) \propto \left(\frac{1}{(2\pi)^{d/2} \mid \epsilon_1 \mid^{1/2}}\right)^{n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(\mathrm{x}_i - \mu)^{\top} \epsilon_1^{-1} (\mathrm{x}_i - \mu)\right)$$

And since $\left(\frac{1}{(2\pi)^{d/2}|\epsilon_1|^{1/2}}\right)^{n}$ depends on neither X nor μ, $\left(\frac{1}{(2\pi)^{d/2}|\epsilon_1|^{1/2}}\right)^{n}$ is omitted:

$$p(\mathrm{X} \mid \mu) \propto \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(\mathrm{x}_i - \mu)^{\top} \epsilon_1^{-1} (\mathrm{x}_i - \mu)\right)$$

Now considering:

$$p(\mu \mid \mathrm{X}) \propto p(\mathrm{X} \mid \mu) \cdot p(\mu)$$

$p(\mathbf{X} \mid \boldsymbol{\mu}) \cdot p(\boldsymbol{\mu})$ can be replaced considering the proportionality equations above:

$$p(\mu \mid \mathrm{X}) \propto \exp\left(-\frac{1}{2}(\mu - \mu_0)^{\top} \epsilon_0^{-1} (\mu - \mu_0)\right) \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\mathrm{x}_i - \mu\right)^{\top} \epsilon_1^{-1} (\mathrm{x}_i - \mu)\right)$$

Combining the exponentials:

$$p(\mu \mid \mathrm{X}) \propto \exp\left(-\frac{1}{2}\left[(\mu - \mu_0)^{\top} \epsilon_0^{-1} (\mu - \mu_0) + \sum_{i=1}^{n}(\mathrm{x}_i - \mu)^{\top} \epsilon_1^{-1} (\mathrm{x}_i - \mu)\right]\right)$$

Consider that $\bar{\mathrm{x}} = \frac{1}{n}\sum_{i=1}^{n} \mathrm{x}_i$, as n represents the number of samples observed. Thus:

$$\sum_{i=1}^{n}(\mathrm{x}_i - \mu)^{\top} \epsilon_1^{-1} (\mathrm{x}_i - \mu) = n(\bar{\mathrm{x}} - \mu)^{\top} \epsilon_1^{-1} (\bar{\mathrm{x}} - \mu)$$

Therefore, the exponential is updated:

$$p(\mu \mid \mathrm{X}) \propto \exp\left(-\frac{1}{2}\left[(\mu - \mu_0)^{\top} \epsilon_0^{-1} (\mu - \mu_0) + n(\bar{\mathrm{x}} - \mu)^{\top} \epsilon_1^{-1} (\bar{\mathrm{x}} - \mu)\right]\right)$$

$$p(\mu \mid \mathrm{X}) \propto \exp\left(-\frac{1}{2}[(\mu-\mu_0)^\top \epsilon_0^{-1}(\mu-\mu_0) + n\mu^\top \epsilon_1^{-1}\mu - 2\mu^\top(n\epsilon_1^{-1}\bar{\mathrm{x}})]\right)$$

Expanding the exponent, grouping terms dependent on μ and summing terms not dependent on μ as constants:

$$p(\mu \mid \mathrm{X}) \propto \exp\left(-\frac{1}{2}[\mu^\top(\epsilon_0^{-1} + n\epsilon_1^{-1})\mu - 2\mu^\top(\epsilon_0^{-1}\mu_0 + n\epsilon_1^{-1}\bar{\mathrm{x}}) + \text{const}]\right)$$

Let $\mathrm{A} = \epsilon_0^{-1} + n\epsilon_1^{-1}$ and $\mathrm{B} = \epsilon_0^{-1}\mu_0 + n\epsilon_1^{-1}\bar{\mathrm{x}}$, and constants removed:

$$p(\mu \mid \mathrm{X}) \propto \exp\left(-\frac{1}{2}[\mu^\top(A)\mu - 2\mu^\top(\mathrm{B})]\right)$$

Which, to complete the square and represent a multivariate normal probability distribution function can be rewritten as:

$$p(\mu \mid \mathrm{X}) \propto \exp\left(-\frac{1}{2}[(\mu - \mathrm{A}^{-1}\mathrm{B})^\top \mathrm{A}(\mu - \mathrm{A}^{-1}\mathrm{B}) + \text{const}]\right)$$

Where the constants representing summed terms not dependent on A or B are removed:

$$p(\mu \mid \mathrm{X}) \propto \exp\left(-\frac{1}{2}[(\mu - \mathrm{A}^{-1}\mathrm{B})^\top \mathrm{A}(\mu - \mathrm{A}^{-1}\mathrm{B})]\right)$$

Thus, the covariance of the posterior distribution =

$$\epsilon_{post} = \mathrm{A}^{-1} = (\epsilon_0^{-1} + n\epsilon_1^{-1})^{-1}$$

And the mean of the posterior distribution =

$$\mu_{post} = \mathrm{A}^{-1}\mathrm{B} = (\epsilon_0^{-1} + n\epsilon_1^{-1})^{-1}(\epsilon_0^{-1}\mu_0 + n\epsilon_1^{-1}\bar{\mathrm{x}})$$

Or:

$$\mu_{post} = \mathrm{A}^{-1}\mathrm{B} = \epsilon_{post}(\epsilon_0^{-1}\mu_0 + n\epsilon_1^{-1}\bar{\mathrm{x}})$$

Thus, the posterior distribution's means and covariance have been inferred via the prior distribution and observed distribution to be used as explained in experimental procedure sections.

There is however, one assumption that this derivation makes that could skew results, or affect the reputability of conclusions made at the end of the experimental process: All data, prior, observed and posterior are assumed to follow a multivariate normal distribution. Setting the prior and observed distributions as multivariate normal limits the risk definition to a single scenario, when in reality another distribution may be better suited to model the returns. This may introduce estimation error and uncertainty once again to proposed portfolio weights obtained by mean variance optimization operations detailed above.

3.4.3 Mathematical Background: Benefits of Bayesian Inference

Despite the potential risk of the assumption made above, utilizing Bayesian theory to infer a posterior distribution estimation has the following benefits:

Incorporation of Prior Knowledge: Bayesian inference allows for the integration of prior beliefs or knowledge into the analysis, which is beneficial when data is limited or historical data may not fully represent current conditions. In contrast, Markowitz optimization relies primarily on historical data without explicitly incorporating subjective beliefs or external information. For the projects purpose, the prior beliefs will be taken as the mean and variance of assets of the 5 year period preceding the new observed data.

Flexibility with Small Samples: Bayesian methods often yield more reliable estimates with smaller sample sizes compared to frequentist methods, such as Markowitz optimization. This advantage is particularly pronounced when informative priors are available or when data is sparse. Markowitz optimization can be less robust with smaller datasets due to its reliance on historical data quality and quantity.

Quantification of Uncertainty: Bayesian inference inherently quantifies uncertainty by providing a posterior distribution for the parameters of interest. This capability is crucial in financial decision-making to assess the range of potential outcomes and their associated probabilities. In contrast, Markowitz optimization typically provides a single optimal portfolio without directly addressing uncertainty unless additional methods or assumptions are applied.

This project will use Bayesian inference via the formulas derived above, and evaluate how portfolio generation was impacted, and how its subsequent performance on the dataset was affected.

3.5.1 Principal Component Analysis

Principal Component Analysis (PCA) is a robust statistical method applied in portfolio optimization to simplify financial data by reducing its dimensionality, uncover the key factors influencing asset returns, and enhance risk management.

This project leverages PCA on a returns matrix, in an attempt to reduce noise represented by un-trendlike changes in the returns of the asset. This should make the covariance matrices and expected returns between periods of asset returns be more resistant to small changes in the return matrix and more focused on the more important variations of returns. This would allow for underlying trends to be identified more accurately by the estimators for use in the Mean-Variance Framework.

The next section details how PCA is used to denoise a matrix:

3.5.2 Principal Component Analysis

The following steps were used to denoise a matrix:

PCA is typically performed on standardized data. To achieve this, the mean of each variable (column) is subtracted from the dataset to center it around the origin. Let X be the original data matrix with n observations and p variables. The standardized data matrix X_s is computed as:

$$\mathrm{X}_s = \mathrm{X} - \mathrm{e}_n \mu^T$$

where e_n is an $n \times 1$ vector of ones and μ is the mean vector of the columns of X.

Then, the Covariance Matrix C of the Standardized data is calculated:

$$\mathrm{C} = \frac{1}{n-1} \mathrm{X}_c^T \mathrm{X}_c$$

Post which, Eigen Decomposition was performed to obtain the eigenvalues and eigenvectors of the return matrix:

$$\mathrm{C} = \mathrm{V} \Lambda \mathrm{V}^T$$

Where V is a matrix of C's eigenvectors and Λ is the diagonal matrix of eigenvalues. The diagonal is taken as it is asserted that there is no correlation between the eigenvectors.

From this, to retain a certain amount of variance explained by each eigenvector's associated eigenvalue, the top k eigenvectors that explain a cumulative variance threshold are retained:

$$\mathrm{V}_k = \text{matrix of eigenvectors explaining a certain threshold of variance.}$$

Using this new matrix of eigenvectors, the standardized data matrix is projected to the reduced eigenvector space:

$$\mathrm{Z}_k = \mathrm{X}_s \mathrm{V}_k$$

Where Z_k is he matrix of k principal components representing a threshold value of total variance.

However, the mean variance framework used to optimize portfolios expects a matrix with dimensions equivalent to the returns matrix. Thus, the matrix Z_k is projected back to the original space using the retained eigenvectors:

$$\mathrm{X}_s^{\text{denoised}} = \mathrm{Z}_k \mathrm{V}_k^T$$

Finally, the standardization is reversed by adding the mean vector μ back to the reconstructed data $\mathrm{X}_s^{\text{denoised}}$:

$$\mathrm{X}^{\text{denoised}} = \mathrm{X}_s^{\text{denoised}} + \mathrm{e}_n \mu^T$$

To summarize, the denoised data matrix $\mathbf{X}^{\text{denoised}}$ is obtained by projecting the centered data onto the top k principal components, reconstructing the data in the reduced space, and then adding back the mean vector. This process filters out noise by removing the components associated with smaller eigenvalues, which typically correspond to less significant patterns or noise in the data.

3.5.3 Benefits of PCA denoising:

There are many theoretical advantages to denoising a returns matrix with PCA:

Improved Risk Management:

- Noise Reduction: PCA denoising helps in removing noise from the return series of assets, leading to more stable and robust covariance matrices. This reduction in noise allows for better estimation of risks associated with the assets.
- Stable Covariance Matrix: A more stable covariance matrix improves the reliability of portfolio optimization models, particularly those that are sensitive to the estimation of covariances, such as Mean-Variance Optimization (MVO).

Enhanced Signal Extraction:

- Focus on Principal Components: By concentrating on the most significant principal components, PCA denoising enhances the extraction of true underlying signals from the data. This helps in identifying the main factors driving asset returns, which can be critical for strategic asset allocation.

- Reduction of Spurious Correlations: PCA helps in filtering out random noise and spurious correlations that do not contribute to the true underlying structure of the data, leading to more accurate asset relationships and better decision-making.

3.6.1: Combining Bayesian Inference and PCA

Via Bayesian Inference, the project hopes to generate estimators of the expected returns and covariance risk matrix that are more representative of more recent historical trends, as evidenced by the formulas derived in section 3.4.2. This should make the estimators more robust and alleviate uncertainty, as instead of an averaged sample estimator in the case of Markowitz estimators, Bayesian estimators are constantly learning and updating themselves based on the presence of new observed trends.

This can also be combined with PCA. As the project considers monthly returns over 10 years for 10 assets, there are exactly 1200 data points to consider. Some of these points however, may be anomalous and unrelated to driving factors under return trends. Using PCA and setting a threshold variance level, the anomalous reading can be filtered out. This should provide a more robust estimator for the risk and return as the filtered values are more akin to establishing a market trend.

The Bayesian inference can be split into two: the prior estimators and the observed estimators. From these estimators, a posterior distrubution is retrieved, and from it, the posterior estimators. In the experiment design below, the project considers the prior estimator to be akin to Markowitz estimators: A 5 year window before the observed data. The observed data will be represented by the year just before the buy period. Thus, in effect, the Bayesian inference also benefits from one additional year of data, as Markowitz estimators would only consider the 5 year period before the buy point.

This raises the question however, where can PCA be implemented? It was decided that PCA would be used to filter the prior returns matrix, and not the observed matrix as the observed matrix is only 10x10x12 data points long. Thresholding the variance there may result in key data points being lost, however, this also raises the concern as to whether without filtering, the observed data year with anomalous data points would greatly vary the posterior distrubution. This is something that is evaluated in the experimental results.

Also of note is the risk estimator takes the form of a 10x10 matrix. As the experiments use 4 different types of estimator generation, at any point, there would be four 10x10 covariance matrices to compare. This would be time and space consuming, so recalling the risk definition in the Markowitz framework:

$$risk = x^T \epsilon x$$

It can be asserted that the risk consists of added up covariances and variances of each asset scaled by its weightage. Thus, only the sum of rows of the covariance matrix will be considered, and called the cumulative variance and by extension, the cumulative risk of an asset.

The next section explores the experiments proposed, and their design.

4.1 Experiment Aims and Design:

Considering the above mathematical background, many questions were posed as to the effectiveness of these mathematical techniques on the returns and risk of generated portfolios:

1. How would PCA subsampling on the Returns Matrix affect the output portfolio weights of the mean-variance optimization? Would it increase estimated/actual returns and decrease estimated/actual risk due to a reduction in noise?
2. How would Bayesian inferred posterior expected returns and risk estimators be different to standard Markowitz estimators? would it be beneficial from an investor looking to reduce risk and increase returns given limited financial data?
3. Can PCA subsampling and Bayesian Inference be combined to affect expected risk and return estimators and ensure that inputs into the mean-variance optimization framework ensure that the algorithm is more robust and provides a more accurate estimate of trends in returns and volatility of each asset? Post which, In a rolling investment scenario, can doing this increase the average monthly returns and decrease the risk?

Considering the above questions, two experiments were designed:

4.2 Experiment 1 Aims and Design:

This experiment will consider an investor looking to maximize the sharpe ratio of his portfolio over the year 2016 to 2017. Available to him are the returns of assets from the beginning of 2010 to the end of 2016.

This experiments analysis are not concerned with the potential return/risk of portfolios generated in this time frame, but rather the behaviour of genesis of parameter estimators via the four methods detailed below:

The following investment strategies will be evaluated:

1. Derivation of portfolio weights via Mean-Variance Optimization with estimators for expected returns and risk being sampled from the returns of the assets over 2011 to 2016. This follows Markowitz's portfolio theory and serves as the baseline for the experiment.
2. Derivation of portfolio weights via Mean-Variance Optimization with estimators for expected returns and risk being sampled from the PCA denoised returns of the assets over 2011 to 2016. This is similar to Markowitz' portfolio theory but evaluates the impact of PCA denoising over various cumulative variance (of Principal Components) thresholds on the expected returns and risk estimators. The cumulative variance thresholds investigated will be 95%, 90% and 85%.
3. Derivation of portfolio weights via Mean-Variance Optimization with estimators for expected returns and risk being sampled from the returns of the assets over 2010 to 2015, and then updated via Bayesian inference with 2016's expected returns and risk. This will evaluate the impact of Bayesian inference on the portfolio weights, and the expected returns and risk.
4. Derivation of portfolio weights via Mean-Variance Optimization with estimators for expected returns and risk being sampled from the PCA denoised returns of the assets over 2010 to 2015, and then updated via Bayesian inference with 2016's expected returns and risk. This will evaluate the impact of Bayesian inference on the portfolio weights, and the expected returns and risk. The threshold for the cumulative variance will be the best performing threshold from the 2nd Derivation.

It will be assumed that the investor wishes to maximize the returns/risk ratio $\frac{R_p}{\sigma_p}$ of the portfolio, and has 1000$ to spread over assets as suggested by derived portfolio weights via the methods above. The evolution of the 1000$ will be tracked across the monthly returns for 2016-2017.

4.3 Experiment Aims and Design:

To emulate typical SIP funds, this experiment will focus on how the generated parameters are varied in a dynamic management scenario.

Consider this timeline:

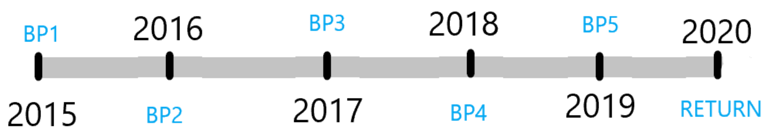


Figure 4: timeline for buypoints Experiment 2

Figure 4 refers to a 5 year dynamic scenario where an investor invests in the beginning of 2015 based on initial portfolio weights given data, and utilizing the same capital under gain or loss, reshuffles the portfolio weights at BP2, BP3, BP4 and BP5 based on that points generated portfolio weights which are affected by variations in historical data. The table below details what historical data will be available to the algorithms:

Buy Point	**BP**	**Markowitz**	**PCA**	**Bayesian**	**Bayesian + PCA**
2015	**BP1**	2010-2015	2010-2015	2010-2015	2010-2015
2016	**BP2**	2011-2016	2011-2016	2010-2015, 2016	2010-2015, 2016
2017	**BP3**	2012-2017	2012-2017	2011-2016, 2017	2011-2016, 2017
2018	**BP4**	2013-2018	2013-2018	2012-2017, 2018	2012-2017, 2019
2019	**BP5**	2014-2019	2014-2019	2013-2018, 2019	2013-2018, 2019

Figure 5: buy points and data available

Here, the Buy Points Correspond to the year over which the portfolio weights are held, and the last 4 sections detail the beginning of the year to the beginning of the end year over which information is available. For instance, 2014-2019 would represent information available from 01/01/2014 till 01/01/2019.

Similar to experiment 1, the following portfolio derivations will be assessed:

1. Derivation of portfolio weights via Mean-Variance Optimization with estimators for expected returns and risk being sampled from the returns of the assets over data available corresponding to each buy point. This follows Markowitz's portfolio theory and serves as the baseline for the experiment.
2. Derivation of portfolio weights via Mean-Variance Optimization with estimators for expected returns and risk being sampled from the PCA denoised returns of the assets over information available at the corresponding buy point. This is similar to Markowitz' portfolio theory but uses PCA at a threshold decided by experiment 1 to denoise the returns matrix.
3. Derivation of portfolio weights via Mean-Variance Optimization with estimators for expected returns and risk being sampled from the returns of the assets over information comprising of the prior buy periods 'Markowitz's information, representing the prior belief of estimators. This is then updated via Bayesian inference with the data of the year immediately preceding the buy point (observed data.) From this, a posterior estimator will be derived and used to infer portfolio weights

4. Derivation of portfolio weights via Mean-Variance Optimization with estimators for expected returns and risk being sampled from the returns of the assets over information comprising of the prior buy periods 'Markowitz's information, but denoised with PCA at experiment 1's chosen threshold, representing the prior belief of estimators. This is then updated via Bayesian inference with the data of the year immediately preceding the buy point (observed data.) From this, a posterior estimator will be derived and used to infer portfolio weights

It will be assumed that the investor wishes to maximize the returns/risk ratio $\frac{R_p}{\sigma_p}$ of the portfolio, and has 1000$ to spread over assets as suggested by derived portfolio weights via the methods above. The evolution of the 1000$ will be tracked across each successive buy point, with the capital for the next buy point being the capital with excess returns from the evolution of returns during the previous buy period.

5.1.1 Experiment 1 Results:

This section will go over portfolio weight selection and its subsequent performance for the buy period 2016-2017.

MATLAB's quadprog() function was used to replicate the Mean-Variance framework and generate weights based on an inputted matrix corresponding to the returns of the data.

For the PCA component, MATLAB's pca() function was used to reduce the returns matrix to a user chosen threshold variance. This reduced returns matrix was then directed to the quadprog() function above

For the Bayesian inference component, the returns matrix was split into corresponding prior and observed data, and then from the formulas established in section 3.4.2, the posterior distribution was inferred, and posterior estimators were obtained. This was fed into the Mean-Variance function described above.

For the PCA reduced Bayesian Inference component,

To begin with, Markowitz Mean Variance Optimization is considered:

6.1.2 Markowitz Optimized weights for buy period '16-'17

Table x presents the expected returns estimators and cumulative risk for each asset from the associated covariance matrix. This table is the average returns and risk for 2011-2016.

Parameter	LMT	MSFT	APPL	AMZN	V	MCD	WFC	CMCSA	MS	C

mean	0.019963	0.013409	0.016312	0.025314	0.026128	0.007869	0.010189	0.017338	0.007566	0.005232
Cum Var	0.009293	0.017052	0.014637	0.016514	0.011076	0.006207	0.012561	0.016765	0.032962	0.029825
Cum Risk	0.096402	0.130583	0.120984	0.128507	0.105241	0.078782	0.112078	0.12948	0.181555	0.1727
return/risk	0.207084	0.102682	0.134828	0.196988	0.24827	0.09988	0.090913	0.133902	0.041675	0.030295

From this, the Mean-Variance framework detailed in Section 5 returned the following portfolio weights:

	LMT	**MSFT**	**APPL**	**AMZN**	**V**	**MCD**	**WFC**	**CMCSA**	**MS**	**C**
Marko	**0.4014**	**0.001**	**0.0368**	**0.1238**	**0.4366**	**0.0003**	**0**	**0**	**0**	**0**

This portfolio weight spread came with an expected return of 0.231 and expected risk 0.0367 yielding an expected sharpe ratio of 0.6312.

These weights suggest that splitting mainly between LMT and V would be the most beneficial, with the both of them comprising 0.8380 of the budget weight. This tracks with the return/risk ratio above in table x, with both of them having a ratio of 0.208 and 0.248 respectively.

The performance of the portfolio weights are as follows. Please note that the full progression of each individual asset can be found in sheet 2 of excel main.xlsx in the github:

Price	Marko
31/12/2015	1000
01/01/2016	952.2609792
01/02/2016	942.6569987
01/03/2016	988.1702331
01/04/2016	1019.214223
01/05/2016	1049.66265
01/06/2016	1042.038925
01/07/2016	1083.325713
01/08/2016	1083.89524
01/09/2016	1102.935778
01/10/2016	1105.684577
01/11/2016	1103.124385
01/12/2016	1080.529731

Returns	Marko
01/01/2016	-0.04774
01/02/2016	-0.01009
01/03/2016	0.048282
01/04/2016	0.031416
01/05/2016	0.029874
01/06/2016	-0.00726
01/07/2016	0.039621
01/08/2016	0.000526
01/09/2016	0.017567
01/10/2016	0.002492
01/11/2016	-0.00232
01/12/2016	-0.02048
mean	0.006824
var	0.000763
risk	0.027626

Thus, with an initial investment of 1000 dollars at the beginning of 2016, the investor would be left with 1080.52 dollars, a return of 8.052%. The average monthly returns were found to be 0.0068 with a risk (represented as standard deviation) of 0.0263, yielding a sharpe performance ratio of 0.246, which was very different to the expected 0.6312. This is now considered the baseline for further results in the experiment.

6.1.3 PCA reduced weights for buy period '16-'17

This section examines the impact of PCA at threshold variance levels of 95%, 90% and 85%. To begin with, the effect of PCA is examined on the returns of a single asset: AMZN. A single asset was considered as much like figure 2 in Section 2.1, the entire plot of all asset returns would be far too visually complex.

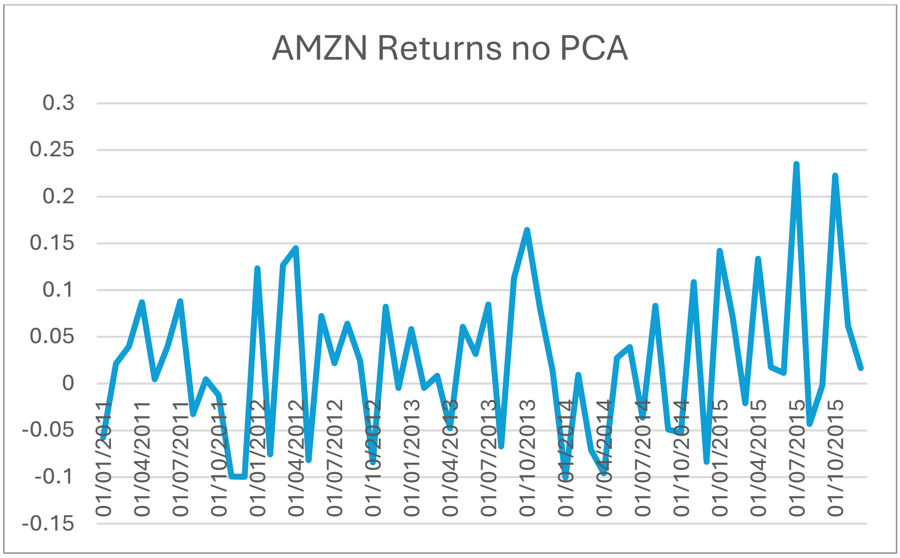


Figure 6: AMZN returns 11-16 no PCA

Figure 6 presents the returns for Amazon over the beginning of 2011 to the beginning of 2016. Not much can be inferred but on a visual basis, no observable trend can be distinguished.

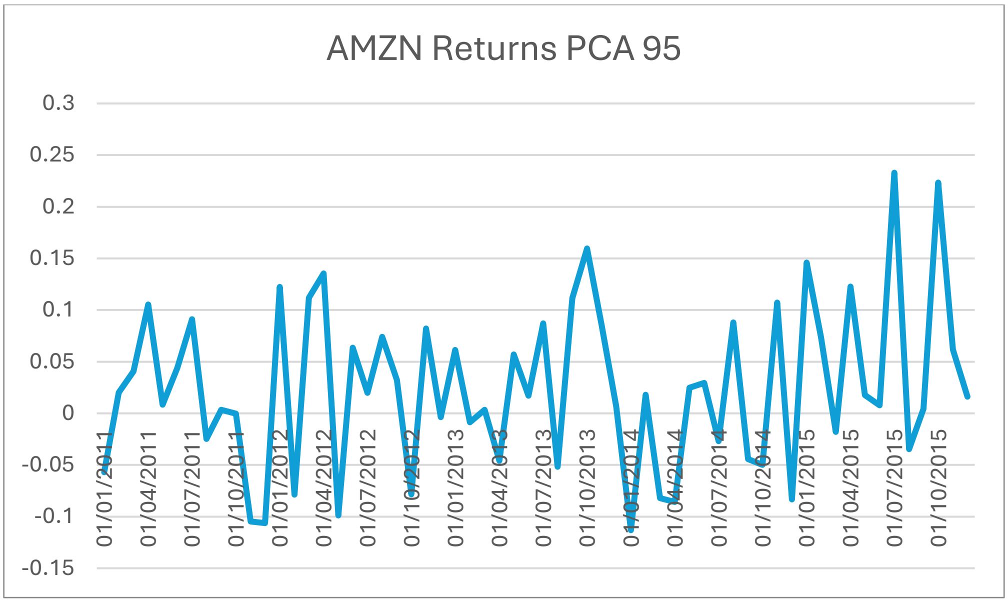


Figure 7: AMZN returns 11-16 PCA 95

At a threshold level of 95%, PCA begins to have some effect. The values of returns are not considered here as the differences were found to be minute, but the first two peak between 01/2011 and 10/2011 show some variance, the first peak is increased and the second is decreased, suggesting that there were both anomalous positive returns during the first year: 2011. Also to be considered is that PCA was done on the returns matrix as a whole, and AMZN is only one of ten assets. These plots are just mainly visual references to see the effect PCA on the returns matrix has on a singular asset.

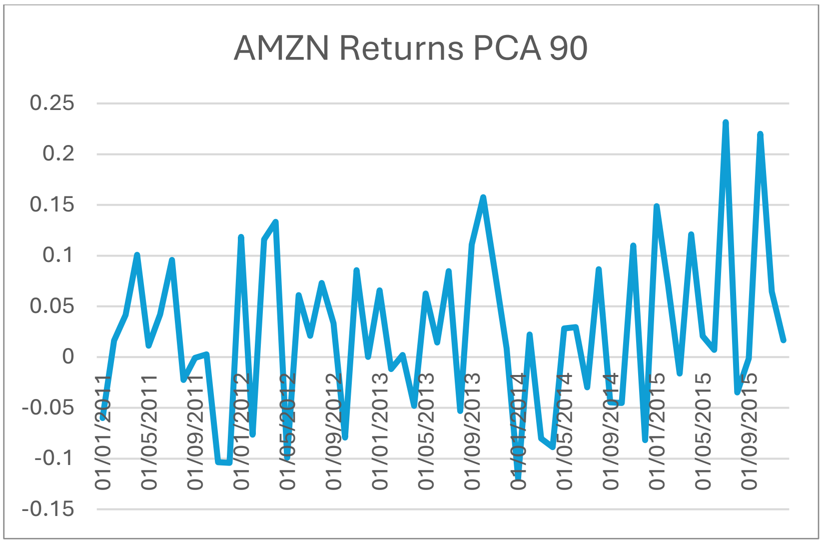


Figure 8: AMZN returns 11-16 PCA 90

At PCA 90, there was no visual correspondence of any return changes across 01/2011 to 01/2016.

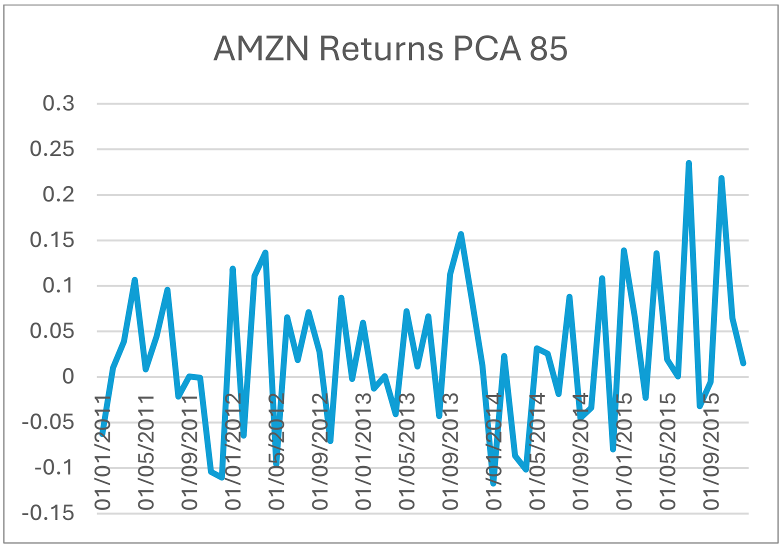


Figure 9: AMZN returns 11-16 PCA 85

At PCA 85, there were some changes seen from PCA 90. Most notably, the peaks at 01/2015 to 05/2015 were attenuated and brought in line with each other, implying the return at 01/2015 was anomalous.

Evaluating the change in data is another interesting aspect of the effectiveness of PCA.

Table 2: PCA 95

Parameter	LMT	MSFT	APPL	AMZN	V	MCD	WFC	CMCSA	MS	C
mean	0.019963	0.013409	0.016312	0.025314	0.026128	0.007869	0.010189	0.017338	0.007566	0.005232
cum var	0.00921	0.017085	0.014449	0.016661	0.011062	0.006105	0.012879	0.016895	0.032637	0.029499
cum risk	0.095968	0.130708	0.120203	0.129077	0.105174	0.078133	0.113484	0.12998	0.180658	0.171753
return/risk	0.20802	0.102583	0.135704	0.196118	0.248427	0.100709	0.089786	0.133387	0.041882	0.030462

Comparing this with the earlier Markowitz estimators:

Parameter	LMT	MSFT	APPL	AMZN	V	MCD	WFC	CMCSA	MS	C
mean	0.019963	0.013409	0.016312	0.025314	0.026128	0.007869	0.010189	0.017338	0.007566	0.005232
Cum Var	0.009293	0.017052	0.014637	0.016514	0.011076	0.006207	0.012561	0.016765	0.032962	0.029825
Cum Risk	0.096402	0.130583	0.120984	0.128507	0.105241	0.078782	0.112078	0.12948	0.181555	0.1727
return/risk	0.207084	0.102682	0.134828	0.196988	0.24827	0.09988	0.090913	0.133902	0.041675	0.030295

Not much change is inferred, however, the cumulative variances seem to have been affected, increasing across the board except in some cases.

Below are the estimator tables for PCA 90 and 85:

Table 3: PCA 90

Parameter	LMT	MSFT	APPL	AMZN	V	MCD	WFC	CMCSA	MS	C
mean	0.019963	0.013409	0.016312	0.025314	0.026128	0.007869	0.010189	0.017338	0.007566	0.005232
cum var	0.009238	0.017125	0.014429	0.016671	0.011018	0.006103	0.012849	0.016968	0.032584	0.029484
cum risk	0.096113	0.130864	0.120119	0.129115	0.104966	0.078122	0.113354	0.130263	0.18051	0.17171
return/risk	0.207706	0.102462	0.135799	0.19606	0.248919	0.100723	0.089889	0.133097	0.041916	0.03047

Table 4: PCA 85

Parameter	LMT	MSFT	APPL	AMZN	V	MCD	WFC	CMCSA	MS	C
mean	0.019963	0.013409	0.016312	0.025314	0.026128	0.007869	0.010189	0.017338	0.007566	0.005232
cum var	0.009237	0.017142	0.014425	0.016667	0.011029	0.006094	0.012845	0.016966	0.032578	0.029486
cum risk	0.096108	0.130926	0.120104	0.1291	0.10502	0.078066	0.113336	0.130253	0.180493	0.171716
return/risk	0.207718	0.102413	0.135816	0.196083	0.248792	0.100796	0.089903	0.133107	0.04192	0.030468

And correspondingly, the portfolio weights generated for PCA 95, 90, 85:

Table 5: PCA portfolio weights 16-17

	LMT	MSFT	APPL	AMZN	V	MCD	WFC	CMCSA	MS	C
Marko	0.4014	0.001	0.0368	0.1238	0.4366	0.0003	0	0	0	0
PCA 95	0.4011	0.0005	0.0381	0.1218	0.4384	0.0002	0	0	0	0
PCA 90	0.3665	0	0.0397	0.1167	0.4771	0	0	0	0	0
PCA 85	0.2992	0.0439	0.0283	0.0772	0.5513	0	0	0	0	0

As seen above, as the PCA threshold increased, weights for LMT decreased from 0.4014 down to 0.2992 (Markowitz weights are essentially PCA 100.) This is interesting as though the expected return didn't change, the cumulative risk did, going from 0.09597 to 0.096109. This corresponded to a drop in return risk ratio from 0.20802 to 0.207718. Conversely, V had an increasing risk/return ratio corresponding to a decrease in cumulative risk going from 0.105241 to 0.105200. this resulted in a increase of the risk/return ratio from 0.24827 to 0.248792, resulting in an increase of portfolio weight from 0.4366 to 0.5513.

Below are the tracked investment for Markowitz, PCA 95, PCA 90, PCA 85:

Table 6: returns Markowitz, PCA, 16-17

Price	Marko	PCA 95	PCA 90	PCA 85
31/12/2015	1000	**1000.1**	1000	999.9
01/01/2016	952.2609792	**952.562407**	952.4595925	957.0818289
01/02/2016	942.6569987	943.0405868	941.4259716	941.3085782
01/03/2016	988.1702331	988.6247082	987.9644388	989.6975094
01/04/2016	1019.214223	1019.335434	1016.638336	1009.282764
01/05/2016	1049.66265	1049.682114	1046.815364	1038.013533
01/06/2016	1042.038925	1041.91773	1034.776861	1017.14295
01/07/2016	1083.325713	1083.21943	1077.048113	1062.6167
01/08/2016	1083.89524	1083.855963	1080.608987	1071.736946
01/09/2016	1102.935778	1102.825938	1100.585719	1089.818453
01/10/2016	1105.684577	1105.689155	1102.642543	1094.187472
01/11/2016	1103.124385	1103.05254	1094.645113	1078.265911
01/12/2016	1080.529731	1080.543796	1074.989891	1064.979886

Interestingly, PCA 95 performed only extremely marginally better than Markowitz, whilst PCA 90 and PCA 95 both performed worse with returns of 1074.99 and 1064.98 respectively. It should be noted there were some slight rounding errors, namely in PCA.95 and PCA 85 both of which went over and under the 1000$ budget by 10 cents. This immediately discredits PCA 95 performing better than PCA as the net gain was only 2 cents over Markowitz, however

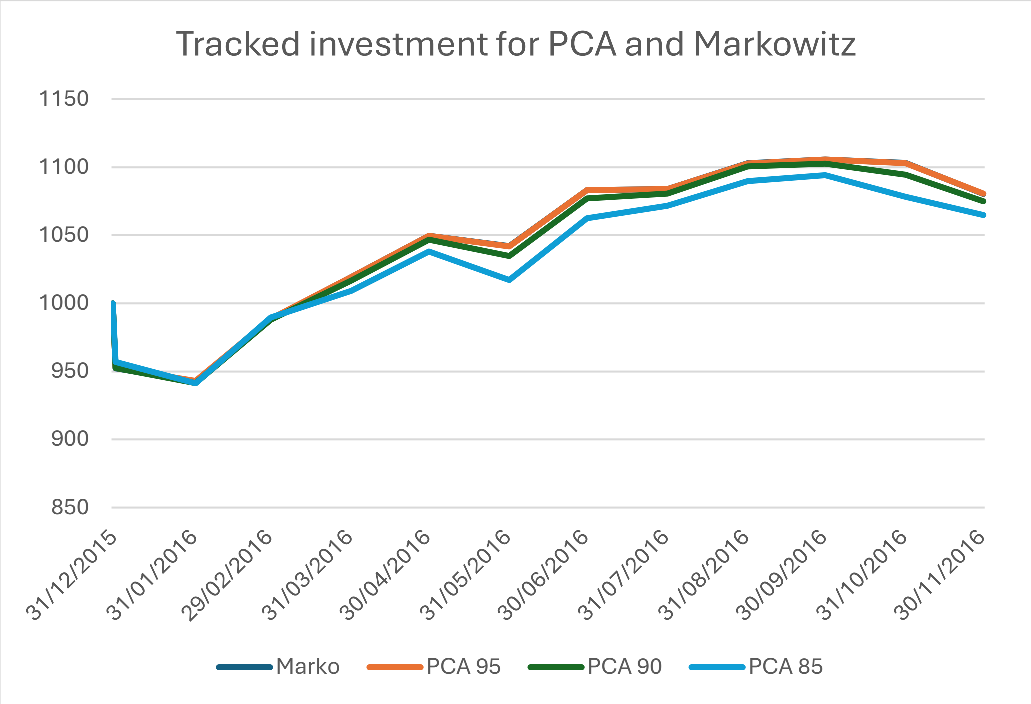
The expected returns, risk and sharpe ratio taken from the mean-variance framework were as follows:

	Expected Return	**expected risk**	**expected sharpe**
Marko	0.023173	0.036706	0.631313682
PCA 95	0.023173	0.036673	0.631881766
PCA 90	0.023384	0.036916	0.633438076
PCA 85	0.023384	0.036106	0.64764859

And the subsequent monthly return performance was as follows:

Returns	Marko	PCA 95	PCA 90	PCA 85
01/01/2016	-0.047739021	-0.04753284	-0.04754	-0.04282
01/02/2016	-0.01008545	-0.009996007	-0.01158	-0.01648
01/03/2016	0.048281861	0.048337391	0.049434	0.051406
01/04/2016	0.03141563	0.031064089	0.029023	0.019789
01/05/2016	0.029874414	0.029771044	0.029683	0.028467
01/06/2016	-0.007263024	-0.007396891	-0.0115	-0.02011
01/07/2016	0.039621157	0.039640078	0.040851	0.044707
01/08/2016	0.000525721	0.00058763	0.003306	0.008583
01/09/2016	0.01756677	0.017502302	0.018487	0.016871
01/10/2016	0.002492256	0.002596255	0.001869	0.004009
01/11/2016	-0.002315481	-0.00238459	-0.00725	-0.01455
01/12/2016	-0.020482418	-0.020405868	-0.01796	-0.01232
mean	**0.006824368**	**0.006815216**	**0.006402**	**0.005629**
var	**0.00076318**	**0.000759207**	**0.000782**	**0.000792**
risk	**0.02762572**	**0.027553704**	**0.027964**	**0.02814**
return/risk	**0.247029503**	**0.247343012**	**0.228921**	**0.20004**

Once again, as was the case with Markowitz estimators, the realised sharpe ratios (return/risk) were significantly lower than the expected sharpe ratios. PCA 95 was the best performing sharpe ratio portfolio, with a ratio of 0.24734 as opposed to Markowitz's 0.24702. As such, it will be taken forward for subsequent combination with Bayesian Inference.



Above is the evolution of the 1000 dollar investment visualized over 2016-17. This trend line is relatively similar between all portfolios with the exception of PCA 85, which seems to deviate and project lower post 05/2016.

6.1.4 Bayesian inferred weights for buy period '16-'17

Now, Bayesian Inference is examined for the buy period 2016-17. It is different in approach as it uses the 5 year period 2010-2015 beginning to represent prior beliefs, and then the entirety of 2015 as the observed to infer new estimators based on updated beliefs.

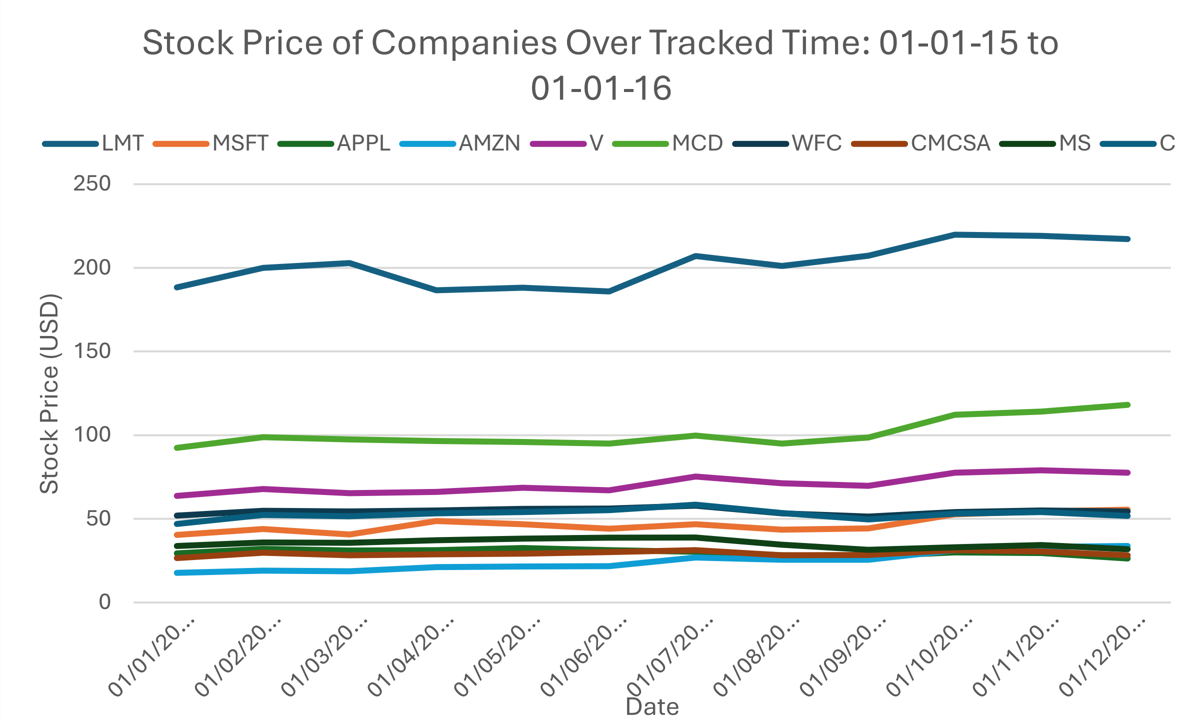
Thus, the standard estimators for 2010-2015 are examined:

Parameter	LMT	MSFT	APPL	AMZN	V	MCD	WFC	CMCSA	MS	C
mean	0.017149	0.010252	0.026419	0.018542	0.021395	0.00746	0.01276	0.024005	0.01157	0.012352
cum var	0.011457	0.019086	0.017583	0.019155	0.013095	0.004846	0.020259	0.02062	0.036888	0.034018
cum risk	0.107036	0.138151	0.132601	0.138403	0.114435	0.069614	0.142333	0.143597	0.192063	0.184439
return/risk	0.160214	0.07421	0.19924	0.133973	0.186962	0.107161	0.08965	0.167171	0.060239	0.066971

And the Observed data from 2016-17:

Parameter	LMT	MSFT	APPL	AMZN	V	MCD	WFC	CMCSA	MS	C
mean	0.011216	0.019459	-0.00216	0.070607	0.015601	0.020595	6.47E-05	-0.00015	-0.01426	-0.00132
cum var	0.015398	0.042469	0.020078	0.035687	0.02832	0.021553	0.019716	0.032835	0.027677	0.033115
cum risk	0.124087	0.20608	0.141699	0.18891	0.168287	0.146811	0.140414	0.181203	0.166363	0.181976
return/risk	0.090384	0.094427	-0.01527	0.373759	0.092702	0.14028	0.000461	-0.00081	-0.08571	-0.00724

Notice in the Observed data, APPL, CMCSA, MS and C log negative monthly expected returns:



Though the above plot all show all stock prices increasing from the start point in 01/15 to the end point in 12/15, the middle period from 04/2015 to 08/2015 experienced many drops. This is further corrobated upon examination of sheet 2 in excel main.xlsx in the github.

2015 was a very tumultuous year for the NYSE market, driven by multiple factors. From the guardian, the Greece debt default seemed to be a major driving factor, requiring US government bailout [10].

Nevertheless, it would be interesting to see how given this new information the posterior predictors derivation.

Recall from section 3.4.2:

$$\mu_{post} = \mathrm{A}^{-1}\mathrm{B} = (\epsilon_0^{-1} + n\epsilon_1^{-1})^{-1}(\epsilon_0^{-1}\mu_0 + n\epsilon_1^{-1}\bar{\mathrm{x}})$$

Or:

$$\mu_{post} = \mathrm{A}^{-1}\mathrm{B} = \epsilon_{post}(\epsilon_0^{-1}\mu_0 + n\epsilon_1^{-1}\bar{\mathrm{x}})$$

Where ϵ_{post} and μ_{post} are inferred from 2010-2015 prior distribution μ_0, ϵ_0 and observed distribution μ_1, ϵ_1.

Considering such, the following estimators where inferred:

Parameter	LMT	MSFT	APPL	AMZN	V	MCD	WFC	CMCSA	MS	C
mean	0.008756	0.012872	-0.00093	0.062665	0.011927	0.017099	-9.09E-04	-0.00121	-0.01385	-0.00251
cum var	0.000982	0.002601	0.001386	0.002269	0.001847	0.001314	0.001333	0.002155	0.001938	0.002244
cum risk	0.031344	0.050995	0.037235	0.04763	0.042971	0.036248	0.036508	0.046419	0.044028	0.047366
return/risk	0.27935	0.252422	-0.02488	1.315668	0.277568	0.471733	-0.0249	-0.0261	-0.31456	-0.05307

This was a very curious result, as it heavily favours AMZN for investment. The belief of return for AMZN was 0.062665, dwarfing the next best LMT at 0.008756. the return/risk was massive for AMZN, projecting 1.31 monthly returns. Predictably, the weightage outputted was as follows.

	LMT	MSFT	APPL	AMZN	V	MCD	WFC	CMCSA	MS	C
Bayesian	0	0	0	1	0	0	0	0	0	0

Everything in AMZN. However, this goes against the diversification principle that Markowitz put forth. If AMZN did crash, then the entire investment would be lost. The cumulative risk also projected is still relatively high at 0.04763 compared to LMT's 0.031344. This is potentially a limitation of using the sharpe ratio to select a portfolio. In such cases, where extreme portfolio weights are skewed due to estimators predicting massive returns for a particular asset, it might require investor intervention to spread the budget across the assets and not just chance it on one.

Still, in the interest of finding out the evolution of the budget with this weightage, the returns for 2016-17 were found via excel:

Price	Marko	Bayesian
31/12/2015	1000	1000
01/01/2016	952.2609792	868.4845681
01/02/2016	942.6569987	817.4703172
01/03/2016	988.1702331	878.3086229
01/04/2016	1019.214223	975.8837075
01/05/2016	1049.66265	1069.390053
01/06/2016	1042.038925	1058.781758
01/07/2016	1083.325713	1122.682692
01/08/2016	1083.89524	1137.995895
01/09/2016	1102.935778	1238.825947
01/10/2016	1105.684577	1168.563031
01/11/2016	1103.124385	1110.491418

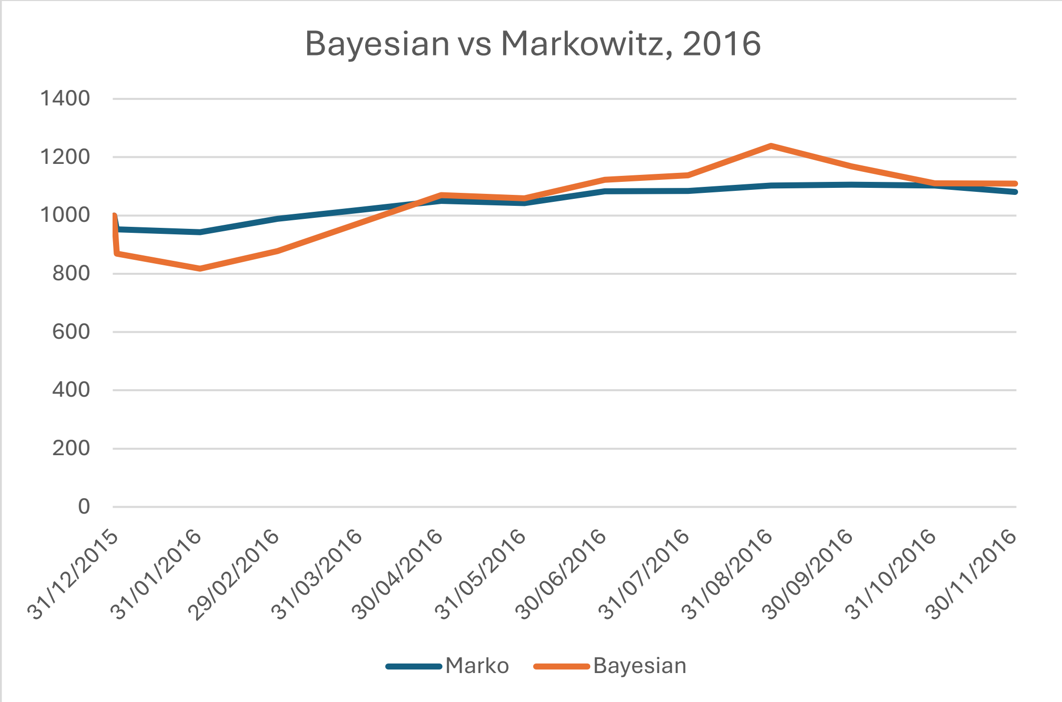
01/12/2016	1080.529731	1109.455746

From the above table, the portfolio performed much better, realising 109.46 dollars in profit over Markowitz's 80.53. the return of the year for the Bayesian portfolio was 10.95% compared to 8.05% achieved by the Markowitz portfolio. However, hindsight is also 20/20, and upon comparing the month on month returns however, a different picture was painted:

Returns	Marko	Bayesian
01/01/2016	-0.04774	-0.13152
01/02/2016	-0.01009	-0.05874
01/03/2016	0.048282	0.074423
01/04/2016	0.031416	0.111094
01/05/2016	0.029874	0.095817
01/06/2016	-0.00726	-0.00992
01/07/2016	0.039621	0.060353
01/08/2016	0.000526	0.01364
01/09/2016	0.017567	0.088603
01/10/2016	0.002492	-0.05672
01/11/2016	-0.00232	-0.04969
01/12/2016	-0.02048	-0.00093
mean	**0.006824**	**0.011368**
var	**0.000763**	**0.005793**
risk	**0.027626**	**0.076111**
return/risk	**0.24703**	**0.149354**

Though the monthly return was significantly higher in the Bayesian Portfolio at 0.011368 compared to Markowitz's 0.006824, the risk experienced was also incredibly high at 0.076111 compared to Markowitz's 0.024703. This shows that the realised Sharpe ratio for the Bayesian portfolio was lower than Markowitz's, at 0.14935 to 0.24703 respectively.

This is better visualized by tracking the portfolio valuation through 2016-17:



Here, the variance experienced by the Bayesian portfolio was extremely significant, dropping heavily initially due to an initial -0.131512 loss experienced by the portfolio, and subsequently AMZN. Perspective is very important here. An investor who views the performance year on year on this portfolio would be very pleased, after all it achieved a 10.95% return on investment. But an investor who is actively monitoring his portfolio may be dissuaded after viewing the first few months 2016, especially considering the 13% loss experienced in Jan 2016.

PCA could prove useful here. By filtering the prior data, a more robust and trendlike distribution could achieved, thus reducing the inherent projected risk/return put forth by the posterior distribution estimator. This is explored in the next section.

6.1.5 PCA95-Bayesian inferred weights for buy period '16-'17

Recall that in section 6.1.3, it was decided PCA at a threshold of 95% cumulative variance would be used. Considering such, the denoised 2010-2015 estimators were retrieved:

Parameter	LMT	MSFT	APPL	AMZN	V	MCD	WFC	CMCSA	MS	C
mean	0.017149	0.010252	0.026419	0.018542	0.021395	0.00746	1.28E-02	0.024005	0.01157	0.012352
cum var	0.011455	0.019088	0.01757	0.019159	0.013094	0.004847	0.020265	0.020608	0.036779	0.034123
cum risk	0.107029	0.138159	0.132553	0.138416	0.114428	0.06962	0.142355	0.143553	0.191778	0.184723
return/risk	0.160224	0.074206	0.199312	0.133961	0.186973	0.107152	0.089636	0.167221	0.060329	0.066868

Compared to the standard 2010-2015 estimators:

Parameter	LMT	MSFT	APPL	AMZN	V	MCD	WFC	CMCSA	MS	C
mean	0.017149	0.010252	0.026419	0.018542	0.021395	0.00746	0.01276	0.024005	0.01157	0.012352
cum var	0.011457	0.019086	0.017583	0.019155	0.013095	0.004846	0.020259	0.02062	0.036888	0.034018
cum risk	0.107036	0.138151	0.132601	0.138403	0.114435	0.069614	0.142333	0.143597	0.192063	0.184439
return/risk	0.160214	0.07421	0.19924	0.133973	0.186962	0.107161	0.08965	0.167171	0.060239	0.066971

It is seen that have the most significant increases in the return/risk ratio reasoned similarly to section 6.1.3

The observed data was not subject to PCA in anticipation of potential loss of crucial trends experienced in 2015. Thus, it was the same as the observed data table in section 6.1.4.

The estimators inferred given the denoised prior data were as follows:

Parameter	LMT	MSFT	APPL	AMZN	V	MCD	WFC	CMCSA	MS	C
mean	-0.00241	-0.00659	0.019409	0.053528	0.000366	0.009521	-6.10E-03	-0.00452	-0.01636	-0.0188
cum var	0.000379	0.001307	0.0011	0.001378	0.001191	0.000555	0.001055	0.001285	0.001761	0.001706
cum risk	0.01948	0.036153	0.033171	0.037115	0.034512	0.023553	0.032476	0.035851	0.041965	0.0413
return/risk	-0.12377	-0.18233	0.585122	1.44221	0.010611	0.404261	-0.18794	-0.12598	-0.38979	-0.45517

Recall the posterior estimators without PCA:

Parameter	LMT	MSFT	APPL	AMZN	V	MCD	WFC	CMCSA	MS	C
mean	0.008756	0.012872	-0.00093	0.062665	0.011927	0.017099	-9.09E-04	-0.00121	-0.01385	-0.00251
cum var	0.000982	0.002601	0.001386	0.002269	0.001847	0.001314	0.001333	0.002155	0.001938	0.002244
cum risk	0.031344	0.050995	0.037235	0.04763	0.042971	0.036248	0.036508	0.046419	0.044028	0.047366
return/risk	0.27935	0.252422	-0.02488	1.315668	0.277568	0.471733	-0.0249	-0.0261	-0.31456	-0.05307

This is alarming, as APPL has suddenly become a potentially worthwhile investment, going from a -0.02488 predicted return/risk rate to a dramatic 0.585122. The difference between the risk/return rate for the 2010-2015 denoised and standard data was very minor: 0.199312 and 0.19924.

This result was very anomalous, and could not be explained mathematically. It was concluded that the error was technical, and that MATLAB may have struggled to take the inverse of the denoised data covariance matrix, resulting in an inaccurate posterior covariance matrix, and subsequently, an inaccurate posterior expected returns vector.

Predictably, the weights assigned given the posterior estimators above provided some weight to APPL:

	LMT	**MSFT**	**APPL**	**AMZN**	**V**	**MCD**	**WFC**	**CMCSA**	**MS**	**C**
Bayesian PCA 95	0	0	0.028	0.6947	0	0.2765	0	0	0	0

The results, though faulty, still provide some interesting insights:

Price	Marko	Bayesian	Bayesian 95
31/12/2015	1000	1000	999.2
01/01/2016	952.2609792	868.4845681	918.9295459
01/02/2016	942.6569987	817.4703172	867.8935267
01/03/2016	988.1702331	878.3086229	933.3001547
01/04/2016	1019.214223	975.8837075	998.924699
01/05/2016	1049.66265	1069.390053	1055.143362
01/06/2016	1042.038925	1058.781758	1042.615018
01/07/2016	1083.325713	1122.682692	1083.001547
01/08/2016	1083.89524	1137.995895	1089.484906
01/09/2016	1102.935778	1238.825947	1160.678159
01/10/2016	1105.684577	1168.563031	1105.467014
01/11/2016	1103.124385	1110.491418	1080.002284
01/12/2016	1080.529731	1109.455746	1086.426743

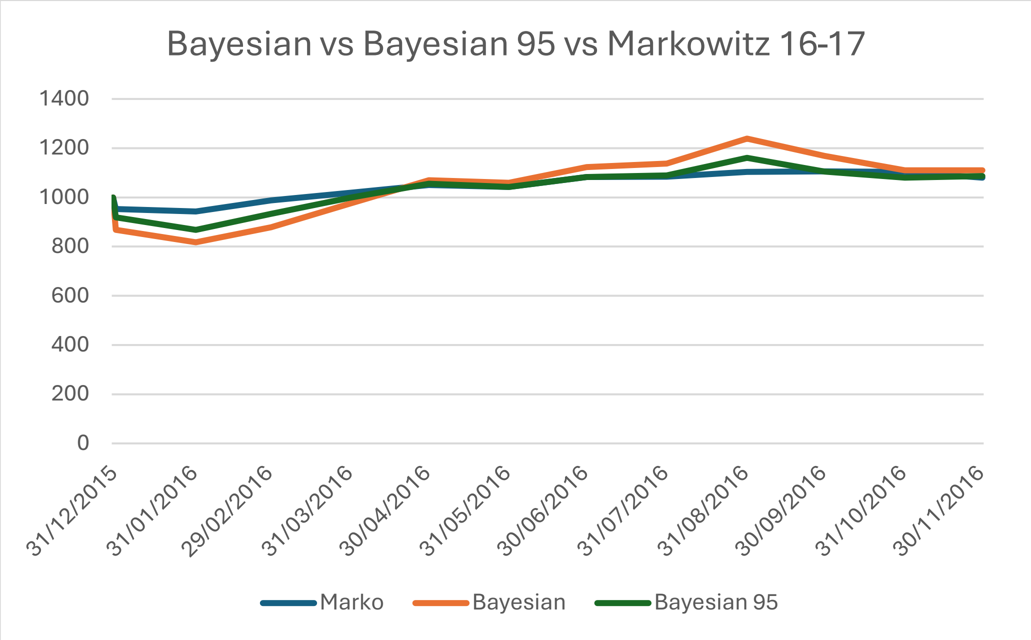
Despite incorrectly predicting APPL (and MCD) to be worthwhile investments, the PCA95 Bayesian Portfolio still had a higher return, achieving 5.13 dollars more than Markowitz. However, this cannot be attributed to more accurate and robust estimators, as the estimators themselves are anomalous and unexplainable.

Still, examining the returns results proves that diversification is important when it comes to risk mitigation:

Returns	Marko	Bayesian	Bayesian 95
01/01/2016	-0.04774	-0.13152	-0.08033
01/02/2016	-0.01009	-0.05874	-0.05554
01/03/2016	0.048282	0.074423	0.075363
01/04/2016	0.031416	0.111094	0.070315
01/05/2016	0.029874	0.095817	0.056279
01/06/2016	-0.00726	-0.00992	-0.01187
01/07/2016	0.039621	0.060353	0.038736
01/08/2016	0.000526	0.01364	0.005986
01/09/2016	0.017567	0.088603	0.065346
01/10/2016	0.002492	-0.05672	-0.04757
01/11/2016	-0.00232	-0.04969	-0.02304
01/12/2016	-0.02048	-0.00093	0.005949
mean	**0.006824**	**0.011368**	**0.008302**
var	**0.000763**	**0.005793**	**0.002843**
risk	**0.027626**	**0.076111**	**0.053324**
return/risk	**0.24703**	**0.149354**	**0.155687**

Compared to the Bayesian Portfolio, the risk was much lower for the PCA95 Bayesian Portfolio at 0.053324 compared to the formers 0.076111. This asserts that spreading out an investment over various assets (even if they don’t perform well) will mitigate the volatility of the portfolio.

This is further reinforced by visualising the portfolio valuation through 2016:



Here, the Dark green line is the faulty PCA 95 Bayesian Model. It constantly lives within the bounds of the Bayesian Portfolio, which consisted of the AMZN asset. This is because while the portfolio consisted of 0.6947 asset weight AMZN, the remaining 0.3153 weightage was spread across other assets that were not as risky as AMZN.

Unfortunately, as the PCA 95 Bayesian Portfolio was faulty, it was omitted from the dynamic investment scenario. Thus, for the next section, the dynamic portfolio management scenario, Bayesian Portfolio generation, Markowitz generation and PCA 95 generation will be investigated under conditions described in section 4.3.

6.2.1 Experiment 2 Results:

This section considers an dynamic investment scenario. An investor looking to maximize his sharpe ratio for his portfolio has 1000 dollars to spread across the 10 assets. The initial buy point is at the beginning of 2015, the second buy point is at 2016, where portfolio weights will be reoptimized given new data, and accordingly, the third, fourth and fifth buy points are in the beginning of 2016, 2017, 2018 and 2019:

Buy Point	BP	Markowitz	PCA 95	Bayesian
2015	**BP1**	2010-2015	2010-2015	2010-2015

2016	**BP2**	2011-2016	2011-2016	2010-2015, 2016
2017	**BP3**	2012-2017	2012-2017	2011-2016, 2017
2018	**BP4**	2013-2018	2013-2018	2012-2017, 2018
2019	**BP5**	2014-2019	2014-2019	2013-2018, 2019

The above table highlights information available to the portfolio generators at each buy point. The data corresponding to each buy point was fed into the parameter estimators, and the following portfolio weights were found:

Markowitz:

Markowitz	**BP**	**LMT**	**MSFT**	**APPL**	**AMZN**	**V**	**MCD**	**WFC**	**CMCSA**	**MS**	**C**
2015	**BP1**	0.2674	0.0000	0.1734	0.0184	0.1717	0.2331	0.0000	0.1359	0.0000	0.0000
2016	**BP2**	0.4014	0.0010	0.0368	0.1238	0.4366	0.0003	0.0000	0.0000	0.0000	0.0000
2017	**BP3**	0.5074	0.0753	0.0000	0.1336	0.1587	0.0000	0.0000	0.1245	0.0005	0.0000
2018	**BP4**	0.5557	0.1722	0.0000	0.0826	0.0554	0.0936	0.0000	0.0000	0.0404	0.0000
2019	**BP5**	0.1267	0.1758	0.0398	0.1027	0.2287	0.3263	0.0000	0.0000	0.0000	0.0000

PCA 95:

PCA 95	**BP**	**LMT**	**MSFT**	**APPL**	**AMZN**	**V**	**MCD**	**WFC**	**CMCSA**	**MS**	**C**
2015	**BP1**	0.2849	0.0000	0.1629	0.0219	0.1423	0.2394	0.0000	0.1486	0.0000	0.0000
2016	**BP2**	0.4011	0.0005	0.0381	0.1218	0.4384	0.0002	0.0000	0.0000	0.0000	0.0000
2017	**BP3**	0.5041	0.0784	0.0000	0.1356	0.1521	0.0000	0.0001	0.1297	0.0000	0.0000
2018	**BP4**	0.5466	0.1665	0.0000	0.0745	0.0567	0.1037	0.0000	0.0000	0.0519	0.0000
2019	**BP5**	0.1233	0.1740	0.0380	0.1014	0.2370	0.3264	0.0000	0.0000	0.0000	0.0000

Bayesian:

Bayesian	**BP**	**LMT**	**MSFT**	**APPL**	**AMZN**	**V**	**MCD**	**WFC**	**CMCSA**	**MS**	**C**
2015	**BP1**	0.2674	0.0000	0.1734	0.0184	0.1717	0.2331	0.0000	0.1359	0.0000	0.0000
2016	**BP2**	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2017	**BP3**	0.3636	0.1786	0.0000	0.0008	0.1088	0.0000	0.0000	0.2789	0.0694	0.0000
2018	**BP4**	0.0976	0.2560	0.0096	0.0004	0.0475	0.0750	0.1175	0.1695	0.0000	0.2269
2019	**BP5**	0.0000	0.9564	0.0000	0.0000	0.0000	0.0436	0.0000	0.0000	0.0000	0.0000

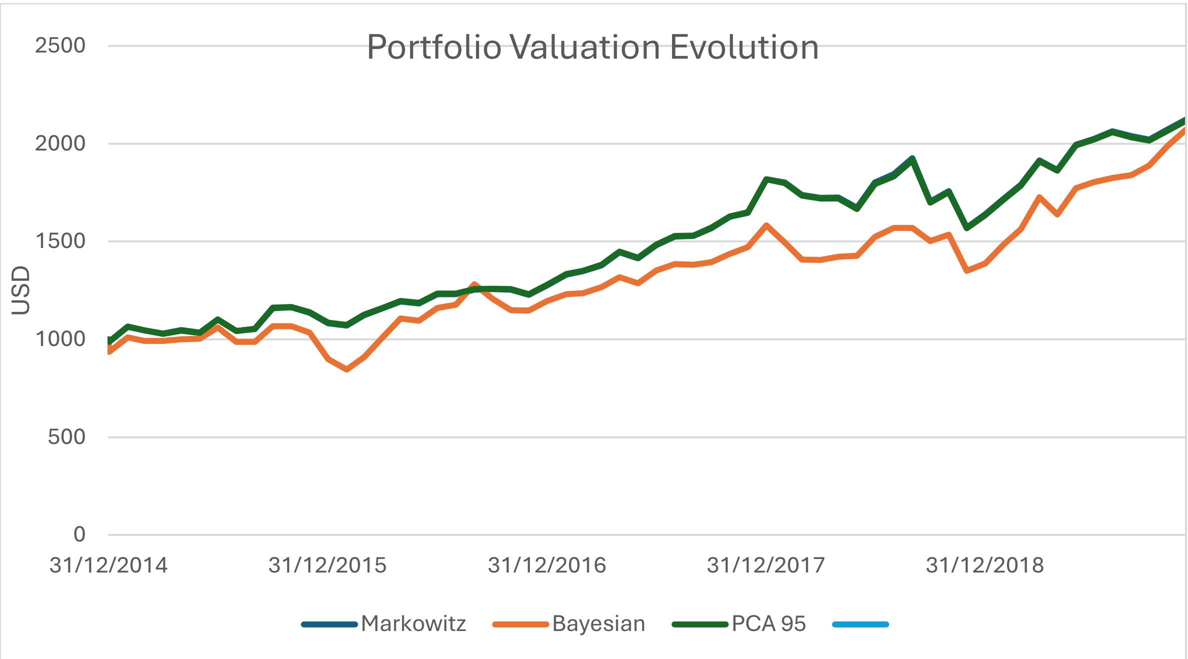
The Markowitz and PCA 95 Portfolios very closely resembled each other, whilst the Bayesian portfolio was widely different. Note that the Bayesian BP1 portfolio is the same as the Markowitz BP1 portfolio. This is because at the time, there was no new observed data for a posterior distribution to be inferred.

However, due to the widely different parameter estimation technique Bayesian inference used, in the BP2 and BP4 periods, expected returns for an asset often came up negative. This is because more importance was given to the most recent observed data, and in these cases, 2016 and 2019 respectively, the markets faced economic turmoil due to various extermal factors; the returns were much lower during these years than the ones preceding it. This can be seen in the excel main.xlsx file. As such, the mean-variance optimization program, limited by the no short selling rule, could only allocate weightage to those with positive expected posterior returns, resulting in widely undiversified portfolio weightages.

6.2.1 Portfolio Value Tracking:

From the Portfolios, the full tracked portfolio values from the beginning of 2015 to the end of 2020 assuming an initial investment of 1000$ are listed in Appendix section C.

The following plot shows the evolution of the 1000 dollars from 2015 to 2020:



Please note that the Markowitz and PCA 95 line were essentially identical. The Bayesian portfolio evolution however, was full of ups and downs and generally performed worse than the Markowitz portfolio.

The monthly returns were also tracked and are listed in Section D of the Appendix.

From the monthly returns, the following table were retrieved:

	Markowitz	PCA 95	Bayesian
Average Monthly returns	1.348%	1.341%	1.355%
Variance	0.0016911	0.0016646	0.0026648
Risk	0.0411224	0.0407994	0.0516221
Sharpe Ratio	0.3276879	0.3286905	0.2624484
ROI at 2016	119.26%	119.56%	110.61%
ROI at 2017	118.01%	118.02%	133.11%
ROI at 2018	141.07%	140.43%	124.94%
ROI at 2019	90.14%	89.94%	87.65%
ROI at 2020	129.50%	129.46%	149.28%
Overall ROI	212.23%	211.64%	207.09%

7.1 Evaluations and Conclusions

8.1 Further work

[1] C. M. Cap, “BRK-B Market Capitalization,” [Online]. Available: https://companiesmarketcap.com/berkshire-hathaway/marketcap/.

[2] A. Hayes, “Investopedia,” 2023. [Online]. Available: https://www.investopedia.com/terms/f/financial-market.asp.

[3] K. Axelton, “Experianhttps://www.experian.com/blogs/ask-experian/what-is-financial-portfolio/,” 2022. [Online].

[4] Y. Finance, “Histroical Stock Prices,” [Online]. Available: These ten stocks were chosen as they branch from infrastructure (such as amazon and ford) to Healthcare (Pfizer and Moderna) to Technology (such as Microsoft and Apple) to hospitality (Mcdonalds, Starbucks) and Social Media (Meta and Netflix.) Having a wi.

[5] IBM, “Monte Carlo Simulation,” 2017. [Online]. Available: https://www.ibm.com/topics/monte-carlo-simulation.

[6] H. Markowitz, “Portfolio Selection,” 1952. [Online]. Available: https://www.math.hkust.edu.hk/~maykwok/courses/ma362/07F/markowitz_JF.pdf.

[7] S. S. A. Maiyahi, “Eficient Frontier Curve,” [Online]. Available: https://www.linkedin.com/pulse/efficient-frontier-curve-sultan-saif-al-maiyahi/.

[8] Investopedia, “risk free asset definitions,” [Online]. Available: https://www.investopedia.com/terms/r/riskfreeasset.asp#:~:text=key%20takeaways,compensated%20for%20taking%20a%20chance..

[9] U. o. Pennsylvannia, "Proof of Bayes Theorem," [Online]. Available: https://www.hep.upenn.edu/~johnda/Papers/Bayes.pdf.

[10] T. Guardian, "Five factors that shook the world's markets in 2015," [Online]. Available: https://www.theguardian.com/business/2015/dec/27/five-factors-shocked-stock-markets-2015.

[11] N. L. F. Medicine, "The origins of Covid 19, a brief overview," [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9874793/#:~:text=The%20most%20widespread%20coronavirus%20to,disease%20(COVID-19)..

[12] Forbes, "The aftermath and impact of covid-19 on stock markets," [Online]. Available: https://www.forbes.com/sites/theyec/2023/02/10/the-aftermath-and-impact-of-covid-19-on-stock-markets/.

[13] BBC, "Coronavirus: How the pandemic has changed the world economy," [Online]. Available: https://www.bbc.co.uk/news/business-51706225.

[14] N. Hoang, "PORTFOLIO OPTIMIZATION METHODS: THE MEAN-VARIANCE APPROACH," 2019. [Online]. Available: https://egrove.olemiss.edu/cgi/viewcontent.cgi?article=2060&context=hon_thesis.

10.1 Appendix:

A: Github Repo

B: Mathematical Background: Analytical Solving of Minimization Problem

This section evaluates how the minimization problem stated can be solved via quadratic algebra and Lagrange Multipliers.

Consider the minimization problem:

$$minimize\ x^T \epsilon x$$

Subject to:

$$e^T x = 1$$

$$\mu^T x = \mu_p$$

Where x represents the vector of asset weights, ϵ represents the covariance of returns, or risk in this constant, e a vector of ones in the same dimensions as x, μ representing a vector of expected returns for each asset and μ_p representing a target portfolio return.

This problem can be solved by introducing Lagrange Multipliers for the constraints:

$$\lambda(e^T x - 1) = 0$$

$$\gamma\left(\mu^T x - \mu_p\right) = 0$$

And so, considering the new Lagrange multipliers, A Langrangian function can be created:

$$\mathcal{L}(x, \lambda, \gamma) = x^T \epsilon x + \lambda(\mathrm{e}^T x - 1) + \gamma(\mu^T x - \mu_p)$$

This can now be differentiated with regard to the three unknowns: x, λ, γ and the resulting partial differential equations set to 0:

Derivative with respect to x:

$$\frac{\partial \mathcal{L}}{\partial x} = 2\epsilon x + \lambda \mathrm{e} + \gamma \mu = 0$$

Derivative with respect to λ:

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \mathrm{e}^T x - 1 = 0$$

Derivative with respect to γ:

$$\frac{\partial \mathcal{L}}{\partial \gamma} = \mu^T x - \mu_p = 0$$

These equation represent a system of linear equations:

$$\begin{gathered} 2\epsilon x + \lambda \mathrm{e} + \gamma \mu = 0 \\ \mathrm{e}^T x = 1 \\ \mu^T x = \mu_p \end{gathered}$$

Representing x in terms of λ, γ from the first equation in the system:

$$x = -\frac{1}{2} \epsilon^{-1} (\lambda \mathrm{e} + \gamma \mu)$$

Substituting the new x into the second equation in the system:

$$\mathrm{e}^T \left(-\tfrac{1}{2} \epsilon^{-1} (\lambda \mathrm{e} + \gamma \mu)\right) = 1 \qquad \text{(a)}$$

And simplifying:

$$\mathrm{e}^T \epsilon^{-1} \mathrm{e} \cdot \lambda + \mathrm{e}^T \epsilon^{-1} \mu \cdot \gamma = -2 \qquad \text{(b)}$$

Substituting the new x into the third equation in the system:

$$\mu^T(-\frac{1}{2}\epsilon^{-1}(\lambda\mathrm{e} + \gamma\mu)) = \mu_p$$

And simplifying:

$$\mu^T\epsilon^{-1}\mathrm{e} \cdot \lambda + \mu^T\epsilon^{-1}\mu \cdot \gamma = -2\mu_p \qquad \text{(c)}$$

Let $A = \mathrm{e}^T\epsilon^{-1}\mathrm{e}$, $B = \mathrm{e}^T\epsilon^{-1}\mu$, $C = \mu^T\epsilon^{-1}\mu$, equations (b) and (c) become:

$$A\lambda + B\gamma = -2$$

$$B\lambda + C\gamma = -2\mu_p$$

Solving for λ, γ:

$$\lambda = -2\frac{C - B\mu_p}{AC - B^2}$$

$$\gamma = -2\frac{A\mu_p - B}{AC - B^2}$$

Substituting new values for λ, γ back into equation (a) for x:

$$x = \frac{1}{2}\epsilon^{-1}(\frac{C - B\mu_p}{AC - B^2}\mathrm{e} + \frac{A\mu_p - B}{AC - B^2}\mu)$$

Thus, optimal portfolio weights for x can be found to minimize risk $x^T\epsilon x$ while achieving a portfolio return μ_p given ϵ.

C: Portfolio Progression