The Link Between Market Correlation Structure and the Performance of Risk-Based Portfolios

Nathan Potgieter[a]

[a] *Stellenbosch University, Stellenbosch, South Africa*

Abstract

This work uses Monte Carlo methods to design and simulate financial market returns for five distinctive markets types, with each market type possesing a unique correlation structure. The equal weight, minimum variance, inverse variance, equal risk contribution and maximum diversification risk-based portfolios are evaluated in each of the simulated markets. The relative performance of each of the portfolios are compaired within market types and the relationship between each portfolio's return characteristics and the market covariance structure is evaluated. **FINDINGS**

Keywords: Monte Carlo, Risk-based Portfolios, Portfolio Selection, Copula

JEL classification L250, L100

1. Introduction

Since Harry Markovitz's (1952) seminal work on mean-variance portfolios scholars from around the globe have been aspiring to develop a robust algorithm capable of situating a portfolio on the efficient frontier *ex ante*. There are now a wide array of available portfolio algorithms raging from simple heuristic based approaches to advanced mathematical algorithms based on quadratic optimization, random matrix theory and machine learning methods; with many more are still to come.

Unfortunately, portfolio optimisers of Markowitz (1952) mean-variance type suffer from seveer sensitivity issues, where slight changes in their expected return input cause large changes in optimal portfolio weights. This is exacerbated by the fact that expected returns are notoriously difficult, if not impossible, to accurately forecast (De Prado, 2016). Due to this issue, this work focuses solely on the so-called risk-based portfolio, defined by De Carvalho *et al.* (2012a) as "systemic quantitative approaches to portfolio allocation" that solely rely on views of risk when allocating capital. These strategies do not require expected return forecasts and are therefore said to be more robust to esti-

Email address: `19959672@sun.ac.za` (Nathan Potgieter)

Contributions:

The author would like to thank Nico Katzke for helping me puzzle and prod my to to the eventual completion of this research project.

mation error. Despite their sole focus on risk mitigation empirical back tests have shown that they often perform surprising well, form a total return standpoint (Choueifaty *et al.*, 2013).

Rather than using the standard empirical approach to evaluating said strategies this work opts to use Monte Carlo simulation methods. This also allows for an investigation of the link between the markets covariance structure and portfolio performance. Monte Carlo methods prove to be invaluable in answering this question as they allow for the creation of *ad hoc* markets with predetermined risk return characteristics, and hence leave no uncertainty regarding the composition of the market. This creates an environment ideal for experimentation as the researcher has complete control over the market and can therefore adjust the independent variable, in this case the markets correlation structure, and observe the response in the dependent variable, which in this study is the portfolio return series. Monte Carlo methods are also beneficial here because they enable the effective reduction of noise in portfolio performance measures.

The risk-based portfolios evaluated in this work include the naive equal weight (EW), minimum variance (MV), inverse variance (IV), equal risk contribution (ERC) and the maximum diversification (MD) portfolios. Section 2 lays out this works aims and objects. Section 3 provides a review of the relevant literature; this includes some general issues plaguing the field of portfolio optimization, the rational and theoretical underpinnings behind the five risk-based portfolios, their relative performance in empirical back tests and finally the importance of using Monte Carlo methods in finance. Section 4 discusses this works methodology, Section 5 provides and discusses the results, and finally Section 6 concludes.

2. Aims and Objectives

This work aims to use Monte Carlo Methods to uncover the relationship between a market's correlation structure and the risk return properties of various risk-based portfolio algorithms. This is achieved through the following objectives:

1. Design and create four distinctive *ad hoc* correlation matrices and estimate one empirical correlation matrix, each describing a market of 50 assets with a unique correlation structure. These market range from those possessing no clusters to those exhibiting hierarchical clustering. All other risk characteristics remain equal between market types.
2. Use these five correlation matrices as the key input in their own Monte Carlo simulation. Where each correlation matrix corresponds to a unique market type. The markets will each be built with a student t multivariate distribution with 4.5 degrees of freedom. The individual asset's univariate distributions will each be normally distributed with means and standard deviations calibrated with S&P500 data. Each of the five market types will be simulated 10 000 times across 500 periods, or approximately two years of trading daily return data.

3. Calculate the returns obtained from various risk based portfolio's in each of the simulated markets. Use the first **250** periods to estimate an out of sample covariance matrix and calculate portfolio weights. Conduct periodic rebalancing every 50 periods, each time looking back *250* periods, recalculating the covariance matrix and the new portfolio weights. Repeat this process until all *500* simulated periods have been considered.
4. Calculate the average Sharp ratio, standard deviation, downside deviation, value at risk (VaR), effective number of constitutes (ENC) and effective number of bets (ENB) for each portfolio across the 10 000 markets for each market type.
5. Compare the relative performance of each portfolio within each of the market types and evaluate how the relative performance of each portfolio is effected by the change in market correlation structure.

3. Litrature Review

3.1. A Review of Portfolio Optimisation Algorithms

3.1.1. Introduction

This literature review will cover some common issues found in the literature surrounding portfolio optimization, the five risk-based portfolios evaluated in this work, their respective performance in both empirical back tests and Monte Carlo studies and finally the importance of using Monte Carlo methods within the field of finance.

3.1.2. Common Issues Portfolio Optimizers

When operating in sample, Portfolio optimization tends to be a perfect science, but out of sample it becomes more of an art form where it is often preferable to use heuristic over hard rules. This section highlights some general issues, highlighted within the portfolio optimization literature, that tend to worsen their performance out of sample.

Firstly, mean-variance optimisers, like those introduced by Markowitz (1952), rely heavily on the accuracy of their expected return forecasts. Small changes in their expected return input can lead to large changes in portfolio weights (De Prado, 2016). Since in practice expected returns are extremely difficult to accurately estimate, this issue serves as a major hindrance to their wide spread use. Due to this the so-called risk based portfolios that intentionally avoid using expected return forecasts have garnered a lot of attention (Maillard, 2010).

However, these risk based portfolios are not void of issues. The quadratic programming methods used in many portfolio optimisers, including Markovitz's (1952) mean variance and some risk-based portfolios, require the inversion of some positive-definite covariance matrix. This requirement for positive definiteness can cause issues as covariance matrices estimated on empirical data are sometimes not positive definite, in which case their inverse does not exist and these portfolios don't have solutions (De Prado, 2016). A common method to get around this issue is to simply compute the nearest positive definite matrix and use that instead (Bates & Maechler, 2019; Higham, 2002).

The covariance estimation step is particularly susceptible to measurement error if the underlying covariance matrix suffers from a high condition number (Zhou *et al.*, 2019). A condition number is defined as the absolute value of the ratio between a covariance matrix's largest and smallest eigenvalues (Bailey & Lopez De Prado, 2012; De Prado, 2016). The condition number is smallest in diagonal matrices (they have a condition number of 1) and increases as more correlated variables are added. When working with high condition number matrices a small change in a single entry's estimated covariance can greatly alter its inverse, which in turn can effect the portfolio weights (De Prado, 2016). This is related to Markowitz's curse which De Prado (2016) summerised by stating that "the more correlated investments are, the greater is the need for a diversified portfolio—and yet the greater are that portfolio's estimation errors".

For a sample with a given number of periods, larger dimension covariance matrices are prone to more noise in estimation. This is essentially due to a reduction in degrees of freedom as a sample of at least $1/2N(N+1)$ independent and identically distributed (iid) observations are required to estimate an $N \times N$ covariance matrix (De Prado, 2016: 60)]. Furthermore, financial market covariance structures tend to vary over time and have been know to change rapidly during so-called regime changes (De Prado, 2016). This exacerbates the issue of requiring a large number of observations when estimating the covariance matrix, since passed data may not be a good refection of the future and looking further into the passed increases this likelihood.

3.1.3. Risk Based Portfolio's

This section reviews the intuition and technical underpinnings within the literature surrounding risk-based portfolios. Those discussed here include the equal weight (EW), minimum variance (MV), inverse volatility (IV), equal risk contribution (ERC) and the maximum diversification (MD) portfolios. The EW is a simple heuristic approach, the minimum variance is more akin to a Markovitz (1952) mean variance portfolio, while the inverse-variance (IV), equal risk contrition (ERC) and maximum diversification (MD) are quite similar in that they all assume that adequate diversification can be obtained by allocating equal risk to each investible security.

3.1.3.1. Naive Equal Weight (EW). Perhaps the oldest and most simple portfolio diversification heuristic involves holding a weight of $1/N$ of the N total available assets (DeMiguel *et al.*, 2009). In other words this strategy can be described as putting an equal number of eggs into each available basket. It doesn't require any historical data when allocating capital and doesn't involve any form of optimization (DeMiguel *et al.*, 2009). This portfolio is commonly called the equal weight or 1/N portfolio, however its failure to recognize the importance of both asset variance and the covariance between assets has resulted in it also being referred to as the naive portfolio. Meanwhile its simplicity means that it has been widely used as a benchmark. Equal weighting is optimal from a mean variance standpoint when there is no correlation between securities and each possesses the same variance. In which case, the EW is theoretically equivalent to the MV portfolio.

3.1.3.2. Minimum Variance (MV). Portfolio optimisers designed to exhibit the minimum variance have in recent years garnered a lot of attention. This can in large part be attributed to their tendency to achieve surprisingly high returns and low variance in historical back tests (Clarke *et al.*, 2011). Their excelent performance has been attributed to the empirical phenomena that low volatility stocks tend to earn returns in excess of the market, and high beta stocks tend not to be rewarded by higher returns (Clarke *et al.*, 2011; Fama & French, 1992). Interesting this later finding is contrary to traditional financial economic theory which predicts an asset's expected return to be proportional to its market beta (i.e. undiversifiable risk) (Perold, 2004).

The minimum variance portfolio selects security weights such that the resulting portfolio corresponds to that with the lowest possible in sample volatility. Therefore, it has the lowest expected volatility and is, in theory, the safest/least risky portfolio (De Carvalho *et al.*, 2012b). Its primary input is a variance covariance matrix, which it uses to minimize aggregate portfolio volatility. This is accomplished by over-weighting low volatility and low correlation securities (De Carvalho *et al.*, 2012b). Interestingly, it is the only portfolio of the efficient frontier that does not depend on expected return forecasts (De Prado, 2016).

Let $\sum$ indicate the markets variance covariance matrix and $w = \{w_i, ..., w_N\}$ be a vector of length N containing individual security weights. The vector containing the MV portfolio's weights can now be described as (De Carvalho *et al.*, 2012b):

$$w^* = arg\min(w' \textstyle\sum w) \quad s.t. \ \textstyle\sum_i^N w_i = 1$$

In some studies the minimum variance (MV) portfolio has been found by to earn cumulative returns equal to or slightly greater than market capitalization weighted portfolio's whilst maintaining a consistently lower variance and achieving a noticeable improvement in downside risk mitigation (Clarke *et al.*, 2011). The MV portfolio can therefore work well out of sample, but if left unrestricted is known to build highly concentrated portfolio's (De Prado, 2016). Its sole objective to minimize portfolio volatility is likely the the primary reason for this. When near the trough of its objective function it to achieves minor reductions in *ex ante* volatility by greatly favoring a small number of low volatility/correlation securities (De Prado, 2016: 68)]. This tendency to produce highly concentrated portfolio's can be costly out of sample as the portfolio does not sufficiently diversify its idiosyncratic risk. It puts too many eggs in too few baskets. In practice This issue can be countered by applying cleaver maximum and minimum portfolio weight constraints.

3.1.3.3. Inverse-Varience (IV) Weighting. The IV portfolio, referred to as the equal-risk budget (ERB) portfolio in De Carvalho *et al.* (2012a), aims to allocate an equal risk budget to each investible security (De Carvalho *et al.*, 2012a). Where the risk budget is defined as the the product of a security's weight and volatility. If σ_i is defigned as security i's volatility, then risk the portfolio risk budget can be equally distributed across N securities by setting security weights as:

$$w_{iv} = (\frac{1/\sigma_1}{\sum_{j=1}^{N} 1/\sigma}, ..., \frac{1/\sigma_N}{\sum_{j=1}^{N} 1/\sigma})$$

This indicates that each securities weight is directly proportional to the inverse of its variance, thereby demonstrating why this is called the IV portfolio. The IV portfolio allocates capital based solely on security variance and is therefore oblivious to the covariance between its constitutes. De Carvalho *et al.* (2012a) found that, if all securities posses the same sharp ratio and their correlation coefficients are all equal, then the IV portfolio is efficient from a mean variance stand point and obtains the highest possible sharp ratio.

3.1.3.4. Equal Risk Contribution (ERC). The ERC portfolio is similar to the IV, however when balancing risk contributions it does account for the covariance between securities (De Carvalho *et al.*, 2012a). The basic idea behind the ERC is to weight the portfolio such that each security contributes equally to overall portfolio risk, this should in turn maximize risk diversification (Maillard, 2010). In practice the ERC acts similar to a weight constrained MV portfolio, with constraints preventing high levels of portfolio concentration. Following Maillard (2010), the weights of an ERC portfolio $x = (x_1, x_2, ..., x_n)$ consisting of n assets can be calculated as follows:

let σ_i^2 resemble asset i's variance, σ_{ij} the covariance between asset i and j and $\sum$ be the markets variance covariance matrix. Portfolio risk can now be written as $sigma(x) = \sqrt{x^T \sum x} = \sum_i \sum_{j \neq i} x_i x_j \sigma_{ij}$ and the marginal risk contribution, $\partial_{x_i}\sigma(x)$, can then be defined as:

$$\partial_{x_i}\sigma(x) = \frac{\partial\sigma(x)}{\partial x_i} = \frac{x_i\sigma_i^2 + \sum_{j \neq i} x_j \sigma_{ij}}{\sigma(x)}$$

Therefore, $\partial_{x_i}\sigma(x)$ refers to the change in portfolio volatility resulting from a small change in asset i's weight (Maillard, 2010). ERC uses this definition to guide central objective to equate the risk contribution across each of the n asset. There is no closed form solution describing the weights of the ERC portfolio, however, if we define $(\sum x)_i$ as the i^{th} row resulting from the product of $\sum$ with x and note that $\partial_{x_i}\sigma(x) = (\sum x)_i$, then the optimal weight for the long only ERC can be described as those that satisfy the following statement (see Maillard (2010), p. 4-7 for more detail):

$$x^* = \{x\ \epsilon[0,1]^n : \sum x_i = 1, x_i \times (\sum x)_i = x_j \times (\sum x)_j\ \forall\ i,j\}$$

Maillard (2010) proved mathematically that the ERC portfolio's *ex ante* volatility is always some where between those of the EW and MV portfolio's. De Carvalho *et al.* (2012a) found that, if all securities posses the same sharp ratio , then the ERC and ERB have identical portfolio weights. If in addition the correlation coefficients between all securities are equal, then the ERC and ERB merge into the EW portfolio and each are mean variance efficient with the maximum attainable sharp ratio (De Carvalho *et al.*, 2012a).

3.1.3.5. Maximum Diversification (MD). Choueifaty & Coignard (2008) originally designed the MD portfolio to maximize some diversification ratio (DR), which he defined as the sum of each securities risk bucket divided by portfolio volatility (De Carvalho *et al.*, 2012a). If we define $w = (w_1, ...w_N)^T$ as a vector of portfolio weights, V as a vector of asset volatilities and $\sum$ as the covariance matrix. Then the DR can be expresses as:

$$DR = \frac{w'.V}{\sqrt{w'Vw}}$$

Much like the IV and ERC portfolios, the MD portfolio attempts to diversify its portfolio by allocating equal risk to each security (Choueifaty & Coignard, 2008). The MD portfolio accomplishes this by over-weighting low volatility securities and those that are less correlated (De Carvalho *et al.*, 2012a). For further detail regarding the theoretical results and properties of the MD portfolio see Choueifaty & Coignard (2008: 33–35).

3.2. Empirical Backtests and Monte Carlo Findings

Choueifaty *et al.* (2013) conducted an empirical back test comparing the relative performance of numerous portfolio optimisers between 1999 and 2010. They used historical data from the MSCI world index and considered the largest 50% of assets at each semi-annual rebalance date. To reduce the noise in estimation, at each rebalance date covariance matrices were estimated using the previous years worth of data (Choueifaty *et al.*, 2013). These were then used as the primary inputs in estimating the respective long-only portfolio weights. The MV portfolio achieved an annual return of 6.7% and outperformed the ERC and EW portfolio's who returned 6.3% and 5.8% respectively. Unsurprisingly, the MV portfolio possessed the lowest daily volatility (10%) followed by the ERC and then the EW portfolio's (with 12.9% and 16.4% respectively). Accordingly the MV portfolio scored the highest sharp ratio (0.36) followed by the ERC and EW portfolio's (0.24 and 0.16 respectively). According to De Carvalho *et al.* (2012a) the performance of the EW portfolio primarily depends on the premium on small-capitalization stocks, thereby suggesting that the relatively poor performance of the EW portfolio in Choueifaty *et al.* (2013), can be attributed to the relatively poor returns achieved by the smaller stocks in the MSCI world index.

Despite the simplistic nature of the EW portfolio empirical studies, like those by DeMiguel *et al.* (2009), who compared the EW to the mean-variance, MV and Bayes-Stein portfolios, tend to find a statistically insignificant difference in Sharp ratio between the naive portfolio and those of the more advanced portfolio optimisers. In this study the EW also performed surprisingly well from a total return perspective.

Due to the aforementioned issues surrounding estimation error in a market's covariance matrix Ardia *et al.* (2017) set out to evaluate the impact of covariance matrix misspecification on the properties of risk based portfolio's. They used Monte Carlo methods to build six distinctive investment universes, each with a unique, variance/correlation structure and a varying number of assets. Numerous covariance matrix estimation techniques were then estimated on the simulated data, one of which serving as the benchmark. They then used the simulated data and the various covariance matrices to access the impact of alternative covariance specifications on the performance of the MV, IV, ERC and MD portfolio's. The ERC and IV portfolios were found to be "relatively robust to covariance misspecification", the MV was found to be sensitive to misspecification in both the variance and covariance and the MD portfolio was found to be robust to misspecification in the variances but sensitive to misspecification in the covariances (Ardia *et al.*, 2017: 1).

3.3. Monte Carlo Methods in Portfolio Optimisation

Ever since the pioneering age of computers people have shown a keen interest in leveraging their ability to perform rapid calculations to conduct randomized experiments (Kroese *et al.*, 2014: 1). The core of Monte Carlo simulation is in the creation of random objects and/or processes using a computer. There are a number of reasons for doing this, but the primary one used in this work and thereby discussed in this review is of the sampling kind (Kroese *et al.*, 2014). This typically involves the modeling of some stochastic object or process, followed by sampling from some probability distribution and the manipulating said sample through some deterministic process such that the result mimics the true underlying process. The primary idea behind Monte Carlo simulation is to repeat this simulation process many times so that interesting properties can be uncovered through the law of large numbers and central limit theorem.

A financial application of this can be found in Wang *et al.* (2012) who designed a Monte Calro procedure that (1) models both the time-series and cross-section properties of financial market returns, this involves the estimation of a random term's probability distribution function (pdf) using extreme value theory. And (2) sampling from the modeled process to produce an ensemble of market returns, with each exerting the same risk properties. The simulated data can then be used in risk management and/or the pricing of financial securities (Kroese *et al.*, 2014; Wang *et al.*, 2012). This unique ability to generate a large number of counterfactuals for an asset market with a known risk structure has made it a uniquely powerful tool in accessing the properties of portfolio optimization algorithms (Bailey & Lopez De Prado, 2012).

Glasserman (2013) is a useful source for understanding the methods and applications of Monte Carlo methods in finance.

4. Methadology

This work used Monte Carlo simulation methods to investigate the link between a markets correlation structure and the relative performance of the EW, MV, IV, ERC and MD portfolios. To avoid possible confusion note the following terminology:

- The term market refers to a set daily returns for a number of assets. e.g. the daily returns for each of the JSE ALSI constitutes between 1 January 2019 and 1 January 2020. Since this is a Monte Carlo study, thousands of markets are simulated, they can therefore be thought of as a single observation from a population of markets.
- The term market type refers to a population of markets each with the same specified risk characteristics.

The R package MCmarket was used to simulate 10 000 markets from five separate market types, with the correlation structure being the only differentiating factor between between market types. (Potgieter, 2020). Four of the correlation structures/matrices were designed *ad hoc*, while the fifth was estimated using S&P 500 data. These correlation matrices range from one exhibiting no correlation (i.e. a diagonal matrix) to one with a hierarchical clustering structure (see 4.1).

The long only EW, MV, IV, ERC and MD portfolios were then back tested on the simulated markets (Section 4.3) and portfolio analytics were calculated and aggregated across the 10 000 markets (Section 4.4). Finally, the portfolio metrics are compared within market types across portfolios and within portfolios, across market types.

4.1. Correlation Structures

This section describes the composition and attributes of the four *ad hoc* correlation matrices used in this study (section 4.1.1) as well as the methodology behind the estimation of the empirical correlation matrix (section \ref{emp). Each of the five matrices top 10 eigenvalues are listed in Table 5.6.

4.1.1. Ad Hoc

This section describes the four *ad hoc* 50 by 50 correlation matrices used as the key inputs in their respective Monte Carlo simulations. See Figure 4.1 for a graphical representation of each correlation matrix. Note that the *gen_corr* function from the R package *MCmarket* was used in the construction of the four *ad hoc* matrices (Potgieter, 2020).

The first and most simplistic of the four matrices is a diagonal matrix (see Diagonal Matrix in Figure 4.1). It describes a market with a zero correlation coefficient between each asset. Each of its 50 eigenvalues are equal to 1 (Table 5.6), it has no risk clusters and has plenty scope for diversification. This correlation matrix has the lowest possible condition number of 1. The Monte Carlo data set constructed using this matrix will be referred to as Market 1.

The second matrix (labeled No Clusters in Figure 4.1) has no risk clusters but describes a market with significant correlation between its constituents. Each asset has a correlation of 0.9 with is closest neighbor (i.e. Asset 1 and 2, 5 and 6 and 11 and 12 each have a pairwise correlation coefficient of 0.9). Correlations then diminish exponentially by the absolute distance between the two assets (i.e. the correlation between Asset 1 and 5 is $0.9^{|1-5|} = 0.6561$). Its has a large first eigenvalue of 15.93, but they quickly diminish such that its 9th largest eigenvalue is less than 1 at 0.79 (Table 5.6). This correlation matrix has the highest condition number of 302.4. The Monte Carlo data set constructed using this matrix will be referred to as Market 2.

The third matrix (labeled Five Clusters in Figure 4.1) contains five distinctive non-overlapping risk clusters. Assets within the same cluster have a pairwise correlation coefficient of 0.6 while those that are not in the same cluster are uncorrelated. ItThis correlation matrix has a condition number of 16.1. The Monte Carlo data set constructed using this matrix will be referred to as Market 3.

The final *ad hoc* correlation matrix has three layers of overlapping risk clusters. The first layer has 10 distinctive clusters, within which assets have a correlation coefficient of 0.7. The second layer has four clusters where assets that are not in same first layer cluster have a correlation coefficient of 0.5. Assets that are in the same third layer cluster but not clustered in layers one and two have a correlation coefficient of 0.3. Finally, those who do not share any cluster have a correlation coefficient of 0.05. Its largest eigenvalue is 14.36, but they diminish fairly quickly as its third largest is only 3.3 (Table 5.6). This correlation matrix has a condition number of 47.9. The Monte Carlo data set constructed using this matrix will be referred to as Market 4.

Table 5.6 provides a list of the ten largest eigenvalues for each of the five correlation matrices.

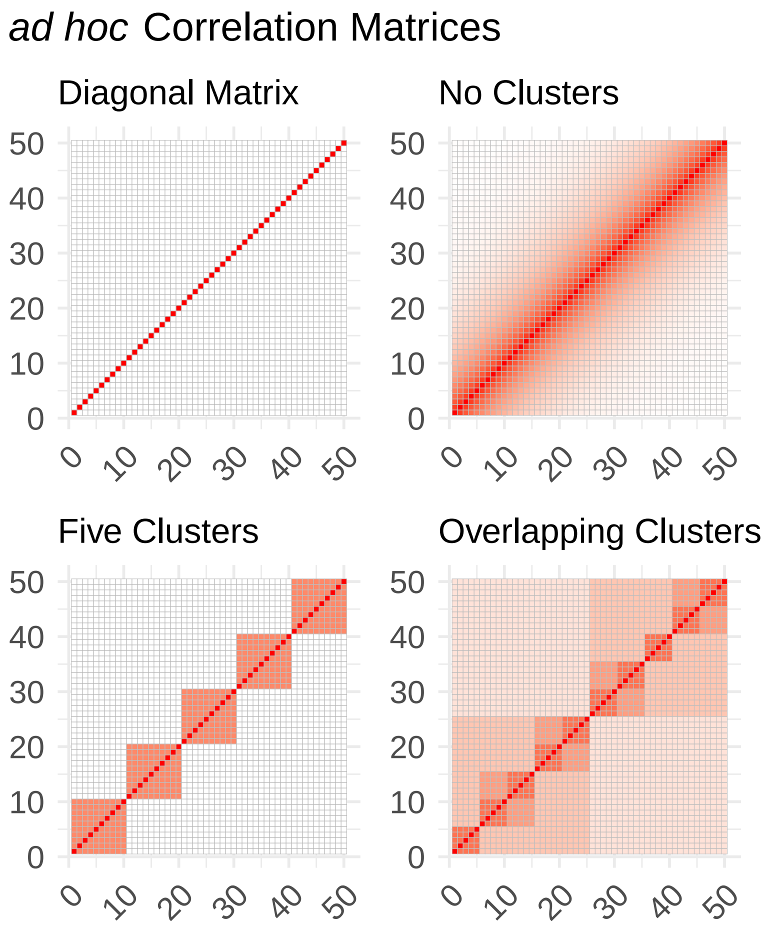

Figure 4.1: Correlation Matricies

4.1.2. Emperical

The empirical correlation matrix used in this study was estimated from the daily returns of a random subset of 50 of the largest (by market capitalization) 100 S&P 500 stocks between 1 January 2016 and 1 January 2021. The market capitalizations were measured as of 12 January 2020.

The markets covariance matrix was then estimated using the *fit_mvt* function from the R package *fitHeavyTail* (Palomar & ZHOU, 2020). This covariance estimation method uses maximum likelihood estimation and generalized expectation maximization to fit a multivariate t-distribution the a matrix of asset returns (Liu & Rubin, 1995). This procedure found that the multivariate t distribution with 4.43 degrees of freedom and the correlation matrix shown in Figure 4.2 best approximated the return series.

Note that in Figure 4.2 assets were ordered by hierarchical clustering so the reader could more easily visualize the risk clusters. The correlation matrix's largest eigenvalue is 18.6 and they quickly diminish to below zero by its 8th largest eigenvalue (Table 5.6). This correlation matrix has a condition number of 211.1. The Monte Carlo data set constructed using this matrix will be referred to as Market 5.

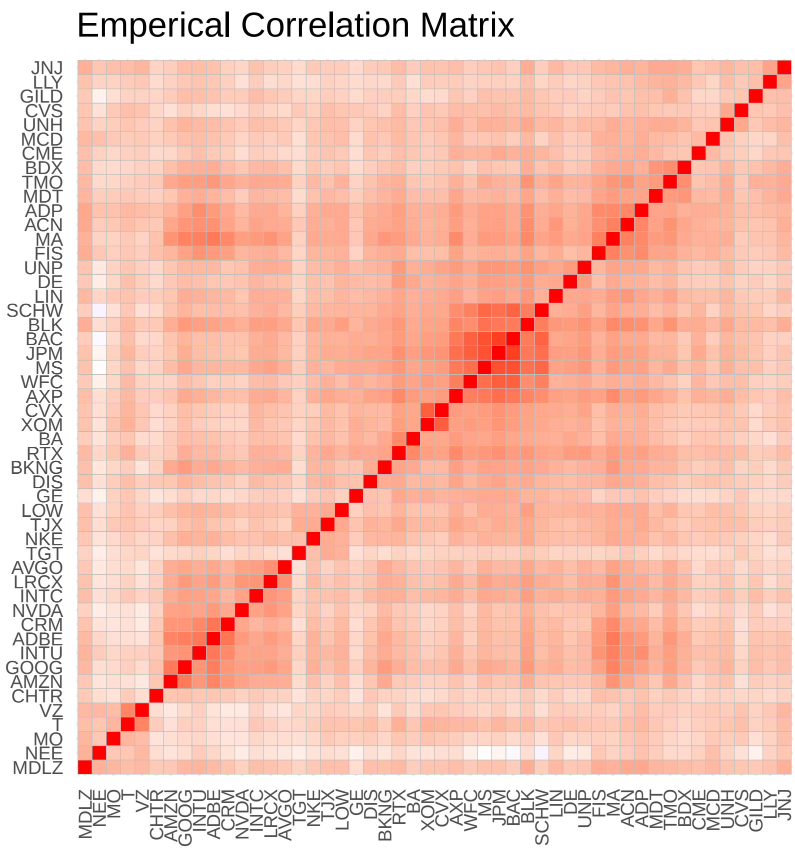

Figure 4.2: Emperical Correlation Matrix

Table 4.1: Eigenvalues

Diagonal	No Clusters	Five Clusters	Overlapping Clusters	Emperical
1	15.93	6.4	14.36	18.6
1	10.38	6.4	6.89	3.09
1	6.28	6.4	3.3	2.52
1	3.92	6.4	3.3	1.4
1	2.59	6.4	2.49	1.28
1	1.81	0.4	2.46	1.21
1	1.32	0.4	1.3	1.03
1	1.01	0.4	1.3	0.92
1	0.79	0.4	1.3	0.88
1	0.64	0.4	1.3	0.83

4.2. Monte Carlo

This section outlines the process behind the Monte Carlo simulations performed as part of this study.

A generalized version of the Monte Carlo procedure developed in Wang *et al.* (2012) was used to simulate five distinctive market types. This framework was build into the R package *MCmarket* which was used to conduct this project's Monte Carlo simulations (Potgieter, 2020). The following briefly describes the process:

An Elliptical t copula with 4.5 degrees of freedom is used, in conjunction with a 50 by 50 correlation matrix (Section **??**), to simulate 500 random uniformly distributed draws (corresponding to 500 trading days or approximately two years worth of trading days) across the 50 assets. The uniformly distributed observations were then transformed into normally distributed observations, via the inverted normal cumulative distribution function (Potgieter, 2020; Wang *et al.*, 2012: 3). This process was repeated 10 000 times for each of the five correlation matrices/ market types set out in section 4.1. The correlation matrix is the only distinguishing factor between market types as all other factors remain equal. All in all, this process created 5 data sets each containing 10 000 markets with 50 assets and 500 periods.

The expected returns and standard deviation of these simulated variables were calibrated using a random subset of 50 of the largest 100 S&P500 stocks (discussed in Section **??**) between 1 January 2020 and 1 January 2021. Maximum likelihood estimation was used to fit the multivariate t distribution to the return series using the method developed by Liu & Rubin (1995). This produced a series of estimated means and variances which were used to calibrate the expected returns and standard deviations of the simulated variables (see Table **??**msd)).

4.3. Back Tests

This section describes the back testing procedure used to calculate the returns obtained by the EW, MV, IV, ERC and MD portfolios. The process described relates to a single market and was therefore applied to each of the 50 000 markets simulated in this study.

To remain consistent with the literature as well as the mandate for the majority of portfolio managers, a long-only weight constraint was applied to all portfolio's. In addition a constraint limiting the maximum weight applied of a single security to 10% is also applied. This prevents some portfolios from building unreasonably highly concentrated holdings, while remaining flexible enough to punish those who do so. These constraints therefore act to provide a fair playing ground for the portfolio's to compete. The back testing procedure works as follows:

The first 250 periods, approximately equivalent to one years worth of daily return data, are used to fit a multivariate t distribution via maximum likelihood and estimate a covariance matrix (Liu & Rubin, 1995). Interestingly, the identifying assumption used in this covariance matrix estimation method, that is that the data comes from a multivariate t distribution, is correct by definition since this is the distribution used to simulate the data set. The estimated covariance matrix is then used as the sole input when calculating the weights for each of the respective risk-based portfolios. Each portfolio holds these weights over the next 50 periods, when they are rebalanced by looking back 250 periods, calculating the covariance matrix and the new portfolio returns. This process is repeated until all periods in the data set are exhausted. Since there are 500 periods in each market, each portfolio is weighted 5 times and 250 periods of daily returns are calculated for each portfolio.

4.4. Portfolio Analytics

This section describes the portfolio performance and concentration metrics used to evaluate and compare portfolios. The Sharp ratio is used to evaluate the risk adjusted return, while the standard deviation, downside deviation and value at risk are used to access portfolio risk. Finally, the effective number of constitutes, calculated as the inverse of the Herfindahl-Hirschman index (HHI), and the effective number of bets, calculated following Meucci (2010), are used to compare diversification between portfolios (Rhoades, 1993).

Since Markowitz (1952) variance of returns has been the standard measure for risk in the financial industry (**???**). With the standard deviation simply being the square-root of the variance, it too is widely used. Standard deviation also benefits due to its relative ease in interpretation. Standard deviation is also key in calculating the next two portfolio performance metrics described in this study, namely the Sharp ratio and value at risk (VaR).

The Sharp ratio is a measure of a portfolio's risk adjusted returns. Generally speaking, the Sharp ratio is calculated by dividing the portfolio return by some measure of portfolio risk, it is therefore interpreted as the return per unit of risk. In this work standard deviation is used as the measure of risk.

The 95% VaR is another risk metric used to evaluate portfolio risk performance in this study. It is one of the financial industry standards for measures for downside risk and can be interpreted as the maximum return expected from in the worst 5% of scenarios (Peterson & Carl, 2020). That is, in the worst 5% of scenarios, one should expect to loose at least this amount. The particular version of VaR used here is the Gaussian VaR, which is calculated by assuming that returns are normally $N(\mu, \sigma)$ distributed, where μ and σ are estimated using historical data. The probability distribution assumptions allows one to attach a probability values to possible future portfolio returns. This assumption can be dangerous in practice, however in this study it correct by definition as the return series were each simulated to be normally distributed. It should therefore result in an accurate estimate of downside risk.

The HHI estimates portfolio concentration by by summing the the portfolio weights squared (Rhoades, 1993). A portfolio with small weights allocated evenly across a large number of securities will have an HHI of approximately zero, while a portfolio with all its capital invested in a single security will have the maximum HHI of 10000. The effective number of constitutes (ENC) can then be calculated as the inverse of the HHI, where an equally weighted portfolio with have an ENC equal to the number of securities and more concentrated portfolio's will have a ENC less than the number of securities. Weight based measures of portfolio diversification are severely limited in that they are oblivious to covariation between portfolio components. The ENC can therefore be misleading in financial applications where portfolio components are known to exhibit significant dependence.

Meucci (2010) attempted to rectify this issue when he introduced a new method to evaluate portfolio diversification that considers the portfolio's risk structure. He used a principle component (PC) approach to estimate the total number of orthogonal bets within a portfolio, which he simply referred to as principle portfolios. With this he estimated a portfolio diversification distribution using the percentage of total portfolio variation attributed to each principle portfolio. The effective number of orthogonal bets (ENB) can then be calculated as the dispersion of the diversification distribution (Meucci, 2010: 10).

5. Results and Discussion

Note that the markets simulated using the diagonal correlation matrix described in Section 4.1.1 will hence forth be referred to as Market 1. Similarly, the markets simulated using the no cluster, five clusters, overlapping clusters and empirical correlation matrices will be respectively referred to as Market 2, Market 3, Market 4 and Market 5. Therefore, each of the Markets 1 - 5 contain a unique correlation structure.

5.1. Comparing Portfolios Within Market Types

5.1.1. Market 1

The portfolios compared here were estimated on the markets simulated using the diagonal correlation matrix (Figure 4.1). The portfolios average Sharp ratio (Sharp), standard deviation (SD), downside deviation and VaR across the 10 000 simulated markets are shown in Table 5.1.

The EW portfolio performed the best overall as it achieved the highest average Sharp ratio and the lowest average standard deviation, downside deviation and VaR. The IV portfolio was a close second with the second highest Sharp ratio, tied for lowest standard deviation, and achieved the second lowest average downside deviation and VaR. The ERC is ranked the third best, followed by the, MD and finally the MV.

With a diagonal matrix characterizing the market correlation structure structure it is reasonable to expect that, the EW and IV to perform well. the asset variance allocation is fairly evenly distribution then the EW, MV, IV, ERC and MD will produce fairly similar.

Table 5.1: Market 1 Risk Metrics

Metric	EW	MV	IV	ERC	MD
Sharp	0.19353	0.13501	0.18341	0.15876	0.14284
SD	0.00386	0.005	0.00386	0.00447	0.00494
Downside Deviation	0.00235	0.00318	0.00236	0.00279	0.00312
VaR	-0.00545	-0.00735	-0.00548	-0.00645	-0.00722

Table 5.2: Market 1 Portfolio Entropy Metrics

Metric	EW	MV	IV	ERC	MD
ENC	50	25.3	47.1	30.1	29.1
ENB	17.5	22.8	20.1	18.9	21

5.1.2. Market 2

The portfolios compared here were estimated on the markets simulated using the no clusters correlation matrix (Figure 4.1). The portfolios average Sharp ratio (Sharp), standard deviation (SD), downside deviation and VaR across the 10 000 simulated markets are shown in Table 5.3.

On average the EW, MV, IV and ERC portfolios performed very similarly according to the standard deviation, downside deviation and VaR metrics. Out of these four portfolios the ERC attained the highest sharp ratio and narrowly achieved the lowest scores across three risk measures. Despite performing the worst from a risk perspective, the MD portfolio attained the highest overall Sharp ratio. Thereby indicating that the MD managed to attain significantly higher average returns compared to the other portfolio's. On the other hand, despite performing well from a risk perspective, the MV portfolio attained a Sharp ratio significantly lower than the other portfolios

Table 5.3: Market 2 Risk Metrics

Metric	EW	MV	IV	ERC	MD
Sharp	0.05171	0.03482	0.04955	0.05312	0.06388
SD	0.01451	0.01434	0.01436	0.01431	0.0156
Downside Deviation	0.00983	0.00985	0.00974	0.00968	0.01047
VaR	-0.02273	-0.02273	-0.02252	-0.02238	-0.02424

5.1.3. Market 3

The portfolios compared here were estimated on the markets simulated using the five clusters correlation matrix (Figure 4.1). The portfolios average Sharp ratio (Sharp), standard deviation (SD), downside deviation and VaR across the 10 000 simulated markets are shown in Table 5.4.

Despite there being significant and distinct risk clusters in the markets

Table 5.4: Market 3 Risk Metrics

Metric	EW	MV	IV	ERC	MD
Sharp	0.08001	0.04974	0.07655	0.07605	0.06638
SD	0.0094	0.01049	0.0093	0.00935	0.01065
Downside Deviation	0.00623	0.00712	0.00618	0.00622	0.00714
VaR	-0.01437	-0.01639	-0.01425	-0.01433	-0.01646

5.1.4. Market 4

Table 5.5: Market 4 Risk Metrics

Metric	EW	MV	IV	ERC	MD
Sharp	0.05409	0.03229	0.05195	0.0518	0.04758
SD	0.01395	0.01456	0.01378	0.01379	0.01486
Downside Deviation	0.00944	0.01002	0.00933	0.00934	0.0101
VaR	-0.02181	-0.02312	-0.02156	-0.02159	-0.02334

5.1.5. Market 5

Table 5.6: Market 5 Risk Metrics

Metric	EW	MV	IV	ERC	MD
Sharp	0.05025	0.0396	0.0478	0.04794	0.0442
SD	0.01508	0.0176	0.01504	0.01531	0.01746
Downside Deviation	0.01023	0.01204	0.01022	0.01041	0.0119
VaR	-0.02371	-0.02795	-0.02368	-0.02412	-0.02764

5.2. Comparing Portfolios Across Market Types

5.2.1. Naive

5.2.2. Minimum Variance

5.2.3. Inverse Volatility

5.2.4. Equal Risk Contribution

5.2.5. Maximum Diversification

5.3. Discussion

6. Conclusion

I hope you find this template useful. Remember, stackoverflow is your friend - use it to find answers to questions. Feel free to write me a mail if you have any questions regarding the use of this package. To cite this package, simply type citation("Texevier") in Rstudio to get the citation for Katzke (2017) (Note that united references in your bibtex file will not be included in References).

References

Ardia, D., Bolliger, G., Boudt, K. & Gagnon-Fleury, J.-P. 2017. The impact of covariance misspecification in risk-based portfolios. *Annals of Operations Research*. 254(1-2):1–16.

Bailey, D.H. & Lopez De Prado, M. 2012. Balanced baskets: A new approach to trading and hedging risks. *Journal of Investment Strategies (Risk Journals)*. 1(4).

Bates, D. & Maechler, M. 2019. *Matrix: Sparse and dense matrix classes and methods*. ed. [Online], Available: https://CRAN.R-project.org/package=Matrix.

Choueifaty, Y. & Coignard, Y. 2008. Toward maximum diversification. *The Journal of Portfolio Management*. 35(1):40–51.

Choueifaty, Y., Froidure, T. & Reynier, J. 2013. Properties of the most diversified portfolio. *Journal of investment strategies*. 2(2):49–70.

Clarke, R., De Silva, H. & Thorley, S. 2011. Minimum-variance portfolio composition. *The Journal of Portfolio Management*. 37(2):31–45.

De Carvalho, R.L., Lu, X. & Moulin, P. 2012a. Demystifying equity risk–based strategies: A simple alpha plus beta description. *The Journal of Portfolio Management*. 38(3):56–70.

De Carvalho, R.L., Lu, X. & Moulin, P. 2012b. Demystifying equity risk–based strategies: A simple alpha plus beta description. *The Journal of Portfolio Management*. 38(3):56–70.

DeMiguel, V., Garlappi, L. & Uppal, R. 2009. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The review of Financial studies*. 22(5):1915–1953.

De Prado, M.L. 2016. Building diversified portfolios that outperform out of sample. *The Journal of Portfolio Management*. 42(4):59–69.

Fama, E.F. & French, K.R. 1992. The cross-section of expected stock returns. *the Journal of Finance*. 47(2):427–465.

Glasserman, P. 2013. *Monte carlo methods in financial engineering*. ed. Vol. 53. Springer Science & Business Media.

Higham, N.J. 2002. Computing the nearest correlation matrix—a problem from finance. *IMA journal of Numerical Analysis*. 22(3):329–343.

Katzke, N.F. 2017. *Texevier: Package to create elsevier templates for rmarkdown*. ed. Stellenbosch, South Africa: Bureau for Economic Research.

Kroese, D.P., Brereton, T., Taimre, T. & Botev, Z.I. 2014. Why the monte carlo method is so important today. *Wiley Interdisciplinary Reviews: Computational Statistics*. 6(6):386–392.

Liu, C. & Rubin, D.B. 1995. ML estimation of the t distribution using em and its extensions, ecm and ecme. *Statistica Sinica*. 19–39.

Maillard, T., Roncalli. 2010. The properties of equally weighted risk contribution portfolios. *The Journal of Portfolio Management*. 36(4):60–70.

Markowitz, H. 1952. Portfolio selection. *The Journal of Finance*. 7(1):77–91.

Meucci, A. 2010. Managing diversification.

Palomar, D.P. & ZHOU, R. 2020. *fitHeavyTail: Mean and Covariance Matrix Estimation under Heavy Tails*. ed. [Online], Available: https://CRAN.R-project.org/package=fitHeavyTail.

Perold, A.F. 2004. The capital asset pricing model. *Journal of economic perspectives*. 18(3):3–24.

Peterson, B.G. & Carl, P. 2020. *PerformanceAnalytics: Econometric tools for performance and risk analysis*. ed. [Online], Available: https://CRAN.R-project.org/package=PerformanceAnalytics.

Potgieter, N. 2020. *MCmarket: Package to simplify the monte carlo simulation of financial markets*. ed. Stellenbosch, South Africa: Stellenbosch University.

Rhoades, S.A. 1993. The herfindahl-hirschman index. *Fed. Res. Bull.* 79:188.

Wang, P., Sullivan, R.N. & Ge, Y. 2012. Risk-based dynamic asset allocation withExtreme tails and correlations. *The Journal of Portfolio Management*. 38(4):26–42.

Zhou, R., Liu, J., Kumar, S. & Palomar, D.P. 2019. Robust factor analysis parameter estimation. In ed. Springer *International conference on computer aided systems theory*. 3–11.

Appendix

Table 6.1: Asset Means and Sd's

Asset	Mean	Sd	Asset...	Mean...	Sd...
Asset_1	-0.00011	0.03277	Asset_26	0.00247	0.02671
Asset_2	0.00023	0.02314	Asset_27	-0.00132	0.04888
Asset_3	-3e-05	0.02359	Asset_28	0.00085	0.02387
Asset_4	-0.0018	0.03654	Asset_29	0.00024	0.01613
Asset_5	-0.00036	0.02925	Asset_30	0.00086	0.01884
Asset_6	0.00052	0.02396	Asset_31	0.00118	0.03147
Asset_7	0.00058	0.02451	Asset_32	0.00053	0.02882
Asset_8	0.00099	0.02287	Asset_33	0.00166	0.02253
Asset_9	0.00057	0.02929	Asset_34	0.00179	0.03002
Asset_10	0.00049	0.01736	Asset_35	0.00235	0.02574
Asset_11	-0.00088	0.01907	Asset_36	0.00157	0.02307
Asset_12	5e-05	0.0301	Asset_37	0.0017	0.02566
Asset_13	-0.00095	0.03185	Asset_38	0.00157	0.02262
Asset_14	-1e-04	0.02111	Asset_39	0.00054	0.03116
Asset_15	0.00102	0.02555	Asset_40	0.00242	0.02494
Asset_16	0.00174	0.02582	Asset_41	0.00246	0.02186
Asset_17	0.00241	0.03639	Asset_42	0.00113	0.01951
Asset_18	0.00033	0.03967	Asset_43	0.00181	0.02099
Asset_19	0.00102	0.02636	Asset_44	-0.00045	0.02297
Asset_20	0.00173	0.02697	Asset_45	4e-05	0.02354
Asset_21	-0.00034	0.01405	Asset_46	0.00485	0.03553
Asset_22	0.0011	0.0233	Asset_47	0.00064	0.02275
Asset_23	0.00236	0.02663	Asset_48	1e-04	0.02713
Asset_24	-0.00055	0.03195	Asset_49	0.00071	0.02186
Asset_25	-0.00283	0.03227	Asset_50	-0.00203	0.03131

6.0.1. Hierarchical Risk Parity (HRP)

The maximum drawdown is calculated by first, calculating portfolio cumulative returns and the maximum cumulative return achieved. The maximum drawdown is then the maximum amount that the cumulative return dips below its maximum, it is measured as a percentage of the maximum cumulative return (Peterson & Carl, 2020).

Due to the multitude of robustness issues related to traditional portfolio optimisers, De Prado (2016) developed a new approach incorporating machine-learning methods and graph theory (**???**). De Prado (2016) argues that the "lack of hierarchical structure in a correlation matrix allows weights to vary freely in unintended ways" and that this contributes to the instability issues. His HRP algorithm requires only a singular co-variance matrix and can utilize the information within without the need for the positive definite property (De Prado, 2016). This procedure works in three stages:

De Prado (2016) carried out an in-sample simulation study comparing the respective allocations of the long-only minimum variance, IVP and HRP portfolios using a co-variance matrix using a condition number that is "not unfavourable" to the minimum variance portfolio. The simulated data consisted of 10000 observations across 10 variables. The following findings were made: The minimum variance portfolio concentrated 92.66% of funds in the top 5 holdings and assigned a zero weight to 3 assets. *Conversly*, HRP only assigned 62.5% of its funds to the top 5 holdings (De Prado, 2016). The minimum variance portfolio's objective function causes it to build highly concentrated portfolio's in favor of a small reduction in volatility; the HRP portfolio had only a slightly higher volatility (De Prado, 2016). This apparent diversification advantage achieved by the minimum variance portfolio is rather deceptive as the portfolio remains highly susceptible to idiosyncratic risk incidents within its top holdings (De Prado, 2016). This claim was further validated by the finding that HRP achieved significantly lower out of sample variance compared to the minimum variance portfolio.

Appendix A

Some appendix information here

Appendix B